Household Hedging Motives and Financial Risk-Taking: Evidence from Europe

Azar Aliyev

Abstract

Financial economists have suggested how hedging motives related to human capital and real estate should affect financial risk-taking of a utility maximizing household. Using rich cross-sectional data from the second wave of the Eurosystem Household Finance and Consumption Survey, I analyze how such motives relate to the propensity to participate in the markets for risky assets and the risky share – portion of financial portfolio allocated to the risky assets. By matching households with the long quarterly time series of aggregate labor income on the country-sector level and housing related variables on the country level, I am able to compute key explanatory variables. While some findings are consistent with the normative predictions, several empirical regularities create a puzzle. In particular, effects on the propensity to participate often differ from effects on the risky share of participants. I find that while size of the human capital encourages participation, it makes participants have lower risky shares. Higher correlation between risky asset, proxied by the local equity indices, and labor income discourages participation but has no significant effect on the risky share among participants. Volatility of the labor income also discourages risk-taking. As for housing, renters are less likely to be participants and have smaller risky shares compared to homeowners consistent with expectations. Moreover, as normative theory would suggest, renters have higher risky shares in countries with higher rent risky asset correlation while owners have lower risky shares in countries with higher house price risky asset correlation. At the same time, in terms of participation decision these two relations are reversed. Also surprisingly, prospective house buyers are found to have lower risky shares for higher levels of house price correlation with the risky asset.

STOCKHOLM SCHOOL OF ECONOMICS

December 9, 2019

Contents

1. Introduction

One of the key economic decisions households have to make is to choose how risky their financial portfolios should be. Apart from the risk-return tradeoff of the financial assets and the general risk preference of an investor, crucial considerations have to be made. Human capital, which can be viewed as an asset entitling its owner to the future stream of labor income, as well as real estate are the key assets which can belong to a household in addition to its financial wealth. Hence, risks associated with labor income and housing should be important determinants of investor's financial risk-taking. Such risks refer not only to the general volatility but also to the interaction between the key non-financial assets and the financial portfolio.

If labor income is deterministic, human capital, except for being non-tradable, is conceptually equivalent to a riskless bond and hence it should incentivize risk-taking in the financial portfolio. More generally, lower the covariation between your labor income and the risky financial assets, higher the hedging motive of investing into these assets. Do households employed in sectors with different cyclicality respond to this motive? How do households react to their general labor income risk? To consume the accumulated wealth, homeowners might want to liquidate their real estate in the future. Is the cross-sectional variation in the correlation between house prices and financial assets affecting the risk-taking decision of households? Are prospective real estate buyers (who effectively have the opposite situation to homeowners) encouraged to invest more in the risky assets if the correlation between house prices and financial assets is higher? Finally, how is the decision of renters affected by the hedging motives?

Understanding sources of heterogeneity in risk-taking among investors on the micro-level can have important implications for general equilibrium asset pricing. From individual perspective, this area of research can help households make better decisions. In their equilibrium setting, Davis and Willen (2000) show that, thanks to labor income risk sharing capacity, financial markets offer sizable welfare gains for individuals. As governments around the world transfer more responsibility of portfolio constructions (e.g. retirement savings) on individuals and costs of investing in risky financial assets falling, understanding choices of individuals in this domain becomes essential. Moreover, there is potential gap between practices of the rapidly growing industry of financial advice and the academia.

For the last few decades the questions listed above have been extensively explored from the theoretical front (see Section 2.1). Some exciting findings have also been made from the empirical perspective (see Section 2.2). Using rich data from the second wave of the Eurosystem Household Finance and Consumption Survey (HFCS), I contribute to the understanding of how financial risk-taking of households is actually related to their human capital and labor income risk as well as real estate and housing risk. In addition to HFCS, I use long quarterly time series of aggregate labor income and real estate related variables which I am able to relate to households in my sample on the country and country-sector level. I proxy the risky financial asset held by a household with the local equity index and refer to the portion of financial portfolio invested in the risky asset as risky share.

In my empirical analysis, I control for the effect of other variables known to be important determinants of risk-taking. Some of these variables (e.g. wealth, education and gender) are also likely to capture heterogeneity in participation costs. HFCS allows me to control for the general risk preference of a household. In addition, given that past financial experience of an investor is an important driver of her behavior, I control for the past performance of the risky asset.

I estimate human capital by using information on cross-sectional age variation in the labor income and time series trends of the aggregate level labor income and find that higher levels of human capital stimulate participation but decrease the risky share among participants. The correlation between labor income (more precisely, log growth in the real labor income) and the risky asset (more precisely, log excess return on the risky asset) as well as general risk of the labor income (standard deviation of log growth in the real labor income) discourage participation as one would expect. Risky share of participants is only affected by the general labor income risk. In addition, the correlation of the labor income and risky asset times the human capital to financial wealth ratio was found to be an important determinant of risk-taking. As human capital itself, this term positively affects the propensity to participate but decreases the risky share of participants even when controlled for its components.

In terms of real estate and housing risk, some of the finding are even more surprising. First, as expected, renters take less financial risk compared to homeowners everything else equal. This finding is robust to various specifications. Second, also consistent with expectations, renters have higher risky shares for a higher levels of correlation between rent (more precisely, real log growth in rent) and risky asset while homeowners have lower risky shares for a higher levels of correlation between house price (more precisely, real log growth in house price) and risky asset. Third, the same two variables have an opposite effect on the propensity to participate. Finally, prospective house buyers react negatively to increasing correlation between house price and risky asset again contrary to expectations.

The alternating effects possess a puzzle since such behavior is hard to rationalize with normative models. Even though my findings do not necessarily support any causal links, the fact that such empirical regularities are present suggests that households own sub-optimal portfolios given their hedging motives. Hence, attempts should be made to encourage households optimally consider their hedging motives as welfare losses can be substantial.

With households from 20 European countries in the sample, my study contributes to the understanding of actual risk-taking behavior of households by presenting broad evidence on the effects of hedging motives related to human capital and real estate. It is one of the few studies finding empirical relation between labor income and housing risk. To the best of my knowledge, this is the first empirical study that focuses specifically on the effects of risks associated with housing on the financial risk-taking of households.

The rest of the paper is organized as follows. Section 2 reviews main theoretical and empirical findings in this line of research. Section 3 discusses data used in the study and methodology for construction of variables. Section 4 reports findings of the empirical analysis and Section 5 concludes.

2. Literature Review

First, I review the literature of normative theoretical findings. Second, I discuss main empirical findings of studies exploring the question of financial risk-taking of households. While the field behavioral finance has been successful in explaining investment decisions of individuals (see Barberis and Thaler (2003) for an overview), it does not offer theories of optimal portfolio construction. Understanding behavioral mechanisms behind nonnormative choices can help investors abandon their biases and converge to decision that are best for them. I refer to some of the behavioral concepts when discussing the results.

2.1. Theoretical Findings of Normative Models

Markowitz (1952) has shown how an investor that cares only about mean and variance of returns on her financial portfolio should choose optimally in a single-period model. It is important to recognize that while this static framework might be appropriate to describe the optimal behavior for a short-term investor, for a long-term investor more complex and a multi-period framework is needed. This is because long-term investor cares about her living standards that can be supported by the portfolio composition rather than her financial returns in isolation. As a result, interactions between labor income, aspects of housing and financial assets become important in addition to the risk-return tradeoff of the financial portfolio and the general risk preference of an investor.

For the sake of analyzing risk-taking, portfolio construction problem is often reduced to a decision on the mix of a risky and a risk-free asset. Henceforth, I refer to the proportion of the risky asset in the financial portfolio as a risky share and denote it as w. The general theoretical approach in this line of household finance literature is to model the preferences over consumption of a household with an intertemporal utility function. The optimal risky share over the life-cycle of an investor is then found by maximizing this utility function subject to the budget constraint of an investor. Most of the models discussed here assume finite investment horizon. While investment practitioners might defy the notion of a utility function, it offers a convenient way to capture important characteristics of individual investors such as impatience and risk aversion.

Seminal life-cycle model by Merton (1969), in which the consumption and portfolio selection problem is solved in continuous time, demonstrates the restrictive conditions under which the optimal risky share decision for the long-term investor is same as for the short-term one. In his framework, neither labor income nor housing related expenses enter the budget constraint and an investor derives utility only from consumption. In the case of constant relative risk aversion (CRRA) utility function, the optimal risky share is[1]:

$$w_t^* = \frac{E_t(r_{t+1}{}^e)}{\gamma Var_t(r_{t+1}{}^e)} \tag{1}$$

where γ is the Arrow-Pratt degree of relative risk aversion (from the CRRA function), $E_t(r_{t+1}{}^e)$ is the conditional expectation of the excess log return on the risky asset (e.g. expected equity risk premium), $Var_t(r_{t+1}{}^e)$ is the conditional variance of this return. If returns are independent and identically distributed[2] then equation (1) can be rewritten as:

$$w_t^* = \frac{E(r^e)}{\gamma Var(r^e)} = \overline{\omega} \tag{2}$$

Equation (2) suggests that the risky share should be proportional to the expected excess return of the risky asset and inversely proportional to the investor's degree of risk aversion and variance of excess return on the risky asset. Surprisingly, equation implies that the risky share should not change over time (as individual ages) and is independent of the wealth level. Hence, I call it $\overline{\omega}$. Campbell and Viceira (2002) show that a similar solution can be derived in a discrete setting.

Next, I discuss models which incorporate presence of the labor income.

2.1.1. Human Capital and Labor Income Risk

As stated above, Merton (1969) assumes labor income is absent. In reality aside from financial portfolio an important asset belonging to a household is human capital. That is human capital can be thought of as implicit holding of an asset paying labor income. Hence, value of human capital is the present value of the expected future labor income. Since in less than perfect capital markets labor income cannot be insured or capitalized due to borrowing constraints (resulting from moral hazard[3]), human capital is important for the risky share decision and can rationalize age varying portfolio composition.

[1] Indeed, same optimal risky share is derived in a basic static mean-variance framework under simple constant absolute risk aversion preferences.

[2] See Campbell and Viceira (2002) chapter 3 and 4 for the discussions of the effect of time-varying expected excess returns on optimal portfolio choice for long-term investors.

[3] See the discussion in Campbell and Viceira (2002) p. 139

In the household finance literature labor income is often taken exogenously (i.e. agents do not make the work-leisure decision).[4] Moreover, some other decisions available to individuals such as investing in human capital through education, affecting risk characteristics of the labor income or determining retirement age are also assumed to be exogeneous.

Deterministic Labor Income

To start with, one can incorporate presence of labor income in the same framework by assuming labor income is deterministic (i.e. riskless) over the life-cycle. In this special case, it is relatively easy to see that the optimal risky share is[5]:

$$w_t^* = \left(1 + \frac{HC_t}{W_t}\right) \times \frac{E(r^e)}{\gamma Var(r^e)} = \left(1 + \frac{HC_t}{W_t}\right) \times \bar{\omega} \tag{3}$$

where HC_t is human capital and W_t is financial wealth at time t. From equation (3), one interesting observation is that in the early stage of the life-cycle it is optimal to have a higher risky share in your financial portfolio since you hold claims on riskless labor income and presumably did not accumulate a lot of financial wealth. As you age and your human capital decreases, you should tilt your portfolio towards the risk-free asset. In other words, since human capital is riskless it encourages risk-taking when it is abundant. Same intuitive results are suggested by even more complex numerically solved models such as one developed by Cocco, Gomes, and Maenhout (2005) in which authors incorporate mortality risk, short-sale and borrowing constraints.

The total risk of the labor income can be decomposed in to two components: 1) risk driven by the covariation with the risky asset and 2) idiosyncratic (i.e. "background") risk.

In a relatively simple, two-period model, Viceira (1998) shows that under general conditions, increase in the background risk reduces the optimal risky share. Viceira (2001)

---

[4] See Bodie, Merton, and Samuelson (1992) for the model with endogenous labor income in the portfolio choice problem. One of the conclusions is that investors should invest more in the risky asset if they can change their labor income in response to circumstances endogenously.

[5] To do so, one should recognize that in this case human capital is future labor income discounted at the risk-free rate (same rate as offered by the available risky-free asset), define the variable $\widehat{w_t}$ – the share of total wealth invested in risky asset, and rewrite the budget constraint to see that it is same as in the initial problem but with renamed variables (total wealth instead of financial wealth and $\widehat{w_t}$ instead of w_t). Finally, equation (3) is derived using equation (2) and the definition of $\widehat{w_t}$.

uses infinite horizon model[6] to demonstrate effect of the background risk on the risky share and finds that increase background risk reduces optimal risky share. However, under reasonable parameters this effect is very limited.

Covariation Risk

Most important risk of the labor income is its undesirable covariation with the financial portfolio. One can consider another extreme in which labor income is perfectly correlated with the risky financial asset. That is when there is no background risk and the labor income risk is driven solely by the covariation with risky asset. In this second special case, the optimal risky share becomes[7]:

$$w_t{}^* = \frac{E(r^e)}{\gamma Var(r^e)} - \frac{HC_t}{W_t} \times \left(1 - \frac{E(r^e)}{\gamma Var(r^e)}\right) = \overline{\omega} - \frac{HC_t}{W_t} \times (1 - \overline{\omega}) \tag{4}$$

Equation (4) suggests that an intuitive conjecture that if human capital behaves like risky financial asset it should decrease the risky share. In other words, risky human capital crowds out investments in risky financial assets. For households with such human capital, one should expect to see lower risky share for higher levels of human capital.

General Case

Campbell and Viceira (2002) show that, in a two-period model where investor makes a portfolio allocation decision in period one and consumes all of her wealth in period two (same model as one developed by Viceira (1998)), the optimal portion of financial wealth to be invested in the risky asset in period one is:

$$w_t{}^* = \frac{1}{\rho} \times \overline{\overline{\omega}} - \frac{1-\rho}{\rho} \times \beta_{HC} \tag{5}$$

where $\overline{\overline{\omega}}$ is the discrete time analog of $\overline{\omega}$, β_{HC} is loosely speaking sensitivity of labor income to the risky financial asset and ρ is the portion of financial wealth in the financial wealth plus human capital in the steady state. Equation (5) can be rewritten as:

$$w_t{}^* = \left(1 + \frac{HC_t}{W_t}\right) \times \overline{\overline{\omega}} - \frac{HC_t}{W_t} \times \beta_{HC} \tag{6}$$

---

[6] In such models, a positive probability of retirement (i.e. permanent drop of labor income to some constant) in each period allows the analysis of horizon effect.

[7] Equation is derived in the similar manner as equation (3)

Equation (6) suggests that more sensitive labor income is to the risky financial asset, lower the risky share should be.[8] Intuitively, more cyclical labor income should discourage financial risk-taking. Henceforth, instead of β_{HC}, I use *CorHC* (i.e. correlation between labor income growth and excess return on the risky financial asset) since this parameter is most often used in the more complex models.

Volatility of the labor income growth (general income risk) affects the value of the human capital and hence the optimal risky share. Campbell and Viceira (2002) show that higher general income risk would decrease ρ in equation (5) and consequently discourage risk-taking if *CorHC* is set to zero.

Introduction of stochastic and non-tradable labor income into finite horizon life-cycle models with more than two periods complicates them to the extend which does not allow the models to be solved even approximately using loglinearization technique. In general, findings of such models are consistent with the predictions of simple models discussed above. Campbell and Viceira (2002) solve such finite horizon model for representative households of 3 different education groups calibrated to the U.S. Panel Study of Income Dynamics (PSID) data where the solution relies heavily on numerical dynamic programming methods. Among other assumptions, the model assumes no borrowing is allowed.

First of all, authors find that for a typical young investor, it is optimal to invest all of her financial wealth into risky asset (since irrespective of education level, *CorHC* is found to be low enough in the U.S. population). This finding is hard to confront with the data since most of the young investors do not participate in the stock market at all (see Section 2.2). The risky share then declines over the life-cycle as financial wealth becomes relatively big. Interestingly, self-employed individuals have high correlation of the labor income with the risky assets and general risk of the labor income[9], consequently it is optimal for them to invest much less in the risky asset.

Cocco et al. (2005) – also discussed in the previous sections – develop an extension of such model. Among other findings, authors contribute by explaining the low participation by the young investors thanks to incorporated (empirically estimated) probability of a huge negative shock to the labor income and allowing for borrowing but with a penalty of default. Because of endogenous borrowing decision, young investors become net borrowers (which

[8] Observe that if $\beta_{HC} = 0$ we are back to equation (3) and if $\beta_{HC} = 1$ we are back to equation (4).
[9] This empirical finding is originally due to Heaton and Lucas (1999).

does not happen in the model discussed previously) in order to optimally smooth consumption and as a result of the default penalty do not invest in the risky asset.

2.1.2. Real Estate and Housing Risk

Yao and Zhang (2004) incorporate the housing decision into the life-cycle model to study the implications of the tenure status as well as housing hedging motives on the optimal portfolio risky share. Housing services provide utility (with Cobb-Douglas type of preference for ordinary consumption and consumption of housing services) but at the same time affect the budget constraint of the household. In their model, an investor can either buy a house of the desired size (which requires a combination of a down payment and a mortgage) or rent it. An owner can also sell her house and become a renter at any time period. Since price of the rent is a constant fraction of the price of the house, in this model rent has the same stochastic properties (e.g. correlation with the risky asset) as house price. Authors realistically account for the illiquidity of the housing market by introducing a transaction cost associated with selling a house. Furthermore, households face an exogeneous moving shock (having to move to another location for exogeneous reasons) which makes renting more attractive since renters can move without incurring any costs.

Authors find that, under their data-matching baseline parameters of the model, when households are indifferent between owning and renting, owners have a higher risky share compared to renters since real estate offers a diversification to risky asset and acts as a buffer to shocks in the financial portfolio and labor income hence motivating household to take more risk in their financial portfolio.[10] Another finding of Yao and Zhang (2004) is that increasing the covariation between the return on housing and the risky asset makes homeowners decrease their risky shares and makes the renters increase their risky shares. The former effect is due to the fact that owners have a long position in the real estate and thus increasing the covariation between housing and risky assets reduces the diversification benefits of housing. The latter effect is due to the fact that renters are effectively holding a short position in the real estate because rent is a fraction of real estate price and it is an

[10] This finding might appear to be contradicting to the previous studies on housing by Cocco (2004) as well as Flavin and Yamashita (2002) which find that housing risk will crowd out investment in stocks. These two papers studying the effect of housing on risk-taking discuss investments in risky asset as a proportion of net worth (risky asset, home equity and riskless asset). While Yao and Zhang (2004) also show that share of stocks in the net worth will be smaller for owners than renters, it is the first study that demonstrates how housing can increase the proportion of liquid wealth invested in risky asset.

expense for them. Moreover, renters are perspective future buyers which means that positive covariation between risky asset and housing offers a hedge for the future expense of real estate purchase.

2.1.3. Other Factors

Under reasonable calibrations, models discussed above predict a negative effect on the risky share as the level of financial wealth goes up. This effect is reversed with the introduction of habit (or subsistence consumption level). A common way to introduce habit in a life-cycle model is to assume that utility is a function of the difference between a habit and consumption in a given period (e.g., see Brunnermeier and Nagel, 2008). As a result, when level of financial wealth is low, household is, loosely speaking, afraid of its consumption to approach its habit and hence takes less risk in the financial portfolio. Measures of habit itself are therefore expected to affect risk-taking negatively.

Finally, higher leverage implies a larger probability of hitting the borrowing constraint in the future and thus ending up in a suboptimal portfolio allocation. In addition, highly levered households face the risk of hitting the costly default. For these reasons leverage should discourages financial risk-taking (e.g., see Cocco et al., 2005)

2.2. Empirical Findings

Until early 2000's, empirical findings on how individual investors actually allocate their wealth between risky and riskless assets was lagging behind the theoretical literature and was rather fragmented.

A stylized fact about the empirical research in this line of literature is that participation rates (proportion of investors holding some risky asset) and risky shares (proportion of risky assets in financial wealth) of investors are very low. That is, most of the models discussed in the previous section predict a much higher risky shares and participation rates than observed in the data. While typical participation rates are around 25-35% in Europe and 40-50% in the U.S., average risky share among all investors is close to 10% in Europe and 20-30% in the U.S. (e.g., see Guiso and Sodini, 2013)

In the sample of Swedish households, Betermier, Jansson, Parlour, and Walden (2012) show that general labor income risk (measured as wage volatility) is an important determinant of risk-taking. In particular, authors show that an increased general labor income risk decreases risky share. Similarly, using the U.S. National Longitudinal Survey of Youth (NLSY), Angerer and Lam (2009) document that as permanent income risk increases, investors decrease their risky shares.

More recent findings are provided by Bonaparte, Korniotis, and Kumar (2014). Using the Dutch National Bank (DNB) Household Survey panel data, find that individuals with greater correlation of income growth with local stock market (proxying the risky asset) have a greater risky share (proportion of mutual funds or direct investments in stocks). Same individuals are also more likely to participate in the stock market. In addition, when controlled for the main variables, authors find that general labor income risk (standard deviation of labor income growth) discourages participation and decreases the risky share among participants. The limitation of the analysis by Bonaparte et al. (2014) is that while controlling for labor income and age, authors do not explicitly estimate human capital and its impact on risk-taking.

Another prominent research on empirical regularities is one by Calvet and Sodini (2014). Authors employ the panel data of twins in Sweden over the 1999 to 2002 period together with a longer panel data on financial portfolio, labor income and other variables. Motivated by theories of habit and subsistence consumption, the main focus of the paper is understanding the sensitivity of risky share with respect to wealth levels. Despite of the fact that measures of risk aversion γ, are not available in the data set, thanks to the ability to match individuals with their twins, authors control for genetic fixed effects across individuals. In contrast to Bonaparte et al. (2014), authors estimate human capital for each individual and document a positive effect of human capital on the risky share of participants. However, human capital was not found to be significant once controlled for twin pair fixed effects. Additionally, measures of general income risk, leverage and measures of subsistence levels and habit are found to affect risky share negatively. Ownerships of commercial real estate and private business are also found to crowd out risk-taking. Residential real estate, however, is found to be positively affecting the risky share among participants.

Interestingly and contrary to Bonaparte et al. (2014), authors do not find support for the risky share being negatively related to the sensitivity of income innovation with respect to portfolio returns despite having substantial variation of this variable across individuals. At the same time, authors find the ratio of human capital to wealth to be insignificant by itself despite establishing the importance of financial wealth and human capital individually.

Even more recently, Fagereng, Gottlieb and Guiso (2017) study risk-taking over time in order to test the prediction of life-cycle models that risky share should decrease over the life-cycle if human capital is riskless. Authors employ the Norwegian Tax Registry panel data and find that those households who participate in the risky asset market, tend to invest relatively a lot in the risky asset when young and gradually decrease their risky share as they age. This supports the models in which human capital is assumed to have relatively low correlation with the risky asset. At the same time, average market participation is found to be a hump-shaped function of age.

In his sample, Vestman (2018) documents that participation rates are lower for renters compared to homeowners by 30 and 40 percentage points in the U.S. and Sweden respectively.

Overall, as discussed in the introduction, there is lack of empirical evidence on the effect of hedging motives, and especially those related to housing risk, on financial risk-taking of investors.

3. Data and Methodology

In this section I discuss the datasets used in the study, definition and construction of required variables and provide a descriptive summary statistics of variables used for the analysis.

3.1. Sources of Data

As mentioned earlier, the main data employed in the analysis is a cross-section of households. In addition, I use time series of labor income, equity indices, real estate related indices and other data.

3.1.1. Cross-section

My cross-sectional data comes from the second wave of the Eurosystem Household Finance and Consumption Survey (HFCS) conducted by European Central Bank (ECB) between 2013 and the first half of 2015. These data became available by December 2016.

HFCS consists of individual level data with 210,662 individuals and household level data with 84,597 unique households. Each individual belongs to a single household. For every household, there is uniquely identified reference person according to UN/Canberra definition[11] to which I refer to as household head henceforth.

On the household level, the survey provides rich set of variables related to financial assets and liabilities, real assets, private businesses and other. On the individual level, household members report their demographics (all individuals), employment status and labor income (individuals aged 16 or more). Employed but not self-employed individuals are referred to as employees.

All financial amounts are reported in EUR. The survey covers 20 countries: 18 eurozone countries – Austria (AT), Belgium (BE), Cyprus (CY), Estonia (EE), Finland (FI), France (FR), Germany (DE), Greece (GR), Ireland (IE), Italy (IT), Latvia (LV), Luxembourg (LU), Malta (MT), the Netherlands (NL), Portugal (PT), Slovakia (SK), Slovenia (SL), and Spain (ES) as well as Hungary (HU) and Poland (PL). Table A1 of the Appendix reports number of households per each country.

In sum, the main advantage of the HFCS data is threefold: 1) detailed information on assets (including real and financial assets) and labor income at the same time, 2) broad coverage across many countries with the same methodologies and definitions and 3) big number of households present in the sample.

Although survey based data are often blamed for measurement errors, one can argue that since self-perception of stocks and flows is what drives the economic decisions, it is important to study how households behave based on this information. In other words, if rationality means a combination of having correct beliefs and making normatively acceptable choices based on those beliefs, my work can be thought of as a test for the latter

[11] The person is identified by sequentially testing the following statements until only one member remains: 1) one of the partners in a registered or de facto marriage, with dependent children; 2) one of the partners in a registered or de facto marriage, without dependent children; 3) a lone parent with dependent children;
4) the person with the highest income; and finally, 5) the oldest person. Dependent children are defined as all persons aged 0-15 and non-working person aged 16-24 and living with a parent.

building block. On the other hand, since there are still unobserved variables (e.g. household's assessment of risky returns), potentially erroneous beliefs of households can be affecting the risk-taking.

3.1.2. Labor Income Time Series

For time series of labor income, I use Compensation per Employee variable obtained from Eurostat and ECB calculations based on Eurostat data which originates from National accounts[12]. Compensation per employee is defined as average remuneration of employees which includes gross wages and salaries as well as any additional payments such as overtime or bonuses and it is denominated in domestic currencies. The variable is available on a quarterly level for all HFCS countries except for Luxemburg and Portugal. For each country, the variable is reported per 11 economic sectors (except for Belgium – 10 and Italy – 1). Hence, in total I have 187 country-sector pairs (CS pairs henceforth). The classification of sectors is based on NACE Rev 2[13] which is the same classification according to which employees in HFCS report the economic sector of their company. The description of each economic sector is reported in Table A2 of the Appendix.

For most of the countries the time series start from 1995 Q1 (except for Germany – 1991 Q1, Finland – 1990 Q1, France – 1980 Q1, Malta and Poland – 2000 Q1) and last until 2019 Q2. For all of the CS pairs, time series exhibit strong seasonal variation. In Figure A1 of the Appendix, I plot these time series for 6 CS pairs for descriptive purposes.

Although aggregated on country-sector level, my labor income data is a much longer time series compared to, for example, Bonaparte et al. (2014), where authors have on average 10 annual observations per individual. Moreover, annual data hides seasonal variation in labor income which might be an important risk characteristic. The disadvantage of my labor income variable is that it hides variation on individual level caused by lay-offs due to the definition of compensation per employee variable. This issue is discussed further in Section 3.3.

[12] Dataset title: National accounts, Main aggregates (Eurostat ESA2010 TP, table 1)

[13] NACE is a statistical classification of economic activities in the European Union; the term NACE is derived from the French Nomenclature statistique des activités économiques dans la Communauté européenne. Rev 2 stands for second review.

3.1.3. Equity, Real Estate, and Other Time Series

Using Bloomberg database, I obtain quarterly time series of log returns on 20 indices specific to every country in HFCS. These returns are in domestic currencies and include both capital gains and dividend yields. Data are collected since the inception of a given index. Table A3 reports names of indices together with their brief description in the Appendix. As will be discussed later, local equity indices will serve as a proxies for risky assets held by households.

For all 20 countries I collect quarterly data on log changes of House Price Index reported by Eurostat. The index captures transaction prices of the residential real estate properties (including flats, detached houses, terraced houses, etc.) purchased by households. These data are available starting from 2005 Q2 for most of the countries until 2019 Q2 (except for Greece – 2011 Q2). In addition, I collect quarterly log changes in Actual Rental Index from Eurostat for all 20 countries. The index is a part of the Harmonized Index of Consumer Prices and it refers to the rent actually paid by tenants or subtenants on their main or secondary residence. For most of the countries, the time series start from 1996 Q2 and last until 2019 Q2 for all of the countries.

From International Monetary Fund, I collect Consumer Price Index on a quarterly basis for all 20 countries in order to compute quarterly log growth in the price level of consumption basket. For all countries, span of the time series is sufficient to cover fully my labor income and real estate time series.

Finally, from database of OECD, I collect quarterly time series of short-term interest rates for all 20 countries which refer to the period averages of the yields on three-month treasury bills. I use this variable as a proxy for risk-free rate when computing excess returns.

3.2. Definition and Construction of Variables

In this section I define simple variables and discuss my methodology for construction of more complex variables used in the study. Following the related literature, I assume that risk-taking decisions are made on the household level. Hence my analysis is performed on the household level data and the ultimate goal is to construct variables for every household.

3.2.1. First Set of Variables

The first set of variables consists of variables initially available or calculated directly for the household as a whole.

Household's risk-free wealth is defined as sum of value of sight accounts[14] (current accounts, draft accounts, or checking accounts), value of bank deposits (saving accounts, time deposits or certificates of deposit), and value of money market mutual funds. Risky wealth is defined as sum of value of directly held stocks, directly held bonds and risky mutual funds (i.e. mutual funds not predominantly investing in money markets). Financial wealth (*W*) is a gross measure of value of liquid assets and it is defined as sum of risk-free wealth, risky wealth, total value of all voluntary pension accounts or whole life insurances (i.e. sum over household members), amount owed to household (i.e. loans to relatives or friends or other private loans), value of any additional assets in managed accounts and value of any additional financial assets. Following Calvet and Sodini (2014), I define risky share (*w*) as risky wealth divided by sum of risky and risk-free wealth. Participation dummy (*PartD*) is equal to one for households with positive risky share and zero otherwise. I refer to the households with a positive risky share as participants.

Risk aversion (*RA*) variable takes integer values between 1 and 4 from least to most risk averse investment attitudes. It is based on a question that elicits general financial risk preference of household head and his or her spouse (if any).

Since I do not observe the actual risky financial asset held by the household, I proxy it with the local equity index following Bonaparte et al. (2014). For every household, I compute the average annualized excess return and variance of the excess return of the local equity index for 5 years before a given household was interviewed ($E(r^e)$ and $Var(r^e)$). Although expected return and variance calculated based on past 5 years of data are clearly not precise estimates, households can still rely heavily on the recent history of the local stock market (e.g., see Malmendier and Nagel, 2011).

Residential real estate wealth (*RRE*) is defined as sum of the value of household's main residence (partial value for partial owners) and sum of values (partial values for partial owners) of other real estate properties (which can include other apartments, maisonettes, villas, terraced houses, garages, etc.) not for business activities (i.e. either used for holidays,

[14] I set value of sight accounts with negative balance to zero. Leverage is considered separately.

other private use, vacant or other non-commercial use). Commercial real estate wealth (*CRE*) is defined as value of other real estate property owned directly by the household (which can include offices, hotels, other commercial buildings, farms, land, etc.) used for business activities. Valuables wealth (*Vlb*) is defined as total value of cars and other vehicles and valuables (such as jewelry, works of art, antiques, etc.).

Private equity wealth (*PE*) is defined as net asset value of household's share in non-publicly listed businesses (both in which some individuals of the household are self-employed or taking active role in managing it and all other businesses). Entrepreneur dummy (*EntrD)* takes value of one if any of the household members received self-employment income (which can be negative) from the family business in the last 12 month and zero otherwise.

Total debt is defined as sum of outstanding balances of all mortgages, credit lines and overdrafts and any other loans. Leverage ratio (*LR*) is defined as total debt divided by sum of financial wealth, residential real estate wealth, commercial real estate wealth and valuables wealth.

Rental dummy (*RntD*) takes value one if household rents main residence and zero otherwise. For every household, I compute correlation of the rent with the risky asset (*CorRnt*). More precisely, the correlation is computed between the log real growth of the Actual Rental Index of household's country and the log excess return of the local equity index. Purchase home dummy (*PurchHD*) is equal to one if one of the reasons why household saves is reported as to buy own home and zero otherwise. Similar to *CorRnt*, for every household, I compute correlation of the house price with the risky asset (*CorHP*). More precisely, the correlation is computed between the log real growth of the House Price Index of household's country (return on housing) and the log excess return of the local equity index.

Finally, to control for habit or subsistence consumption level, for every household I calculate number of household members (*NHM*). The choice of the variable is consistent with the previous empirical research.

3.2.2. Second Set of Variables

Here I discuss variables which are first defined or constructed on the individual level and are then calculated for a household as a whole.

Demographic Characteristics

For demographic characteristics, education (*Educ*) variable ranges from 1 to 4 from smallest to greatest levels of highest education completed by household head[15]. Age (*Age*) variable measures age of the household head and falls in one of 15 age brackets. Unemployment dummy (*UnemD*) takes value of one if household head is unemployed and zero otherwise. Retirement dummy (*RetD*) is equal to one if household head is retired and zero otherwise. Gender dummy (*GenD*) takes value of one if household head is male and zero otherwise.

Labor Income Risk

Labor income is defined as gross cash employee income. It includes regular salaries or wages, any overtime or vacation payments, tips, bonuses etc. Although theoretical models discussed in Section 2.1 presume disposable labor income, it is not uncommon to use gross measure of labor income for the sake of empirical analysis (e.g., see Betermier et al., 2012). Henceforth, $(L_t)_i$ refers to the real labor income of individual *i* for year *t* in the prices of 2013 Q1.

As mentioned earlier, employees (and persons on a temporary leave also referred to as employees henceforth) report their economic sector of employment according to the NACE Rev 2 classification[16]. For every such individual, I compute correlation of the labor income with the risky asset. More precisely, I compute correlation between log growth of quarterly real labor income of individual's country-sector pair and log excess return of the equity index of country corresponding to the individual.[17] All time series are in domestic currencies. Some technical discussion on this variable is postponed to the Appendix Section Properties of Labor Income Risk Variables.

In addition, I calculate general risk of the labor income for every employee. More precisely, I compute the standard deviation of quarterly log growth of real labor income of the CS pair an employee belongs to.

For self-employed individuals or individuals which are not employed, labor income risk variables are set to zero. For individuals not in the labor force, labor income can be viewed as a constant stream of zeros and hence both of the labor income risk variables are

[15] 1 – Primary education, 2 - Lower secondary, 3 – Upper secondary and 4 – Tertiary; based on International Standard Classification of Education.

[16] In case a company is a diversified company, economic sector of the division is reported.

[17] Note that, according to Campbell and Viceira (2002), correlation between the aggregate component of shocks to the real labor income and the risky asset is the variable important for the risk-taking (see p. 177).

indeed equal to zero. I control for unemployed, retired and self-employed household heads using dummies discussed above.

On the household level, correlation of the labor income with the risky asset (*CorHC*) and general risk of the labor income (*StdL*) are defined as income weighted averages of the corresponding variables of household head and his or her spouse (if present)[18]. If household head does not have a spouse then the household level variable is set to the household head's individual level variable.

Human Capital

Next, I discuss my methodology for estimating human capital. For non-employees human capital is set to zero. I control for unemployed, retired and self-employed household heads in my analysis. For employees, the first step is to estimate the stream of expected future real labor income. Since I do not observe historical labor income stream on individual level, my aim is to estimate it using information on the time series of labor income on the aggregate country-sector level, cross-sectional variation of labor income of individuals, age of the individual and her observed point in time level of the labor income. To start with, for every country-sector pair *K,* I forecast the real labor income time series variable based on the annual frequency. The forecasts are made for post 2018 period using all the information before 2018. Let $f^{(K)}(t)$ be the level of aggregate real labor income of a CS pair *K* for year *t*. I estimate the ARMA(*p,q*) model of the form:

$$f^{(K)}(t) = \alpha + \sum_{a=1}^{p} \varphi_a f^{(K)}(t-a) + \sum_{b=1}^{q} \theta_b \epsilon_{t-b} + \epsilon_t \tag{7}$$

Since $f^{(K)}(t)$ is not a stationary variable due to clear trending patterns, equation (7) is estimated after taking the first difference of $f^{(K)}(t)$ (i.e. using ARIMA(*p,1,q*) model). For every CS pair, the model is selected based on the AIC. For most CS pairs the drift coefficient was estimated to be positive. The coefficient was estimated to be equal to 0 mostly for sectors of Cyprus, Greece, Ireland and Poland. In total fours sectors of Cyprus, Greece and Spain were forecasted to have a decreasing real labor income.

[18] In rare cases when both spouses are employees but labor income in not available for either one of them, simple average is used. There are no households for which labor income risk measure is not available for both spouses.

The next step is to incorporate the observed cross-sectional variation in labor income and estimate the labor income profile for every employee. If there was no birth-year (cohort) or country-sector effect on the labor income, one could find the approximate shape of the labor income as a function of age by simply regressing observed point in time labor income of different individuals on their age. However, the observed trending pattern of aggregate labor income time series as well as variation in levels of aggregate labor income across CS pairs suggests that both effects are present. To capture country-sector effect, I perform the estimations for each CS pair separately. To account for the cohort effect, I make several assumptions. Before stating these assumptions, I first write process of annual expected real labor income as a sum of two components:

$$E(L_t)_{i \in K} = f^{(K)}(t) + g^{(K)}(t, R_i)$$

where $E(L_t)_{i \in K}$ is the expected real labor income of person *i* belonging to CS pair *K* at year *t* and $g^{(K)}(t, R_i)$ is the difference between $E(L_t)_{i \in K}$ and aggregate level real labor income $f^{(K)}(t)$. I assume that this difference is a function of time and the reference birthday year R_i of person *i.* Next, I write $g^{(K)}(t, R_i)$ as a sum of two components:

$$g^{(K)}(t, R_i) = a^{(K)}(R_i) + h^{(K)}(t - R_i)$$

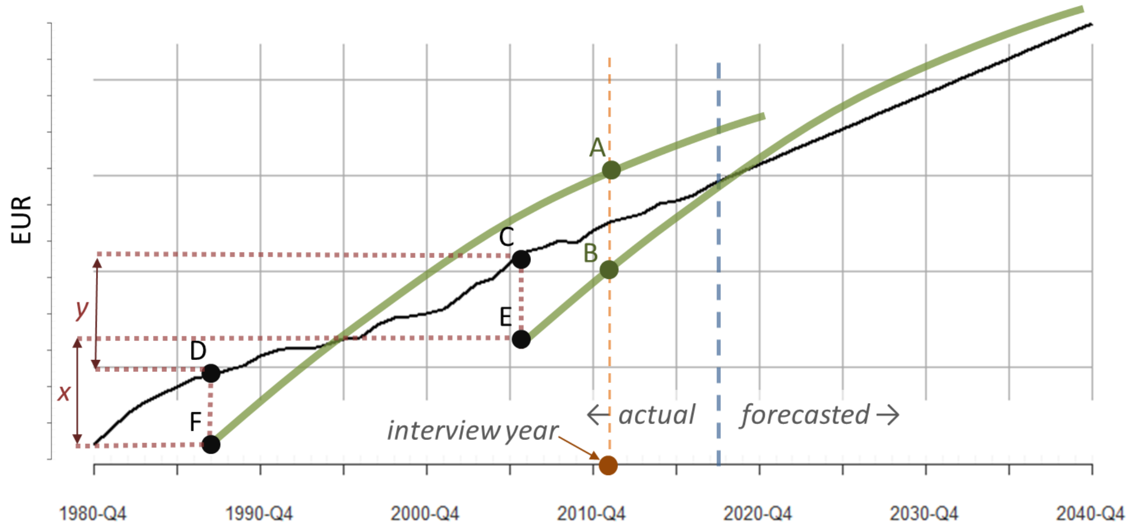
where $a^{(K)}(R_i)$ is time-invariant constant which captures cohort effect and $h^{(K)}(t - R_i)$ is a function of age which captures career progression. That is, I effectively am assuming that for any two individuals belonging to the same CS pair, the shape (bot not the level) of the labor income profile is the same. I write expected real labor income as:

$$E(L_t)_{i \in K} = f^{(K)}(t) + a^{(K)}(R_i) + h^{(K)}(t - R_i) \tag{8}$$

Next, I assume that $\forall i, j \in K,\ a^{(K)}(R_i) - a^{(K)}(R_j) = f^{(K)}(R_i) - f^{(K)}(R_j)$. That is, I assume that the level differences of labor income profiles result only from the differences in country-sector level aggregate labor income at the reference years of given individuals. Consider Figure 1 for illustrative representation of the assumptions made. In black solid line, I depict actual and forecasted annual aggregate labor income for the information and communication sector of France. Consider two individuals belonging to this CS pair – one in the earlier and the other in the later stage of the life-cycle. I depict their hypothetical labor income profiles before retirement with green curves. Points A and B are observed at the time of the interview for the older and the younger individual respectively. The first assumption is represented by shapes of labor income profiles being the same for both individuals. Next, let the reference birthday year R be the year of the 16th birthday (i.e. the first year an individual earns labor income). The assumption on the nature of cohort effect can effectively be described as length of distance *x* being equal to length of distance *y*. Given the available data, I do not observe points F and E and hence I cannot find the length of distance *x*. What I do observe, however, is points C and D.

Figure 1. *Illustrative Figure*

This figure is meant to describe the assumptions made for the estimation of labor income profiles. The black solid line is the actual (to the left of dashed vertical blue line) and forecasted (to the right of dashed vertical blue line) annual aggregate labor income of information and communication sector of France. The dashed vertical orange line represents the point in time when two hypothetical individuals belonging to this CS pair were interviewed. Green curves represent their labor income profiles before retirement. The two assumptions are represented by the shapes of the green curves and lengths of distances x and y being the same (for R *= 16th birthday year).*



Now, consider individuals $i, j \in K$. The difference between their expected real labor incomes at the time of the interview[19] $t*$, according to equation (8) is:

$$E(L_{t^*})_{i\in K} - E(L_{t^*})_{j\in K} = f^{(K)}(t^*) + a^{(K)}(R_i) + h^{(K)}(t^* - R_i) - f^{(K)}(t^*) - a^{(K)}(R_j) - h^{(K)}(t^* - R_j)$$

[19] Labor income always refers to the year preceding the year when interviews in a given country have started, Hence, for any two individuals in the same country, labor income always refer to the same year.

Using the assumption on $a^{(K)}(R_i) - a^{(K)}(R_j)$ and writing $(L_{t^*})_{i\in K}$ instead of $E(L_{t^*})_{i\in K}$ the equation above can be rewritten as:

$$h^{(K)}(t^* - R_i) - h^{(K)}(t^* - R_j) = (L_{t^*})_{i\in K} - (L_{t^*})_{j\in K} - \left[f^{(K)}(R_i) - f^{(K)}(R_j)\right] \quad (9)$$

By keeping *j* fixed[20] and calculating the quantity of equation (9) for different $i's \in K$, this variable can be thought of as a function of age of individual i. Since I can observe all the variables on the right-hand side of equation (9) and given that the quantity depends only on the age of individual i at the time of the interview, for every CS pair *K,* I estimate this quantity as a function of age of individual i. I call this function $\delta^{(K)}(\cdot)$ and write:

$$\delta^{(K)}(age_i) = h^{(K)}(t^* - R_i) - h^{(K)}(t^* - R_j) \quad (10)$$

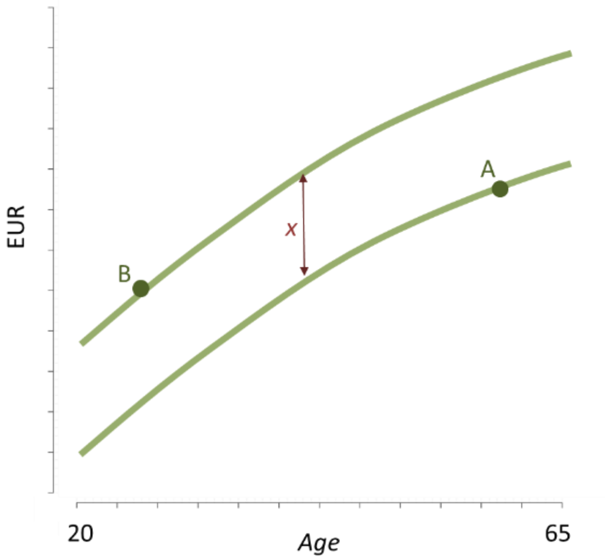
The function is estimated for non-retired employees by keeping *j* fixed and regressing the known quantity of the right-hand side of equation (9) on the age of individual i. Taking into account the fact that year 2000 is the first year when all aggregate level labor income time series are available, I choose R to be the 50th birthday year[21]. CS pairs with less than 15 observations are dropped from the estimation. There are 10 such pairs – almost all are real estate activities sectors. In addition, I require that there are at least 2 observations in each of the following age brackets: ≤ 35; > 35 and ≤ 45; > 45. All CS pairs pass this requirement. Finally, I exclude CS pairs for which $\widehat{\delta^{(K)}}(age_i)$ is a downward slopping function of age. There are 9 such pairs.

In sum, the assumption on the nature of cohort effect, allows me to estimate the difference between individual level labor income and aggregate level labor income that is due to the age (career) effect. As can be illustrated by Figure 2, the idea is to adjust the observed levels (more precisely the differences) of labor incomes by the assumed cohort effects (more precisely the differences in cohort effects) before analyzing the relationship of labor income with respect to age.

[20] I arbitrarily choose *j* to be the employee belonging to the country-sector pair *K* appearing first in my dataset.
[21] Employees with a birthyears smaller than 1950 are dropped from the estimation in order for $f^{(K)}(R_i)$ to be defined for every *i.*

Figure **A2** of the Appendix compares the scatter plots of $(L_{t^*})_i$ and $\delta^{(K)}(age_i)$ with respect to age for the wholesale/retail trade and vehicle repair economic sector of Austria.

Figure 2. *Illustrative Figure*



This simple figure continues the illustration of Figure 1. The two green curves correspond to the labor income profiles before retirement of the two individuals belonging to the same CS pair considered in Figure 1. This figure is meant to demonstrate how adjustment to the level of observed individual labor incomes of these two individuals (more precisely adjustment to the differences in their labor incomes) by the cohort effect x aligns unobserved labor income profiles (more precisely disentangles the difference due to career from the difference due to cohort effect). Point B observed for the older person can be moved down by distance x to become part of the hypothetical labor income profile of the younger person. Doing this for all persons in the same CS pair allows for the estimation of the shape of the labor income profile.

Next, based on equation (8), I write:

$$E(L_{\Delta t+t^*})_{i\in K} - E(L_{t^*})_{i\in K} = f^{(K)}(\Delta t + t^*) + a^{(K)}(R_i) + h^{(K)}(\Delta t + t^* - R_i) - f^{(K)}(t^*) - a^{(K)}(R_i) - h^{(K)}(t^* - R_i)$$

where t^* is the year of the interview for CS pair K. By writing L_{t^*} instead of $E(L_{t^*})$ and by adding and subtracting $h^{(K)}(t^* - R_j)$ the equation above can rewritten as:

$$E(L_{\Delta t+t^*})_{i\in K} = (L_{t^*})_i + f^{(K)}(\Delta t + t^*) - f^{(K)}(t^*) + h^{(K)}(\Delta t + t^* - R_i) - h^{(K)}(t^* - R_j) - \left[h^{(K)}(t^* - R_i) - h^{(K)}(t^* - R_j)\right]$$

Observe that $h^{(K)}(t^* - R_i) - h^{(K)}(t^* - R_j)$ is known and it does not depend on Δt. Hence I denote it as c_i. I substitute $h^{(K)}(\Delta t + t^* - R_i) - h^{(K)}(t^* - R_j)$ and $f^{(K)}(\Delta t + t^*)$ with their estimates and write:

$$E(L_{\Delta t+t^*})_{i\in K} = (L_{t^*})_i + \widehat{f^{(K)}}(\Delta t + t^*) - f^{(K)}(t^*) + \widehat{\delta^{(K)}}(age_{i,\Delta t+t^*}) - c_i \qquad (11)$$

Equation (11) is then used to estimate labor income profile for every employee before the retirement.[22] The equation has an intuitive interpretation: after the interview, the labor income is expected to grow thanks to the developments in the country-sector level and thanks to the average patterns of career progression of individuals in a given country-sector pair.

Since there can be negative shocks to the $f^{(K)}(t)$, equation (11) can predict negative $E(L_{\Delta t+t^*})$ for individuals with very small L_{t^*}. In this case $E(L_{\Delta t+t^*})$ is set to 0. This can be thought of as a temporary exit from the labor force.[23]

Finally, I make the second adjustment to the estimated labor income profile and before computing human capital. Observe that in equation (9) I am writing $(L_{t^*})_i$ instead of $E(L_{t^*})_i$. However, $(L_{t^*})_i = E(L_{t^*})_i + \epsilon_i$, where ϵ_i can result from characteristics such as level of education, skills or gender. Since $\delta^{(K)}(age_i)$ is estimated for individuals with different levels of such characteristics, $\widehat{\delta^{(K)}}(age_i)$ might under or overestimate the level of the labor income for a given individual. In other words, equation (11) might predict a slightly disruptive continuation (in the form of a jump) of the labor income profile compared to the observed labor income at the date of the interview. To take this into account, I adjust the level of the predicted real labor income profile in the following fashion:

$$E(L_{\Delta t+t^*})_i' = E(L_{\Delta t+t^*})_i + [(L_{t^*})_i - 2 \times E(L_{1+t^*})_i + E(L_{2+t^*})_i]$$

Human capital is then calculated for all household heads and if present for their spouses according to the present value formula:

$$HC_{t^*,\ i\in K} = \sum_{\Delta t=1}^{T_i} \frac{E(L_{\Delta t+t^*})_{i\in K}}{(1+d)^{\Delta t}} \tag{12}$$

where T_i denotes the difference between life expectancy of person i and the age of person i at time of the interview t^*. Life expectancy is taken to be 79 for all males and 84 for all females. In the equation (12), $E(L_{\Delta t+t^*})_i$ takes the values predicted by equation (11) with the two adjustments for the pre-retirement years of a given individual. In the fashion of Cocco et al. (2005), after the retirement year, $E(L_{\Delta t+t^*})_i$ is assumed to be equal to the

[22] For Poland and Hungary, $\widehat{f^{(K)}}(\Delta t + t^*) - f^{(K)}(t^*)$ is first converted to EUR at a fixed exchange rate prevailing at the time of the interviews in these countries.

[23] There are 33 employees for whom equation (11) predicts negative $E(L_{\Delta t+t^*})$.

constant fraction of the last real labor income before retirement (i.e. I assume a constant replacement ratio). The retirement age is equal to the expected retirement age reported by a given individual[24]. Replacement ratio is assumed to be 60%. Following Calvet and Sodini (2014), I assume the discount rate d to be a constant 3%. The discount rate is likely to depend on the general labor income risk which I control for in my analysis.

Human capital of the household (*HC*) is defined as the sum of human capital of household head and, if there is any, of his or her spouse.

Appendix Table A4 summarizes the definitions of all variables discussed in this section grouped by their themes.

3.3. Summary Statistics

As was mentioned earlier, there are in total 210,662 individuals in my sample. There are 35,363 persons aged below 16 (as noted earlier, for these persons only demographic information is provided). Out of the persons aged 16 or higher, 81,674 are employees not on temporary leave and self-employed persons while 1,183 are employees on temporary leave (i.e. on sick, maternity or other kind of leave and are planning to return to work. Such individuals are also referred to as employees and human capital as well as labor income risk variables are computed for them as well). Out of the employed ones, 69,156 are employees (121 out of which did not provide their economic sector of employment). Table A5 displays number of employees in my sample in each economic sector.

Before summarizing variables on the households level, I discuss the computed labor income risk and housing risk variables on the country and economic sector level. Table 1 summarizes the two labor income risk variables on the economic sector level. In general, sector level cross-country averages of correlation of the labor income with the risky financial asset is relatively low for all of the sectors (ranging from -0.042 to 0.068). As mentioned earlier, this variable does not directly capture labor income variation due to lay-offs of employees. However, if lay-offs tend to happen when average labor income is low, this

[24] For individuals who report their expected retirement age above their life expectancy, it is equated to their life expectancy. For the small fraction of individuals who did not report their expected retirement age, it is equated to 65. For individuals who reported their expected retirement age below their age at the time of the interview, expected retirement age is equated to their age at the time of the interview. Only a small fraction of the employee household heads or their spouses need one of the 3 adjustments.

variable proxies labor income risk that is bigger in magnitude. Not surprisingly, the public sector (OTQ) has the lowest correlation with the risky financial asset (-0.042) although it is not the sector with smallest general labor income risk (0.105). Sector of activities related to infrastructure (BTE) and manufacturing sector (C) have the highest correlation with the risky asset (0.061 and 0.068 respectively). Real estate sector (L) has the highest general income risk (0.155) and a relatively high correlation with the risky financial asset (0.060).

***Table 1**. Summary of Labor Income Risk Variables on the Economic Sector Level*

The table summarizes the two labor income risk variables used in the study by economic-sector level cross-country averages. Table A2 of the Appendix describes sector codes while Table A4 of the Appendix provides definitions of the variables. All computations are on a quarterly basis and variables are not annualized.

Sector Code	A	BTE	C	F	GTI	J	K	L	M_N	OTQ	RTU
CorHC	0.037	0.061	0.068	0.045	0.026	0.014	0.049	0.060	0.034	-0.042	0.039
StdL	0.121	0.089	0.093	0.103	0.074	0.095	0.110	0.155	0.082	0.105	0.086

Table 2 summarizes labor income risk and housing risk variables on the country level. Average correlation of labor income with the risky asset is relatively high in countries like Belgium, Italy and Netherlands and relatively low in countries like Cyprus, Greece and Slovakia. It varies from -0.072 in Cyprus to 0.188 in Netherlands. The general risk of labor income is relatively high in Belgium (0.159), Greece (0.222) and Italy (0.179) and relatively low in Estonia (0.058), Ireland (0.054) and France (0.047). Housing risk variables exhibit substantial cross-country variation. While correlation of rent with the risky asset is -0.145 in Austria, -0.203 in Hungary and -0.163 in Luxemburg, it is 0.139 in Estonia, 0.279 in Ireland and 0.188 in Portugal. At the same time, correlation of house price with the risky asset is relatively high in Cyprus (0.288), Estonia (0.336) and Finland (0.417) and relatively low in Austria (-0.188), Italy (-0.312) and Malta (-0.161). Discussion on the correlation between rents and house prices is postponed to Section 4.2. House prices are in general volatile in Austria (0.051), Estonia (0.058) and Latvia (0.059).

Appendix Table A6 and Figure A3 summarize historical performance of the country specific equity indices used in the study.

Table 3 reports summary statistics of household level variables used in the analysis.

***Table 2**. Summary of Labor Income Risk and Housing Risk Variables on the Country Level*

The table summarizes labor income and housing risk variables on the country level. For labor income risk variables (CorHC and StdL) the values displayed are cross-sector averages. Table A4 of the Appendix provides definitions of the variables. All computations are on a quarterly basis and variables are not annualized.

Country Code	AT	BE	CY	DE	EE	ES	FI	FR	GR	HU
CorHC	0.056	0.099	-0.072	0.061	-0.052	0.060	0.001	0.030	-0.026	0.011
StdL	0.065	0.159	0.112	0.100	0.152	0.058	0.070	0.047	0.222	0.105
CorRnt	-0.145	0.102	0.065	0.035	0.139	0.041	-0.118	-0.044	0.052	-0.203
CorHP	-0.188	0.124	0.288	0.052	0.336	0.171	0.417	0.175	0.037	0.143
StdHP	0.051	0.013	0.025	0.014	0.058	0.024	0.012	0.016	0.030	0.026
Country Code	IE	IT	LU	LV	MT	NL	PL	PT	SI	SK
CorHC	0.071	0.110	NA	0.081	0.071	0.188	0.065	NA	-0.020	-0.035
StdL	0.054	0.179	NA	0.094	0.060	0.133	0.066	NA	0.063	0.125
CorRnt	0.279	-0.048	-0.163	-0.002	0.074	-0.126	-0.043	0.188	-0.102	0.102
CorHP	0.213	-0.312	0.048	0.493	-0.161	0.065	0.160	0.336	0.005	0.364
StdHP	0.035	0.010	0.015	0.059	0.040	0.017	0.044	0.020	0.024	0.032

Before computing the summary statistics, all nominal variables as well as leverage ratio have been winsorized at the 99th percentile in order to avoid few outlier observations affecting the results. While around 26% of households in my sample have some risky wealth in their financial portfolios, the average risky share is only around 11%. If one restricts the sample to participants only, the average risky share is around 46%. Human capital is around 500,000 euro for the average household. However, since it is equated to zero for non-employees, more than 40,000 households have a zero human capital. Hence, average human capital for households where at least one of the spouses is employee is around 1,200,000 euro. Compared to the average financial wealth of 57,433 euro, this emphasizes the importance of human capital. The low averages of income risk variables suggest that human capital behaves more like a risk-free asset rather than a risky asset at least for the average household. Around 5% of the households heads are unemployed while 32% are retired. Around 17% of households have some members depending on the income from family business. On the scale from 1 to 4, average risk aversion is at 3.68 but there is a substantial

Table 3. *Summary Statistics of Household Level Variables*

The table reports summary statistics for household level variables used in the study. The data are from second wave of Eurosystem Household Finance and Consumption Survey. All nominal amounts are in EUR. Summary statistics have been computed after winsorizing all nominal variables and leverage ratio (LR) at the 99th percentile. Column 5 of the table (Number of Zeros) reports number of observations of the corresponding variable equal to zero. Total number of observations in the sample is 84,597. Definitions of all variables are reported in Table A4. of the Appendix.

Variable	Mean	Standard Deviation	Number of Observations	Number of Zeros
Regressands				
w	0.109	0.251	73,952	56,536
PartD	0.258	0.438	76,957	56,536
Regressors				
Human capital and labor income risk				
HC	514,479	846,934	74,617	42,113
CorHC	0.002	0.055	80,608	45,593
StdL	0.027	0.042	80,608	45,593
UnemD	0.051	0.219	84,597	80,320
RetD	0.323	0.468	84,597	57,290
EntrD	0.169	0.374	84,597	70,337
Risk aversion and past equity returns				
RA	3.680	0.590	84,239	0
$E(r^e)$	0.007	0.095	84,597	0
$Var(r^e)$	0.059	0.042	84,597	0
Financial and business assets				
W	57,433	161,074	84,583	6,367
PE	28,041	130,630	81,831	70,107
Real assets and housing risk				
RRE	181,867	249,587	84,596	19,433
Vlb	12,685	20,980	84,591	14,504
CRE	20,183	87,984	84,597	77,228
RntD	0.204	0.403	84,596	67,361
CorRnt	-0.007	0.127	84,597	0
PurchHD	0.101	0.302	67,255	60,447
CorHP	0.147	0.209	84,597	0
StdHP	0.023	0.013	84,597	0
Leverage				
LR	0.227	0.639	84,548	47,017
Habit or subsistence consumption				
NHM	2.490	1.332	84,597	0
Demographic characteristics				
Educ	2.899	1.012	84,241	0
Age	53	16	84,573	0
GenD	0.625	0.484	84,597	31,701

variation across households (0.59). Average value of the residential real estate is around 182,000 euro and around 236,000 euro for households with some residential real estate. Around 20% of the households in my sample rent their main residence while 10% of the households are planning to purchase a house. The leverage ratio is at 23% for the average household. Average household size is at 2.5 members, average age of the household head is 53 years and 62.5% of the household heads are males. Overall the HFCS is representative of population of the covered countries.

4. Results

In this section, I discuss findings of the study. Given that many households are not participants, it is natural to analyze the effect of regressors on the propensity to participate in addition to the effects on the risky share. Hence, using the whole sample, I analyze the empirical relation between the participation dummy and regressors using probit regressions and the relation between the risky share and regressors using tobit regressions as in Bonaparte et al. (2014). In addition, I use OLS to understand the risky share decision of participants only as in Calvet and Sodini (2014). Motivating by the theoretical findings discussed in Section 2.1 and following the existing empirical literature, I control for variables like financial wealth, education, number of household members and other variables that were found to be important determinants of risk-taking. First, I discuss findings related to human capital and dimensions of labor income risk and then findings related to real estate and dimensions of housing risk.

4.1. Human Capital and Labor Income Risk

Table 4 reports estimates of probit, tobit and OLS regressions focused on the baseline analysis of the effect of human capital and labor income risk on propensity to participate and the risky share. Focused discussion on the theoretical expectations is in Section 2.1.1.

First, my results suggest that human capital is an important determinant of the financial risk-taking. Surprisingly, human capital has a positive effect on the propensity to participate but negative effect on the risky share among participants. Tobit estimates reveal that risky share decision among both participants and non-participants is dominated by the participation decision. This effects are highly statistically significant under any specification. The magnitude of the effect decreases once I control for the unemployed and retired household heads as well as entrepreneurial households. Considering the summary statistics reported in Table 3, effect of human capital is also economically meaningful. In particular, given a one million euro rise in the human capital, household is 9% more likely to be a participant controlling for all other variables. On the other hand, among participants a one million euro rise in the human capital is expected to decrease the risky share by 2.4 percentage points. This finding is contrary to results of Calvet and Sodini (2014) which suggest that human capital increases the risky share of participants in the Swedish population.

Second, I find that correlation of labor income with the risky asset has an expected negative effect on the participation decision. This is in line with the findings of Bonaparte et al. (2014). Once controlled for the other determinants, one standard deviation increase in the correlation, decreases the propensity to participate by 2 percentage points (= 0.364 × 0.055 × 100). Thus, in economic terms, effect of the correlation of labor income with the risky asset is found to be smaller compared to Bonaparte et al. (2004) in which comparable figure would be around 11 percentage points. The correlation variable, however, does not significantly affect the risky share decision of the participants (the positive coefficient in model (7) arises due to positive correlation of *CorHC* with the $Var(r^e)$ which is omitted at first). This result is in line with the findings of Calvet and Sodini (2014). Tobit regression estimates are dominated by the participation decision.

Third, similar finding is made for the general risk of labor income. While more volatile labor income makes a household less likely to be a participant, it does not significantly affect the risky share decision of already participating households (the negative coefficient in model (7) arises because *StdL* happens to be positively correlated with $Var(r^e)$ which is omitted at first). The effect on the participation decision is consistent with findings of Bonaparte et al. (2014). As discussed earlier, since my estimate of human capital assumes same discount rate for all households, general risk of labor income is likely to affect risk-taking through human capital. In particular, general risk of labor income is likely to decrease the value of the human capital which is found to affect the propensity to participate positively. Hence, more volatile labor income has an expected negative effect on the propensity to participate.

Although households with unemployed household heads on average have a higher risky share, in the whole sample, contrary to Bonaparte et al. (2014), I find that unemployment decreases the likelihood of household's participation. Households with a retired household head are more likely to participate and have a higher risky share on average. This result is consistent with expectations since, everything else equal, retired investors have a safer labor income and hence should be encouraged to participate more. As expected, already participating entrepreneurial households have a lower risky share although the effect is not significant. On the other hand, in terms of participation decision, entrepreneurial households are more likely to be a participants. This result is unexpected since such households presumably have a riskier labor income and should be discouraged to participate. However, such households might be participating more due to unobserved overconfidence which can be correlated with entrepreneurship.

It can be concluded that while human capital and labor income risk has a mostly expected effect on the participation decision, participating households choose their risky shared contrary to intuition. Everything else equal, increase in the human capital decreases the risk share as if human capital was risky. However, notice that general risk of the labor income does not positively affect the risky share which would be the case if human capital was risky. This facts combined suggest that participating households choose their risky shares sub-optimally.

Results on the main control variables like financial wealth, risk aversion and leverage ratio are broadly consistent with the existing literature. Wealthier, less risk averse and households with lower leverage are more likely to be participants while wealthier and less risk averse households are also expected to have a lower risky shares. As expected, more volatile past returns on the risky asset negatively affect the propensity to participate and the risky share of participants. Interestingly, I find that while average past 5-year returns on the risky asset encourages participation it also decreases the risky share among participants. Although this measure is a very noisy estimate of the expected returns, the former finding is in line with expectations. The latter finding might suggest the presence of a well-known disposition effect (e.g., see Barberis and Thaler, 2003 p. 50-51). That is, everything else equal, households which presumably experienced higher returns on their financial portfolios in the past few years, have sold off some of the risky asset to cash-in the "wins".

***Table 4**. Human Capital and Labor Income Risk Baseline Regression Estimates*

The table reports estimates of probit (column 1-3), tobit (column 4-6) and OLS (column 7-9) regressions focused on the baseline analysis of the effect of human capital and labor income risk on propensity to participate (PartD, columns 1-3) and the risky share (w, columns 4-9). All standard errors (reported in parentheses below the parameter estimates) refer to the heteroskedasticity consistent standard errors. Probit regression estimates are estimates of marginal effects. All nominal variables are in euro million. Estimates for intercept and additional control variables (PE, RRE, Vlb, CRE, RntD, NHM, Educ, GenD) have been suppressed. "Controls" row indicates which models include the additional control variables. N stands for number of observations. R^2 is the McFadden pseudo R squared for probit and the adjusted R squared for OLS regressions. The data are from second wave of Eurosystem Household Finance and Consumption Survey. Definitions of all variables are reported in Table A4 of the Appendix.

	PartD			*w*					
	Probit			Tobit			OLS		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
HC	0.139***	0.124***	0.091***	0.206***	0.148***	0.109***	-0.040***	-0.024***	-0.024***
	(0.002)	(0.003)	(0.003)	(0.007)	(0.008)	(0.009)	(0.003)	(0.003)	(0.003)
CorHC	-0.247***	-0.468***	-0.364***	-0.535***	-0.719***	-0.602***	0.229***	-0.012	-0.005
	(0.031)	(0.033)	(0.032)	(0.091)	(0.085)	(0.097)	(0.035)	(0.034)	(0.037)
StdL	-1.622***	-1.040***	-0.966***	-2.802***	-1.177***	-1.241***	-0.362***	-0.030	-0.029
	(0.052)	(0.062)	(0.060)	(0.244)	(0.238)	(0.263)	(0.079)	(0.082)	(0.085)
UnemD		-0.079***	-0.057***		-0.164***	-0.121*		0.040**	0.039*
		(0.009)	(0.008)		(0.061)	(0.070)		(0.020)	(0.021)
RetD		0.043***	0.028***		0.124***	0.104***		0.044***	0.042***
		(0.005)	(0.005)		(0.018)	(0.022)		(0.007)	(0.008)
EntrD		0.068***	0.021***		0.151***	0.080***		-0.010	-0.005
		(0.006)	(0.006)		(0.015)	(0.019)		(0.006)	(0.007)
W		1.712***	1.221***		1.275***	0.918***		0.207***	0.195***
		(0.116)	(0.099)		(0.019)	(0.023)		(0.009)	(0.012)
RA		-0.118***	-0.100***		-0.233***	-0.218***		-0.063***	-0.063***
		(0.003)	(0.003)		(0.008)	(0.009)		(0.004)	(0.004)
$E(r^e)$		0.572***	0.432***		0.931***	0.775***		-0.484***	-0.475***
		(0.029)	(0.027)		(0.098)	(0.108)		(0.037)	(0.039)
$Var(r^e)$		-0.533***	-0.567***		-0.539**	-0.690**		-1.683***	-1.658***
		(0.057)	(0.053)		(0.258)	(0.285)		(0.105)	(0.109)
LR		-0.037***	-0.021***		-0.073***	-0.048*		0.003	0.002
		(0.003)	(0.004)		(0.021)	(0.025)		(0.006)	(0.007)
Controls	No	No	Yes	No	No	Yes	No	No	Yes
N	66,453	66,162	63,598	64,288	64,015	61,611	15,635	15,575	14,111

R^2	0.048	0.256	0.280			0.026	0.112	0.107
Note:								*p<0.1; **p<0.05; ***p<0.01

Next, I consider an alternative model specification. Motivated by equation (6), correlation of the labor income with the risky asset times the human capital to wealth ratio should be an important determinant of risk-taking.[25] Table 5 reports estimates of probit, tobit and OLS regressions focused on the alternative analysis of the effect of human capital and labor income risk on propensity to participate and the risky share. I control for human capital, correlation of the labor income with risky asset and wealth, in order to isolate the effect of the term from the known relations discussed earlier.

Results suggest that the $CorHC \times HC/W$ term is an important determinant of risk-taking. The term increases the propensity to participate while it negatively affects the risky share among participants. In probit and OLS, the coefficient is highly statistically significant. Tobit specifications are dominated by participation decision but coefficient of the term is only significant when not controlled for its components. Based on models (1) and (7), one standard deviation increase in the $CorHC \times HC/W$ term (0.026 for the scaled version), increases the propensity to participate by 6 percentage points (=2.326 × 0.026 × 100) while it decreases the risky share by 3.9 percentage points (= 1.502 × 0.026 × 100). There is, however, almost no change in the R squared when comparing models (3) and (9) of Table 4 and Table 5.

This findings resemble the results of the baseline analysis discussed above. From equation (6) it follows that for a human capital with sufficiently low levels of risk, the $CorHC \times HC/W$ term would positively affect risk-taking which is what can be observed in the probit and tobit regressions for the whole sample. Participating households, however, choose their risky shares contrary to this prediction.

4.2. Real Estate and Housing Risk

Before discussing main results on the real estate and housing risk, I discuss the relation between house prices and rent in my sample. As Section 3.3 has illustrated, second

[25] Household with no financial wealth have been assigned a notional amount of 1 euro. The $\mathrm{CorHC} \times \mathrm{HC/W}$ term has been winsorized at 1st and 99th percentile.

***Table 5**. Human Capital and Labor Income Risk Alternative Regression Estimates*

The table reports estimates of probit (column 1-3), tobit (column 4-6) and OLS (column 7-9) regressions focused on the alternative analysis of the effect of human capital and labor income risk on propensity to participate (PartD, columns 1-3) and the risky share (w, columns 4-9). All standard errors (reported in parentheses below the parameter estimates) refer to the heteroskedasticity consistent standard errors. Probit regression estimates are estimates of marginal effects. All nominal variables are in euro million. The $CorHC \times HC/W$ term has been scaled by 1/10,000. Estimates for intercept and additional control variables (PE, RRE, Vlb, CRE, RntD, NHM, Educ, GenD) have been suppressed. "Controls" row indicates which models include the additional control variables. N stands for number of observations. R^2 is the McFadden pseudo R squared for probit and the adjusted R squared for OLS regressions. The data are from second wave of Eurosystem Household Finance and Consumption Survey. Definitions of all variables are reported in Table A4 of the Appendix.

	PartD			*w*					
	Probit			Tobit			OLS		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\frac{CorHC \times H}{W}$	2.326***	1.783***	1.722***	1.525***	1.026	0.959	-1.502***	-2.259***	-2.257***
	(0.221)	(0.294)	(0.311)	(0.470)	(0.676)	(0.785)	(0.513)	(0.539)	(0.544)
StdL	-0.095	-3.190***	-3.195***	-0.397**	-1.135***	-1.203***	-0.965***	-0.047	-0.047
	(0.123)	(0.188)	(0.196)	(0.162)	(0.239)	(0.264)	(0.064)	(0.081)	(0.085)
HC		0.387***	0.305***		0.147***	0.108***		-0.024***	-0.024***
		(0.009)	(0.009)		(0.008)	(0.009)		(0.003)	(0.003)
CorHC		-1.595***	-1.354***		-0.785***	-0.664***		0.027	0.038
		(0.102)	(0.108)		(0.089)	(0.101)		(0.035)	(0.038)
W		5.351***	4.109***		1.276***	0.918***		0.207***	0.195***
		(0.088)	(0.096)		(0.019)	(0.023)		(0.009)	(0.012)
UnemD		-0.270***	-0.210***		-0.165***	-0.122*		0.040**	0.039*
		(0.034)	(0.036)		(0.061)	(0.069)		(0.019)	(0.021)
RetD		0.131***	0.094***		0.124***	0.104***		0.044***	0.042***
		(0.016)	(0.018)		(0.018)	(0.022)		(0.007)	(0.008)
EntrD		0.203***	0.069***		0.151***	0.080***		-0.010	-0.005
		(0.016)	(0.018)		(0.015)	(0.019)		(0.006)	(0.007)
RA		-0.368***	-0.338***		-0.233***	-0.218***		-0.063***	-0.063***
		(0.010)	(0.010)		(0.008)	(0.009)		(0.004)	(0.004)
$E(r^e)$		1.787***	1.451***		0.929***	0.774***		-0.483***	-0.474***
		(0.077)	(0.081)		(0.099)	(0.108)		(0.036)	(0.039)
$Var(r^e)$		-1.620***	-1.868***		-0.506*	-0.662**		-1.678***	-1.653***
		(0.180)	(0.187)		(0.259)	(0.286)		(0.104)	(0.108)
LR		-0.113***	-0.071***		-0.072***	-0.047*		0.001	0.001
		(0.011)	(0.012)		(0.021)	(0.025)		(0.006)	(0.007)

Controls	No	No	Yes	No	No	Yes	No	No	Yes
N	66,452	66,162	63,598	64,288	64,015	61,611	15,635	15,575	14,111
R^2	0.001	0.257	0.280				0.014	0.114	0.109
Note:									*p<0.1; **p<0.05; ***p<0.01

moments of house prices and rents are not always varying similarly across countries. Table 6 plots correlation between house prices (more precisely, log housing returns) and rents (more precisely, log growth of the rents). Even though this correlation is substantially positive for most of the countries, the average correlation of 29% is far from the 100% correlation which would be implied by the assumption of the typical theoretical framework (see discussion in Section 2.1.2).

***Table 6**. Correlations Between House Prices and Rents*

The table shows correlation between house prices (more precisely, log housing returns) and rents (more precisely, log real growth of the rents) across 20 HFCS countries. Sources of data are described in Section 3.1.3.

AT	BE	CY	DE	EE	ES	FI	FR	GR	HU
0.090	0.100	0.373	0.247	0.584	0.371	0.255	0.106	0.657	0.447
IE	IT	LU	LV	MT	NL	PL	PT	SI	SK
0.610	-0.048	0.042	0.526	0.066	0.304	0.355	0.441	0.337	-0.112

Table 7 reports estimates of probit, tobit and OLS regressions focused on the analysis of the effect of real estate and housing risk on propensity to participate and the risky share. Among other variables, I control for the human capital and the two labor income risk variables. Focused discussion on the theoretical expectations is in Section 2.1.2.

Consistent with expectations, I find that renters are less likely to be participants and among participating households, renters have a lower risky share. This result is highly statistically significant under all specifications. The effect is also economically significant. In particular, renters are 10% less likely to be participants everything else equal. Although the difference in participation rates (14% for renters and 30% for owners) is not as extreme, it is consistent with Vestman (2018). Among participants, renters have on average 5.7 percentage points lower risky share. Observe that, since I control for the house purchase motive, this effect is solely due to renting. Households planning to purchase a house are more likely to be participants. Among participants, however, the plan to purchase a house does not affect the risky share decision by itself.

Conditional on being a renter of the main residence, household is surprisingly less likely to be a participant as correlation between the rent and risky asset goes up. In fact, as discussed earlier, one would expect the opposite relation. In OLS, however, once controlled for all other variables, increase in the correlation of rent and the risky asset increases the risky share of renting participants. In economic terms, one standard deviation increase in *CorRnt*, makes household 3% (= 0.237 × 0.127 × 100) less likely to be a participant conditional on being a renter of its main residence. For participants, on the other hand, one standard deviation increase in *CorRnt*, increases the risky share by 2 percentage points (= 0.158 × 0.127 × 100) conditional on being a renter.

Interestingly, I find a similar pattern for the house price hedging motive. Conditional on the ownership, households are more likely to participate as correlation between house price and the risky asset increases. As for the rent hedging motive, this result is opposite to expectations. Among participants, however, as one would expect, the variable affects the risky share negatively conditional on the ownership. Coefficients of the tobit regression on the whole sample are dominated by the participation decision. In economic terms, one standard deviation increase in *CorHP*, makes household 3.3% (= 0.158 × 0.209 × 100) more likely to be a participant, conditional on being an owner of its main residence. For participants, one standard deviation increase in *CorHP*, increases the risky share by 3.7 percentage points (= 0.176 × 0.209 × 100) conditional on ownership.

In addition, I find that more volatile house prices make households less likely to be participants and make participating households decrease their risky shares conditional on the ownership. That is, as expected, risky residential real estate crowds out risky financial asset ownership. Conditional on participation and planning to purchase a house, households have a lower risky shares as correlation between house prices and the risky asset goes up. This result is contrary to expectations since risky asset offers a good hedge when the correlation of the house price (future expense for the potential buyers) and the risky asset is high and should hence encourage ownership of the risky asset instead of discouraging it.

The value of the residential real estate owned by a household increase its propensity to participate but decreases the risky share of the participants while controlling for all other assets of the household. One million euro increase in the residential real estate wealth

***Table 7**. Real Estate and Housing Risk Regression Estimates*

The table reports estimates of probit (column 1-3), tobit (column 4-6) and OLS (column 7-9) regressions focused on the analysis of the effect of real estate and housing risk on propensity to participate (PartD, columns 1-3) and the risky share (w, columns 4-9). All standard errors (reported in parentheses below the parameter estimates) refer to the heteroskedasticity consistent standard errors. Probit regression estimates are estimates of marginal effects. All nominal variables are in euro million. Estimates for intercept and additional control variables (HC, CorHC, StdL, UnemD, RetD, EntrD, PE, Vlb, NHM, Educ, GenD) have been suppressed. "Controls" row indicates which models include the additional control variables. N stands for number of observations. R^2 is the McFadden pseudo R squared for probit and the adjusted R squared for OLS regressions. The data are from second wave of Eurosystem Household Finance and Consumption Survey. Definitions of all variables are reported in Table A4 of the Appendix.

	PartD			*w*					
	Probit			Tobit			OLS		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RntD	-0.170***	-0.108***	-0.099***	-0.536***	-0.308***	-0.292***	-0.094***	-0.073***	-0.057***
	(0.003)	(0.004)	(0.005)	(0.033)	(0.033)	(0.040)	(0.010)	(0.011)	(0.012)
PurchHD	0.057***	0.031***	0.029***	0.083***	0.034	0.043	-0.031***	-0.027**	-0.015
	(0.007)	(0.007)	(0.008)	(0.030)	(0.029)	(0.034)	(0.010)	(0.010)	(0.012)
$RntD \times CorRnt$	-0.256***	-0.139***	-0.237***	-0.575**	-0.249	-0.396	0.022	0.069	0.158**
	(0.029)	(0.032)	(0.032)	(0.223)	(0.234)	(0.263)	(0.064)	(0.064)	(0.068)
$OwnD \times CorHP$	0.224***	0.240***	0.158***	0.350***	0.313***	0.165***	-0.310***	-0.225***	-0.176***
	(0.010)	(0.009)	(0.011)	(0.037)	(0.038)	(0.049)	(0.012)	(0.014)	(0.017)
$OwnD \times StdHP$	-5.022***	-3.618***	-3.162***	-11.31***	-7.159***	-6.704***	-1.240***	-0.703***	-0.461*
	(0.135)	(0.132)	(0.139)	(0.706)	(0.694)	(0.804)	(0.242)	(0.259)	(0.280)
$PurchHD \times CorHP$	-0.027	0.0004	-0.044	-0.178	-0.113	-0.236*	-0.164***	-0.124***	-0.104**
	(0.022)	(0.023)	(0.028)	(0.113)	(0.111)	(0.135)	(0.039)	(0.040)	(0.044)
RRE		0.181***	0.138***		0.367***	0.287***		-0.043***	-0.030**
		(0.010)	(0.012)		(0.023)	(0.030)		(0.011)	(0.013)
CRE		0.065**	0.037		0.128**	0.099		-0.043*	-0.015
		(0.031)	(0.032)		(0.053)	(0.060)		(0.024)	(0.026)
W		1.190***	1.085***		1.155***	1.058***		0.241***	0.213***
		(0.085)	(0.090)		(0.028)	(0.033)		(0.014)	(0.016)
RA		-0.124***	-0.101***		-0.270***	-0.235***		-0.052***	-0.060***
		(0.003)	(0.003)		(0.009)	(0.011)		(0.004)	(0.004)
$E(r^e)$		0.530***	0.251***		1.068***	0.540***		-0.329***	-0.401***
		(0.021)	(0.023)		(0.105)	(0.127)		(0.039)	(0.046)
$Var(r^e)$		0.298***	-0.085*		0.457	-0.264		-1.156***	-1.361***
		(0.043)	(0.047)		(0.285)	(0.350)		(0.115)	(0.130)
LR		-0.022***	-0.031***		-0.067**	-0.078*		-0.015*	0.010

		(0.003)	(0.004)		(0.030)	(0.041)		(0.008)	(0.009)
Controls	No	No	Yes	No	No	Yes	No	No	Yes
N	64,540	64,358	53,457	64,202	64,030	53,173	13,688	13,661	11,059
R^2	0.042	0.24	0.269				0.051	0.101	0.113
Note:									*p<0.1; **p<0.05; ***p<0.01

increases the propensity to participate by 13.8 percentage points and the risky share by 3 percentage points among participants. On the other hand, commercial real estate does not have a statistically significant effect on risk-taking when controlled for all other variables. The effect of other variables is similar to results discussed in Section 4.1.

The ignorance of some of the hedging motives can be explained by bounded rationality concept (e.g., see Barberis and Thaler, 2003). While investors might understand that they need to consider the interactions between their labor income as well as housing related expenses with the risky asset, due to complexity of the problem, an investor might think about the returns on the financial portfolio in isolation. Such principle, for example, motivates the choice of the utility function used in the explanation of equity premium puzzle by Barberis, Huang, and Santos (2001).

Overall, while some of the results are consistent with expectations, most of the empirical regularities related to real estate and housing risk are rather puzzling. While these results need not be of a causal nature, the observed relations point to the sub-optimal portfolio construction and a potentially substantial welfare losses. For example, consider two otherwise identical households, one future house buyer and one with no plans to buy a house. My results suggest that if the former household is from a country with a higher correlation between house prices and the risky asset, it is likely to have a higher risky share compared to the latter household. This indicates that at least one of these two households has a sub-optimal portfolio because risky asset offers a hedge for the prospective house buyer and hence, it should be invested more heavily in that risky asset.

5. Conclusion

Motivated by the predictions of optimal portfolio choice life-cycle models, I study empirical relations between household hedging motives and financial risk-taking. Such hedging motives should occur due to the presence of human capital and labor income risk as well as real estate and housing risk. I use rich cross-sectional data from the second wave of the Eurosystem Household Finance and Consumption Survey (HFCS). The survey was conducted between 2013 and 2015 by ECB among 20 European countries. This cross-sectional data is used in combination with long time series of labor income and real estate related variables. The risky financial asset held by households is proxied by the local equity indices. With the probit regressions I analyze the effects on the propensity to participate (i.e. own some risky asset) and with tobit and OLS I analyze the effects on the risky share (share of the risky assets in household's financial wealth) while controlling for the other variables known to be important determinants of risk-taking.

I find that households with a bigger human capital are more likely to be participants while at the same time, human capital decreases the risky share of participants. As one would expect, higher correlation between the risky asset and labor income reduces the chances of a given household to be a participant. However, this labor income risk does not affect the risky share of the participating households. General risk of the labor income affects risk-taking negatively and unambiguously.

For the real estate and housing risk, findings are more puzzling. First, renting a house compared to owning it, decreases risk-taking under different specifications. Moreover, participating renters behave consistently with expectations in terms of responding to the correlation between rent and the risky asset: increase in this variable makes renters increase their risky shares everything else equal. Participation decision of renters, however, is affected negatively by the same variable. For owners, the pattern is similar. While participating owners have lower risky shares for higher correlation between house price and the risky asset, increase in this variable makes owners more likely to be participants in the whole sample. Finally, contrary to expectations, prospective house buyers are found to have lower risky shares for higher levels of correlation between house price (which is a future expense for them) and the risky asset. Tobit regression estimates are almost always dominated by the participation decisions due to relatively small share of participants in my sample.

While some of the results are consistent with the theoretical literature, several findings, especially those with alternating patterns, are puzzling. These empirical regularities suggest that households are selecting sub-optimal allocation to the risky asset given their human capital and real estate hedging motives. Attempts should be made to help households achieve optimality in their portfolio in this respect. Understanding of how financial literacy affects the responsiveness to the hedging motives might provide more support for the sub-optimality of the current portfolio choices given the risks associated with labor income and housing.

References

Angerer, X., & Lam, P. S. (2009). Income risk and portfolio choice: An empirical study. *The Journal of Finance*, *64*(2), 1037-1055.

Barberis, N., Huang, M., & Santos, T. (2001). Prospect theory and asset prices. *The quarterly journal of economics*, *116*(1), 1-53.

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, *1*, 1053-1128.

Betermier, S., Jansson, T., Parlour, C., & Walden, J. (2012). Hedging labor income risk. *Journal of Financial Economics*, *105*(3), 622-639.

Bodie, Z., Merton, R. C., & Samuelson, W. F. (1992). Labor supply flexibility and portfolio choice in a life cycle model. *Journal of economic dynamics and control*, *16*(3-4), 427-449.

Bonaparte, Y., Korniotis, G. M., & Kumar, A. (2014). Income hedging and portfolio decisions. *Journal of Financial Economics*, *113*(2), 300-324.

Brunnermeier, M. K., & Nagel, S. (2008). Do wealth fluctuations generate time-varying risk aversion? Micro-evidence on individuals. *American Economic Review*, *98*(3), 713-36.

Calvet, L. E., & Sodini, P. (2014). Twin picks: Disentangling the determinants of risk-taking in household portfolios. *The Journal of Finance*, *69*(2), 867-906.

Campbell, J. Y., & Viceira, L. M. (2002). *Strategic asset allocation: portfolio choice for long-term investors*. Oxford University Press, USA.

Cocco, J. F. (2004). Portfolio choice in the presence of housing. *The Review of Financial Studies*, *18*(2), 535-567.

Cocco, J. F., Gomes, F. J., & Maenhout, P. J. (2005). Consumption and portfolio choice over the life cycle. *The Review of Financial Studies*, *18*(2), 491-533.

Davis, S. J., & Willen, P. (2000). *Using financial assets to hedge labor income risks: estimating the benefits*. working paper, University of Chicago.

Fagereng, A., Gottlieb, C., & Guiso, L. (2017). Asset market participation and portfolio choice over the life-cycle. *The Journal of Finance*, *72*(2), 705-750.

Flavin, M., & Yamashita, T. (2002). Owner-occupied housing and the composition of the household portfolio. *American Economic Review*, *92*(1), 345-362.

Guiso, L., & Sodini, P. (2013). Household finance: An emerging field. In *Handbook of the Economics of Finance* (Vol. 2, pp. 1397-1532). Elsevier.

Guiso, L., Paiella, M., & Visco, I. (2006). Do capital gains affect consumption? Estimates of wealth effects from Italian households' behavior. *Long-run growth and short-run stabilization: essays in memory of Albert Ando. Cheltenham: Edward Elgar*, 46-82.

Heaton, J., & Lucas, D. (1999). Stock prices and fundamentals. *NBER macroeconomics annual*, *14*, 213-242.

Malmendier, U., & Nagel, S. (2011). Depression babies: do macroeconomic experiences affect risk taking?. *The Quarterly Journal of Economics*, *126*(1), 373-416.

Markowitz, H. (1952). Portfolio selection. *The journal of finance*, *7*(1), 77-91.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, 247-257.

Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of political Economy*, *112*(S1), S164-S187.

Vestman, R. (2018). Limited Stock Market Participation Among Renters and Homeowners. *The Review of Financial Studies*, *32*(4), 1494-1535.

Viceira, L. M. (1998). Optimal Consumption and Portfolio Choice for Long-Horizon Investors. *PhD thesis, Harvard University*

Viceira, L. M. (2001). Optimal portfolio choice for long-horizon investors with nontradable labor income. *The Journal of Finance*, *56*(2), 433-470.

Yao, R., & Zhang, H. H. (2004). Optimal consumption and portfolio choices with risky housing and borrowing constraints. *The Review of Financial Studies*, *18*(1), 197-239.

Appendix

***Table A1**. Number of Households per Country*

The table reports number of households per each country in the Eurosystem Household Finance and Consumption Survey.

Country Code	AT	BE	CY	DE	EE	ES	FI	FR	GR	HU
Number of Households	2,997	2,238	1,289	4,461	2,220	6,106	11,030	12,035	3,003	6,207
Country Code	IE	IT	LU	LV	MT	NL	PL	PT	SI	SK
Number of Households	5,419	8,156	1,601	1,202	999	1,284	3,455	6,207	2,553	2,135

***Table A2**. Description of Economic Sectors*

The table lists economics sectors codes and their description. Economic sectors are classified according to NACE Rev 2. Aggregate labor income time series as well as reported sector of employment of individuals in HFCS data are based on this classification.

Sector Code	Sector Description
A	Agriculture, forestry and fishing
BTE	Mining and quarrying; manufacturing; electricity, gas, steam and air conditioning supply; water supply; sewerage, waste management and remediation activities
C	Manufacturing only
F	Construction
GTI	Wholesale and retail trade; repair of motor vehicles and motorcycles; transportation and storage; accommodation and food service activities
J	Information and communication
K	Financial and insurance activities
L	Real estate activities
M_N	Professional, scientific and technical activities; administrative and support service activities
OTQ	Public administration and defense; compulsory social security; education; human health and social work activities
RTU	Arts, entertainment and recreation; other service activities; activities of households as employers; undifferentiated goods- and services-producing activities of households for own use; activities of extraterritorial organizations and bodies

__Table A3__. Description of Equity Indices Used in the Study

Country Code	Bloomberg Ticker	Brief Index Description
AT	ATX	Capitalization-weighted index of the most heavily traded stocks on the Vienna Stock Exchange. The equities use free-float adjusted shares in the index calculation.
BE	BEL20	Free float market capitalization weighted index that reflects the performance of the 20 largest and most actively traded shares listed on Euronext Brussels and is the most widely used indicator of the Belgian stock market.
CY	CYSMMAPA	Major stock market index which tracks the performance of all companies listed on the Cyprus Stock Exchange
EE	TALSE	Capitalization weighted chain-linked index which includes all the shares listed on the Main & Secondary lists on the Tallinn Stock Exchange. The aim of the index is to reflect the current status & changes on the Tallinn market.
FI	HEX	Includes all the shares listed on the Helsinki Stock Exchange. The aim of the index is to reflect the current status and changes in the market. The HEX Index is broken down using the ICB Classification as of February 1, 2012.
FR	CAC	Free float market capitalization weighted index that reflects the performance of the 40 largest and most actively traded shares listed on Euronext Paris and is the most widely used indicator of the Paris stock market.
DE	DAX	Index of 30 selected German blue chip stocks traded on the Frankfurt Stock Exchange. The equities use free float shares in the index calculation.
GR	ASE	Capitalization-weighted index of Greek stocks listed on the Athens Stock Exchange.
HU	BUX	Capitalization-weighted index adjusted for free float. The index tracks the daily price only performance of large, actively traded shares on the Budapest Stock Exchange. The shares account for 58% of the domestic equity market capitalization.
IE	ISEQ	Benchmark value-weighted stock market index composed of companies that trade on Euronext Dublin.
IT	FTSEMIB	Consists of the 40 most liquid and capitalized stocks listed on the Borsa Italiana. Foreign shares are eligible for inclusion. Secondary lines are not eligible for inclusion. The calculation and methodology is unchanged from S&P MIB Index.
LV	RIGSE	All-share index consisting of all the shares listed on the Main & Secondary lists on the Riga Stock Exchange in Latvia with exception of the companies where a single shareholder controls at least 90% of the outstanding shares. The aim of the index is to reflect the current status & changes on the Riga market.
LU	LUXXX	Weighted index of the most capitalized (by free-float) and liquid Luxembourg stocks.
MT	MALTEX	Capitalization weighted index encompassing all shares traded on the Stock Exchange of Malta.
NL	AEX	Free float market capitalization weighted index that reflects the performance of the 25 largest and most actively traded shares listed on Euronext Amsterdam and is the most widely used indicator of the Dutch stock market.
PL	WIG20	Modified capitalization-weighted index of 20 Polish stocks which are listed on the main market.
PT	PSI20	Free float market capitalization weighted index that reflects the performance of the 20 largest and most actively traded shares listed on Euronext Lisbon and is the most widely used indicator of the Portuguese stock market.
SK	SKSM	Capital-weighted total return index that compares the market capitalization of a selected set of shares with the market capitalization of the same set of shares as of a given reference day.
SI	SBITOP	Slovenian blue-chip index. It is a free-float capitalization-weighted index comprising the most liquid shares traded at Ljubljana Stock Exchange. Each stock's weighting is capped at 30%.
ES	IBEX	Official index of the Spanish Continuous Exchange. The index is comprised of the 35 most liquid stocks traded on the Continuous market. The equities use free float shares in the index calculation.

Table A4. *Definitions of Variables Used in the Study Grouped by Themes*

The table provides brief description of the variables used in the study. Methodologies for constructing these variables are discussed in Section 3.2 in more detail.

Variable	Definition
Regressands	
w	Share of risky and risk-free financial assets invested in risky financial assets
PartD	One if *w* is positive and zero otherwise
Regressors	
Human capital and labor income risks	
HC	Sum of human capital of household head and, if present, of the spouse. Human capital of an individual is the present value of expected future real labor income.
CorHC	Income weighted average of correlation variable of household head and, if present, the one of the spouse. Correlation is between log growth of real labor income of the individual's country-sector pair and log excess returns of the equity index of country corresponding to the individual.
StdL	Income weighted average of standard deviation of real labor income log growth of the household head and, if present, the one of the spouse.
UnemD	One if household head is unemployed and zero otherwise.
RetD	One if household head is retired and zero otherwise.
EntrD	One if any of the household members received some self-employment income from the family business in the last 12 month and zero otherwise.
Risk aversion and past equity returns	
RA	Risk aversion variable ranging from 1 to 4 (from less to most risk averse) and measuring general willingness to take financial risk.
$E(r^e)$	Average annualized excess return on local equity index for 5 years before the interview.
$Var(r^e)$	Annualized variance of excess returns on local equity index for 5 years before the interview.
Financial and business assets	
W	Gross value of all financial (liquid) assets.
PE	Net asset value of household's share in non-publicly listed businesses.
Real assets and housing risks	
RRE	Value of household's main residence (partial value for partial owners) and value (partial value for partial owners) of other residential real estate properties.
Vlb	Value of cars and other vehicles and valuables.
CRE	Value of household's commercial real estate properties.
RntD	One if household rents main residence and zero otherwise.
CorRnt	Correlation between log excess returns of the local equity index and the log real growth of the Actual Rental Index of household's country.
PurchHD	One if one of the reasons why household saves is buying own home and zero otherwise.

CorHP	Correlation between log excess returns of the local equity index and the log real growth of the House Price Index of household's country.
Leverage	
LR	Leverage ratio defined as total debt divided by total wealth.
Habit or subsistence consumption	
NHM	Number of household members.
Demographic characteristics	
Educ	Ranges from 1 (primary) to 4 (tertiary) levels of highest education completed by household head.
Age	Age of the household head which falls in one of 15 age categories.
GenD	One if household head is male and zero otherwise.

Table A5. *Number of Employees in Economic Sectors*

The table reports number of employees per each economic sector across countries in the Eurosystem Household Finance and Consumption Survey.

Sector Code	A	BTE	C	F	GTI	J	
Number of Persons	1,482	2,329	9,285	3,989	14,550	2,482	
Sector Code	K	L	M_N	OTQ	RTU	NA	Total
Number of Persons	3,063	460	5,077	21,184	5,134	121	69,156

***Table A6**. Summary of Performances of Country Specific Equity Indices*

The table displays annualized average of excess log returns as well as variance and standard deviation of these returns for 20 country specific indices in the 2004 Q4 – 2015 Q1 period. For every country, the index used is described in Table A3. Figure A3. Graphical Representation of Performances of Country Specific Equity Indices provides graphical representation of past performances.

Country Code	AT	BE	CY	DE	EE	ES	FI	FR	GR	HU
$E(r^e)$	0.026	0.049	-0.256	0.089	0.057	0.067	0.063	0.048	-0.099	-0.023
$Var(r^e)$	0.089	0.041	0.243	0.044	0.110	0.040	0.056	0.038	0.124	0.086
$Std(r^e)$	0.298	0.202	0.493	0.209	0.332	0.200	0.236	0.194	0.352	0.293
Country Code	IE	IT	LU	LV	MT	NL	PL	PT	SI	SK
$E(r^e)$	0.012	0.002	0.024	-0.020	0.045	0.055	0.024	0.0005	-0.009	0.016
$Var(r^e)$	0.074	0.054	0.063	0.074	0.036	0.041	0.046	0.055	0.075	0.039
$Std(r^e)$	0.272	0.232	0.251	0.272	0.189	0.204	0.213	0.234	0.274	0.197

***Figure A1.** Examples of Nominal Labor Income Time Series for 6 Country-Sector Pairs*

Amounts are in EUR. The naming convention for country sector pairs is as follows: country code followed by a full stop followed by economic sector code. Economic sector codes are described in Table A2.

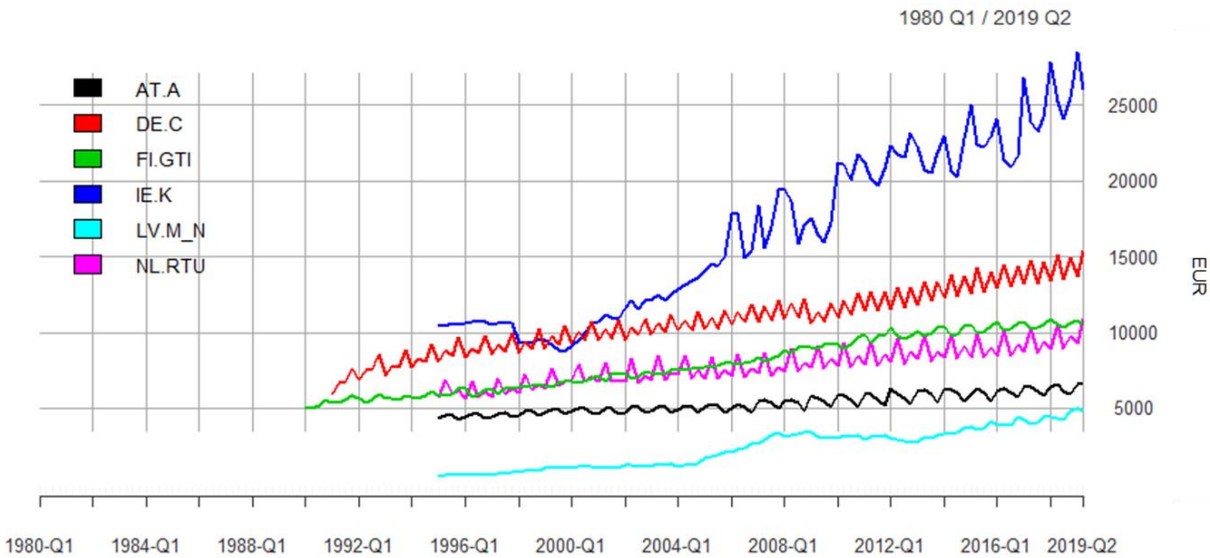


***Figure A2**. Example of the Nominal Labor Income and Values of Function $\delta^{(K)}(age_i)$*

The two figure are meant to contrast labor income observed for the year of the interview $(L_{t^*})_i$ and values of function $\delta^{(K)}(age_i)$ plotted against age of individuals for the wholesale/retail trade and repair of vehicles economic sector of Austria.

Labor income observed for the year of the interview - $(L_{t^*})_i$

Values of function $\delta^{(K)}(age_i)$

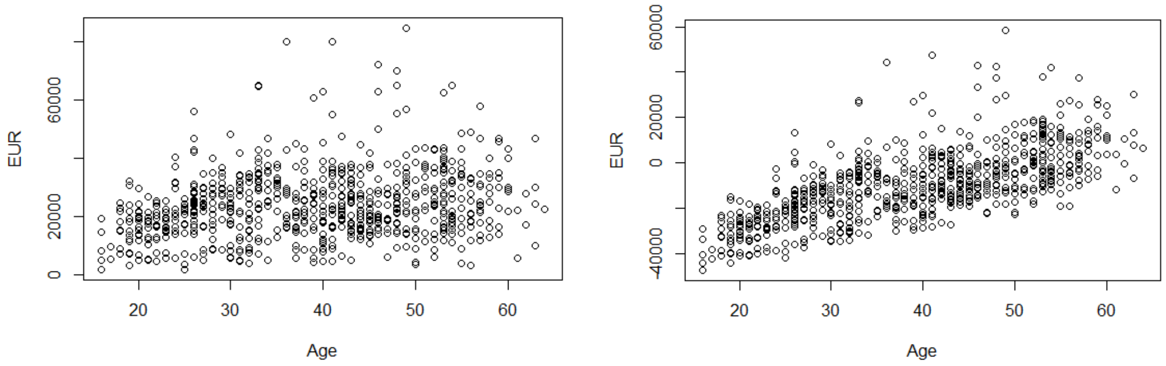
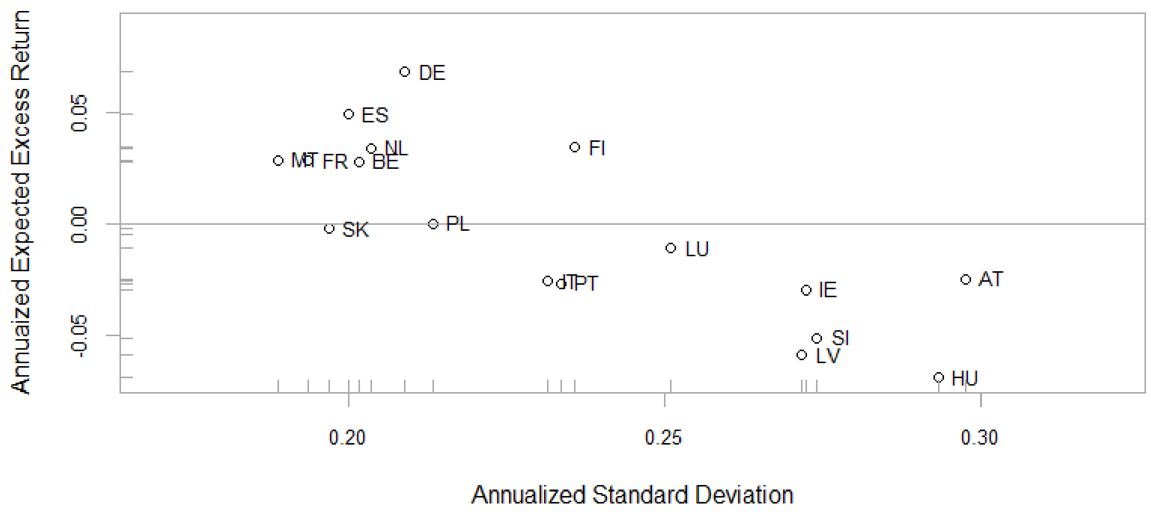


***Figure A3**. Graphical Representation of Performances of Country Specific Equity Indices*

For every country, the figure plots annualized average of excess log returns and standard deviation of these returns for the country specific equity index in the 2004 Q4 – 2015 Q1 period. The figure does not include Greece and Cyprus. For every country, the index used is described in Table A3. Table A6 summarizes the past performance.



Properties of Labor Income Risk Variables

The point of this section is to prove that if a labor income risk variable is equal among individuals belonging to the same country-sector pair then this labor income risk variable under my definition is an unbiased estimator of the true labor income risk variable.

First, I discuss the general income risk variable defined as the standard deviation of log growth of real labor income of the CS pair an employee belongs to. Consider a CS pair *K* with *n* individual. Let $\Delta l^{(K)}$be the observed log growth of real labor income of the CS pair *K* and Δl_i be the unobserved log growth of real labor income of the person *i*. My definition of $\Delta l^{(K)}$implies:

$$Var\left(\Delta l^{(K)}\right) = w_1^2 Var(\Delta l_1) + \cdots + w_n^2 Var(\Delta l_n) + 2w_1 w_2 \rho_{1,2}\sqrt{Var(\Delta l_1)Var(\Delta l_2)} + \cdots$$
$$+\ 2w_{n-1}w_n\rho_{n-1,n}\sqrt{Var(\Delta l_{n-1})Var(\Delta l_n)}$$
$$\text{and } w_1 + w_2 \ldots + w_n = 1$$

Using $Var(\Delta l_i) = Var(\Delta l)$ and $\rho_{i,j} = 1 \quad \forall i,j \in K$ we can rewrite the equation above as:

$$Var\left(\Delta l^{(K)}\right) = Var(\Delta l)\ [w_1^2 + \cdots + w_n^2 + 2w_1 w_2 + \cdots 2w_n w_{n-1}\ ]$$
$$= Var(\Delta l)\ [w_1(w_1 + \cdots + w_n) + \cdots + w_n(w_1 + \cdots + w_n)] = Var(\Delta l)$$
$$=>\ \sigma\left(\Delta l^{(K)}\right) = \sigma(\Delta l_i)\ \forall i \in K$$

That is, given that general labor income risk is same for all individuals of a given CS pair and assuming that correlation between log growth of real labor income between all pairs of individuals in the same CS pair is equal to 1, the standard deviation of every individual's real labor income log growth is equal to the standard deviation of my observed real labor income log growth of a given CS pair. If second assumption is violated, the relationship between $\sigma\left(\Delta l^{(K)}\right)$ and $\sigma(\Delta l_i)$ will inversely depend on the square root of *n.*

Next, I discuss the correlation variable defined as correlation between log growth of real labor income of individual's CS pair and log excess return of the equity index of country corresponding to the individual. Let *r* denote log excess return of the given equity index. Then my definition of $\Delta l^{(K)}$implies:

$$Cov\left(\Delta l^{(K)}, r\right) = w_1 Cov(\Delta l_1, r) + \cdots + w_n Cov(\Delta l_n, r)$$

Using $Cov(\Delta l_i, r) = Cov(\Delta l, r) \;\; \forall i \in K$ we can write $Cov\left(\Delta l^{(K)}, r\right) = Cov(\Delta l, r)$. Now, using the two assumptions in the previous proof we get:

$$Cor\left(\Delta l^{(K)}, r\right) = \frac{Cov\left(\Delta l^{(K)}, r\right)}{\sigma(\Delta l^{(K)}) \times \sigma(r)} = \frac{Cov(\Delta l, r)}{\sigma(\Delta l) \times \sigma(r)} = Cor(\Delta l, r)$$

$$=> \; Cor\left(\Delta l^{(K)}, r\right) = Cor(\Delta l_i, r) \; \forall i \in K$$

That is, given that correlation variable is same for all individuals of a given CS pair, my observed correlation variable of the CS pair is equal to the true correlation variable of individuals belonging to this CS pair.