Hedging, Familiarity and Portfolio Choice

Massimo Massa
INSEAD

Andrei Simonov
Stockholm School of Economics

We exploit the restrictions of intertemporal portfolio choice in the presence of nonfinancial income risk to test hedging using the information contained in the actual portfolio of the investor. We use a unique data set of Swedish investors with information broken down at the investor level and into various components of investor wealth, income, and demographic characteristics. Portfolio holdings are identified at the stock level. We show that investors do not hedge but invest in stocks closely related to their nonfinancial income. We explain this with familiarity, that is, the tendency to concentrate holdings in stocks to which the investor is geographically or professionally close or that he has held for a long period. We show that familiarity is not a behavioral bias, but is information driven. Familiarity-based investment allows investors to earn higher returns than they would have otherwise earned if they had hedged.

One of the main questions in finance is what determines portfolio choice. The theoretical literature proposes hedging as one important driving motive: investors hold risky financial assets in order to offset their non-financial income risk. This stands in stark contrast to anecdotal evidence suggesting that investors, far from selecting an optimal portfolio, pick individual stocks on the basis of heuristics and stock familiarity. The lack of good-quality data on stock holdings, broken down at the investor level, and the scarcity of information about investors' overall assets, wealth, and income have made it almost impossible to test the competing explanations. We bridge this gap by using portfolio data to assess the extent to which investors actively hedge non-financial income risk. Our contribution is along four dimensions.

First, we design and implement a first micro-based test of hedging. As we will explain in more detail in the following section, the standard tests

We thank for the helpful comments of J. Campbell, J. Cocco, B. Dumas, F. Gomes, C. Harvey (the Editor), J. Heaton, M. Lettau, A. Lynch, S. Ludvigson, P. Maenhout, T. Moskowitz, J. Shanken, P. Veronesi, L. Viceira, A. Vissing-Jørgensen, M. Weber, two anonymous referees, and the participants of the Summer Financial Markets Symposium at Gerzensee, 2003, WFA meeting and the NBER Asset Pricing Summer Institute. We are grateful to Sven-Ivan Sundqvist for numerous helpful discussions and for providing us with the data. Andrei Simonov acknowledges financial support from the Stockholm Institute for Financial Research and Jan Wallander och Tom Hedelius Stiftelse. We also thank Jean Cropper for editorial assistance. All the remaining errors are ours. Address correspondence to M. Massa, Finance Department, INSEAD, Boulevard de Constance, 77305 Fontainebleau Cedex, France, or email: massimo.massa@insead.edu.

doi:10.1093/rfs/hhj013 Advance Access publication January 19, 2006

of hedging—based on the relationship between the investment in risky assets and the correlation between investor nonfinancial income and the market portfolio—do not have power and provide indeterminate results if investors can choose among many risky assets and do not hold the market portfolio. Our test—based on *multiasset* intertemporal portfolio choice in the presence of nonfinancial income risk—does not face this problem as it directly exploits data on the actual portfolio composition of the investor. Information detailed at the stock level allows us to identify the degree to which an investor tilts his portfolio away from the market portfolio and to use this tilt to test for hedging. We show that investors, in general, do not deliberately hedge. Quite the contrary, they tilt their portfolios toward stocks that are more correlated with their nonfinancial income. We also report evidence of how demographic, professional, and wealth heterogeneity affect this tilt.

The second main contribution is our analysis of what induces investors to behave in a way opposite to the one posited by hedging. We document that this is directly related to "familiarity," that is, the tilt to invest in stocks that are geographically and professionally close to the investor or that have been held for a long period. We show that the effect of familiarity is strong enough to more than offset the hedging motive, inducing an overall portfolio choice skewed toward familiar stocks. Moreover, familiarity affects the overall risk-taking of the investors.

Third, we investigate the nature of familiarity and we show that familiarity-driven investment is a rational response to information constraints as opposed to a behavioral heuristic. We identify classes of investors differentially informed and subject to different types of "familiarity shocks" and trace how sensitive their portfolio choice is to familiarity. We show that the sensitivity changes with the degree of informativeness of the investors and with their exposure to familiarity shocks. Familiarity mostly affects the less informed investors. When an investor has been subject, in the immediate past, to a shock affecting the source of familiarity (change of profession, relocation, change of employment status), the sensitivity of the investor to familiarity decreases.

Finally, we document how different strategies—that is, hedging and familiarity-based investment—differ in terms of their profitability/costs. The portfolios of investors who hedge are characterized by lower returns than those of familiarity-based investors. This is consistent with an information-based nature of familiarity.

Our approach also has two additional merits. It is, to our knowledge, the first test of hedging that properly accounts for the lack of financial diversification. Indeed, it explicitly controls for the fact that investors hold a very undiversified portfolio. Moreover, it distinguishes hedging from other portfolio-choice motives—such as speculative investment ("myopic portfolio choice")—and controls for spurious cross-sectional correlation—such as the one between nonfinancial income and overall stock returns that naturally exists before any portfolio choice is made.

Our results are robust and improve on the existing literature thanks to the use of a new and unique data set that combines, for the first time, individual portfolio holding data with comprehensive information on *all* the components of nonfinancial household wealth. We are able to inspect the *individual components of the investor's overall portfolio and relate them to his nonfinancial income*.

The data set contains a representative sample of the Swedish population and has information on the wealth of the investors, broken down into their components (cash, equity holdings, mutual funds, real estate, loans, bonds, and other assets). We also have available the income, wealth, and the tax position of the investors as well as detailed information on their demographic and employment characteristics. The richness of our data allows us to overcome some of the main limitations that have affected the empirical studies of portfolio choice: the lack of information on individual portfolio composition, the use of aggregated data, the problem of inference based on survey data without a proper panel structure, and the lack of information about real estate.

We improve along four dimensions. We use *individual portfolio data* to quantify the extent to which investors hedge. We exploit information on demographic and employment characteristics of the investors as well as information on their level and sources of wealth and income to study *heterogeneity in investor behavior*. We use, for the first time, a panel in which the *same investors are traced over time*, in terms of both portfolio choices and income, wealth, demographic, and occupational characteristics. The *complete panel dimension* also allows us to control for past portfolio choices and return-related strategies such as trend-chasing, momentum, and so on. Finally, we use a very broad and comprehensive measure of total wealth that also properly accounts for *real estate*. This allows us to explicitly consider the correlation between real estate and other sources of income of the investor.

The remainder of the article is articulated as follows. In Section 1, we describe our approach and contribution in detail. In Sections 2 and 3, we describe the data sets we use, the construction of the variables and the main econometric issues. We report the main findings in Section 4. A brief conclusion follows.

1. A New Test of Hedging

There exists a vast theoretical literature that analyzes portfolio choice in the presence of nonfinancial income. Nonfinancial income may include labor income (Davis and Willen 2000; Haliassos and Michaelides 2002) and entrepreneurial income (Heaton and Lucas 2000; Polkovnichenko

2002). Models may rely on the standard intertemporal Merton (1971) framework (Telmer 1993; Viceira 2001) as well as include limited horizon (Dammon, Spatt, and Zhang 2001) and life-cycle considerations (Campbell *et al*., 2001; Gomes and Michaelides, 2002).

At the empirical level, however, a definitive assessment of the way investors react to non-financial risk has, until now, eluded the literature (Heaton and Lucas 2000; Vissing-Jørgensen 2002). This is due to the fact that the literature has not used individual portfolio data to test for hedging. All the empirical tests have been, rather, based on theoretical restrictions defined in a *single-asset framework where the market portfolio is the only asset investors can choose to hedge non-financial income risk*.

In fact, in the presence of *multiple risky assets*, the standard tests of hedging which relate holdings of risky assets to the nonfinancial risk of the investor (i.e., the variance of non-financial income and its correlation to financial income) are indeterminate and do not have any power. Indeed, an investor may hedge his nonfinancial income by increasing his holdings of the assets that are negatively related to his nonfinancial income, as well as by reducing his holdings of the assets that are positively related to his nonfinancial income.

For example, if an investor has available two stocks, one positively correlated to his nonfinancial income and one negatively correlated to it, he may either buy the stock that is negatively correlated with his nonfinancial income or sell the one that is positively related to it. Therefore, the decision to hedge may actually involve either an *increase* in risky assets (that are negatively correlated with nonfinancial income) or a *reduction* in risky assets (that are positively correlated with nonfinancial income). This implies that a negative correlation between the share of the portfolio invested in risky assets and the correlation between financial and nonfinancial income is consistent with *both* investors buying risky assets to hedge and investors selling risky assets to enact a familiarity-based strategy. Moreover, the standard tests counterfactually assume that the investor holds a well-diversified portfolio, the market.

One way out is to redefine the tests by directly exploiting the information contained in the actual portfolio of the investor. A direct inspection of the way the investor tilts his portfolio in order to hedge allows us to formulate two hypotheses. Let us consider them separately.

Hypothesis 1. *The direction of the tilt in the portfolio risk profile.*

The risk profile of the financial portfolio should be tilted toward assets with a negative correlation with the nonfinancial income of the investor and away from assets with a positive correlation.

The intuition is that holding financial assets allows the investor to hedge only if such assets are negatively related to his nonfinancial risk. Ideally, if the investor wants to hedge, he will choose stocks that are negatively related to his nonfinancial risk. This will induce a negative correlation between the investor's financial and nonfinancial income.

However, a possible criticism of quantifying hedging on the basis of the actual correlation between financial and nonfinancial income is that this measure may be affected by the preexisting correlation between stocks and an investor's nonfinancial income. For example, it is possible that the investor's income is negatively related to the average stock available on the market. In this case, a negative correlation between the nonfinancial income of the investor and his portfolio would not be evidence of deliberate hedging if such a correlation *is less negative than that between the investor's nonfinancial income and the market portfolio*. In other words, the investor would actually be increasing the exposure to his nonfinancial risk by deliberately holding stocks more related to his nonfinancial income.

We therefore need a measure that proxies for hedging or the extent to which investors actively pursue a negative correlation between financial and nonfinancial income that differs from the one embedded in the correlation between the investor's nonfinancial income and the market. We will call this measure "index of hedging." It quantifies the extent to which the investor's portfolio differs from the market portfolio in terms of correlation (covariance) with his nonfinancial risk. In the following we will consider indexes of hedging built using differences in correlations and indexes of hedging built using differences in covariances.[1] We consider two alternative indexes of hedging:

$$\Gamma = \mathrm{Corr}_{y,m} - \mathrm{Corr}_{y,p} \text{ and } \Delta = \mathrm{Cov}_{y,m} - \mathrm{Cov}_{y,p}. \tag{1}$$

These measures track, at the investor level, the difference between two correlations (covariances). The first is the correlation (covariance) between his nonfinancial income and the overall *market* portfolio ($\mathrm{Corr}_{y,m}$ or $\mathrm{Cov}_{y,m}$). The second is the correlation (covariance) between his nonfinancial income and his financial portfolio ($\mathrm{Corr}_{y,p}$ or $\mathrm{Cov}_{y,p}$). The correlation (covariance) between the investor's nonfinancial income and the market portfolio represents the extent to which holding the market portfolio would help the investor diversify away his nonfinancial risk. It is a benchmark that can be used to assess the actual strategy of the investor. $\Gamma(\Delta)$ captures the contribution of the portfolio choice to the

[1] The advantage of using the correlation is that it is measure free, that is, it is a standardized variable that is not affected by the size of the investment. This is particularly useful if we want to assess the impact of hedging on portfolio choices over time and across investors. On the other hand, the standard intertemporal portfolio model with multiple assets provides sharper restrictions when these are cast in terms of covariances (i.e., restriction H2). In the Appendix, we will provide a more elaborate discussion of this point.

reduction of the investor's overall risk. It is positive in the case of hedging, so that restriction H1 requires $\Gamma > 0$ ($\Delta > 0$). We will construct indexes of hedging for both sources of nonfinancial risk: labor and entrepreneurial risk as well as for total income.

Hypothesis 2A. *The determinants of the tilt in the portfolio risk profile.*

The tilt in the risk profile of the financial portfolio should be positively related to the variance of the nonfinancial income and to the covariance between the different sources of nonfinancial income.

If we assume that the investor has two sources of nonfinancial income (Y_z and Y_x, or labor and entrepreneurial income), and a level of wealth W, this restriction can be expressed as

$$\Delta_z = \frac{Y_z}{W}\mathrm{Var}_{Y_Z} + \frac{Y_x}{W}\mathrm{Cov}_{Y_Z,Y_x} + \frac{1}{W}\Theta_z, \tag{2}$$

where $\Theta_z = -(Y_z + Y_x)\sum_{j=1}^{n}\Omega_{S_j} \times \mathrm{Cov}_{S_j,Y_z}$, $\Omega_{S_j} = \frac{\left(\mu_{s_j}-r\right)'}{(1-\gamma)\sigma^2_{s_j}} \cdot S_j, \mu_{s_j}$ and $\sigma^2_{s_j}$ are, respectively, the price, the mean, and the variance of the return of the *jth* risky asset. In the Appendix, we report a detailed derivation of Equation 2. The intuition is the following. The investor will actively hedge more (i.e., will increase Δ_z or actively tilt his portfolio toward assets negatively correlated to his nonfinancial income), the higher the risk of the nonfinancial income (Var_{Y_z}) and the higher the covariance between his nonfinancial sources of income $\left(\mathrm{Cov}_{Y_z,Y_x}\right)$. Indeed, if the other nonfinancial sources of income are negatively related to one another, they should already provide a hedge. This should reduce the demand for financial hedging. In other words, the second term captures how much one can hedge one source of nonfinancial income with a different source of nonfinancial income.

Hedging is also related to the assets' mean/variance ratios. An asset that is positively correlated to the nonfinancial income of the investor $\left(\mathrm{Cov}_{S_j,Y_z}\right)$, but has a high expected mean−variance ratio $\left(\Omega_{S_j} = \frac{\text{mean}-\text{variance ratio}}{1-\gamma}\right)$, will reduce hedging. Indeed, in this case, hedging would be expensive as it requires forgoing the gains implied by the high mean/variance ratio. We can recast the testable restriction H2A as

$$\Delta_z = \beta\frac{Y_z}{W}\mathrm{Var}_{Y_z} + \gamma\frac{Y_x}{W}\mathrm{Cov}_{Y_z,Y_x} + \zeta\frac{1}{W}\Theta_z + \boldsymbol{\delta}\mathbf{F}, \tag{3}$$

where $\mathbf{F}$ is a vector of control variables. Hedging requires that $\beta > 0$, $\gamma > 0$ and $\zeta > 0$.

The alternative hypothesis posits that the investor buys more of the stocks that covary with his nonfinancial sources of income. We will rationalize this alternative behavior in terms of "familiarity." What is familiarity? Huberman (2001) argues that there is a "general tendency of people to have concentrated portfolios,...to hold their own company's stock in their retirement accounts...invest in stocks of their home country. Together, these phenomena provide compelling evidence that people invest in the familiar while often ignoring the principles of portfolio theory." In general, familiarity is defined in terms of professional or geographical proximity to the stocks. For example, investors may choose the stocks of the company for which they work because familiarity induces them to optimistically extrapolate past returns (Benartzi 2001). Also, investors may invest in stocks of companies headquartered close to where they live (Coval and Moskowitz 1999; Hau 2001; Huberman, 2001) or in stocks with high brand recognition (Frieder and Subrahmanyam 2002). Can the tilt toward stocks that covary with the investor's nonfinancial income be explained in terms of a tilt toward "familiar" stocks? This question can be addressed by testing whether the deviation from the optimal portfolio in a direction opposite to that of hedging is related to the decision to invest in familiar stocks.

Hypothesis 2B. *The role of familiarity.*

If familiarity affects investor portfolio choice, we expect the index of familiarity to be negatively related to the portfolio tilt.

This can be expressed by expanding Equation 3:

$$\Delta_z = \beta \frac{Y_z}{W} \mathrm{Var}_{Y_z} + \gamma \frac{Y_x}{W} \mathrm{Cov}_{Y_z, Y_x} + \zeta \frac{1}{W} \Theta_z + \boldsymbol{\delta} \mathbf{F} + \boldsymbol{\nu} \boldsymbol{\Psi}, \tag{4}$$

where Ψ is the index of familiarity as defined later. If it induces investors to reduce hedging, we expect $\nu < 0$. Therefore, the testing of restrictions H2A and H2B can be reduced to testing Equation 4.

1.1 The nature of familiarity

Familiarity may be due either to some behavioral heuristic or to better information on the particular stock.[2] Behavioral theories relate the familiarity bias to the findings in psychology that show that human beings use heuristic simplifications in their decision-making process. One of these heuristics is the saliency or availability bias. This is the tendency to focus heavily on information that is salient or is often mentioned rather than

[2] A recent paper (DeMarzo, Kaniel and Kremer 2004) provides an alternative modelization that rationally explains investors' bend for familiarity. In this context, familiarity arises out of investors' desire to hedge "relative consumption" *vis-a-vis* the neighbors.

information that is blended in the background. We will define this hypothesis, entirely grounded on behavioral heuristics, as “pure familiarity.”

The alternative approach is “information-based familiarity.” This states that “investors buy and hold only those securities about which they have enough information” (Merton 1987). That is, investors are either not aware of all the stocks or do not know them well enough to be willing to invest in them. Information about a stock affects the investment decision by altering the perceived expected payoff in a rational portfolio decision.

What differentiates the two hypotheses is the relationship to information. In the case of pure familiarity, investors erroneously rely on what is often mentioned or is closer to them (i.e., geographically or professionally) because this seems more salient and relevant. In the case of information-based familiarity, on the other hand, familiarity is a way of coping with limited information. For example, a behavioral story postulates that geographical proximity is relevant as it makes more salient/available the characteristics of the stocks that are located closer, whereas a limited information story posits that investors are more likely to invest in stocks located near them simply because geographical proximity provides a cheap way of acquiring information.

This has two implications. First, in the case of pure familiarity, access to more information should not alter the bias. In the case of information-based familiarity, however, more information would reduce the bias and induce investors to rely less on “cheap” familiarity-based information. We can exploit this feature in order to determine the nature of familiarity. If familiarity is information-based, it should differ across investors with different degrees of informativeness.

Second, the impact of familiarity on investor’s behavior should change when there is a change in the investor’s proximity to the stock. For example, let us consider an investor living close to a Volvo factory who moves to a location far from Volvo and close to an Ericsson subsidiary. If familiarity (i.e., geographical proximity) is a behavioral bias, the investor should switch from Volvo to Ericsson, but *his overall tendency to invest in closer stocks should not change*. If, however, familiarity is information-based, the mere process of moving should affect the amount of information the investor derives from his geographical proximity to such stocks. Indeed, it will take time for the investor to adjust to the new source of information on Ericsson that his new geographical location entails and, in the meantime, the sensitivity to familiarity should drop.

Therefore, if proximity to a stock is a cheap way of acquiring information, the impact of familiarity should be different and lower if the investor has recently been subject to a shock that has affected his source of familiarity. On the contrary, if familiarity is a behavioral heuristic, prior changes in the proximity to the stock should not matter. We define the events that change the investor's proximity to the stock as "familiarity shocks."

We therefore have available two sets of restrictions. The first restriction links the impact of familiarity on investor behavior to differences in the degree of information of the investors. The second restriction links the impact of familiarity on investor behavior to the exposure to familiarity shocks.

Hypothesis 3. *The nature of familiarity.*

If familiarity (i.e., Ψ) proxies for information, the sensitivity of the investor to familiarity should differ with his degree of informativeness. That is, referring to Equation 4, we have

$$H_0 : \left|\nu_{\text{high info}}\right| = \left|\nu_{\text{low info}}\right|; \quad H_a : \left|\nu_{\text{high info}}\right| \neq \left|\nu_{\text{low info}}\right|. \tag{5}$$

If familiarity proxies for information, the sensitivity of the investor to familiarity should differ with his exposure to familiarity shocks. That is

$$H_0 : \left|\nu_{\text{fam. shocked}}\right| = \left|\nu_{\text{non fam. shocked}}\right|; H_a : \left|\nu_{\text{fam. shocked}}\right| \neq \left|\nu_{\text{non fam. shocked}}\right|. \tag{6}$$

The subscripts "high info" and "low info" identify high-informed and low-informed investors while the subscripts "fam. shocked" and "nonfam. shocked" identify investors who have experienced a familiarity shock. The null of no change in the case of pure familiarity is tested against the alternative of information-based familiarity. If familiarity is information-based, it should differ across investors with different degrees of informativeness (restriction 5) and should change for investors subject to familiarity shocks (restriction 6).

In order to identify the degree of informativeness of the investors, we use their wealth. This variable is strongly related to information and, presumably, independent of their behavioral heuristics. Rational theories have a role for wealth. Higher wealth may relax informational constraints and make it easier to purchase more information. If we assume a standard information technology, a wealthier investor would be willing to spend more to purchase private information on a particular stock than a less wealthy investor, because the relative cost of investing in information decreases with the level of wealth (Peress, 2004). Therefore, a wealthier investor would be less dependent on "cheap" publicly available information, being able to rely on his "private" source. On the contrary, behavioral theories are mute about the role of wealth. In other words, investors are, in general, assumed to suffer from biases, regardless of their level of wealth. Indeed, if we are really dealing with human biases, saliency and behavioral heuristics should equally affect wealthy and non-wealthy investors. The main point of this classification is that investors who are “high wealth” (i.e., have a lot of accumulated wealth in liquid-assets stocks and illiquid assets-real estate) should be more informed to take financial decisions (trade in financial assets). Indeed, they not only have more resources available to pay analysts, but they also rely relatively more on the income that accrues from this wealth—that is, financial income—and should therefore care more about it.

We define as “high wealth investors” all the investors who, *in the previous year*, paid wealth tax. We define as “low wealth investors” all the others. The high wealth investors represent approximately 10% of the overall sample.

A way of assessing the quality of our identification is to look at the profits made by the different classes of investors. In the last section, we will show that high wealth investors consistently make more profits than the low wealth ones. This confirms our identification, as there is a direct mapping between “informed investors” and profits generated on financial assets. That is, the investors we consider as more informed (i.e., high wealth investors) are also those who make more profits. We refer the reader to the last section for a detailed analysis.

This classification has some limitations. Given that the wealth tax is paid on the level of wealth and is not directly related to what we define as nonfinancial income (not including financial and real estate income), it is possible that some of the investors not paying the wealth tax have a nonfinancial income higher than that of some of the investors paying the wealth tax. For example, by looking at the standard deviation columns of Table 1, panel D, and comparing their magnitudes to the ones of the means and medians, it is evident that there are quite some individuals who are defined as low wealth individuals who have a higher labor income than many high wealth individuals. This is especially true for entrepreneurial income. In this case, the individual with highest entrepreneurial income is a low wealth investor. This is due to the fact that entrepreneurial income is a “low wealth source of income” as it is mostly the income from being self-employed (e.g., a small shop owner or carpenter) as well as from a professional activity (e.g., a dentist) that does not require a high stock of wealth/capital.

Finally, we define as familiarity shocks events that change the investor’s proximity to the stocks. Unemployment shocks—that is, having being unemployed in previous years—and professional change shocks—that is, having changed profession in the previous years—are natural candidates as they represent a structural break in the professional life of the investor. Given that our familiarity variables are related to the geographical location of the investor, another potential candidate is the change in the location of the investor. This proxies for “relocation shocks” and is based on the investor having moved at least once in the previous three years. One caveat applies. Holding onto to a stock after moving might simply be an artifact of the investor rebalancing very infrequently.

Table 1
Descriptive statistics of the sample

Variable	Mean	Median	Standard deviation	IQR
Panel A: descriptive statistics of households				
Number of households	292,901	291,913	647	686
Full sample				
Financial wealth	2.98	4.26	2.35	4.96
Real estate wealth	2.12	5.04	3.76	7.62
Capital gains/losses	0.00	−0.02	1.00	0.00
Secondary education	0.43	0.00	0.46	1.00
Higher education	0.31	0.00	0.50	1.00
Ability	0.05	0.00	0.58	0.31
Size of household	2.67	2.00	1.51	3.00
Immigration status	0.14	0.00	0.35	0.00
Age	49	47	17	24
Unemployment risk	0.14	0.15	0.10	0.15
Stockholm dummy	0.20	0.00	0.40	0.00
Sample of stock market participants				
RetPortfolio	0.56	0.23	1.52	0.85
Financial wealth	4.51	4.95	1.47	0.94
Real estate wealth	3.87	5.58	3.24	1.05
Capital gains/losses	0.07	−0.02	2.62	0.02
Secondary education	0.35	0.00	0.44	1.00
Higher education	0.47	0.00	0.48	1.00
Ability	0.08	1.00	0.44	0.29
Size of household	2.78	2.00	1.35	2.00
Immigration status	0.11	0.00	0.32	0.00
Age	54	53	15	22
Unemployment risk	0.12	0.13	0.10	0.17
Stockholm dummy	0.22	0.00	0.42	0.00

Table 1
(continued)

Age	Males (%)	Females (%)	Age of oldest household member (%)
Panel B: age and gender distribution of the sample			
0–19	18.2	17.2	0.5
20–29	4.8	4.9	10.7
30–39	7.1	8.2	21.7
40–49	7.4	7.4	23.6
50–59	5.9	5.3	17.9
60+	6.6	7.2	25.8
Total	49.9	50.2	100

Variable	Representation in the sample (%)	Mean	Median	Standard deviation	IQR
Panel C: wealth and income characteristics of households					
Wealth-tax payers	7.9	360	103	2649	353
Labor income earners	100.0	321	288	238	276
Entrepreneurial income earners	9.8	88	43	173	112
Real estate holders	54.6	449	387	349	340

	Low wealth			High wealth			*t*-test		Median-test	
	Mean	Median	Standard deviation	Mean	Median	Standard deviation	*t*-stat	*p* value	*KS*	*p* value
Panel D: income characteristics for different classes of households										
Labor income	351	323	204	507	439	461	91.67	0.00001	65.02	0.00001
Entrepreneurial income	8	0	50	20.4	0	132	33.54	0.00001	32.40	0.00001
Financial capital gains	2	0	11	16	0	76	66.27	0.00001	46.42	0.00001
Financial capital losses	0.1	0	1.9	1.1	0	16	22.31	0.00001	16.47	0.00001
Net financial gains/losses	2	0	11	15	0	74	63.57	0.00001	44.91	0.00001
Nonfinancial capital gains	8	0	805	121	0	2145	18.45	0.00001	120.00	0.00001

Nonfinancial capital losses	1	0	25	2	0	64	4.99	0.00001	15.08	0.00001
Net financial gains/losses	7	0	805	119	0	2146	18.29	0.00001	118.59	0.00001

Variable	Mean	Median	Standard deviation	IQR
Panel E: conditional moments of income characteristics of households				
Total sample				
Conditional mean of labor income	420	332	503	352
Conditional standard deviation of labor income	461	60	2343	156
Conditional mean of entrepreneurial income	14	0	127	0
Conditional standard deviation of entrepreneurial income	363	0	3391	0
Conditional mean of total income	434	343	522	357
Conditional standard deviation of total income	523	60	3096	160
Sample of nonwealthy households				
Conditional mean of labor income	383	317	348	329
Conditional standard deviation of labor income	405	57	2203	145

Table 1
(continued)

Variable	Mean	Median	Standard deviation	IQR
Conditional mean of entrepreneurial income	12	0	95	0
Conditional standard deviation of entrepreneurial income	203	0	1670	0
Conditional mean of total income	395	322	359	343
Conditional standard deviation of total income	498	58	3218	140
Sample of wealthy households				
Conditional mean of labor income	677	462	1037	517
Conditional standard deviation of labor income	731	91	2934	224
Conditional mean of entrepreneurial income	41	0	331	0
Conditional standard deviation of entrepreneurial income	1170	0	4382	0

Conditional mean of total income	718	488	1048	448
Conditional standard deviation of total income	1001	114	4950	288

Panel A: we report the number of households for each year and the descriptive statistics of household characteristics, for the full sample and for the subsample of stock market participants. We report the logarithm of financial and real estate wealth (in SEK), and the *normalized* capital gains and losses (standarized by subtracting the mean and dividing by the standard deviation of capital gains/losses in overall sample), dummies for secondary education (defined as 1 if the person' highest level of education is either completed or uncompleted high school and 0 otherwise), dummies for high education (defined as 1 if the person has some post-high school education and 0 otherwise), an immigration dummy (it takes the value 0 if all the members of the household are native Swedes and 1 if at least one member of household immigrated), a Stockholm dummy (defined as 1 if the household resides in the capital and 0 otherwise), the size of household and the age of its oldest member. We also report variables to account for the professional ability and unemployment risk of the investor. The first variable is based on the difference between his income and the average income of his profession (normalized by the standard deviation of the income in the profession). The second variable is the one-year-ahead forecast of a linear probability model where the unemployment status (i.e., 1 if unemployed and 0 otherwise) is regressed on demographic variables, measures of income and wealth and regional, geographical and professional dummies. For stock market participants we also report the yearly portfolio return. Panel B: we present the age and gender distribution of the sample. The second and third columns (males and females) are defined at the individual level while column 4 is defined at the household level. Panel C: the column "Representation in the sample" reports the fraction of households in the sample who pay wealth tax, earn labor or entrepreneurial income or hold real estate wealth. The other columns report the statistics (mean, median, standard deviation, and IQR) of, respectively, the value of wealth, labor and entrepreneurial income (gross yearly income), and real estate. The values are defined in thousands of Swedish Krones (SEK). Panel D: we report the statistics on the different sources of income for low and high wealth investors. In particular, we consider the labor income, the entrepreneurial income and the income due to capital gains and losses (not normalized), both financial and nonfinancial. The values are defined in thousands of SEK. We report, for the different classes of investors, the mean, the median, and the standard deviation as well as *t-test* and *Kologorov–Smirnov test* of differences between the two groups. Panel E: conditional moments of labor and entrepreneurial income estimated using Carroll and Samwick (1996) methodology described in Section 3.1. The values are defined in thousands of SEK.

2. Data Description

We use Swedish data. Sweden provides a very good experiment as, contrary to common belief, it has a one of the most flexible labor markets in Europe. For example, the termination notice is the shortest among all the European countries (including the UK). Moreover, unemployment benefits are capped at relatively low level and phased out over time and terminated after 6 months.[3] This makes nonfinancial income risk-hedging more relevant. Data have been collected from different sources.

For each investor, we have detailed information on his individual holdings of stocks (broken down at the stock level), the holdings of mutual funds,[4] bank accounts, real estate, and other types of wealth. These data have been collected by Värdepapperscentralen (VPC), the Security Register Center. The data contain stockholding held directly and on the street name, including holdings of US-listed ADRs. In addition, SIS Ägarservice AB collects information on ultimate owners of shares held through trusts, foreign holding companies, and the like [for details, see Sundin and Sundquist (2002)]. Our data cover the period 1995–2000. Overall, the records provide information about the owners of 98% of the market capitalization of publicly traded Swedish companies.

For each investor, we also have available information on the different sources of income of the investor provided by the fiscal authorities, as well as his demographic and family characteristics. These data come from Longitudinal INdividual DAtaset for Sweden (LINDA). LINDA is a register-based longitudinal data set and is a joint endeavor between the Department of Economics at Uppsala University, The National Social Insurance Board (RFV), Statistics Sweden, and the Ministries of Finance and Labor. LINDA consists of a large panel of individuals and their household members that is representative of the population during the period 1966–2000. For each year, information on all family members of the sampled individuals are added to the data set. Apart from being a panel which is representative of the population in general, the sampling procedure ensures that the data are representative for each year. Moreover, *the same family* is traced over time. This provides a real time series dimension, in general missing in surveys based on different cohorts polled over time.

The variables available include individual background characteristics (sex, age, marital status, country of birth, citizenship, year of immigration, place of residence detailed at the parish level, education, profession,

[3] Also, government tight control and monitoring of unemployed prevents the abuse of the system (Ljungqvist and Sargent 1995).

[4] We have the aggregate value of the money invested in mutual funds. For the purpose of this study, we consider mutual funds as risky assets analogous to stocks and we proxy their return with the market index. Our results are robust to the way we treat the mutual funds. That is, re-estimating our specifications excluding the mutual funds from the set of risky assets we get results consistent with those that are reported.

employment status), housing information (type and size of housing, owner, rental and occupation status, one-family or several-family dwelling, year of construction, housing taxation value) and tax and wealth information. The income and wealth-tax registers include informations on labor income, capital gains and losses, business income and losses, pension contributions, taxes paid, and taxable wealth. A detailed description of the data set is provided by Edin and Fredriksson (2000) and is available on the web site http://linda.nek.uu.se/. We do not have information on the implicit claims on retirement benefits through state-provided pensions. However, it is worth mentioning that the level of these benefits (just like in most European countries) is directly related to the salary level. Therefore, by including the level of the nonfinancial income (wage, salary, and so on), we are implicitly and partially controlling for them.

The tax part deserves more detailed discussion. In Sweden, in addition to usual income taxation, there exists a further wealth tax which is paid by every investor with a net worth in excess of 900,000 SEK (about US$90,000). The taxable wealth includes tax-accessed value of real estate, market value of publicly listed securities, balance of bank accounts, and fair value of valuable possessions (including jewelry, cars, antiques, and so on). For the purpose of this article, we compute the current market value of housing using the tax-accessed value provided by LINDA. We evaluate it at current prices by using the average ratio of market value to tax-accessed value that is provided for each year and country by Statistics Sweden. For the privately held unlimited liability companies, the value of the assets is included in the household's tax return. There is no estimate of market value of privately held limited liability companies that are not listed. However, the data contain an indicator variable for owners of privately held companies. The size of the group is rather small (1.74–1.91% of the sample, according to the year) and is unlikely to affect our estimates in a significant way. Moreover, for the members of the wealthiest 5000 families, we have been able to reconstruct their values and to correctly impute it by using information from SIS Ägarservice AB (Sundin and Sundquist 2002).

This information has been matched *at the individual level* by Statistics Sweden with the data on stockholding provided by SIS Ägarservice AB we mentioned before. This has generated a time series of investment and income for each investor. The combined LINDA/shareholding data set covers the period 1995–2000. The overall sample contains 1,807,602 observations. However, only 1,757,406 observations are used.[5] In addition, we also use 1990–1994 data from LINDA in the implementation of the Carroll and Samwick (1996) procedure to construct the moments of

[5] We excluded observations for households that were in the sample for less than three years and households with the oldest member being younger than 18 years. Also, it is worth noting that we define a shareholder as anyone who has more than SEK 2,000 worth of stock (that is US$200). This is the definition used by Statistics Sweden. Finally, we trimmed the outliers.

conditional nonfinancial income. In Table 1, we report some descriptive statistics. Panel A contains the general demographic characteristics (number of households, members in household, age of the oldest member of household, percentage of the sample with secondary and higher education, etc.). Panels B and C report, respectively, the age and gender distribution of the sample and their wealth and income characteristics defined in terms of wealth, real estate, labor and entrepreneurial income. The statistics are defined at the household level except in Panel B in which the second and third columns (males and females) are defined at the individual level. Panel D reports the statistics on the different sources of income for low- and high wealth investors. In particular, we consider the income from real estate, the labor income, the entrepreneurial income, and the income due to capital gains and losses, both financial and nonfinancial. We report, for the different classes of investors, the mean, the median, and the standard deviation as well as the (mean and median) tests of differences between the two groups.

One point that is worth stressing is the fact that we use data from the stock market bubble period (1995–2000). This might affect the results on hedging as investors may be overall less cautious in their investment strategies, jumping on the bandwagon of global euphoria of the period. This would induce them to hedge less and allocate their investment toward the risky assets where they think they will have better information (i.e., "closer stocks"). However, the availability of a proper panel structure helps us along this dimension in many different ways. First, the data on nonfinancial income span an entire decade (1990–2000) that covers a recession (1990–1994), a recovery (1994–1996), a boom (1997–1999), and the burst of the bubble (2000). This implies that our measures of permanent nonfinancial income, its volatility, and its correlation to financial income are scarcely affected by the bubble.

Second, even if financial data (investor holdings and portfolios) are determined during a bubble period, two caveats apply. First, the bubble in Sweden did not affect all the stocks in the same way, and we are dealing with disaggregated data and information detailed at the stock level. Second, the availability of stock-level information allows us to construct measures of past performance and volatility of the investor portfolio ("momentum/stock performance variables"). These measures are meant to capture the shift in an investor's portfolio because of stock-market changes or to stock tracking, momentum or performance-chasing activity of the investor. They should, as least partially, control for trend-chasing, momentum-investing, and short-term strategies induced by the bubble. The use of these measures makes our results more robust than the equivalent ones based on US data, for which the sample is shorter (in general one or, at best, two nonconsecutive years) and with no panel structure.

For each stock, we have detailed information on the company and the price, volume and volatility at which it trades. In order to derive information on individual security returns and to track the overall market index [stock market index (SIX index], we use the SIX Trust Database. We also use the set of Swedish residential real estate indices computed at the county level. These are based on the resale value of properties (Englund, Quigley, and Redfearn 1998). The consumer confidence index is provided by Statistics Sweden. Geographical coordinates are supplied by Swedish Postal Service and contain latitude and longitude of Swedish Postal Offices (on a three-digit level). We aggregate the information of the individuals at the household level and perform the analysis at the household level.

3. Construction of Variables, Identification, and Econometric Issues

3.1 Income-related variables

Following the standard approach, we specify investors' portfolio policies in terms of their permanent income, that is, the conditional moments of the long-term income as in Heaton and Lucas (2000). The other results are available upon request. In order to construct proxies for permanent nonfinancial income, its variance and its correlation and covariance to the return on the market portfolio and to the return on the portfolio of the investor ($\mathrm{Corr}_{y,m}$, $\mathrm{Corr}_{y,p}$, $\mathrm{Cov}_{y,m}$ and $\mathrm{Cov}_{y,p}$), we use the approach of Carroll and Samwick (1996) and Vissing-Jørgensen (2002). We refer to their papers for a detailed exposition.[6] Following Vissing-Jørgensen (2002), given the potential inaccuracy of estimates based on few observations, we calculate the correlation over the entire sample. We report descriptive statistics of the conditional moments of income characteristics in Table 1, Panel D.

Also, as a robustness check, following Campbell *et al.* (2001), we consider the correlation between nonfinancial income and stock returns lagged one year. In line with the literature, the correlation increases. However, the results of the estimates of the main specifications do not differ from those reported. We consider both total nonfinancial income (i.e., the sum of labor and entrepreneurial income) and labor and entrepreneurial income taken separately.

3.2 Indexes of familiarity

We need a measure that captures the extent to which an investor tilts his portfolio toward assets with which he is more familiar. We will call this

[6] The set of instruments used to estimate conditional income according to this methodology are demographic variables (secondary education, higher education, age, age squared, marital status, size of the household, number of adults belonging to the household), changes in the demographic variables, industry dummies for the company the investor is working for (e.g., oil industry), dummies for the type of profession of the investor (e.g., doctor), immigration status.

"index of familiarity." We consider three indexes of familiarity. The first is related to "professional proximity." It is a dummy taking the value 1 if the investor's profession is in the same area of activity as the company whose stock is under consideration and zero otherwise. We use the one-digit SNI92 codes (similar to SIC codes) to identify the areas of activities. For example, in the case of an investor working in the mining sector holding a stock of a mining company, the dummy would be equal to 1.

The second measure is related to "geographical proximity," that is, the proximity between the residence of the investor and the place where the company is located. We use two different measures: the first one is the logarithm (to base 10) of the inverse of the distance between the ZIP code of the investor and the ZIP code of the closest branch/subsidiary of the company whose stock we consider. As an alternative measure, we use the logarithm of the inverse of the distance between the ZIP code of the investor and the ZIP code of the company headquarters. Given that the results do not differ and the variables are highly collinear, we report only the first specification. These measures are analogous to the one put forward by Coval and Moskowitz (1999) in the study of geographical preferences in mutual fund investment. The greater the value of the variable, the closer is the investor to the stock.

Finally, we may argue that investors are more likely to be informed about the stocks they already own than about stocks that are not yet part of their portfolio. Indeed, once the stock is in the portfolio, investors follow it more closely, reading the reports, paying attention to the earning announcements, and actively purchasing information about it. In other words, stockholding may proxy for selective attention and active purchase of (private) information. We therefore construct a variable that proxies for "holding period" based on the time a stock entered the investor's portfolio.

These measures are constructed at the stock level. They are then aggregated across all the stocks of the investor and weighted by their share in the portfolio. This procedure delivers three measures of familiarity for each investor and time t.

3.3 Control variables

We consider five types of control variables: measures of income and wealth, demographic variables, professional ability and risk, momentum/stock performance variables, and residual control variables.

The *measures of income and wealth* contain the vector of the wealth of each investor at time t, broken down into its individual components (i.e., financial, real estate, and other), as well as measures of income (i.e., labor and entrepreneurial) and overall (i.e., financial and non-financial) capital gains, and losses of the each investor at time t. We also include the correlation between nonfinancial income (both labor income and entrepreneurial income) and real estate.

The *demographic variables* include the profession of the investor, his level of education, broken down into high-school and university level, the age of the oldest member of the family of the investor, and its value squared. This latter variable is consistent with standard results (Guiso and Jappelli 2002; Vissing-Jørgensen 2002), which find a nonlinear relationship between age and the degree of stock-market participation.

We also construct variables to account for the *professional ability and risk* of the investor. A first variable proxies for the ability of the investor in his occupation. This is based on the difference between his income and the average income of his profession. The assumption is that the higher the income of the investor relative to the average income of the other investors in the same area, the higher his ability should be. A second variable is a measure of *unemployment risk* that proxies for the probability of being unemployed in the following year. It is the one-year-ahead forecast of a linear probability model where the unemployment status (i.e., 1 if unemployed and zero otherwise) is regressed on demographic variables, measures of income and wealth, and regional, geographical and professional dummies.

The *momentum/stock performance variables* are meant to capture the shift in the investor's portfolio because of the stock-market changes or to stock tracking, momentum or performance-chasing activity of the investor. They are the return and volatility on the investor's portfolio in the previous twelve months.

The *residual control variables* include standardized levels of debt for the investor (ratio of investor debt to total income and ratio of investor debt to total wealth), the return and volatility on the market portfolio in the previous twelve months, an Index of Consumer Confidence, and a set of dummies that account for the regional location of the investor as well as the industry in which he works. The debt ratios may be considered as proxies for borrowing constraints. Indeed, Hayashi (1985) and Zeldes (1989) define as liquidity constrained the households with low savings or low financial assets. Given that we can directly observe the debt, we can use it as a proxy. We also consider eight geographical areas and eleven industries.

Finally, we also include a Stockholm dummy and a dummy that controls for the immigration status. The Stockholm dummy takes the value of 1 if the investor lives in the capital and 0 otherwise. The immigration status is a dummy that takes the value 0 if all the members of the household are native Swedes and 1 if at least one member of household immigrated.[7] All

[7] We also tried two alternative specifications. In the first one, we used the sum of the immigration statuses of the members of the household. That is, if two members of the household are immigrants, the variable takes value 2. In a second specification, we used the inverse of the number of years since the oldest immigrant in the household arrived in Sweden. These two alternative specifications deliver results that are qualitatively analogous to those reported. These results are available upon request from the authors.

monetary variables (level and variance of nonfinancial income, wealth, etc.), except capital gains/losses, have been transformed into logarithms.

3.4 Econometric issues

We now move on to the econometric issues. Henceforth, we will use the subscript t to refer to the year. We assume that the investment decision takes place in two steps: first, the investor decides whether to enter the stock market (stocks, mutual funds), and then he selects which asset to buy. The decision to enter the market can be described as

$$P_t = \alpha_{00} + \beta_{00} + X_{00,t} + \varepsilon_{00,t}, \tag{7}$$

where P_t is the observed probability of market participation (i.e., $P_t = 1$ if the investor participates in the stock market and zero otherwise). The probability that the investor enters the financial market is modeled as a normal c.d.f. In order to estimate this probability, we consider a bigger data set based on the whole sample universe: i.e., both the households that hold financial assets and those that do not. The variables $X_{00,\ t}$ include the correlation between the different sources of nonfinancial income and the market portfolio, the volatility of the sources of nonfinancial income, and all the aforementioned control variables (i.e., measures of income and wealth, measures of nonfinancial income, demographic variables, professional ability and risk of the investor, the momentum/stock performance variables, and the residual control variables).

The second stage deals with portfolio choice. For expositional purposes, let us define C_t as the generic dependent variable (e.g., Δ in the case of Equation 4) and $\mathbf{X}_{\mathbf{1},t}$ as a generic vector of explanatory variables as defined in the previous restrictions. We therefore have

$$C_t = \alpha_1 + \beta \mathbf{X}_{1,t} + \varepsilon_{1,t}. \tag{8}$$

The identification restrictions require us to use, in the first stage, control variables that do not appear in the second stage.[8] We use Heckman's (1979) two-stage procedure. In the first stage, we estimate stock market participation. In the second stage, we include a variable that accounts for the possibility of selection bias at the first stage. This variable is defined as $\lambda_{i,t}$ ("Heckman's lambda") and controls for the problem of omission of variables due to self-selection. We therefore estimate

$$C_t = \alpha_1 + \beta \mathbf{X}_{1,t} + \theta_1 \lambda_t + \varepsilon_{1,t}. \tag{9}$$

[8] These variables include the set of time and industry dummies as well as the correlations between the market portfolio and investors' sources of nonfinancial income (i.e., labor income, entrepreneurial income, and real estate).

The significance of the values of θ_1 provides a test of the null of no sample selection bias. We will see that, in all the specifications, a high degree of significance of θ_1 suggests that self-selection is indeed important in the sample. Specification 9 is estimated by using two-stage least squares with consistent variance-covariance matrix. We perform the analysis at the household level.

Some of the explanatory variables (i.e., the proxies for hedging $\mathrm{Corr}_{y,p,t}$ and Γ_t) are affected by the investor's choice and are therefore endogenous. Moreover, nonfinancial income itself may be endogenous (Gentry and Hubbard 2002; Moskowitz and Vissing-Jørgensen 2002). To address this issue, we pursue a two-pronged approach. First, we use an instrumental variable methodology (Vissing-Jørgensen 2002). We instrument the potentially endogenous variables (measures of hedging and familiarity, nonfinancial sources of income, lagged dependent variables, etc.) using as instruments a combination of strictly exogenous variables (i.e., demographic variables, industry and time dummies) and the lagged values of the main variables in the different specifications.[9] Alternatively, we modify the estimation of the second stage of Heckman's procedure and perform a robustness check based on the estimation of a system of simultaneous equations. That is, we re-estimate Equation 9 as part of a two-equation system where also the proxies for hedging are jointly determined. The results (not reported) do not differ from those derived from the instrumental variable estimation (reported).

Given that one of the variables—the correlations between financial and non-financial income—is estimated over the entire sample, the usage of a panel estimate may overstate the significance of the results. Therefore, as a robustness check, we also run the case where each investor accounts for one observation. The results (not reported) are consistent with those reported in the text.

Finally, Figure 1 highlights that the distribution of hedging/nonhedging behavior is bimodal. From an economic perspective, if the R^2 is low and the distribution of the dependent variable is bimodal, we may be missing something interesting economically and this could affect the results. To control for it, we include a dummy variable (*Distributional Dummy*). The dummy variable takes the value 1 for the investors for which Γ is greater than 0 and 0 otherwise. This exploits the information contained in the bimodality of the sample.[10]

[9] The endogeneity issue further complicates the task of finding proper instrumental variables, as only strictly exogenous variables or predetermined ones can be used. As Arellano (1989) and Kiviet (1995) showed, lagged values represent predetermined variables, uncorrelated with the residuals, whereas the demographic variables are strictly exogenous in the Granger-Sims sense. The *Adjusted RSquares* of the first stage regressions range between 25% and 68%.

[10] We thank an anonymous referee for alerting us to this problem.

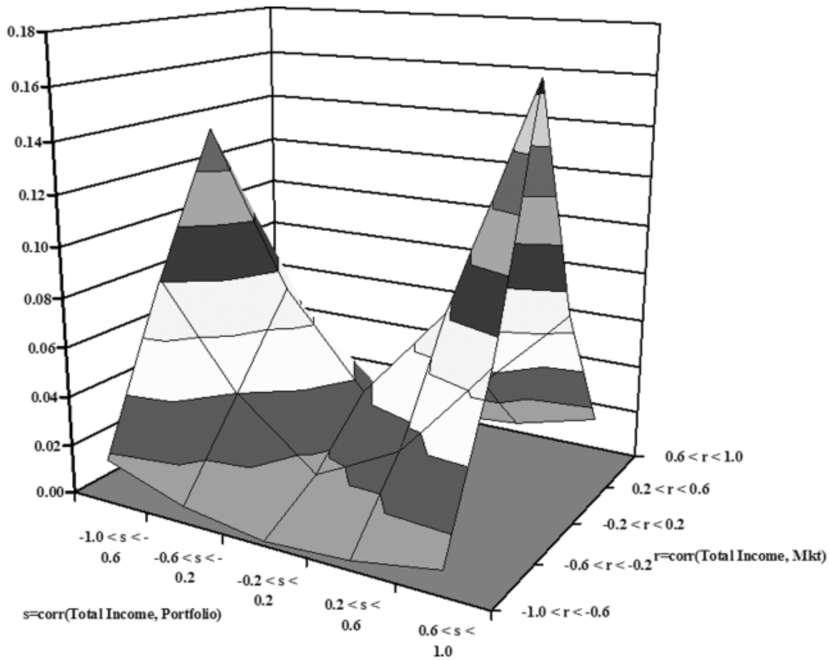

Figure 1
Two-dimensional frequency plots for the correlation between total income and overall stock market Index (SIX Index) and total income and return on household portfolio of risky assets. Percentage of the total sample is along the z-axis, range of values of the correlation coefficients are along the x-axis. We use five ranges for correlations: from –1 to –0.6, from –0.6 to –0.2, from –0.2 to 0.2, from 0.2 to 0.6, and from 0.6 to 1.

4. Main Findings: Hedging Versus Familiarity

We proceed as follows. First, we consider the determinants of portfolio choice and test whether investors select their portfolio so as to hedge nonfinancial income risk. Then, we provide evidence of familiarity, we analyze it, and show its information-based nature. Finally, we analyze the implications of hedging and familiarity-driven investment both in terms of risk taking and in terms of costs and benefits of the two strategies for the investors.

4.1 Preliminary evidence of portfolio choice

As a first step, we relate to the standard literature. We consider a heuristic reformulation, based on portfolio data, of the standard testing equation that links the investment in risky assets to the main moments of financial and nonfinancial income (Heaton and Lucas 2000; Vissing-Jørgensen, 2002):

$$h_t = \alpha_0 + \beta_0 \mathrm{var}_{y,t} \times \mathrm{sign}\left(\mathrm{Corr}_{y,p,t}\right) + \gamma_0 \mathrm{Corr}_{y,p,t} + \delta_0 \mathbf{F}_{0,t} + \theta_0 \lambda_t + \mu_0 h_{t-1} + \varepsilon_{0,t}. \tag{10}$$

This specification provides a preliminary evidence of whether investors increase their investment in risky assets to hedge nonfinancial-income risk. If investors buy risky assets to hedge, we expect a positive correlation between nonfinancial and financial income to be associated with a reduction of the investment in risky assets and a negative correlation to be associated with an increase of it. That is, $\gamma_0 < 0$ if $\mathrm{Corr}_{y,p} > 0$ and $\gamma_0 > 0$ if $\mathrm{Corr}_{y,p} < 0$.

An additional restriction is related to the variance of nonfinancial income ($Var_{y,t}$). In the case of stochastic nonfinancial income, the variance of nonfinancial income ($Var_{y,t}$) reduces the investment in risky assets only if the correlation between financial and nonfinancial income is positive or close to zero (Viceira 2001). A negative correlation, on the other hand, makes it more attractive to invest in risky assets as it reduces the risk of the overall portfolio. The incentive is stronger the higher the nonfinancial income risk is. This implies that the test of hedging is a joint one that requires a negative correlation between the investor's financial and nonfinancial income and a positive correlation between the investment in risky assets and the variance of nonfinancial income. The multiplication of nonfinancial risk by the term *sign* ($\mathrm{Corr}_{y,\ m}$) allows us to account for it. The higher the nonfinancial risk, the more it is worth investing in risky assets if this allows the investor to diversify (i.e., $Corr_{y,p} < 0$) and the less it is worth investing if this, instead, increases the overall risk (i.e., $Corr_{y,p} > 0$). That is, if investors perceive the investment in financial assets as a way of hedging nonfinancial income, we expect $\beta_0 < 0$.

As in the traditional literature, we consider a specification based on the percentage value of the investment in risky assets (stocks and mutual funds) over overall wealth (Risky Share) and a specification based on the dollar value of the investment (Risky Value). The results are reported in Table 2. We consider separately the case where the correlation between financial and nonfinancial income is positive from the case where it is negative.

The results based on the correlations between financial and nonfinancial income strongly reject the hypothesis that investors buy to hedge. Indeed, γ_0 is always positive in the case of a positive correlation and negative or not significant in the case of a negative correlation. This holds for both the specification based on the percentage investment and the one based on the dollar value of the holdings. Moreover, the results are robust across investors, regardless of their wealth level.

Table 2
Portfolio choice: investment in risky assets

	All households				Low wealth households				High wealth households			
	Risky share		Risky value		Risky share		Risky value		Risky share		Risky value	
Variable	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	*t*-stat
Var(income)*sign[corr (income,m)]	−0.01	−2.77	−0.11	−4.36	−0.01	−1.82	−0.03	−3.34	0.28	2.69	5.07	2.95
Corr(income, portfolio)$^{+}$	0.15	8.45	1.44	8.15	0.15	7.04	3.22	6.89	0.48	2.47	1.02	2.69
Corr(income, portfolio)$^{-}$	−0.05	−2.74	−0.33	−2.12	−0.08	−3.76	−1.29	−3.08	0.22	1.65	0.43	1.80
Control variables												
Intercept	−0.08	−0.37	−3.89	−1.99	0.01	0.03	1.37	1.48	−0.32	−1.53	−3.54	−0.91
Income (level)	−0.01	−0.60	0.03	0.31	0.03	3.72	0.29	8.73	0.16	2.56	2.99	2.81
Corr(income, real estate)	0.05	2.35	1.10	5.49	0.07	3.57	0.69	7.57	0.23	3.18	4.23	3.63
Ret	−0.01	−2.97	−0.01	−3.61	0.00	−0.18	0.00	−2.16	0.00	0.66	0.00	0.18
Portfolio												
Financial wealth	−0.04	−4.52	0.21	2.97	−0.04	−4.32	0.10	2.89	−0.01	−2.28	−0.03	−0.38
Real estate wealth	−0.05	−36.24	0.01	2.40	−0.05	−46.63	0.00	0.68	−0.01	−9.43	−0.01	−0.47
Capital gains/losses	0.01	3.90	0.03	3.61	0.00	2.92	0.02	3.52	0.00	−0.41	−0.01	−0.44
Secondary education	0.50	4.55	0.43	5.38	0.39	3.58	1.77	5.04	0.01	0.48	0.10	0.51
Higher education	0.63	4.72	0.52	5.43	0.47	3.68	2.11	5.10	−0.01	−0.30	−0.22	−0.61
Ability	0.08	5.25	0.88	7.62	0.06	4.07	0.41	8.24	0.04	3.69	0.80	4.23
Size of household	−0.04	−4.45	−0.16	−2.51	−0.03	−3.90	−0.07	−2.30	0.01	2.78	0.36	4.32
Immigration status	−0.03	−3.58	−0.14	−1.93	−0.01	−2.44	−0.05	−2.54	−0.01	−1.60	−0.26	−1.60
Age	−0.02	−4.91	−0.01	−4.39	−0.01	−3.32	−0.03	−3.33	0.01	3.76	0.17	4.59
Age2	0.18	5.12	0.14	5.58	0.01	3.51	0.05	4.43	−0.04	−3.10	−0.90	−3.63

Unemployment risk	−0.24	−2.34	−2.34	−3.08	−0.14	−2.78	−1.25	−5.08	0.12	3.71	2.40	4.04
Stockholm dummy	−0.04	−4.40	−0.27	−4.36	−0.02	−2.80	−0.08	−3.54	0.00	0.10	0.06	0.62
Lagged dependent variable	0.10	15.51	−0.04	−3.28	0.07	14.70	0.01	2.71	0.57	27.17	−0.07	−1.99
Lambda	0.13	2.58	1.40	3.27	0.11	2.16	0.45	2.06	−0.17	−1.60	−1.56	−1.85
Residual control variables	Yes		Yes		Yes		Yes		Yes		Yes	
Adj R^2	0.54		0.22		0.61		0.10		0.49		0.18	

We report the estimates of the determinants of the investment in risky financial assets. The dependent variable is either the percentage of the investment in risky assets over overall wealth (risky share) or the dollar value of the investment (risky value). The positive and negative correlations ("+" and "–") are separately considered. We consider different types of control variables: *measures of income and wealth, demographic variables, professional ability and risk, momentum/stock performance variables, and residual control variables*. The *measures of income and wealth* contain the vector of the wealth of the *ith* investor at time *t*, broken down into its individual components (i.e., logarithm of financial wealth, logarithm of real estate wealth), as well as the total nonfinancial income (i.e., sum of labor and entrepreneurial) and the overall (i.e., financial and nonfinancial) net of capital gains and losses of the *ith* investor at *t*. We also include the correlation between nonfinancial income and real estate. All monetary variables (level and variance of nonfinancial income, wealth) except capital gains/losses, have been transformed into logarithms. The *demographic variables* include the profession of the investor, his level of education, broken down into high school (*secondary education*) and university level (*higher education*), the number of members of the family (*size of household*), the age of the oldest member of the family of the investor and its value squared (*age* and *age2*). We also construct variables to account for the professional ability (*ability*) and unemployment risk (*unemployment risk*) of the investor. The first variable is based on the difference between his income and the average income of his profession. The second variable is the one-year-ahead forecast of a linear probability model where the unemployment status (i.e., 1 if unemployed and 0 otherwise) is regressed on demographic variables, measures of income and wealth and regional, geographical and professional dummies. The *momentum/stock performance variable* is the return of the investor's portfolio in the previous twelve months. The *residual control variables* include the standardized levels of debt for the investor (ratio of investor debt to total income and ratio of investor debt to total wealth), the return and volatility on the market portfolio in the previous twelve months, an index of consumer confidence and a set of dummies that account for the regional location of the investor as well as the industry in which he works. We also include eight geographical are as and ten industries. We also include a *Stockholm dummy* and a dummy that controls for the immigration status (*immigration status*). The Stockholm dummy takes the value of 1 if the investor lives in the capital and 0 otherwise. The immigration status is a dummy that takes the value 0 if all the members of the household are native Swedes and 1 if at least one member of household immigrated. *Age2* is divided by 1000. Estimates are performed using 2SLS.

A 1% increase in the positive correlation between financial and nonfinancial income increases the amount invested in risky assets of 3.4% for the overall sample (7.7% for the low wealth investors and by 2.38% for the high wealth investors). A 1% increase in the negative correlation between financial and nonfinancial income decreases the amount invested in risky assets of 0.8% overall (by 3% for the low wealth investors and not significant for the high wealth investors). Another way of reading these results is that the investors who invest more in risky assets are those who have chosen a portfolio composition more positively related to their nonfinancial income. This also provides first evidence of familiarity.

The effect of nonfinancial income risk (β_0) is negative for the low wealth investors and positive for the high wealth investors. This is robust across specifications (i.e., both in the case of the level of holdings and in the case of share in the portfolio). The results for the low wealth investors may be interpreted as weak evidence in favor of hedging as well as they may suggest that investing in risky assets is perceived as a risky strategy.

Overall, these findings provide a first evidence against hedging. One important feature of the data is that we can measure holding changes at the individual investor level. Analyzing directly the changes takes advantage of the panel dimension. To exploit this information, we re-estimated specification 10 using as a dependent variable the change in holdings. The results (not reported) are consistent with those reported in the text.

4.2 Portfolio choice: the risk profile of the financial portfolio

4.2.1 Hypothesis H.1: the tilt in the risk profile. The weakness of the previous findings is related to the fact that standard tests do not account for the investors tilting their portfolio away from the market portfolio. We now analyze this tilt in more detail. We start by studying the correlation (covariance) between the portfolio of the investor and his total nonfinancial income and comparing it to the correlation (covariance) between his total nonfinancial income and the market.

An overall view is provided by Figure 1. This is a 3D frequency distribution where the *y*- and *x*-axes are the correlation between nonfinancial income and the market portfolio and the correlation between nonfinancial income and the investor's portfolio, respectively. These are constructed in the following way. First, we calculate for each investor the correlation of his total income, as defined above, with either the market portfolio (return the value-weighted Swedish SIX index) or the investor's portfolio. Then, we construct the frequency distributions of such correlations.

Figure 1 shows that if we focus on the correlation between nonfinancial income and the market, most of the mass of the distribution is centered around zero. This reflects the standard results in the literature that find a very low or zero correlation between financial and nonfinancial income. However, if we focus on the correlation between nonfinancial income and the investor's portfolio, the results change drastically. Now, most of the mass of the distribution is in the tails. This suggests that most of the investors tilt their portfolio toward a correlation between financial and nonfinancial income different from the one that the market portfolio would generate.

Moreover, the distribution based on the correlation between nonfinancial income and the investor portfolio is more skewed to the right. That is, investor portfolios are more positively related to investor income than the market portfolio is. To more properly quantify these claims, we consider some descriptive statistics. In Table 3, we report descriptive statistics on the average value of the correlations between nonfinancial income and either investor portfolios or aggregate stock market and tests of the difference between them. We also report our indexes of hedging.

The comparison of the correlation with the portfolio and the correlation with the market confirms the tilt away from hedging. Indeed, while the correlation of labor income with the market is negative (–0.01 for the low wealth investors and –0.04 for the high wealth investors), the correlation of labor income with the investor's portfolio is positive (0.03 for the low wealth investors and 0.05 for the high wealth investors). In other words, investors construct their portfolios so as to actually turn their natural negative correlation with the market into a positive one. The differences between statistics based on the market portfolio and statistics based on the investor portfolio are reflected in the indexes of hedging.

In terms of economic significance, we see that the tilt, if measured in terms of changes in correlation, is of the same order as the original correlation with the market. That is, while, on average, the investors have a correlation between labor (entrepreneurial) income and the market equal to –0.01 (0.03), the tilt in the portfolio is equal to –0.03 (–0.07).

If we consider Δ, we see that the tilt is equal to 2.5% of the value of the risky financial portfolio for the low wealth investors and 0.2% for the high wealth investors. Another way of assessing the economic magnitude of the tilt is to calculate the difference between the entity of the tilt from the market portfolio in case the investor were hedging nonfinancial income risk as theory would require and the actual entity of the tilt.[11] We see that the investors are actually tilting their portfolio away from what theory would require by an amount equal to 41% of the value of the risky financial portfolio for the low wealth investors and 10% for the high wealth investors.

Overall, the limited magnitude of both the correlation with the market portfolio and the tilt provide first evidence against standard theories of hedging and point in the direction opposite to hedging. These findings

[11] In particular, we construct the tilt required by theory as $\Delta_z = \frac{Y_z}{W}\text{var}_{Y_Z} + \frac{Y_x}{W}\text{cov}_{Y_z,Y_x} + \frac{1}{W}\Theta_z$, assuming a coefficinet of risk aversion equal to two.

Table 3
Correlation and indices of hedging (tests of restriction H2A)

	Low wealth					High wealth				t-test of the difference		
Variable	Mean	Median	Standard deviation	IQR	*t*-test mean = 0	Mean	Median	Standard deviation	IQR	*t*-test mean = 0	*t*-stat	*p* value
Corr(labor income, market)	−0.01	0.00	0.31	0.41	−4.86	−0.04	−0.05	0.31	0.41	−16.51	10.82	<0.0001
Corr(labor income, portfolio)	0.03	0.09	0.79	1.77	2.57	0.05	0.10	0.68	1.10	7.44	−2.32	0.020
Corr(labor income, real estate)	−0.06	−0.09	0.47	0.76	−27.81	−0.13	−0.19	0.46	0.74	−39.51	16.71	<0.0001
Corr(entrepreneurial income, market)	−0.07	−0.14	0.61	0.90	−11.01	−0.07	−0.14	0.59	0.90	−9.14	0.09	0.926
Corr(entrepreneurial income, portfolio)	−0.01	0.00	0.84	1.88	−0.29	−0.06	−0.14	0.77	1.60	−3.99	2.10	0.036
Corr(entrepreneurial income, real estate)	−0.18	−0.32	0.64	0.98	−25.24	−0.18	−0.21	0.62	0.86	−21.33	0.40	0.686
Corr(total income, market)	−0.01	0.00	0.39	0.48	−33.63	−0.03	−0.03	0.33	0.43	−32.31	21.02	<0.0001
Corr(total income, portfolio)	0.03	0.09	0.79	1.80	10.40	0.05	0.10	0.68	1.10	15.59	−3.30	0.001
Corr(total income, real estate)	−0.05	−0.08	0.75	0.84	−70.81	−0.12	−0.18	0.48	0.76	−90.56	45.14	<0.0001
Γ_1	−0.03	−0.03	0.82	1.34	−2.31	−0.08	−0.11	0.71	1.07	−10.42	3.36	0.001
Γ_2	−0.07	−0.00	0.88	1.39	−2.71	−0.09	−0.03	0.78	1.10	−4.05	0.70	0.483
Γ_T	−0.02	−0.02	0.78	1.27	−4.60	−0.08	−0.11	0.69	1.02	−23.48	11.73	<0.0001
Δ_1	−0.34	−0.00	4.18	0.85	−7.07	−0.49	−0.07	3.96	0.74	−12.06	2.47	0.014
Δ_2	−0.34	−0.00	2.50	0.49	−0.99	−0.41	−0.00	10.43	0.39	−2.99	0.16	0.866
Δ_T	−0.25	−0.10	5.68	1.05	−10.62	−0.48	−0.29	3.32	1.44	−30.42	12.95	<0.0001

We report statistics of the correlations of nonfinancial income (i.e., labor income, entrepreneurial income, and total income) with financial returns (i.e., portfolio returns and overall stock market returns) and real estate return. The nonfinancial income (total and components) is estimated by using the Carrol and Samwick (1997) and Vissing-Jørgensen (2002) methodology. We report the descriptive statistics separately for high wealth and low wealth households as well as our indexes of active hedging (i.e., Γ_1, Γ_2, and Γ_T and Δ_1, Δ_2, and Δ_T where Γ_1 = corr (labor income, market) – corr(labor income, portfolio), Γ_2 = corr(entrepreneurial income, market) – corr(entrepreneurial incoome, portfolio), and Γ_T = corr(total income, market) – corr(total income, portfolio) and Δ_1 = cov(labor income, market) – cov(labor income, portfolio), Δ_2 = cov(entrepreneurial income, market) – cov(entrepreneurial income, portfolio), and Δ_T = cov(total incoome, market) – cov(total income, portfolio). We report the results of the *t-tests* for each group (where we test for mean = 0 hypothesis) as well as the difference between high and low wealth households. The values of Δs are expressed in tens of thousands of SEK.

may be due either to deliberate behavior or to spurious correlation. We will now use a more formal test based on a model of portfolio choice to address this issue.

4.2.2 Hypothesis H.2: determinants of the tilt in the risk profile. We now focus on restriction H.2, estimated conditional on market participation. The general equation that describes the decision to invest in risky financial assets is

$$\begin{aligned}\Delta_{z,t} =& \alpha_1 + \beta_1 \frac{Y_{z,t}}{W} \mathrm{Var}_{Y_{z,t}} + \gamma_1 \frac{Y_{x,t}}{W} \mathrm{Cov}_{Y_z, Y_{x,t}} + \zeta_1 \frac{1}{W} \Theta_{z,t} + \boldsymbol{\delta}_1 \mathbf{F}_{1,t} \\ &+ \nu_1 \boldsymbol{\Psi}_t + \boldsymbol{\theta}_1 \lambda_t + \varepsilon_{1,t}.\end{aligned} \tag{11}$$

We have different options for the definition of nonfinancial income. We can aggregate the different sources of nonfinancial income of the investor and assume that he just hedges their aggregated value. Alternatively, we can analyze them separately. Finally, we can consider investors who have both sources of income separately from those who have just labor income. We proceed as follows. We start by focusing on the separate sources of income, as Equation 4 would require, and test hedging of labor income and hedging of entrepreneurial income separately. Then, we also consider the case in which the sources of income are aggregated. Finally, as additional robustness check, we also consider the investors who have only labor income. Given that all the results are consistent, we will not report the last ones. We consider two main specifications. All the specifications include the five sets of control variables: measures of income and wealth, demographic variables, professional ability and risk, momentum/stock performance variables, and residual control variables. Hedging requires that $\beta_1 > 0$, $\gamma_1 > 0$, and $\zeta_1 > 0$. Specification I does not include the measures of familiarity (Ψ_t). Specification II also includes the measures of familiarity. Familiarity requires that $\nu_1 > 0$.

The results are displayed in Table 4, Panel A for the case of labor-income risk, Panel B for the case of entrepreneurial income risk, and Panel C for the case of aggregated total income risk. They reject the hypothesis of hedging. For the overall sample, all three coefficients, β_1, γ_1, and ζ_1, are always negative. If we consider labor income, we see that β_1 is always negative and significant for both high- and low wealth investors. This is a particularly strong rejection of hedging, as β_1 represents the very coefficient that captures the direct impact of the riskiness of the *zth* source of income on the portfolio tilt to hedge it. In the case of entrepreneurial income, β_1 is still negative and significant for the low wealth investors and becomes positive and significant for the high wealth investors. Also the coefficients γ_1 and ζ_1 do not support the hypothesis of hedging. The coefficient ζ_1 is always negative for both types of nonfinancial income

Table 4
Determinants of the tilt in the risk profile

	All households				Low wealth households				High wealth households			
	I		II		I		II		I		II	
Variable	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	t-stat	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	*t*-stat
Panel A: determinants of Δ_1 *(hedging labor income risk)*												
$Y_L\text{var}(Y_L)$	−0.25	−11.80	−0.26	−12.53	−0.33	−14.73	−0.32	−14.20	−9.03	−2.98	−9.01	−2.69
$Y_E\text{cov}(Y_E, Y_L)$	−1.40	−17.38	−1.39	−17.38	−1.41	−16.13	−1.35	−15.01	−4.91	−1.91	−4.23	−1.36
Θ_L	−0.78	−12.62	−0.77	−12.56	−0.81	−11.90	−0.76	−10.84	−0.79	−3.51	−0.78	−3.10
Geographical proximity			−3.81	−9.40			−3.88	−5.79			−7.66	−5.15
Professional proximity			−2.84	−4.20			−2.54	−2.30			0.90	0.57
Holding period			−6.74	−7.12			−6.55	−4.64			−3.12	−2.51
Control variables												
Intercept	−89.76	−10.56	−80.38	−9.41	−124.49	−12.23	−114.29	−10.88	−58.30	−1.82	−67.00	−1.83
Corr(labor income, real estate)	2.68	23.75	2.81	23.91	2.99	21.31	3.09	21.39	4.48	2.26	2.68	1.27
Corr(entrepreneurial income, real estate)	−0.19	−1.07	−0.29	−1.66	−0.01	−0.04	−0.10	−0.43	−1.13	−3.06	−1.24	−3.06
RetPortfolio	0.09	1.30	0.03	0.48	−0.26	−3.38	−0.29	−3.66	0.00	−0.79	0.00	−0.93
Financial wealth	1.50	10.22	1.82	12.43	2.02	11.10	2.02	11.02	0.41	1.25	0.73	1.87
Real estate wealth	0.17	5.89	0.20	7.06	0.24	7.55	0.25	7.51	−0.41	−2.59	−0.45	−2.56
Capital gains/losses	0.44	10.37	0.42	10.20	0.61	10.51	0.58	9.86	0.10	1.04	0.22	2.07
Secondary education	−0.88	−2.84	−0.58	−1.93	−0.01	−0.02	0.02	0.04	−1.36	−0.75	0.38	0.19
Higher education	−0.71	−1.89	−0.56	−1.52	0.10	0.20	0.03	0.07	−0.89	−0.43	0.07	0.03
Ability	0.76	3.03	1.25	5.00	1.50	4.62	1.49	4.54	0.26	0.23	0.79	0.64
Size of household	1.21	11.81	1.08	10.66	1.68	13.22	1.58	12.13	−0.18	−0.46	−0.68	−1.51
Immigration status	−1.55	−7.61	−1.78	−8.78	−2.22	−8.63	−2.27	−8.50	1.72	2.01	2.13	2.17
Age	0.72	6.67	0.78	7.47	0.61	5.31	0.58	4.97	1.79	1.39	0.46	0.32
Age2	−0.76	−7.40	−0.79	−7.97	−0.69	−5.98	−0.64	−5.48	−1.51	−1.39	−0.38	−0.31
Unemployment risk	−20.25	−9.59	−24.62	−11.53	−30.36	−10.27	−31.13	−10.36	1.06	0.10	−14.73	−1.20
Stockholm dummy	−0.63	−4.25	−0.82	−3.11	−1.49	−7.09	−0.86	−4.11	−0.25	−0.46	0.23	0.30
Distributional dummy	13.20	7.90	24.17	10.39	26.27	9.06	30.49	7.63	5.30	1.84	39.11	5.21
Lambda	7.88	9.48	9.95	11.96	12.21	10.87	12.08	10.78	5.11	1.45	5.04	1.35
Residual control variables	Yes		Yes		Yes		Yes		Yes		Yes	
Adj R^2	0.02		0.03		0.02		0.03		0.02		0.03	

Panel B: determinants of Δ_2 *(hedging entrepreneurial income risk)*												
$Y_E \mathrm{var}(Y_E)$	−0.43	−31.59	−0.43	−29.13	−0.22	−9.25	−0.22	−7.73	0.86	12.14	0.71	7.79
$Y_L \mathrm{cov}(Y_E, Y_L)$	−0.47	−3.39	−0.43	−2.88	−0.30	−2.58	−0.27	−1.89	−7.82	−3.07	−5.97	−1.90
Θ_E	−0.34	−46.84	−0.35	−42.78	−0.23	−20.98	−0.23	−17.45	−0.61	−28.45	−0.62	−23.58
Geographical proximity			−5.28	−5.52			−0.26	−0.23			−8.19	−5.20
Professional proximity			−6.14	−3.41			−4.80	−2.10			−4.42	−1.47
Holding period			−6.05	−2.40			−5.88	−2.05			−20.19	−3.90
Control variables												
Intercept	−16.94	−0.80	12.93	0.53	41.30	4.34	−40.61	−2.00	−31.03	−1.53	17.57	0.65
Corr(labor income, real estate)	−0.27	−1.05	0.20	0.65	−0.02	−0.09	−0.25	−0.86	−0.77	−1.30	−0.37	−0.42
Corr(entrepreneurial income, real estate)	−2.13	−5.08	−2.30	−5.04	−1.52	−3.91	−1.61	−3.39	−2.98	−4.70	−2.77	−3.57
RetPortfolio	−0.06	−0.37	−0.19	−1.14	0.04	0.27	0.01	0.06	0.00	−0.80	0.00	−0.55
Financial wealth	−0.78	−2.35	−0.83	−1.88	−1.20	−3.91	−1.06	−2.81	0.10	0.31	−0.18	−0.44
Real Estate wealth	−0.31	−4.70	−0.30	−3.97	−0.20	−3.74	−0.13	−1.86	−0.22	−1.77	−0.31	−1.96
Capital gains/losses	0.24	2.47	0.22	2.08	0.02	0.20	−0.01	−0.06	0.77	4.20	0.76	3.38
Secondary education	−4.85	−6.54	−4.49	−5.55	−2.88	−3.82	−2.93	−3.18	−1.87	−1.24	0.19	0.10
Higher education	−5.03	−5.74	−4.99	−5.22	−2.86	−3.41	−2.72	−2.66	−0.40	−0.21	1.40	0.60
Ability	−2.60	−3.86	−2.22	−3.01	−2.65	−3.99	−2.83	−3.49	−2.57	−3.16	0.47	0.41
Size of household	−0.54	−2.22	−0.88	−3.30	−0.75	−4.02	−0.78	−3.53	−0.50	−1.52	−0.99	−1.41
Immigration status	−0.20	−0.41	−0.38	−0.70	0.83	1.89	−0.19	−0.33	−1.21	−1.37	0.35	0.31
Age	0.53	2.14	0.72	2.65	−0.04	−0.19	−0.19	−0.79	−0.07	−0.13	−0.56	−0.87
Age2	−0.61	−2.58	−0.74	−2.89	−0.01	−0.06	0.12	0.49	−0.09	−0.20	0.30	0.54
Unemployment risk	18.17	3.69	15.69	2.77	22.64	6.22	11.04	2.24	25.40	3.14	2.34	0.21
Stockholm dummy	0.09	0.27	0.78	1.90	0.04	0.13	−0.41	−1.00	2.96	2.42	3.29	3.49
Distributional dummy	12.59	2.91	38.12	6.29	1.80	3.93	0.81	0.74	14.46	4.14	44.66	4.00
Lambda	−4.79	−2.49	−3.96	−3.43	−7.99	−4.36	−6.32	−2.79	7.14	2.33	20.63	2.90
Residual control variables	Yes		Yes		Yes		Yes		Yes		Yes	
Adj R^2	0.01		0.02		0.01		0.02		0.01		0.02	
Panel C: Determinants of Δ_T *(hedging total income risk)*												
$Y_T \mathrm{var}(Y_T)$	−0.09	−5.32	−0.11	−6.39	−0.14	−9.63	−0.16	−8.78	1.27	12.47	1.15	8.69
Θ_T	−0.35	−9.07	−0.37	−9.38	−0.07	−2.38	−0.06	−2.05	−5.85	−8.62	−5.20	−5.56
Geographical proximity			−3.36	−7.54			−1.79	−2.74			−5.54	−6.67
Professional proximity			−3.35	−4.49			−0.70	−0.62			−5.72	−3.68
Holding period			−8.74	−8.58			−7.32	−4.98			−7.89	−2.48
Control variables												
Intercept	−134.19	−15.60	−135.11	−14.78	−136.26	−28.31	−181.04	−18.75	−22.43	−1.09	3.11	0.11
Corr(total income, real estate)	2.48	20.83	2.42	18.65	2.64	21.43	2.57	17.02	2.61	7.40	3.88	6.52
RetPortfolio	0.06	0.90	0.03	0.45	0.00	−0.02	−0.05	−0.64	−0.04	−0.59	−0.22	−2.68

Financial wealth	1.33	8.88	1.69	10.57	1.79	11.61	2.04	10.73	−0.18	−1.03	0.26	1.08
Real estate wealth	0.14	4.73	0.20	6.46	0.19	6.82	0.27	7.75	−0.16	−2.01	−0.23	−2.14
Capital gains/losses	0.39	8.87	0.36	8.00	0.59	11.69	0.56	9.09	0.03	0.36	0.13	1.04
Secondary education	−0.83	−2.60	−0.53	−1.61	−0.90	−2.51	−0.31	−0.69	−2.36	−2.54	−1.46	−1.24
Higher education	−0.90	−2.30	−0.59	−1.45	−0.54	−1.30	0.01	0.02	−2.93	−2.69	−1.48	−1.06
Ability	1.59	6.22	2.09	7.70	1.57	7.58	2.50	8.82	−0.03	−0.05	0.98	1.35
Size of household	1.48	14.00	1.41	12.74	1.66	18.98	1.96	16.08	0.47	1.50	−0.50	−1.17
Immigration status	−1.80	−8.55	−2.22	−9.96	−2.22	−10.02	−2.91	−10.35	−0.02	−0.04	−0.04	−0.06
Age	0.56	5.07	0.60	5.22	0.66	6.87	0.54	4.56	1.17	2.18	1.56	2.32
Age2	−0.63	−5.99	−0.64	−5.89	−0.75	−7.72	−0.63	−5.33	−1.08	−2.39	−1.38	−2.43
Unemployment risk	−26.39	−12.29	−31.83	−13.83	−27.34	−16.91	−38.42	−15.89	−1.01	−0.21	−1.99	−0.27
Stockholm dummy	−0.75	−4.87	−0.14	−0.81	−1.06	−6.72	−1.23	−5.63	0.15	0.42	1.79	3.43
Distributional dummy	24.96	14.99	27.44	10.66	23.32	11.98	23.04	6.29	1.45	0.52	29.69	4.80
Lambda	7.34	8.65	9.15	10.07	10.75	11.50	12.27	10.67	−2.92	−1.99	8.14	2.01
Residual control variables	Yes		Yes		Yes		Yes		Yes		Yes	
Adj R^2	0.01		0.02		0.01		0.02		0.01		0.02	

This table reports estimates of the determinants of Δ_1 (panel A), Δ_2 (panel B), and Δ_T (panel C), where Δ_1, Δ_2, and Δ_T are defined in Table 3. Θ_L is $(Y_L + Y_E)\Sigma_j\Omega_{sj} \times \mathrm{cov}(\mathrm{ret}_j, Y_L)$, Θ_E is $-(Y_L + Y_E)\Sigma_j\Omega_{Sj} \times \mathrm{cov}(\mathrm{ret}_j, Y_E)$ and $\Omega_{sj} = (\mu_{sj} - r)' \big/ \left[(1-\gamma)\sigma_{sj}^2\right]$. μ_{sj} and σ_{sj}^2 are, respectively, the mean and the variance of the return of the *jth* risky asset while $\mathrm{cov}(Y_L, Y_E)$ is the covariance between labor and entrepreneurial income. The Δs and Ys are measured in thousands of SEK. The familiarity variables are *professional proximity, geographical proximity,* and *holding period*. They are constructed as follows. For each stock in the portfolio we identify three measures of familiarity and then we aggregate them for each investor on the basis of his portfolio composition (i.e., using as weights the value of the portfolio holding). *Professional proximity* is a dummy taking the value 1 if the investor's profession is in the same area of activity as the company whose stock is under consideration and zero otherwise. We use the one-digit SNI92 codes (similar to SIC codes) to identify the areas of activities. *Geographical proximity* is the logarithm of the inverse of the distance between the ZIP code of the investor and the ZIP code of the closest branch/subsidiary of the company whose stock we consider. *Holding period* is based on the time a stock entered the investor's portfolio. Each index of familiarity is constructed weighting the measures for each investor on the basis of his portfolio composition. The *control variables* are defined as in Table 2. *Distributional dummy* is a dummy that takes the value 1 if $\Gamma_i, > 0$ and 0 otherwise (i = total income in the case the dependent variable is Δ_T, labor income in the case the dependent variable is Δ_I, and entrepreneurial income in the case the dependent variable is Δ_2).

and across different specifications, whereas γ_1 is either negative or not significant.

The results are robust across specifications and for different definitions of non-financial income—that is, aggregate, labor, and entrepreneurial income considered separately. They also hold for the case in which we focus on investors who have just labor income.

Regarding the economic significance, let us consider the case of total income. We can represent it either in terms of changes in the correlations or directly considering the changes in Δs. In particular, a 1SD increase of the term that accounts for the direct effect of nonfinancial income variance $\left(\frac{Y_{z,t}}{W}\text{var}_{Y_{z},t}\right)$ raises the correlation between nonfinancial and financial income by 5% (8%) for the overall sample (low wealth investors). This corresponds in an increase of the tilt toward more correlated stocks (Δ) equal to 0.51 (0.68) standard deviation of the average value of the tilt for the overall sample (low wealth investors). This translates into portfolio reallocation toward assets correlated with the nonfinancial income equal to 15% (31%) of the risky portfolio for the overall sample (low wealth investors).[12]

A 1SD increase of the term that accounts for the "myopic portfolio" decision $\left(\Theta_{z,t}\right)$ raises the correlation between nonfinancial and financial income by 4% (5%) for the overall sample (low wealth investors). This corresponds to an increase of the tilt toward more correlated stocks (Δ) equal to 0.37 (0.47) standard deviation of the tilt for the overall sample (low wealth investors). This translates into a portfolio reallocation toward assets correlated with the nonfinancial income equal to 14% (17%) of the risky portfolio for the overall sample (low wealth investors).

These findings suggest that we can reject the hypothesis of hedging, overall, for different classes of investors and for both labor and entrepreneurial income. Investors characterized by a negative correlation between their labor income and the market tend to increase their loadings on risky assets that are more closely correlated to their income than the market as a whole. In the case of the low wealth investors, it seems that they systematically act in a fashion opposite to that required by hedging. This is more consistent with the alternative hypothesis of familiarity.

Let us now consider familiarity. The results show that familiarity affects investors' decision to hedge labor income as well as entrepreneurial income. There is a strong negative correlation between active hedging and two measures of familiarity: geographical proximity and the holding period-based proximity. This holds regardless of the level of wealth of the investors and across specifications. These findings confirm the intuition that investors deliberately behave in a way opposite to that required by

[12] The reallocation is defined as the ratio between the change in Δ (which is defined in SEK) and the value of the portfolio invested in risky assets (defined in SEK).

hedging because they want to invest in stocks with which they are familiar. This induces a more positive correlation with their nonfinancial income.

Also, professional proximity is related to hedging both labor and entrepreneurial income risk. This finding can be explained in two ways. On the one hand, the mere fact of working in a specific industry provides the investors with information about the stocks of companies operating in that industry. On the other hand, the decision not to hedge is affected by company-specific constraints such as belonging to a stock plan that invests in the company's stocks.

In terms of the economic magnitude of the impact, an increase in the distance from the company by a factor of 10 (i.e., a move from a 100 km distance to a 10 km distance) increases the correlation between nonfinancial and financial income by 2% (2%) for the overall sample (low wealth investors). The same reduction in distance from the company raises the tilt toward more correlated stocks (Δ) by 0.21 (0.22) standard deviation for the overall sample (low wealth investors). This corresponds to a portfolio reallocation toward assets correlated with the nonfinancial income equal to 7% (10%) of the risky portfolio for the overall sample (low wealth investors).

An increase in the professional proximity by 1SD increases the correlation between nonfinancial and financial income by 0.06% (0.04%) for the overall sample (low wealth investors). This implies an increase in the tilt toward more correlated stocks (Δ) equal to 0.0006 (0.0004) standard deviation for the overall sample (low wealth investors). This corresponds to a portfolio reallocation toward assets correlated with the nonfinancial income equal to 0.2% (0.2%) of the risky portfolio for the overall sample (low wealth investors).

Finally, an increase in the holding period by one year increases the correlation between nonfinancial and financial income by 5% (6%) for the overall sample (low wealth investors). This implies an increase in the tilt toward more correlated stocks (Δ) by 0.17 (0.16) standard deviation for the overall sample (low wealth investors). This corresponds to a portfolio reallocation toward assets correlated with the nonfinancial income equal to 8% (7%) of the risky portfolio for the overall sample (low wealth investors). These findings provide evidence of a sizable economic impact of geographical and holding period-based proximity. Professional proximity, instead, even if statistically significant has a low ecomonic magnitude. Previous literature (Coval and Moskowitz 1999; Hau 2001) have already provided some evidence of a geographical proximity bias, possibly related to having better information about local companies. Here, we clearly identify the main dimension of proximity relevant to investment—geographical proximity as opposed to professional proximity—relating it to the decision to hedge nonfinancial income.

It is interesting to note that θ_1 is always significant and positive in the case of labor income and significant and negative in the case of entrepreneurial income. This suggests that the selection bias due to endogenous stock market participation is significant. The omission of λ would mean that the least squares methodology overestimates the value of the relationship between the hedging index and the nonfinancial income variables in the case of labor income hedging and underestimates it in the case of entrepreneurial income hedging. In other words, the people who participate are more likely to hedge labor income and less likely to hedge entrepreneurial income. This bias is properly accounted for by including λ_t.

Finally, three points are worth stressing. First, our results are not affected by the share of the labor force employed by the government. A significant proportion of Swedish households is employed by the government (either national or municipal). This suggests that their incomes are much less variable and less correlated with the return on the stock market and the risk of unemployment different. Therefore, as an additional robustness check, we also perform the analysis on two separate subsamples: government employees and nongovernment employees.[13] The results are reported in Table 5, Panels A for the case of the private sector and Panel B for the case of the public sector. The findings are consistent with the previous results. Investors employed in the private sector do not hedge, and the decision to invest in assets with payoffs closely related to their income is strongly affected by familiarity. Public employees, on the other hand, seem to be less affected by familiarity. Indeed, both low- and high wealth investors do not seem to be affected by any proximity variable.

Second, our results are robust to the way we treat the mutual funds. To guarantee this, we re-estimate our specifications for the subset of investors who do not hold funds, and the results are consistent with those that are reported.

Third, taking full advantage of the data, we also analyze whether when an investor buys a new stock or closes his position in a stock he increases or decreases the correlation between his nonfinancial wealth and financial wealth. Essentially, we test whether the stocks that the investor initiated (terminated) during a given year are the ones that increase (decrease) the correlation between financial and nonfinancial wealth. This provides an additional insight and a powerful robustness check. To address this issue, we directly focus our analysis at the level of the individual stock and we consider the following simplified reduced form:

[13] We define as public sector households those whose share of income coming from occupations with SNI codes between 75,000 and 91,999 is in excess of 50%. This provides approximately 192,000 investors belonging to the private sector and 58,000 belonging to the public sector. We focus only on labor income as entrepreneurial income is very rare among public sector employees.

Table 5
Additional evidence

	All households				Low wealth households				High wealth households			
	I		II		I		II		I		II	
Variable	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	*t*-stat
Panel A: determinants of Δ_1 for *private sector employees*												
$Y_L$var(Y_L)	−0.16	−7.20	−0.17	−8.07	−0.21	−11.22	−0.21	−10.95	−6.71	−2.94	−5.56	−2.63
$Y_E$cov(Y_E, Y_L)	−1.50	−14.53	−1.49	−15.21	−1.39	−16.60	−1.32	−15.01	−2.58	−1.15	−1.83	−0.88
Θ_L	−0.94	−11.51	−0.93	−12.00	−0.86	−12.53	−0.81	−11.51	−0.58	−3.45	−0.49	−3.16
Geographical proximity			−3.40	−7.71			−3.93	−6.60			−0.85	−3.15
Professional proximity			−2.63	−3.27			−2.38	−2.22			−0.67	−0.75
Holding period			−5.89	−5.42			−6.81	−5.26			−0.47	−2.85
Control variables	Yes		Yes		Yes		Yes		Yes		Yes	
Adj R^2	0.02		0.03		0.02		0.03		0.02		0.03	
Panel B: determinants of Δ_1 for *public sector employees*												
$Y_l$var(Y_L)	−0.49	−17.67	−0.49	−17.83	−0.65	−21.91	−0.65	−21.97	0.40	0.09	0.43	0.10
$Y_e$cov(Y_E, Y_L)	−0.75	−9.72	−0.76	−9.98	−0.86	−10.83	−0.86	−10.92	−4.65	−1.38	−2.14	−0.67
Θ_L	−0.19	−26.47	−0.19	−26.98	−0.19	−25.60	−0.19	−25.77	−0.26	−0.78	−0.21	−0.69
Geographical proximity			−0.46	−4.15			−0.25	−1.64			−2.02	−1.91
Professional proximity			1.18	0.87			0.60	0.32			−1.63	−0.76
Holding period			−0.08	−0.76			−0.12	−0.92			−3.34	−0.89
Control variables	Yes		Yes		Yes		Yes		Yes		Yes	
Adj R^2	0.02		0.03		0.02		0.03		0.02		0.03	
Panel C: initiations/terminations and increase/decrease the correlation between financial and nonfinancial wealth												
Corr(income, stock return)	5.02	3.60	2.03	4.93	1.66	2.91	0.90	2.60	7.53	3.18	2.57	3.57
Control variables	Yes		Yes		Yes		Yes		Yes		Yes	
Adj R^2	0.01		0.01		0.01		0.01		0.01		0.01	

This table contains estimates of the determinants of the tilt for private sector (panel A) and public sector (panel B) employees. We define as public sector employees those whose share of income coming from occupations with SNI codes between 75000 and 91999 is in excess of 50%. The notations are as in Table 4. We consider labor income. We use the same control variables as in Table 4. The Δs and Ys are measured in thousands of SEK. In panel C, we estimate $\Delta NB_{ijt} = \alpha + \beta C_{ijt} + \gamma D_{ijt} + \epsilon_{ijt}$, where for individual i, stock j at time t, ΔNB_{ijt} represents the net investment in stocks [change in the number of stocks weighed by the stock price at the beginning of the period $(t-1)$] for specification I and net change in number of shares for specification II, C_{ijt} is a variable that reports the correlation between the stock and the nonfinancial source of income of the investor (i.e., total income). D_{ijt} is a vector of control variables defined as in Table 2 and augmented by stock-specific variables. These are the logarithm of market capitalization (size), the market-to-book ratio, leverage, the stock returns in the previous twelve months, the stock bid-ask spread (per share, in SEK), the beta of the stock with respect to the market and the residual risk.

$$\Delta NB_{ijt} = \alpha + \beta C_{ijt} + \gamma D_{ijt} + \varepsilon_{ijt},$$

where for individual i, at time t, ΔNB_{ijt} represents the net investment in stock j, C_{ijt} is the correlation between the stock and the nonfinancial source of income of the investor. In order to make the test clean of price effects, we construct ΔNB_{ijt} as the change in the number of stocks weighed by the stock price at the beginning of the period ($t - 1$). We also consider a specification based on the change in the number of shares. D_{ijt} is a vector of control variables that contains all of the alternative factors that may affect the portfolio choice of the ith investor in the jth stock (e.g., labor income risk, level of wealth, demographic characteristics, etc.). We include control variables that account for the standard portfolio choice in the presence of risky nonfinancial income (i.e., correlation of financial and nonfinancial income, volatility, and conditional value of nonfinancial income) and variables that proxy for the degree of informativeness and sophistication of the investors. These variables allow us to control for the investor's view and to focus directly on the impact of the biases. Moreover, given that now the estimation is performed at the stock level, we also use stock-specific variables as control. These are the logarithm of market capitalization (size), the market-to-book ratio, leverage, the stock returns in the previous 12 month, the stock bid-ask spread (per share, in SEK), the beta of the stock with respect to the market and the residual risk. We consider only the positions that have been initiated/terminated during a given year.

The results are reported in Table 5, Panel C. We focus on the coefficient β. If $\beta > 0$, the investor is deliberately buying shares in order to increase the correlation of his portfolio with nonfinancial income. We consider a specification for the entire sample and one broken down in low- and high wealth investors. The results are consistent with our working hypothesis as $\beta > 0$ is positive and statistically significant. These results are robust across different specifications and across the different samples. They suggest that the stocks that the investor initiated (terminated) during a given year are the ones that increase (decrease) the correlation between financial and nonfinancial wealth and that the tilt in the portfolio is indeed a deliberate act of the investor.

4.3 H3: the nature of familiarity

We now study the nature of familiarity by investigating in more detail the heterogeneity across investors. For brevity, we directly focus on Equation 11 by including interactive dummies to account for cross-sectional differences across investors. We recall that if familiarity is information driven, we expect it to differ across differentially informed investors and to change after familiarity shocks. As explained in Section 4.1, we use the wealth of the investor to identify the informed investors. We estimate

$$\begin{aligned}\Delta_{z,t} =& \alpha_2 + \beta_2 \frac{Y_{z,t}}{W} \mathrm{var}_{Y_{z,t}} + \gamma_2 \frac{Y_{x,t}}{W} \mathrm{Cov}_{Y_z, Y_{x,t}} \\ &+ \zeta_2 \frac{1}{W} \Theta_{z,t} + \boldsymbol{\nu}_2 \boldsymbol{\Psi}(1 - \xi_1) \\ &+ \pi_2 \boldsymbol{\Psi} \xi + \boldsymbol{\delta}_2 \mathbf{F}_{2,t} + \theta_2 \lambda_t + \varepsilon_{2,t}.\end{aligned} \tag{12}$$

In the case of differences in wealth, ξ is a continuous variable based on the amount of wealth of the investor and $\xi_1 = 0$. In the case of familiarity shocks, instead, ξ is a dummy that takes the value 1 if in the previous three years the investor has been subject to a "familiarity shock" and 0 otherwise and $\xi_1 = \xi$.

In particular, in the case of unemployment shocks, ξ is a dummy that takes the value 1 if the investor has been unemployed at least once in the previous three years and 0 otherwise. In the case of professional change shocks, ξ is a dummy that takes the value 1 if the investor has changed his profession and the current profession differs from the previous one and 0 otherwise. In the case of relocation shocks, ξ is a dummy that takes the value 1 if the investor has moved and changed country (municipality) at least once in the previous three years and 0 otherwise.

We first focus on how the impact of familiarity differs across differentially informed investors. We use the level of wealth as a proxy for the degree of informativeness. The results are reported in Table 6, Panel A. In this case ν_2 represents the sensitivity to familiarity of investors in general and π_2 proxies for the differential effect of being more informed (i.e., wealthier). The first thing to note is that the impact of familiarity on the decision not to hedge is decreasing in wealth. This holds for the case of geographic, professional, and holding period-based familiarity. There is evidence of a statistical difference between the two classes of investors. This suggests that wealthier investors rely less on the information based on professional proximity because they can afford better quality information.

Let us now consider the familiarity shocks. In this case, ν_2 represents the sensitivity to familiarity of investors who have not been affected by shocks, whereas π_2 represents the sensitivity to familiarity of investors who have been affected by shocks. As we mentioned before, we focus on three types of shocks: unemployment shocks, changes of profession shocks, and relocation shocks. Given that these shocks are likely to be correlated—that is, a change in profession may induce the investor to move to another town—we expect each shock to affect all the three sources of familiarity and not just the one to which the shock appears to be more related (i.e., the relocation shock should be mostly related to geographically based familiarity). The results are summarized in Table 6,

Table 6
The nature of familiarity

	Labor income		Entrepreneurial income		Total income	
Variable	Value	t-stat	Value	t-stat	Value	t-stat
Panel A: wealth and tilt of the portfolio						
$Y_i\text{var}(Y_i)$	−0.27	−5.78	−0.48	−31.01		
$Y_i\text{cov}(Y_i, Y_j)$	−1.48	−10.99	0.18	0.88		
Θ_I	−1.05	−10.08	−0.36	−43.74		
$Y_T\text{var}(Y_T)$					0.03	0.75
Θ_T					−0.13	−4.35
Geographical proximity	−29.77	−2.65	−57.91	−3.74	−68.33	−7.52
Geographical proximity* WEALTH	3.78	2.14	9.12	3.79	10.22	6.95
Professional proximity	−183.34	−4.30	−359.41	−4.42	−336.18	−2.60
Professional proximity* WEALTH	29.27	4.30	57.24	4.41	116.54	2.88
Holding period	−40.83	−1.99	−119.39	−3.80	−113.41	−6.77
Holding period* WEALTH	4.76	2.04	20.06	3.81	16.96	6.09
Adj R^2	0.02		0.01		0.02	
Panel B: unemployment and tilt of the portfolio						
$Y_i\text{var}(Y_i)$	−0.27	−11.68	−0.42	−32.17		
$Y_i\text{cov}(Y_i, Y_j)$	−1.42	−16.06	−0.49	−3.69		
Θ_I	−0.80	−11.76	−0.34	−47.97		
$Y_T\text{var}(Y_T)$					−0.08	−4.21
Θ_T					−0.28	−5.62
Geographical proximity* (1-UNEMPL)	−4.84	−10.83	−2.62	−4.64	−2.76	−4.34
Geographical proximity* UNEMPL	1.41	2.64	−0.01	−0.03	0.84	4.46
Professional proximity* (1-UNEMPL)	−2.18	−9.40	−0.77	−2.46	−2.52	−4.76
Professional proximity* UNEMPL	2.93	9.36	0.38	1.43	3.43	4.75
Holding period* (1-UNEMPL)	−10.29	−9.00	−6.75	−5.69	−0.29	−2.13
Holding period* UNEMPL	−0.58	−3.17	0.78	2.52	1.61	4.05
Adj R^2	0.02		0.01		0.01	
Panel C: professional change and tilt of the portfolio						
$Y_i\text{var}(Y_i)$	−0.27	−12.26	−0.43	−29.88		
$Y_i\text{cov}(Y_i, Y_j)$	−1.41	−16.54	−0.58	−3.91		
Θ_I	−0.80	−12.10	−0.34	−43.92		
$Y_T\text{var}(Y_T)$					−0.09	−2.40
Θ_T					−0.36	−2.85
Geographical proximity* (1-PCHANGE)	−7.19	−10.13	−4.08	−2.78	−8.42	−9.71
Geographical proximity* PCHANGE	1.07	2.97	6.29	1.66	3.20	7.30
Professional proximity* (1-PCHANGE)	−3.32	−3.83	−4.50	−2.16	−3.72	−3.50
Professional proximity* PCHANGE	15.64	5.76	2.49	2.11	17.78	5.28
Holding Period* (1-PCHANGE)	−14.00	−8.51	−7.34	−2.01	−20.25	−10.10
Holding Period* PCHANGE	2.41	8.41	14.58	1.80	2.90	8.19
Adj R^2	0.02		0.01		0.01	
Panel D: relocation and tilt of the portfolio						
$Y_i\text{var}(Y_i)$	−0.28	−12.75	−0.44	−30.41		
$Y_i\text{cov}(Y_i, Y_j)$	−1.50	−17.20	−0.50	−3.44		
Θ_I	−0.86	−12.76	−0.35	−44.63		

Table 6
(continued)

	Labor income		Entrepreneurial income		Total income	
Variable	Value	*t*-stat	Value	*t*-stat	Value	*t*-stat
$Y_T\text{var}(Y_T)$					−0.05	−2.38
Θ_T					−0.35	−7.80
Geographical proximity* (1-MOVER)	−6.03	−11.80	−7.30	−6.39	−2.25	−6.59
Geographical proximity* MOVER	0.68	1.82	16.64	3.12	−1.00	−4.02
Professional proximity* (1-MOVER)	−3.83	−4.97	−6.82	−3.85	−3.04	−3.74
Professional proximity* MOVER	66.85	6.20	0.03	0.57	20.75	2.36
Holding period* (1-MOVER)	−9.93	−8.37	−8.00	−3.15	−0.93	−2.88
Holding period* MOVER	4.78	8.86	39.69	3.58	0.42	1.66
Adj R^2	0.02		0.02		0.02	

In the case of labor and entrepreneurial income we estimate that

$$\Delta_{z,t} = \alpha_2 + \beta_2\left(\frac{Y_{z,t}}{W}\right)\text{var}\,Y_{z,t}+\gamma_2\left(\frac{Y_{x,t}}{W}\right)\text{cov}\,Y_z,Y_x+\varsigma_2\left(\frac{\Theta_{z,t}}{W}\right)+\nu_2\Psi(1-\xi_1)+\pi_2\Psi\xi+\delta_2F_{2,t}+\theta_2\lambda_t+\varepsilon_{2,t},$$

where $\Delta_{z,\ t}$ = cov(labor income, market) – cov(labor income, portfolio) in the case of labor income and $\Delta_{z,\ t}$ = cov(entrepreneurial income, market) – cov(entrepreneurial income, portfolio) in the case of entrepreneurial income. $Y_{z,\ t}$ is labor (entrepreneurial) income in the case of labor (entrepreneurial) income. In the case of total income we estimate that

$$\Delta_{T,t} = \alpha_2 + \beta_2(Y_{T,t}/W)\text{var}\,Y_{T,t} + \varsigma_2(\Theta_{T,t}/W)+\nu_2\Psi(1-\xi_1) + \pi_2\Psi\xi + \delta_2F_{2,t} + \theta_2\lambda_t + \varepsilon_{2,t},$$

where $\Delta_{T,\ t}$ = cov(total income, market) – cov(total income, portfolio) and $Y_{T,\ t}$ is total income. W is the level of wealth while Θ_Z is $-(Y_L+Y_E)\Sigma_j\Omega_{Sj}\times\text{cov}(\text{ret}_j, Y_Z)$, where Y_Z is labor (entrepreneurial) income in the case of labor (entrepreneurial) income and $\Omega_{\text{sj}} = (\mu_{sj}-r)'\Big/\Big[(1-\gamma)\sigma^2_{sj}\Big]$. μ_{sj} and σ^2_{sj} are, respectively, the mean and the variance of the *jth* risky asset returns while cov(Y_E, Y_L) is the covariance between labor and entrepreneurial income. $\mathbf{F}_2$ is a vector of control variables, Ψ is a vector of proximity variables, and λ_t is Heckman's lambda. The Δs and Ys are measured in thousands of SEK. In the case of differences in wealth, ξ is a continuous variable based on the amount of wealth of the investor and $\xi_1 = 0$. In the case of familiarity shocks, ξ is a dummy that takes the value 1 if in the previous three years the investor has been subject to a "familiarity shock" and 0 otherwise and $\xi_1 = \xi$. In panel A, we use the logarithm of net wealth as a proxy for wealth (i.e., $\xi_1 = 0$, ξ = WEALTH). The estimates are done for the whole sample. We define as familiarity shocks events that change investor's proximity to the stocks. In panel B, we use an unemployment dummy ($\xi_1 = \xi$ = UNEMPL) that takes the value 1 if the investor has been unemployed at least once in the previous three years and 0 otherwise. In panel C, we use a professional change dummy ($\xi_1 = \xi$ = PCHANGE) that takes the value one if the investor has changed profession at least once in the previous three years and the current profession differs from the previous 1 and 0 otherwise. In panel D, we use a relocation dummy ($\xi_1 = \xi$ = MOVER) that takes the value 1 if the investor has moved and changed county (municipality) at least once in the previous three years and the current address differs from the previous one and 0 otherwise. In panels B, C, and D the estimates are done for the full sample. The familiarity variables are *professional proximity, geographical proximity*, and *holding period*. They are constructed as follows. For each stock in the portfolio we identify three measures of familiarity and then we aggregate them for each investor on the basis of his portfolio composition (i.e., using as weights the value of the portfolio holding). *Professional proximity* is a dummy taking the value 1 if the investor's profession is in the same area of activity as the company whose stock is under consideration and 0 otherwise. We use the one-digit SNI92 codes (similar to SIC codes) to identify the areas of activities. *Geographical proximity*, that is the proximity between the residence of the investor and the place where the company is located. We use the logarithm of the inverse of the distance between the ZIP code of the investor and the ZIP code of the closest branch/subsidiary of the company whose stock we consider. *Holding period* is based on the time a stock entered the investor's portfolio. Each index of familiarity is constructed weighting the measures for each investor on the basis of his portfolio composition. *Distributional dummy* is a dummy that takes the value 1 if Γ_i, > 0 and 0 otherwise (i = total income in the case the dependent variable is Δ_T, labor income in the case the dependent variable is Δ_I and entrepreneurial income in the case the dependent variable is Δ_2). We use the same control variables as in the previous tables (omitted for brevity).

Panel B for the case of shocks due to unemployment, Panel C for the case of shocks due to professional change, and Panel D for the case of shocks due to relocation.

The results show that familiarity shocks in general change investors's sensitivity to familiarity. (Unreported) tests show that the values of the coefficients ν_2 and π_2 are statistically different from one another. An investor who has changed profession or has been unemployed recently is less subject to a professional bias. This is expected, as such an investor has not yet had the time to absorb the information embedded in his new profession. Analogously, an investor who has changed his address is less subject to geographical proximity than an investor who has not moved. This is consistent with the intuition that relocation affects the availability of geographically based information. After a move it takes time for the investor to accumulate the same information on closer stocks that he used to have before the move. In many cases, the shocks not only reduce the impact of familiarity but also seem to increase hedging. This is consistent with the fact that shocks increase the informational uncertainty of the investor, raising his desire to hedge. As an additional robustness check we re-estimate the main specification including proxies for the frequency of rebalancing.[14] The results do not differ from the reported ones.

Finally, the comparison between movers and nonmovers is also redone using a specification based on changes in portfolio holdings. In this case, the dependent variable is either the percentage change of the investment in risky assets over overall wealth or the change in dollar value of the investment. The explanatory variables are the familiarity variables, the Γs and the control variables as defined in Table 2. The familiarity variables are interacted with two relocation dummies: the first dummy takes the value 1 if the investor has moved and changed country at least once in the previous three years and the current address differs from the previous one and 0 otherwise. The second dummy is the complement to 1 of the first dummy. The results (not reported) are consistent with the reported ones and show that the impact of familiarity decreases in the presence of relocation shocks.

These findings support the idea that the impact of familiarity depends on the degree of informativeness of the investor. In other words, more informed investors are less affected by familiarity—and on its stability over time—that is, familiarity shocks modify the impact of familiarity on investor behavior. This suggests that the investment choice is driven by the availability of information and that familiarity is a substitute for better information. Its importance decreases when the investor has access

[14] We consider two alternative proxies for rebalancing: either a dummy for whether the investor has ever been rebalancing in the previous two years or a proxy for the number of times the investor has been rebalancing over the year in the different stocks.

to more information or if there is a perturbation to his source of familiarity. These results are also consistent with the findings of Bodnaruk (2003). Analyzing the portfolios of individual investors that changed their place of residence, he finds that the further away investors move from the company's closest establishment, "the more of its stock they abnormally sell relative to the investors that do not move." He also documents "that originally held stocks, which holdings have not increased after the move, are more distant and provide lower return to the investors in their new location than stocks acquired after the move and originally held stocks, which holdings increased after the move."

Finally, it is worth noting that unreported analysis focusing on the wealthiest top 1% of the population shows that almost all dimensions of the previous results are irrelevant to the very rich. This suggests not only that the very rich are less subject to behavioral biases, but also that the standard rational portfolio theory based on hedging nonfinancial income does not explain their behavior.

4.3.1 The aggregate picture. These findings allow us to say something regarding the "aggregate investor." We saw that a random investor does not hedge, but does the wealth-weighted investor hedge? In our sample, the high wealth investors represent a mere 8% overall but on a value basis they represent 73.42% of the assets. Moreover, the low wealth and high wealth investors are almost evenly split between high (in absolute value greater than 0.2) and low (less than 0.2) correlation with the market. What do these elements tell us?

Let us consider total income. In Figure 2, we report the distribution of the shareholdings of the different classes of investors broken down according to investors' measure of hedging (Γ_T). Each histogram represents the value of the stock market held by the investors with a specific tendency to hedge/go for familiar stocks. Investors are grouped according to those who invest in "close" and "distant" stocks (right and left columns, correspondingly). The distributions are reported separately over subsamples of low wealth and high wealth investors (Figures 2*a* and 2*b*, correspondingly) and for the overall sample (Figure 2*c*).

We see that, among the high wealth investors, the distribution is slightly skewed toward investors who hedge. In contrast, for the low wealth investors, the distribution is strongly skewed toward familiarity-driven investors. How does this aggregate? In aggregate, the fact that the high wealth investors own more than 73% of the assets offsets the bend toward familiarity of the low wealth investors. This implies that, even if the low wealth investors represents 92% of the population, their representation at the aggregate level is very tiny. The wealth-weighed investor is less subject to familiarity and behaves more in line with what theory would dictate.

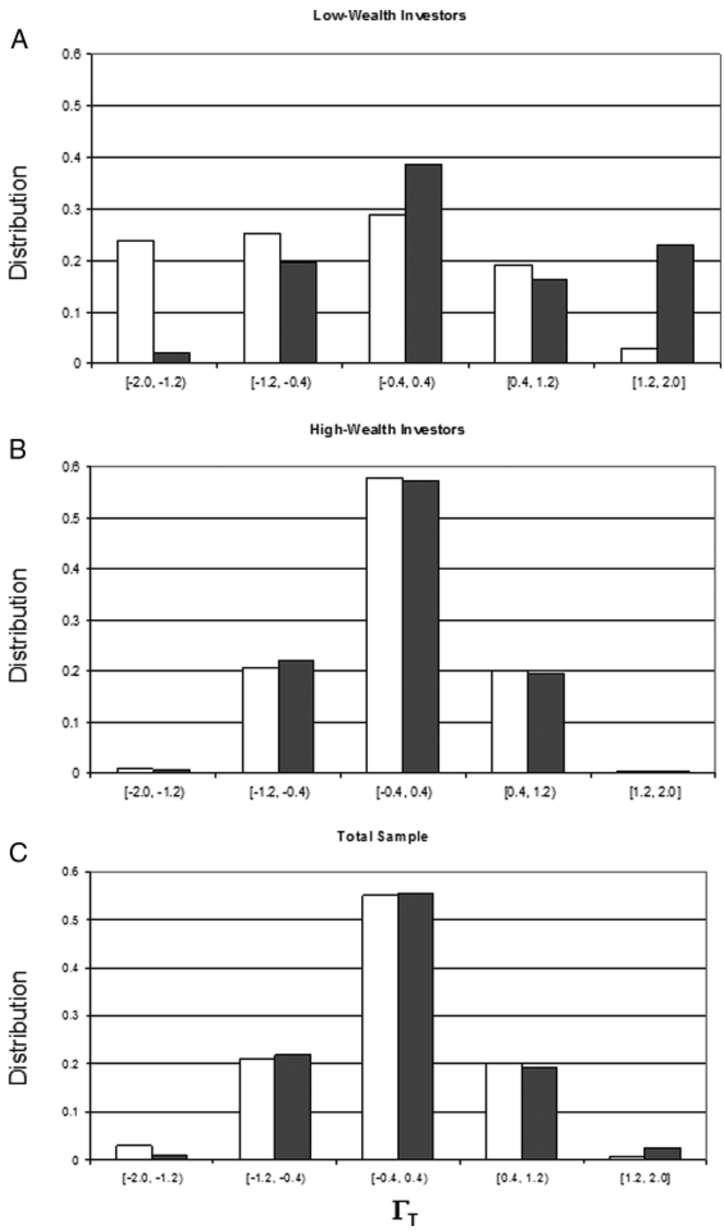

Figure 2
Distribution over index of hedging for total income (Γ_T) *of stock holdings* of investors grouped into the ones who invest in "close" and "distant" stocks (right and left columns, correspondingly). The distributors are reported separately over subsamples of low wealth and high wealth investors (a and b, correspondingly) and for overall sample (c). We use five ranges for Γ_T: from –2 to –1.2, from –1.2 to –0.4, from –0.4 to 0.4, from 0.4 to 1.2 and from 1.2 to 2.

4.4 Cost of hedging/familiarity

The decision to implement hedging as opposed to familiarity-driven strategies should also have direct implications in terms of profitability. We expect that if familiarity is information based, once we control for the level of private information of the investor (e.g., proxied by wealth), familiarity should increase the profitability while hedging, being a sort of costly insurance, should reduce it.

We consider two measures of "profits": the financial gains/losses in the year standardized by the value of the risky assets at the beginning of the year (Π_F) and the change in wealth standardized by the value of wealth at the beginning of the year (Π_w). Financial gains/losses include the realized capital gains/losses and the dividends. Both measures may be subject to criticism and must be taken with a pinch of salt. Financial gains/losses represent an imprecise measure if investors do not turn over their positions regularly. A change in wealth does not account properly for the saving decisions of the investors. Given that both biases may be related to the income of the investor, we include both the level and the variance of investors' income and wealth among the control variables.

Gains and losses are directly reported by the financial intermediary through which the transaction is executed reduces the potential bias due to underreporting of capital gains for tax purposes. Other biases may be the result of the "lock-in effect" that generates different incentives to sell the stocks for investors with different marginal tax rates and clientele effects. However, for all these biases (underreporting, lock-in, and clientele) Yitzhaki (1987) shows that "the likely effects of these biases is to cause an underestimate of the observed differences in rates of return among income classes." Therefore, the bias acts *against* the possibility of actually finding a statistically significant difference between classes.

Descriptive statistics of the profits and their differences between groups of investors and tests of the differences are reported in Table 7, Panels A and B. Bootstrapping has been employed to assess the robustness of the tests. As expected, the high wealth investors make more profits than the low wealth ones. This holds for the statistics based on the mean as well as those based on the median. In Table 7, Panel C, we compare the profits of the investors who hedge (i.e., $\Gamma > 0$) and those who do not hedge (i.e., $\Gamma < 0$). Hedgers, in general, earn lower profits than nonhedgers. This suggests that hedging is indeed expensive as we would expect it to be in equilibrium. Alternatively, this also means that familiarity provides the low wealth investors with a cheap source of information.

Note that this comparison is done *separately* for both the high wealth and the low wealth investors, and, therefore, it is not inconsistent with the fact that the high wealth investors make more profits than do the low wealth ones. It is also interesting to note that the cost of hedging (i.e., the

Table 7
Profits of high wealth and low wealth households

		Mean		Median		*t*-test		Wilcoxon Test		Kolmogorov–Smirnov Test	
Variable	Hedging	Low wealth	High wealth	Low wealth	High wealth	*t* value	*p* value	*Z*	*p* value	K_{sa}	*p* value
Panel A: profit measures											
Π_{Ft}		0.13	0.22	0.00	0.06	41.94	<.0001**	120.39	<.0001**	62.83	<.0001**
Π_{Wt}		0.05	0.10	0.07	0.08	23.43	<.0001**	11.34	<.0001**	16.89	<.0001**
Panel B: profit measures (low versus high wealth households)											
Π_{Ft}	Yes	0.12	0.22	0.00	0.05	36.91	<.0001**	96.40	<.0001**	48.63	<.0001**
Π_{Ft}	No	0.19	0.23	0.05	0.07	7.44	<.0001**	31.92	<.0001**	18.66	<.0001**
Π_{Wt}	Yes	0.04	0.08	0.06	0.07	17.72	<.0001**	3.56	0.0002**	14.76	<.0001**
Π_{Wt}	No	0.09	0.13	0.09	0.10	6.11	<.0001**	2.55	0.0054*	9.01	<.0001**
Variable	Wealthy	Hedgers	Nonhedgers	Hedgers	Nonhedgers	*t* value	*p* value	*Z*	*p* value	K_{sa}	*p* value
Panel C: profit measures (hedgers versus nonhedgers)											
Π_{Ft}	Yes	0.22	0.230	0.05	0.07	3.15	0.0020**	34.40	<.0001**	22.44	<.0001**
Π_{Ft}	No	0.12	0.194	0.00	0.05	19.99	<.0001**	60.75	<.0001**	32.96	<.0001**
Π_{Wt}	Yes	0.08	0.091	0.07	0.10	19.48	<.0001**	24.09	<.0001**	13.62	<.0001**
Π_{Wt}	No	0.04	0.131	0.06	0.09	7.62	<.0001**	17.47	<.0001**	12.78	<.0001**

We report tests of the difference of two measures of profits for high wealth and low wealth households. The measures are $\Pi_{Wt} = (wealth_t/wealth_{t-1} - 1)$ and $\Pi_{Ft} = [(capital\ gain/losses_t + dividends_t)/Risky\_Assets_{t-1})$, where capital gains and losses are the reported realized gains/losses of the household in risky assets at yeat *t* and $RISKY\_ASSETS_t$ represents the amount invested in risky assets at *t*. Panel A reports *t-tests, Wilcoxon* and *Kolmogorov–Smirnov tests* of the equality of profits between high wealth and low wealth households. For the *t-tests* we used *Satterwhaite* version of the *t-tests* that assumes inequality of variances in two subsamples. In all cases equality of variances was rejected at the 1% level. In panels B and C, households are separated into two groups: the "hedgers" (defined as the households with positive Γ_1) and nonhedgers (households with negative Γ_1). Panel B reports the tests of difference of profits between high and low wealth households after they have been separated into hedgers and nonhedgers. Panel C reports the tests of difference of profits between hedgers and nonhedgers performed separately for high wealth and low wealth households. The asterisks denote that the statistics are robust to bootstrapping on 0.1% ("**") and 1% ("*") level. Bootstrapping is based on 20,000 resamplings.

difference in profits between the hedgers and the nonhedgers) is higher for the low wealth investors than for the high wealth ones. This fits with the intuition that high wealth investors have access to better financial services and advice.

To assess the relationship between profits and hedging/familiarity strategies, we estimate

$$\Pi_t = \alpha_3 + \beta_3\Gamma + \boldsymbol{\gamma}_3\boldsymbol{\Psi}_t + \boldsymbol{\delta}_3\mathbf{F}_t + \theta_3\lambda_t + \varepsilon_{3,t}, \tag{13}$$

where Π_t are the profits realized at time t by the investor and Γ is the index of hedging for the total nonfinancial income. The results are reported in Table 8, Panel A, for the case of profits based on changes of wealth (Π_w) and, Panel B, for the case of profits based on financial gains/losses (Π_F). Note that the fact that we control for the level of wealth (both in the definition of the sample and by including it among the control variables) allows us to condition on the level of private information of the investor. We expect that $\beta_3 < 0$ and $\gamma_3 > 0$.

The results show that there is a negative or nonsignificant correlation between hedging and profits (β_3) and a positive correlation between profits and geographical proximity and holding period proximity (γ_3). This holds, regardless of the level of wealth of the investors, across specifications and for different classes of investors. Professional proximity, on the other hand, does not seem to be related to profits. A 1SD increase in Γ (i.e., an increase in hedging) reduces profits by 0.74% for the overall sample. Similarly, an increase in the proximity to the stock by a factor of 10 (i.e., a move from 100 to 10 km) increases financial profits by 5.80% for the overall sample (by 1.20% for low wealth investors and by 3.50% for high wealth investors). An increase in the holdings period of one year raises financial profits by 3.70% for the overall sample (by 0.40% for low wealth investors and by 4.60% for high wealth investors).

As additional robustness check, we also re-estimate the impact of familiarity on both volatility of profits and the Sharpe ratio of profits. The results (not reported) show that, in the case of low wealth investors, geographical proximity increases both profits and risk. This suggests that the information contained in the proximity to the stock may help increase returns at the expense of taking on more risk. Holding period increases profits and reduces risk. The effect in terms of net-of-risk return is positive. Indeed, both measures of familiarity increase profits net-of-risk (Sharpe ratio). The results hold for both measures of profits (wealth based and financial based). In the case of the high wealth investors, geographical proximity and holding period increase profits, volatility, and the Sharpe ratio only for the case of wealth-based profits, and the impact is lower than the one for low wealth investors.

Table 8
Determinants of Profits

	All households				Low wealth households				High wealth households			
	I		II		I		II		I		II	
Variable	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	*t*-stat	Estimate	*t*-stat
Panel A: determinants of Π_{Ft}												
Γ_T	0.26	0.34	−0.09	−2.76	0.06	1.14	0.08	1.83	−2.96	−2.85	−0.05	−1.21
Geographically proximity			0.58	8.90			0.12	15.61			0.35	2.72
Professional proximity			−0.27	−1.41			0.10	0.72			−0.06	−0.23
Holding period			0.37	6.62			0.04	3.17			0.46	4.18
Income (variance)	0.29	10.04	0.27	7.38	0.30	2.16	0.30	1.86	0.09	1.21	0.02	0.56
Control variables	Yes		Yes		Yes		Yes		Yes		Yes	
Adj R^2	0.01		0.01		0.01		0.01		0.01		0.01	
Panel B: determinants of Π_{Wt}												
Γ_T	−0.10	−3.26	−0.09	−2.88	−0.13	−2.79	−0.13	−2.68	−0.06	−2.95	0.04	1.55
Geographically proximity			0.12	12.47			0.12	8.35			0.53	7.72
Professional proximity			0.16	0.97			0.36	1.29			−0.11	−0.82
Holding Period			0.24	8.98			0.18	8.29			0.31	5.20
Income (Variance)	−0.12	−5.48	−0.12	−5.55	−0.15	−4.98	−0.15	−4.99	−0.05	−2.52	−0.06	−3.11
Control variables	Yes		Yes		Yes		Yes		Yes		Yes	
Adj R^2	0.02		0.02		0.02		0.02		0.02		0.02	

We report estimates for the two measures of profits defined in Table 7. Methodology and variables are identical to those described in Tables 2, 4, and 5. Estimates are based on Heckman correction and 2SLS. All estimates are multiplied by 10. The control variables are the same as the ones used in Table 2. Π_{Ft} and Π_{Wt} are defined as in Table 7 and Γ_T = corr(total income, market) – corr(total income, portfolio). The familiarity variables are defined as in Table 4.

All these findings are consistent with the intuition that familiarity is, for the low wealth investors, a proxy for cheap information that allows the low wealth investors to have higher (both gross and net of risk) profits. This is not the case for the high wealth investors who, presumably, have access to private information. Hedging, rather, appears to be expensive.

Also in this case, it is worth noting that unreported analysis focusing on the wealthiest top 1% of the population finds no significant relation between profits and hedging or familiarity-driven behavior.

5. Conclusion

We study whether investors use their investment in financial assets to hedge their nonfinancial income using a new approach based on the inspection of the relationship between investor nonfinancial and financial income and a portfolio-based index of hedging. We show that investors do not engage in hedging but deliberately tilt their portfolio toward stocks that are most closely related to them. We rationalize this in terms of "familiarity." We then investigate the nature of familiarity and show that it is information based. Familiarity-driven behavior is a way of conditioning on a cheap source of information for the investors.

There is heterogeneity in investor behavior. High wealth investors are more likely to hedge than the low wealth ones. In the case of the high wealth investors, the impact of geographical proximity and holding period on profits is lower and less significant than in the case of low wealth investors. Almost all the dimensions of the results discussed in the article are irrelevant to the very rich.

These findings challenge the standard portfolio theory and, at the same time, generate evidence to extend it, by shedding new light on investor behavior. The identification of the determinant of familiarity has practical relevance, given that the information-based familiarity hypothesis and the pure familiarity hypothesis have different normative and operational implications. Behavioral biases are related to human characteristics and are equally likely to be present in different countries and across markets. Informational constraints and market frictions are, on the other hand, more likely to be affected by institutional as well as endowment differences. If familiarity is information based, we may expect it to lose importance as the degree of sophistication of the investors or their relative wealth increases. Therefore, processes such as globalization and financial integration, by increasing information, should reduce the impact of familiarity on investors' choices and therefore on asset prices.

Appendix

We develop the testable restrictions on the tilt in the portfolio. We rely on the standard literature on portfolio choice with multiple assets. We consider an economy with n risky assets denoted by S and a riskless asset B. The riskless asset earns an instantaneous interest rate $r > 0$ while the risky securities follow a geometric Brownian motion, such that

$$dB = rBdt \text{ and } \boldsymbol{dS} = \mathrm{I}_s \boldsymbol{\mu} dt + \mathrm{I}_s \boldsymbol{\Sigma} \mathbf{dw}, \tag{14}$$

where, $\mathbf{w}$ is an n-dimensional standard Brownian motion, $\mathbf{I}_s$ is an n-dimensional diagonal matrix with the risky securities prices as entries, $\boldsymbol{\mu}$ is an n-dimensional vector of mean returns, and $\boldsymbol{\Sigma}$ is the matrix of diffusion coefficients. We assume $\boldsymbol{\Sigma}$ to be diagonal, that is, markets are complete and each asset loads only on a specific source of uncertainty. The representative investor has other nonfinancial sources of income $(\mathbf{Y})$:

$$d\mathbf{Y} = \mathbf{I}_y \mathrm{a} dt + \mathbf{I}_y s d\mathbf{w}, \tag{15}$$

where $\mathbf{I}_y$ is an n-dimensional diagonal matrix with the value of the income source as entries, $\mathbf{a}$ is an n-dimensional drift vector, and $\mathbf{s}$ is the matrix of diffusion coefficients. We assume that to rule out arbitrage opportunities $(\mathbf{a} - r) = \mathbf{s}\boldsymbol{\Sigma}'(\boldsymbol{\Sigma}\boldsymbol{\Sigma}')^{-1}(\mu - r)$. The investor maximizes utility of terminal wealth, $W(T)$,[15] where the wealth follows: $dW = [\boldsymbol{\theta}'(\boldsymbol{\mu} - r) + W_r]dt + \boldsymbol{\theta}\boldsymbol{\Sigma}\boldsymbol{d}\mathbf{w}$.

The investor is endowed with a HARA utility function $U = \frac{1-\gamma}{\gamma}\left(\frac{y}{1-\gamma}\right)^{\gamma}$ of terminal wealth $y = W + \mathbf{Y}'\mathbf{e}$, where $\mathbf{e}$ is a vector of ones. Let us define $cov_{\mathbf{Y},\, p}$ as the covariance between the return on the financial portfolio of the investor and the rate of change of his nonfinancial income and $cov_{\boldsymbol{Y},\, \boldsymbol{m}}$ as the covariance between the return on the market portfolio and the rate of change of the investor's nonfinancial income. The measure of deviation from the market is $\Delta = cov_{\boldsymbol{Y},\, \boldsymbol{m}} - cov_{\boldsymbol{Y},\, \boldsymbol{p}}$. The covariances between the nonfinancial risk and the market portfolio and the investor's actual portfolio are $cov_{\mathbf{Y},\boldsymbol{m}} = \boldsymbol{\theta}_m \boldsymbol{\Sigma} \mathrm{s}'$ and $cov_{\mathbf{Y},\boldsymbol{p}} = \boldsymbol{\theta}_p \boldsymbol{\Sigma} \mathrm{s}'$, where $\boldsymbol{\theta}'_m$ and $\boldsymbol{\theta}'_p$ are, respectively, the vectors of the proportion of wealth invested in risky assets for the market portfolio in the absence of nonfinancial risk and for the investor's portfolio in the presence of nonfinancial risk. They are

$$\boldsymbol{\theta}'_p = \frac{1}{W}\boldsymbol{\Sigma}^{-2}\left[\frac{W + \boldsymbol{Y}'\boldsymbol{e}}{1-\gamma}(\boldsymbol{\mu} - r) - \boldsymbol{\Sigma}\mathbf{s}\mathbf{Y}\right] \text{and } \boldsymbol{\theta}'_m = \frac{1}{W}\boldsymbol{\Sigma}^{-2}\left[\frac{W}{1-\gamma}(\boldsymbol{\mu} - r)\right]. \tag{16}$$

Therefore, Δ can be rewritten as $\boldsymbol{\Delta} W = \mathbf{Y}'\left\{\mathbf{s}\mathbf{s}' - \mathbf{e}[(\boldsymbol{\mu} - r)'/(1-\gamma)]\boldsymbol{\Sigma}^{-2}\boldsymbol{\Sigma} s\right\}$. Let us assume that there are two nonfinancial sources of income: x and z. The deviation from the market to hedge income y is

$$\Delta_z W = Y_z \mathrm{var}_{Y_z} + Y_x \mathrm{Cov}_{Y_z, Y_x} - (Y_z + Y_x)\sum_{j=1}^{n} \Omega_{S_j} \times \mathrm{Cov}_{S_j, Y_z}, \tag{17}$$

where $\Omega_{S_j} = \left(\boldsymbol{\mu}_{s_j} - r\right)' / (1-\gamma)\boldsymbol{\sigma}^2_{s_j}$ and μ_{s_j} and σ_{s_j} are, respectively, the mean and the variance of the jth risky asset.

---

[15] Alternatively, we can consider the case of the labor income flow. In this case, the investor does not maximize the utility of terminal wealth, but

$$MaxE\left[\int_0^T U(C(s))ds\right], \quad \text{s.t.} E\left[\int_0^T \xi_s C(s)ds\right] \leq \xi_0 W_0 + E\left[\int_0^T \xi_s Y(s)ds\right],$$

where ξ_s is the state price density. The result is analogous to the one we report in Equation 16, but with labor income defined from time t to T. That is, the residual working life of the invetsor.

$$\boldsymbol{\theta}'_p = \frac{1}{W}\boldsymbol{\Sigma}^{-2}\left[\frac{W + \boldsymbol{Y}(\boldsymbol{T} - \boldsymbol{t})'\boldsymbol{e}}{1-\gamma}(\boldsymbol{\mu} - r) - \boldsymbol{\Sigma}\boldsymbol{s}\boldsymbol{Y}(\boldsymbol{T} - \boldsymbol{t})\right]$$

This corresponds to the Carrol and Samwick construction of permanent income.

References

Arellano, M., 1989, "Testing for Autocorrelation in Dynamic Random Effects Models," *Review of Economic Studies*, 57, 127–134.

Benartzi, S., 2001, "Excessive Extrapolation and the Allocation of 401(k) Accounts to Company Stock," *Journal of Finance*, 56, 1747–1764.

Bodnaruk, A., 2003, "Proximity Always Matters: Evidence from Swedish Data," Working Paper, Stockholm School of Economics.

Campbell, J. Y., J. Cocco, F. Gomes, and P. Maenhout, 2001, "Investing Retirement Wealth: A Life-Cycle Model", in John Y. Campbell and Martin Feldstein eds., *Risk Aspects of Investment-Based Social Security Reform*, University of Chicago Press: Chicago, IL, 439–473, 2001.

Carroll, C. D., and A. A. Samwick, 1996, "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 40, 41–71.

Coval, J. D., and T. J. Moskowitz, 1999, "Home Bias at Home: Local Equity Preference in Domestic Portfolios," *Journal of Finance*, 54, 2045–2073.

DeMarzo, P., R. Kaniel, and I. Kremer, 2004, "Diversification as a Public Good: Community Effects in Portfolio Choice," *Journal of Finance*, 59 (4), 1677–1715.

Dammon, R., C. Spatt, and H. Zhang, 2001, "Optimal Consumption and Investment with Capital Gains Taxes," *Review of Financial Studies*, 14, 583–616.

Davis, S., and P. Willen, 2000, "Occupation-Level Income Shocks and Asset Returns," Working Paper, University of Chicago, Illinois.

Edin, P. A., and P. Fredriksson, 2000, "LINDA: Longitudinal Individual Data for Sweden 1960–1997," Working Paper, Uppsala University.

Englund, P., J. M. Quigley, and C. Redfearn, 1998, "Improved Price Indexes for Real Estate: Measuring the Course of Swedish Housing Prices," *Journal of Urban Economics*, 44, 171–196.

Frieder, L., and A. Subrahmanyan, 2002, "Brand Perceptions and the Market for Common Stocks," Working Paper, UCLA.

Gentry, W. M., and R. G. Hubbard, 2002, "Entrepreneurship and Household Savings," Working Paper, Columbia University, New York.

Gomes, F., and A. Michaelides, 2002, "Portfolio Choice with Habit Formation: A Life-Cycle Model with Uninsurable Labor Income Risk," Working Paper, London Business School.

Guiso, L., and T. Jappelli, 2002, "Household Portfolios in Italy," in L. Guiso, M. Haliassos, and T. Jappelli (eds.), *Households Portfolios*, MIT Press, Cambridge, MA.

Haliassos, M., and A. Michaelides, 2002, "Calibration and Computation of Household Portfolio Models," in L. Guiso, M. Haliassos, and T. Jappelli (eds.), *Households Portfolios*, MIT Press, Cambridge, MA.

Hau, H., 2001, "Location Matters: An Examination of Trading Profits," *Journal of Finance*, 56, 1959–1983.

Hayashi, F., 1985, "The Effect of Liquidity Constraints on Consumption: A Cross-sectional Analysis," *Quarterly Journal of Economics*, 10, 183–206.

Heaton, J., and D. Lucas, 2000, "Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk," *Journal of Finance*, 55, 1163–1198.

Heckman, J., 1979, "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153–161.

Huberman, G., 2001, “Familiarity Breeds Investment,” *Review of Financial Studies*, 14, 659–680.

Kiviet, J. F., 1995, “On Bias, Inconsistency and Efficiency of Various Estimators in Dynamic Panel Data Models,” *Journal of Econometrics*, 68, 53–78.

Ljungqvist, L., and T. Sargent, 1995, “The Swedish Unemployment Experience,” *European Economic Review*, 39, 1043–1070.

Merton, R. C., 1971, “Optimum Consumption and Portfolio Rules in a Continuous-Time Model,” *Journal of Economic Theory*, 3, 247–413.

Merton, R. C., 1987, “A Simple Model of Capital Market Equilibrium with Incomplete Information,” *Journal of Finance*, 42, 483–510.

Moskowitz, T., and A. Vissing-Jørgenson, 2002, “The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?,” *American Economic Review*, 92, 745–778.

Peress, J., 2004, “Wealth, Information Acquisition and Portfolio Choice,” *The Review of Financial Studies*, 17, 879–914.

Polkovnichenko, V., 2002, “Limited Stock Market Participation and the Equity Premium,” Working Paper, University of Minnesota, Minneapolis.

Sundin, A., and S.-I. Sundqvist, 2002, *Owners and Power in Sweden’s Listed Companies*, SIS Ägarservice AB, Dagens Nyheter, Stockholm, 1986–2002.

Telmer, C. I., 1993, “Asset Pricing Puzzles and Incomplete Markets,” *Journal of Finance*, 48, 1803–1832.

Viceira, L.M., 2001, “Optimal Portfolio Choice for Long-Horizon Investors with Non-Tradable Labor Income,” *Journal of Finance*, 56, 433–470.

Vissing-Jørgensen, A., 2002, “Towards an Explanation of Household Portfolio Choice Heterogeneity: Non-Financial Income and Participation Cost Structures,” working paper, University of Chicago.

Yitzhaki, S., 1987, “The Reationship between Return and Income," *The Quarterly Journal of Economics*, 102, 78–95.

Zeldes, S. P., 1989, “Consumption and Liquidity Constraints: An Empirical Investigations,” *Journal of Political Economy*, 97, 305–347.