U.PORTO

FEP FACULDADE DE ECONOMIA
UNIVERSIDADE DO PORTO

Cross-Country Heterogeneity in Household Financial Portfolios: Evidence from 22 European Countries in 2021

David Lennart Bohner - up202508842

Master Dissertation supervised by
Professor Ana Paula Serra

Master in Finance

2026

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Professor Ana Paula Serra, for her guidance, constructive feedback, and continued support throughout the development of this dissertation. Her insightful comments at each stage of the research substantially improved the analytical and methodological depth of this work.

I am also grateful to Dr. Ana Isabel Sá for taking the time to share her experience with the HFCS microdata. Her practical insights into working with the dataset provided valuable guidance in the early stages of the empirical analysis and helped me avoid several pitfalls along the way.

I also wish to thank the Faculty of Economics of the University of Porto (FEP) and as well as the Prague University of Economics and Business (VŠE) for the opportunity to pursue this Double Degree programme. The exposure to two distinct academic environments has enriched both the perspective and the rigour of this research.

Finally, I am grateful to the European Central Bank and the Eurosystem Household Finance and Consumption Network for compiling and maintaining the harmonised microdata that underpin this dissertation.

Abstract

European households exhibit substantial heterogeneity in their financial portfolio choices, both in their willingness to participate in risky financial assets and in the channels through which they gain exposure. Using harmonised data from the European Household Finance and Consumption survey (HFCS), conducted in 2021 across 22 European countries, this dissertation analyses three dimensions of household portfolio choice: 1. Cross-country variation in risky asset participation 2. The allocation between directly held shares and intermediated products 3. The consistency of life-cycle patterns across European regions. The analysis of 72,515 households, representing a weighted population of 152 million households yielded three main findings. First, household-level participation is most strongly associated with income, net wealth, and above all education. The wealth gradient for highly educated households is approximately five times stronger than for those with low education. Second, the direct versus indirect split of risky financial assets varies radically, ranging from 23% of direct holdings in the Netherlands to 92% in Cyprus. The variation is largely absorbed by country specific effects rather than by household's characteristics. Third, the hump-shaped age profile often predicted by life-cycle theory is confirmed only in Western and Nordic Europe. Southern Europe exhibits a monotonically increasing pattern and Eastern Europe relatively flat one. The findings provide an updated empirical baseline for cross-country household portfolio research and provide relevant findings for policy debates on retirement security, financial education, and the deepening of European capital markets.

Keywords: Household Finance, Portfolio Choice, Cross-country Comparison, Life-cycle Patterns, HFCS, European Union

Resumo

Existe uma grande heterogeneidade nas carteiras de investimento das famílias nos diferentes países europeus, tanto no que respeita à sua apetência para investir em ativos financeiros com risco, como dos canais escolhidos para realizar esses investimentos. Com base nos dados harmonizados do Inquérito às Finanças e ao Consumo das Famílias (HFCS), realizado em 2021 em 22 países da União Europeia, esta dissertação analisa três dimensões das decisões de investimento das famílias: 1. As diferenças entre países quanto ao investimento das famílas em ativos financeiros com risco; 2. As diferenças entre investimento directo e investimento através de fundos de investimento e outros produtos financeiros afins; 3. A consistência entre as diferentes regiões europeias dos padrões de investimento ao longo do ciclo de vida. A análise de 72 515 agregados familiares, representativos de uma população de 152 milhões de agregados familiares, apresenta três resultados principais. Em primeiro lugar, o investimento das famílias em ativos com risco está fortemente associado ao rendimento, à riqueza líquida e, sobretudo, ao nível de escolaridade. O impacto da riqueza nas famílias com elevado nível de escolaridade é aproximadamente cinco vezes superior ao impacto nas famílias com baixo nível de escolaridade. Em segundo lugar, a repartição entre ativos financeiros de risco detidos diretamente e indiretamente varia de forma acentuada entre países, oscilando entre 23% de investimento direto nos Países Baixos e 92% em Chipre. Estas diferenças são explicadas sobretudo por efeitos específicos de cada país, e não pelas características dos agregados familiares em cada país. Em terceiro lugar, o perfil etário em forma de "U invertido" , previsto pela teoria do ciclo de vida, apenas se observa na região dos países da Europa Ocidental e países Nórdicos. Os países da Europa do Sul apresentam padrões de crescimento monotónico e os da Europa de Leste revelam um padrão sem crescimento significativo ("flat"). Os resultados apresentados estabelecem uma base empírica atualizada para futuros trabalhos de investigação comparada sobre os padrões de investimento das famílias nos diferentes países da União Europeia e apresentam conclusões relevantes para os temas do atual debate europeu sobre adequação e sustentabilidade das pensões, literacia financeira e aprofundamento dos mercados de capitais, enquadrados na agenda da União dos Mercados de Capitais.

Palavras chave: Ativos financeiros das famílias; Carteiras de Investimento; Estudos Comparados; Investimento Finananceiro ao Longo do Ciclo de Vida; HFCS; União Europeia

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Introduction

Household financial portfolios and investment decisions are a fundamental pillar of economic wellbeing, shaping families' ability to smooth consumption over the life cycle, accumulate wealth for retirement, and buffer adverse economic and personal financial shocks. The composition of these portfolios reflects not only individual preferences, market access and possibly net worth, but also gives insights into broader national characteristics on the market environments and democratization of financial markets in which households make their financial decisions. In this context experts see trends such as demographic aging and lacking pension system reforms, which increase the responsibility placed on individuals to secure their personal retirement. Thus, understanding how and why households structure their financial asset holdings has become increasingly important to scientists, policymakers and financial service providers.

The theoretical foundations of household portfolio choice trace back to modern portfolio theory predicting that, given that an equity risk premium exists, rational investors should diversify their holdings to maximize expected returns for a given level of risk (Markowitz, 1952). On top of that life-cycle models suggests households should adjust portfolio composition over time, reducing exposure to risky assets, when approaching retirement age (Merton, 1969; Samuelson, 1969). Empirical evidence on advanced economies, including the Household Financial and Consumption Survey (HFCS), consistently show that real-world financial investment behaviour deviates from those predictions. Observed deviations include no risky asset participation at all, poorly diversified portfolios, often subject to the so-called home-bias and investment behaviour inconsistent with stated risk preferences. Scientists find several reasons for these patterns, highlighting participation costs, information frictions, behavioral biases, and institutional factors that vary substantially across countries shaping real-world portfolio decisions (Christelis et al., 2013; Guiso & Haliassos, 2001; Vissing-Jørgensen, 2002). Failure of participation entails forgoing equity risk premium, significantly reducing the wealth accumulation, subsequently complicating the goal of a safe and comfortable retirement. Understanding the drivers of portfolio heterogeneity is therefore essential not only for household welfare analysis but also for the design of effective macroeconomic and financial policies.

This heterogeneity is particularly visible in the European setting. Europe provides a particularly compelling laboratory for comparative analysis because its member states share many common features operating in a shared European Union's single market with many countries sharing the Euro currency. Opposingly, empirical studies document that cross-country variation in portfolio composition within Europe is often larger than regional variation within the United States, suggesting that national institutions, regulations, and cultural factors exert powerful influences on household financial decisions (Christelis et al., 2013). The magnitude of these cross-country differences is substantial. Stock market participation rates, for instance, vary dramatically across European nations. Risky asset participation rates in 2021 range from below 3 percent in Greece (1.9%) and Latvia (2.5%) to above 30 percent in Luxembourg (33.2%) and Finland (43.2%), a spread of more than 40 percentage points (own calculations based on HFCS Wave 4; see chapter 4 and Figure 1 Panel A). Country-specific factors therefore appear to account for a sizeable share of the variation in household financial behaviour.

Researchers found several specific drivers that can differ in the intensity of their impact on national investment behaviour. Pension system design plays an important role, as efficient and generous public pension schemes can crowd out private financial wealth accumulation and reduce households' incentives to hold risky assets (D'Addio et al., 2020). Further, tax incentives and regulatory frameworks governing financial markets also differ across countries, affecting both the supply of and demand for various financial products. Differences in development and accessibility of financial markets effect the availability of financial intermediaries. Countries with deeper capital markets, broader product availability, and well-developed mutual fund industries offer households more channels through which to gain equity exposure, while those with thinner markets effectively constrain participation regardless of household preferences (Bekhtiar et al., 2019). Macroeconomic experiences leave lasting imprints, as households that have lived through favourable market outcomes tend to increase their risky asset holdings, while crises and high-volatility episodes can persistently depress participation (Ampudia & Ehrmann, 2017). Behavioural variables including financial literacy as well as psychological ones being cultural attitudes towards risk taking and trust in the financial industry contribute to explanatory power. Capturing portfolio heterogeneity across European households therefore requires a framework that takes both supply-side institutions and demand-side household characteristics seriously.

The relevance of these patterns extends beyond academic literature. Household portfolio choice has direct implications for two of the most pressing policy challenges facing European governments. The first is retirement security, including worker-to-retiree ratios, rising old-age dependency, and the gradual reduction of replacement rates in many public pension systems. Policy changes included reducing payout-to-payin ratios and raising retirement ages, while states also encouraged private and occupational pension schemes through decreased tax burdens. When reforms are not binding but simply offer other additional paths of retirement security, the success relies on households' ability and willingness to accumulate adequate private financial wealth, which in turn depends on their portfolio choices. With historically decreasing returns on deposits and government bonds, while real returns are often zero or negative, many households invest in real estate and equity. Cross-country differences in portfolio composition therefore translate directly into differences in retirement preparedness, with consequences for fiscal pressure on public pension systems and for households' welfare in retirement.

Internationally the EU has committed to deepening and integrating the European capital markets. The European Union's Capital Markets Union initiative, launched in 2015 and reaffirmed in subsequent policy documents, aims to create a single market for capital across member states, reducing fragmentation, lowering the cost of capital for firms, and providing households with greater access to diverse investment opportunities. While supply-side reforms can broaden the range of available investment products, their success ultimately depends on the demand side, so if households are willing and able to use these opportunities. The composition of household portfolios, the channels through which households gain equity exposure, and the way these patterns evolve across the life cycle therefore directly determine whether retirement security and capital market integration translate into measurable progress. Understanding these dimensions across European countries is the empirical task of the present dissertation.

Research Gap and Contribution

Despite the volume of research on household portfolio choice, three gaps in the existing literature motivate this dissertation. First, much of the existing cross-country literature relies either on national surveys with limited harmonisation, which complicates direct cross-country comparisons, or on earlier waves of the HFCS conducted before the

COVID-19 pandemic. The most recent wave of the HFCS, fielded around 2021, captures household balance sheets at a moment of exceptional macroeconomic intervention and elevated saving rates, and high inflation, yet remains comparatively less studied.

Second, most of the empirical literature focuses on the extensive margin of participation, that is, on whether households hold risky financial assets at all. Fairly less attention has been paid to how participating households allocate their risky portfolios. The distinction between direct equity holdings on the one hand and intermediated investment vehicles such as mutual funds on the other, carries meaningful economic implications for diversification, transaction costs, and household exposure to information frictions.

Third, age-related heterogeneity in portfolio composition is well documented in single-country studies but is typically treated as a control variable in cross-country work rather than as a central object of analysis, leaving the cross-country consistency of life-cycle patterns underexplored.

In addressing these three gaps, this dissertation makes three corresponding contributions. First, it provides an updated descriptive baseline across 22 European countries using harmonized microdata from the HFCS Wave 4 conducted in 2021. The analysis covers 72,515 households with strictly positive financial assets, of which 17,955 hold at least one risky financial asset and form the basis for the analysis of risky-portfolio composition. Together, these households represent a weighted population of approximately 152 million across Europe. Second, the dissertation jointly analyzes participation in risky assets, the direct-versus-indirect split of risky portfolios, and life-cycle heterogeneity using survey-weighted, multiply-imputed estimation that accounts for the complex sampling design of the HFCS. Finally, the analysis explicitly tests the inter-country consistency of life-cycle patterns by interacting age groups with European macro-regions, an empirical question that previous comparative work has typically left implicit.

Three main findings emerge from this analysis. First, household-level participation in risky assets is strongly associated with income, net wealth, and especially education. Relative to households whose reference person has at most lower-secondary education, tertiary-educated households are roughly sixteen percentage points more likely to participate in risky asset markets, after controlling for income and wealth. Second, conditional on holding risky assets, we observe substantial variation across countries and their proportions of direct and

indirect held investments, ranging from 23% in direct investment in Netherlands to 92% in Cyprus. Among households that participate, those aged fifty-five and above hold approximately ten percentage points less of their risky portfolio in directly held shares, relative to 57 percent in funds for the youngest age group, shifting instead toward intermediated investment vehicles such as mutual funds. Income and education, explain only little of the direct-versus-indirect split, indicating that this allocation is shaped primarily by structural features of national financial systems and by life-cycle considerations rather than by household socio-economic status. Lastly, life-cycle patterns are not consistent across European regions. Age effect in Western and Nordic households is mainly absorbed by income, wealth, and education, while older Southern and Eastern households display systematically higher participation conditional on equal observable characteristics. The relationship between age and risky asset participation is potentially shaped by region-specific institutional and cohort factors.

The dissertation is organized in five chapters. The present chapter has introduced the research topic and motivated its policy and academic relevance. Chapter 2 reviews the theoretical and empirical literature on household portfolio choice and several tangent fields, covering portfolio theory and the participation puzzle, institutional and behavioural determinants of cross-country heterogeneity, distinctions of direct and indirect risky asset holdings, and life-cycle patterns. Chapter 3 describes the HFCS data, the construction of the analytical sample, the key outcome and control variables, and the empirical strategy used to address each research question. Particular attention is given to the survey design, the multiple-imputation framework, and the bootstrap variance estimation method employed. Chapter 4 displays the empirical results structured by the different research questions. Chapter 5 synthesizes the main findings, draws out their policy implications, acknowledges the limitations of the analysis, and outlines directions for future research.

2. Literature Review

The composition of household's financial portfolios has been part of economic research for several decades. While theoretical models prescribe how households should allocate savings across asset classes, empirical observations reveal substantial and persistent deviations from these predictions. This review synthesizes the key theoretical and empirical contributions relevant to this dissertation, organized around six themes. The last sub-chapter summarizes the literature, identifies the research gaps that motivate the present dissertation, and positions the analytical contribution of this dissertation within the field.

2.1. Household Portfolio Choice and Financial Asset Composition

To this day, most of the theories around portfolio choice are rooted in the mean-variance framework, introduced by Markowitz (1952), characterizing investment choices as an optimal trade-off between expected returns and variance. Within this framework, rational investors are expected to diversify their portfolios by utilizing assets with positive risk premiums, increasing the share of risky assets with rising risk tolerance. Extending this framework, Merton (1969) and Samuelson (1969) showed that, under standard preferences, the optimal share of risky assets is constant over the life cycle. Bodie, Merton and Samuelson (1992) included human capital as a bond-like asset, resulting in a downward facing age profile of optimal risk taking. These models imply that most households should hold at least some risky assets and explain the adjustments of portfolio composition as households age and accumulate wealth.

Empirical evidence drastically diverges from these predictions, documenting that households' portfolios are often less diversified and more concentrated than predicted, while overall participation in markets of risky assets is remarkably lower than expected. Studies in the United States showed that even among households of sufficient income and wealth, a large share does not hold stocks (Haliassos & Bertaut, 1995), with similar results across European countries (Gollier, 2001; Guiso, Haliassos, & Jappelli 2003). This discrepancy between theoretical predictions and observed behavior has become known as the portfolio participation puzzle.

One of many attempts to explain discrepancies focused on participation costs, including information acquisition costs, transaction costs, and minimum investment requirements, which may halt households from entering risky asset markets (Vissing-

Jørgensen, 2002). Another prominent study finds reasons for non-participation in the borrowing constraints and liquidity needs, arguing that households facing income uncertainty or limited access to credit may prefer liquid and low-risk assets (Guiso, Haliassos, & Jappelli, 2003; Paxson, 1990). Bertaut (1998) has found that certain education characteristics are positively associated with holdings of risky assets, indicating that more educated households are better equipped to process financial information, decreasing the hurdle to investments in complex instruments. This creates a layer of complexity as heterogeneous risk preferences result in differing portfolio compositions among economically similar households, highlighting that compositions reflect not only constraints but also behavioral factors.

Evidence from the Household Finance and Consumption Survey (HFCS) confirms that household portfolios are highly skewed: a small proportion of households holds the majority of risky financial assets, and there is substantial cross-country heterogeneity in holding patterns (Christelis et al., 2013; Zhan, 2015). These findings highlight that portfolio composition reflects not only individual constraints and preferences but also the broader institutional environment in which households operate.

2.2. Cross-Country Determinants of Portfolio Heterogeneity: Institutional and Macroeconomic Factors

A growing body of literature demonstrates that country-level institutional and macroeconomic factors account for a substantial share of cross-country heterogeneity in household portfolios, beyond what can be explained by observable household characteristics. Christelis, Georgarakos, and Haliassos (2013) show that the economic and regulatory environment, including legal institutions, financial market structure, and regulatory quality, explains more of the international variation in portfolio composition than demographic differences across populations.

Among institutional channels, pension system design is particularly important. D'Addio, Roger, and Savignac (2020) find that mandatory pension wealth crowds out private financial wealth accumulation using HFCS data, with contrasting effects across age groups. Households with larger expected pension entitlements hold less in private financial assets, reducing incentives to participate in risky asset markets. Another major effect is seen in context of Investor protections and shareholder rights. Bekhtiar, Fessler, and Lindner (2019) document that institutional measures such as shareholder protection indices are positively

associated with remaining cross-country differences in risky asset holdings. Stronger legal frameworks increase household confidence in financial markets. Countries with more sophisticated capital markets, broader product availability, and lower transaction costs exhibit systematically higher participation rates, proving the effect of deep and developed domestic financial markets (Arrondel et al., 2014).

Ampudia and Ehrmann (2017) show that households increase risk-taking following favorable stock market experiences, while financial crises and periods of high volatility can persistently depress participation. Macroeconomic experience acts as a slow-moving determinant of inter-country differences next to institutional factors. Even after controlling for the full set of observable household and institutional characteristics, however, large and statistically significant country dummies typically remain. This reflects unmeasured cultural, historical, and institutional differences that continue to shape household investment behaviour.

2.3. Behavioral and Cognitive Determinants

Next to institutional factors, behavioral and cognitive variables play a meaningful role in explaining portfolio heterogeneity. Jappelli and Padula (2015) developed a life-cycle model in which households invest in financial literacy at a cost, and literacy in turn reduces the perceived cost of participating in financial markets. They confirmed that financial literacy is strongly associated with higher participation rates, and that this relationship is weaker in countries with generous public pensions.

Kaustia, Conlin, and Luotonen (2023) provide a comprehensive decomposition of participation determinants across 19 European countries, finding that individual characteristics such as income, wealth, education, risk aversion, explain approximately 50% of the model's predictive power. More precisely, country dummies explain approximately 30%, and behavioral variables such as financial literacy, trust, cognitive ability, approximately 20%. This decomposition confirms that while economic circumstances remain the dominant predictor, behavioral frictions are far from negligible.

Another scientifically well-established field is trust into financial institutions representing a behavioral barrier to participation. Brown, Kontonikas, Montagnoli, Pickard, and Taylor (2023) document a post-crisis "flight to delegation", in which households,

especially ones with higher financial literacy, shifted toward mutual funds rather than direct stock holdings following the 2008 financial crisis. This observation suggests that financial literacy does not simply increase participation, it also shapes the channel through which households gain exposure to risk, with trust mediating the choice between direct and intermediated investment. Kochaniak and Ulman (2020) further document inconsistencies between self-reported risk attitudes and actual portfolio behavior among euro area households, underscoring that behavioral frictions can drive a wedge between stated preferences and observed investment decisions.

2.4. Trends in Risky Asset Holding and Portfolio Composition

Focusing on the smaller part of research concerning non-static models, one can find that they typically rely on repeated cross-sectional data to assess whether changes in financial market conditions are reflected in observable shifts in household portfolios. Although overall participation rates remain substantially below the levels predicted by standard portfolio theory, evidence points to a gradual increase in many advanced economies (Campbell, 2006; Guiso, Haliassos, & Jappelli, 2003). Importantly, these developments are not evenly distributed across households, as wealthier and more educated households accounted for most increases in risky holdings. Financial innovation and declining transaction costs have expanded the range of available investment products, particularly through the growth of collective investment vehicles such as mutual funds. These new products reduce information and monitoring costs and allow households to gain diversified exposure to risky assets without directly holding individual assets. Subsequently, it is not only important to observe if and how many households invest in the stock market, but also what types of product/vehicles they use to gain access. U.S. data documented gradual shifts toward more diversified financial portfolios through mutual funds, with the downside of concentration risks in periods of financial instability (Curcuru et al., 2010).

In the European context, the introduction of HFCS has substantially improved the scope for such research by providing repeated, comparable cross-sections across countries and survey waves (ECB, 2016a; ECB, 2016b), although it is not recommended to compare results from different waves. Existing HFCS-based studies document cross-country differences in risky asset participation and its determinants (Balogh et al., 2019; Zhan, 2015). More recent years have been characterized by the democratization of capital markets,

digitalization of investment access, and rapid inflows into exchange-traded funds (ETFs), developments that possibly have complex implications for household portfolio composition.

2.5. Direct versus Indirect Risky Asset Holdings

The channel through which participation occurs is of importance when aiming to grasp household investment behavior drivers and implications. Households may gain exposure to risky assets either by holding individual securities or baskets of assets directly or by investing indirectly through collective investment vehicles such as mutual funds. Distinguishing between these forms of participation is economically meaningful, as they differ in several characteristics. Direct ownership requires greater information acquisition, monitoring effort, and minimum investment thresholds, while indirect investment through funds offers professional management, diversification, and typically lower barriers to entry though at the cost of management fees and potential agency problems (Guiso, Haliassos, & Jappelli, 2003).

Household survey data show that direct equity holding is more common among wealthier, more educated households and those with higher financial literacy, while indirect holdings through funds are more widespread across the population (Haliassos & Bertaut, 1995). Though indirect investments come across as lowering the barrier to entering the stock market, the effect is minimal, resulting in reduced but not eliminated participation frictions.

Cross-country variation is substantial, in countries with more developed financial markets and deeper capital market infrastructure. Households are more likely to hold both individual equities and investment funds, whereas in less developed markets, risky asset exposure is more limited and often concentrated in indirect forms (Christelis et al., 2013; Guiso, Haliassos, & Jappelli, 2003). Studies using HFCS waves confirm strong correlations between channel choice, wealth, and education, though the distinction between direct and indirect holdings has often been treated as secondary to overall participation (Balogh et al., 2019).

From a methodological standpoint, modelling the channel choice and conditional portfolio shares involves outcomes with different statistical properties, such as binary participation decision and a continuous portfolio share bounded in the unit interval, and the latent allocation between two channels. Linear approaches such as the linear probability model and ordinary least squares with country dummies offer an alternative trade-off,

sacrificing some of the bounded-outcome properties in exchange for direct percentage-point interpretation of marginal effects.

2.6. Life-Cycle Effects and Age Differences in Household Investment Behavior

Another central part of research examines how portfolio choice changes systematically over the life cycle as households age, reflecting shifts in investment horizons, labor income risk, and accumulated wealth. Theoretical models predict that younger households allocate larger shares to risky assets, with reasons being longer investment horizons and higher implicit holding of human capital, while older households typically rebalance toward safer assets approaching retirement age (Bodie, Merton, & Samuelson, 1992; Cocco et al., 2005; Merton, 1969).

While there is a broad consensus supporting age-related patterns in household investment behaviour, empirical work documents irregularities across institutional contexts. Ameriks and Zeldes (2004) document a hump-shaped relationship between age and overall stock market participation, with participation increasing towards early working ages and fading with higher seniority. Studies based on European data validate this pattern even after controlling for income and wealth (Christelis et al., 2013; Guiso, Haliassos, Jappelli, et al., 2003), highlighting the importance of life-cycle considerations when encountering household portfolio data. Researchers have also examined contrasts in investment vehicles across age groups, finding that younger households tend to rely on more simplified or intermediated products, while older households hold investments more directly, reflecting differences in experience, resources, and market exposure (Bertaut & Starr-McCluer, 2002; Zhan, 2015).

A critical challenge is that observed age differences in a single cross-section are mixed with both life-cycle effects (changes within households as they age) and cohort effects (persistent differences across birth cohorts shaped by distinct historical experiences). Ameriks and Zeldes (2004) and Campbell (2006) emphasize that disentangling these two sources of variation requires longitudinal panel data. The cross-sectional structure of any single HFCS wave therefore limits causal inference about life-cycle dynamics and constrains analysis to descriptive interpretation of age patterns. However, the consistency of these patterns across countries, has rarely been examined explicitly in the comparative literature.

2.7. From Literature to Research Questions

The literature on household portfolio choice provides strong theoretical foundations and extensive empirical evidence on the determinants of risky asset holding. Classical portfolio models generate diversification and life-cycle considerations that are systematically violated in the data. More recent contributions extend the status quo by incorporating how financial wealth is allocated across asset classes and investment channels. Participation costs, borrowing constraints, and behavioral frictions explain part of the gap between theory and observation, while institutional factors, pension systems, investor protection, financial market development, and cultural attitudes, account for a substantial share of cross-country heterogeneity. Financial literacy and trust emerge as the most quantitatively important behavioral determinants, together explaining roughly 20% of participation variation across European countries (Kaustia et al., 2023).

Despite these advances, three gaps remain. First, most comparable work relies either on national surveys with limited harmonization or on earlier waves of the HFCS. Wave four, reflect a truly special moment in household finance, yet remains comparatively less studied. Most of the existing literature focuses on the extensive margin of participation, whether households hold any risky assets. Second, while the distinction between direct and indirect risky asset holdings is recognized as economically meaningful, it is rarely examined explicitly. Third, age-related heterogeneity in portfolio composition is well documented but is typically treated as a control variable rather than a central object of analysis, leaving the cross-country consistency of age patterns underexplored.

Addressing these gaps requires a comparative approach that jointly considers portfolio participation, investment channels, and life-cycle heterogeneity across a broad and harmonized set of European countries. This dissertation provides exactly such an analysis, using the most recent HFCS wave (2021) covering 22 European countries and a three-part econometric framework designed to model participation, channel choice, and conditional portfolio shares simultaneously. The three research questions stated below translate these gaps into specific empirical objects of analysis.

RQ1: How does the allocation of financial wealth across major asset classes vary across European countries, and what household characteristics are associated with risky asset participation within and across these countries?

The first research question documents how households in the twenty-two HFCS countries distribute their financial wealth across deposits, listed shares, mutual funds, bonds, and other instruments. It identifies the household-level characteristics, such as income, net wealth, education, employment status, and demographic composition, that are associated with the decision to hold risky financial assets. The analysis distinguishes household-level gradients common across Europe from country-specific levels of participation.

RQ2: Conditional on holding risky financial assets, how do households allocate their risky portfolios between directly held shares and investment funds, and how does this allocation vary across countries and household characteristics?

The second research question narrows the analysis to households that hold risky financial assets, examining how they split their risky portfolios between directly held listed shares and intermediated investment vehicles, principally mutual funds. The two channels imply different levels of diversification, information requirements, transaction costs, and reliance on financial intermediaries. The question identifies whether the direct-versus-indirect split is driven by structural features of national financial systems or by household demographic and economic characteristics.

RQ3: How do risky asset participation rates and risky portfolio compositions differ across age groups, and are these life-cycle patterns consistent across European regions?

The third research question addresses heterogeneity across the life cycle. Theory suggests that households should adjust their exposure to risky assets as they age, reflecting changes in human capital, income risks, and accumulated wealth. Empirical evidence frequently documents departures from this prediction, attributed variously to behavioural inertia, financial illiteracy, and cohort-specific experiences such as exposure to stock-market crises. By splitting the population into age brackets and macro-regions the models test if life-cycle patterns in participation and channel choice are uniform across Europe or if they are shaped by region-specific institutional and cohort factors.

3. Data and Methodology

This chapter describes the data, sample construction, and empirical strategy used to address the three research questions stated in chapter 1. and starts by introducing the Household Finance and Consumption Survey (HFCS) and its data structure. It goes on by detailing the construction of the analytical sample and the definition of outcome and control variables. It proceeds by showcasing, the descriptive analysis and regression specifications estimated for each research question. 3.4 describes the inference framework, combining Rubin's rules for multiple imputation with the rescaling bootstrap of Rao and Wu (1988), and lists the robustness checks performed on the primary specifications. Finally, it closes with the scope conditions and limitations of the design.

3.1. Data and Sample

This section describes the source, structure, and key characteristics of the data underlying the empirical analysis.

3.1.1. HFCS - Survey Design and Coverage

The empirical analysis in this dissertation is based on microdata from the fourth wave of the Household Finance and Consumption Survey, a harmonized survey of households' assets, liabilities, income, consumption, and socio-demographic characteristics. It is conducted by the European Central Bank in cooperation with the national central banks and statistical institutes of participating countries. The fourth and most recent wave contains data from 19 Euro area member states as well as Croatia, Czechia, and Hungary, and involves a sample of 83,162 households. The participating countries are Austria, Belgium, Croatia, Cyprus, Czechia, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Portugal, Slovakia, Slovenia, and Spain.

The survey follows a layered probability sampling design, with deliberate oversampling of wealthy households to ensure sufficient representation of the upper tail of the wealth distribution, which is known to be highly skewed. The survey reference unit is the household, and within each household a financially knowledgeable "reference person" provides most of the answers.

The data collection coincides with a historically exceptional economic moment, since the survey fieldwork largely occurred approximately one year after the onset of the COVID-

19 pandemic. The macroeconomic context surrounding this period is discussed in section 2.4 and should be kept in mind when interpreting the results.

3.1.2. Data Structure

The HFCS User Database is distributed in a household-person relational structure. The core data are provided in several file types. H-files contain household-level variables; P-files contain person-level variables; D-files contain derived household-level aggregates such as total net wealth, total financial assets, and gross income; and a W-file contains replicate weights for variance estimation. All files except the W-file are provided in five versions corresponding to the five imputation implicates (H1-H5, D1-D5, etc.). The country identifier SA0100 and household identifier SA0010 serve as the primary linking keys across files.

Item non-response is handled through multiple stochastic imputation. For every missing value of a wealth, consumption, or income variable, the HFCS supplies M = 5 imputed replacement values. The five resulting datasets (implicates) are point estimates from a chained-equation model with country-specific specifications (ECB, 2020). Point estimates and standard errors reported in chapter 4 are obtained by combining results across the five implicates using Rubin's (1987) rules, as described in 3.4. Variance estimation uses the 1,000 bootstrap replicate weights, generated by the rescaling bootstrap of Rao and Wu (1988), which preserves the complex sampling design of stratification, clustering, and oversampling. All survey-weighted point estimates use the household sampling weight HW0010, which scales the sample to represent the population of households in each country.

3.2. Sample and Variables

3.2.1. Sample Construction

The construction of the analytical sample proceeds in three steps. Firstly, the analysis is restricted to wave four thereby ensuring cross-country comparability and avoid methodological interruptions across waves cautioned against by the ECB. Secondly, households with strictly positive total financial assets were filtered, reducing the population from 83,162 households to 79,742. This excludes households with zero financial assets, typically those without any banking relationship or any deposits, securities or other financial instruments, as well as households with undefined financial assets. Non-existent investment behaviour of these households requires a separate analytical framework. Third, households for which all reference-person characteristics (age, education, employment status, gender)

are missing across all five implicates are dropped, as are households flagged with implausible values on key wealth aggregates. Applying these restrictions the sample shrinks to 72,515 households, representing a weighted population of approximately 152 million households across the twenty-two countries. The analysis of risky-portfolio composition in RQ2 & RQ3 further restrict the sample towards participating households, resulting in 17,955 households and a weighted population of 29 million. Reference-person variables are taken from the survey-defined household reference person throughout.

3.2.2. Outcome Variables

Two outcome variables were constructed from the D-file aggregates. *Risky_part* is a binary indicator set to one if the household holds either listed shares or mutual fund shares, and zero otherwise. It serves as the dependent variable in the participation analyses for RQ1 and the participation component of RQ3.

Direct_share is the share of financial assets held directly in listed shares, defined as

$$\boldsymbol{Direct\_share} = \frac{\textbf{listed shares}}{\textbf{(listed shares + fund shares)}} \tag{3.1}$$

It is defined only for households with strictly positive risky asset holdings and serves as the dependent variable in the channel-choice model (RQ2) and the channel-choice component of RQ3 (Model 3c). A broader risky-asset definition that additionally includes bonds is examined in robustness check R2 (section 3.4.3).

3.2.3. Control Variables

The control vector consists of socio-economic and demographic variables of the reference person and the household, drawn from the H- and D-files. The continuous economic variables are transformed to mitigate skewness. Household income enters as the natural logarithm of gross household income. Household net wealth enters as the inverse hyperbolic sine transformation,

$$\boldsymbol{arsinh(x) = \ln\left(x + \sqrt{(x^2 + 1)}\right)} \tag{3.2}$$

, which behaves like a log transformation at large positive values, equals zero at zero, and is defined for negative values, thereby retaining the full sample.

Reference-person age is categorized using four age brackets, namely under 35, 35-54, 55-64, 65 and above. Education is divided into three categories following the International Standard Classification of Education (ISCED) 2011: low as lower secondary or below; ISCED 0-2, medium as upper secondary and post-secondary non-tertiary; ISCED 3-4, and

high as tertiary; ISCED 5-8. Employment status is divided into 5 brackets: employee, self-employed, retired, unemployed, and other/not in labour force. Female identifier is coded as a binary indicator. Household size is the number of household members. Homeownership is a binary indicator equal to one if the household owns its main residence either outright or with a mortgage.

Country dummies are operationalized individually based on the country SA0100, with Austria as a reference category. For the analysis of regional life-cycle effect heterogeneity (RQ3), countries are grouped into the three macro regions. Western and Nordic Europe consists of Austria, Belgium, Finland, France, Germany, Ireland, Luxembourg, and the Netherlands; Southern Europe comprises Cyprus, Greece, Italy, Malta, Portugal, and Spain; and Eastern Europe comprises Croatia, Czechia, Estonia, Hungary, Latvia, Lithuania, Slovakia, and Slovenia. The grouping follows the convention commonly used in comparative European household-finance research (e.g. Christelis et al., 2013; Bekhtiar et al., 2019). The regional grouping enters as the second dimension of the age-by-region interaction in Models 3b and 3c. Country dummies are retained in these specifications to absorb remaining within-region heterogeneity. Throughout the regression analysis, employment status indicators, household size, and country dummies are included as controls but not separately discussed in the results, as the focus lies on income, wealth, education, age, gender, and homeownership as the primary determinants of interest.

3.3. Empirical Strategy

3.3.1. Descriptive Analysis (RQ1)

To address RQ1, the analysis documents the distribution of financial wealth across major asset classes for all 22 countries. Country-level participation rates and portfolio shares are computed as weighted means, using the HFCS cross-sectional household weight. Inference follows the multiple-imputation and bootstrap framework described in section 3.4.1. Summary statistics are reported for the full sample, separately by country, and by the three macro-regions defined above to highlight regional patterns.

3.3.2. Participation Model - RQ1 Regression & RQ3 Component

The first research question examines the household-level determinants of risky asset participation. Three specifications are estimated, with Model 1a (Table 1 column 1) being a

pooled probit without country dummies to relate participation probability before national differences come into play:

$$\boldsymbol{Pr(Risky\_part_i = 1 \mid X_i) = \Phi(\alpha + X_i\beta)} \quad (3.3)$$

$\Phi(\cdot)$ is the standard normal cumulative distribution function and vector $\mathbf{X}_i$ collects the households control variables described in 3.2.3, namely log income, the inverse hyperbolic sine of net wealth, indicators for the four age groups, three education levels, five employment statuses, female reference person, household size, and homeownership. The probit with country dummies (Model 1b; Table 1 column 2) adds 21 country indicators to absorb country-level heterogeneity, with Austria as the reference. Model 1c with country dummies (Model 1c; Table 1 column 3) replaces the probit link with ordinary least squares:

$$\boldsymbol{Risky\_part_i = \alpha + X_i\beta + \delta_{c(i)} + \varepsilon_i} \quad (3.4)$$

$\delta_{c(i)}$ is the country dummy for the country in which household $\boldsymbol{i}$ resides. Several considerations support the linear probability model (LPM) as the primary specification. On benefit is the easiness of the interpretability of results. The linear coefficients can be understood as direct percentage point changes in the probability of participation. Additionally, it ensures methodological consistency with the comparable specifications used for RQ2 and RQ3, where the dependent variable is either a continuous portfolio share or interaction terms. In non-linear models such as probit, interaction effects are not directly interpretable as marginal effects (Ai and Norton, 2003). Lastly, the LPM produces unbiased coefficient estimates under the standard assumptions, and the parallel probit specification (Model 1b and 3a) are reported as a methodological cross-check confirming that the conclusions do not depend on the linearity assumptions. The share of fitted values falling outside the unit interval is reported in chapter 4 alongside corresponding probit estimates.

The choice of country dummies rather than country-level macroeconomic and institutional controls also deserves explanation. With 22 countries in the sample, identifying the marginal contribution of any single country-level variable would yield very limited statistical power. Furthermore, country-level institutional and macroeconomic indicators tend to be highly collinear with each other and with the regional grouping used in RQ3. Country dummies, by contrast, absorb the full set of country-specific factors, both observed and unobserved, including unmeasured cultural attitudes, trust in financial markets, and past experiences. The empirical focus of the dissertation lies in household-level gradients within and across countries, not in quantifying which specific country-level factors matter most.

3.3.3. Channel-Choice Model RQ2

The second research question focuses on households participating in risky assets and investigates relative differences between how they allocate towards single shares or mutual funds. The dependent variable is *direct_share* and Model 2a (Table 2 column 1) is the base specification, estimated by ordinary least squares with country dummies on the sub-sample of risky-asset holders

$$\boldsymbol{direct\_share_i = \alpha + X_i\beta + \delta_{c(i)} + \varepsilon_i} \quad (3.5)$$

Where X_i is the same control vector as in Model 1c. Although $direct\_share_i$ is bounded in the unit interval, ordinary least squares is preferred to a fractional logit or beta regression because the marginal-effect interpretation of the coefficients is more transparent and the dependent variable is not concentrated at the bounds. Less than five percent of risky-asset holders hold exclusively direct or exclusively indirect risky portfolios. Inference follows the same multiple-imputation and bootstrap procedure as in Model 1c. An extended specification (Model 2b; (Table 2 column 2) adds the interaction between age groups and macro-regions and is identical to the channel-choice model used in RQ3 - Model 3c.

3.3.4. Life-Cycle Heterogeneity across Regions - RQ3

The third research question tests whether life-cycle patterns in participation and channel choice are uniform across European regions or whether they are shaped by region-specific institutional or cohort factors. Three specifications are estimated. The participation model with the age-region interaction (Model 3b, primary, (Table 1 column 1)) is the primary specification, estimated by OLS with country dummies:

$$\boldsymbol{Risky\_part_i = \alpha + \sum_{g=2}^{4} \beta_g AgeGroup_{\{g,i\}} + \sum_{r=2}^{3} \gamma_r Region_{\{r,i\}} + \sum_{g=2}^{4}\sum_{r=2}^{3} \theta_{\{g,r\}} \beta_g AgeGroup_{\{g,i\}} Region_{\{r,i\}} + Z_i\varphi + \delta_{c(i)} + \varepsilon_i} \quad (3.6)$$

$AgeGroup_{\{g,i\}}$ indexes the four age brackets (under 35, 35-54, 55-64, 65 and above), $Region_{\{r,i\}}$ indexes the three macro-regions (Western and Nordic, Southern, Eastern), Z_i is the remaining control vector (log of household income, inverse hyperbolic sine of net wealth, education, employment status, gender indicator, household size, homeownership indicator), and $\delta_{c(i)}$ is the country dummy. The coefficients $\theta_{\{g,r\}}$ on the age-region interactions are the central objects of interest. They capture how the age slope of risky asset participation differs in Southern and Eastern Europe relative to the Western and Nordic baseline, conditional on income, wealth, education, employment, and country. Joint significance of the interaction

terms is tested via a Wald F-test. Model 3a is the probit analogue of Model 3b, reported as a methodological cross-check on the linearity assumption.

The channel-choice analogue (Model 3c) applies the same interacted specification to *direct_share* on the sub-sample of risky-asset holders and is identical to Model 2b above. Together, Models 3a-3c test whether the cross-country consistency of life-cycle patterns differs between the extensive margin of participation and the intensive margin of channel choice. One identification caveat applies to Models 3a-3c. Because country dummies and the regional grouping jointly absorb all country- and region-specific variation, the interaction coefficients $\theta_{\{g,r\}}$ document the existence and magnitude of regional differences in life-cycle patterns but cannot disentangle whether they reflect demand-side household characteristics correlated within regions, such as risk aversion or cohort-specific macroeconomic experiences, or supply-side institutional features such as financial-market depth and pension-system design.

3.4. Inference and Robustness

This section describes how point estimates and standard errors are computed across all five-imputation implicates and 1,000 bootstrap replicate weights. Further, the diagnostic tests and robustness checks applied to primary specifications are outlined.

3.4.1. Combining Multiple Imputation with the Rescaling Bootstrap

Estimates reported in chapter 4 are subject to two estimation uncertainties. Sampling uncertainty, captured by the rescaling bootstrap of Rao and Wu (1988), and imputation uncertainty, captured through Rubin's (1987) combination rules. All models are first estimated separately on each implicate, with data variance computed from 1,000 bootstrap replicate weights drawn form the W-file, resulting in implicate specific point estimates. An average of all implicates is formed, resulting in a total variance consisting of a within-imputation component (the average of the implicate-specific variances) and a between-imputation component (the variance of the implicate-specific estimates), with a finite-sample correction reflecting the limited number of imputations.

The Stata implementation procedures and the mathematical details of Rubin's combination rules and the rescaling bootstrap are reported in Appendix A.

3.4.2. Diagnostics Test

Models are accompanied by three diagnostics to logically frame and validate the results of primary specification. First, joint significance tests on grouped coefficients are computed via "mi test". The Wald F-test on the twenty-one country dummies confirms that country-level heterogeneity contributes meaningfully beyond household characteristics in Models 1b, 1c, and 2a.

The Wald F-test on the six age-by-region interaction terms provides the formal test of cross-regional heterogeneity in life-cycle patterns central to RQ3, applied to Models 3a, 3b, and 3c. A joint test on the four age-group indicators in Model 1c is also reported, since the individual age coefficients are small once income, wealth, and education are controlled for.

As linear probability models can produce fitted probabilities outside the unit interval, the share of values is computed, which did exactly that for models 1c and 3b. A small out-of-bounds share gives meaning to the results and supports the interpretation of LPM coefficients as percentage-point marginal effects, later reported in chapter 4.

Third, the parallel probit specifications used in models 1b and 3a, serve as a methodological cross-check on the primary LPM models. Substantive agreement on sign, significance, and approximate magnitude across the two estimators is reported as evidence that the conclusions do not depend on the linearity assumption of the LPM.

3.4.3. Robustness Checks

Four robustness checks are applied to the primary participation specification (Model 1c). Because the extended specifications share the same covariate set, robustness of Model 1c implies robustness of Models 2a, 2b, 3a, 3b, and 3c. The four robustness checks (R1-R4) are reported in Table 4.

R1 replaces the four discrete age-group indicators with a continuous age polynomial (age and age-squared), verifying that the income, wealth, and education effects do not depend on the discretization of age.

R2 broadens the definition of risky financial assets to additionally include bonds, probing whether the participation gradient is sensitive to the baseline asset definition.

R3 substitutes the inverse-hyperbolic-sine transformation for the natural logarithm of household income, retaining households with non-positive income and confirming that the income gradient is not an artefact of the log-based exclusion.

R4 adds the interaction between education and the inverse-hyperbolic-sine of net wealth, formally testing whether the wealth gradient varies across education levels and providing the empirical basis for the financial-literacy interpretation discussed in section 4.2. Robustness is judged by stability of sign, significance, and approximate magnitude of the three core gradients (income, wealth, tertiary education) across all four checks.

3.5. Scope and Limitations

All following interpretations of the presented results in chapter four are subject to several limitations and scope conditions. The first and most important limitation lies in the set-up, more precisely the HFCS being a single cross-sectional questionnaire. The ECB explicitly cautions on pooling waves due to methodical amendments, restricting the interpretability of results over time and the tracking of individual households. Subsequently, age-related differences in participation and channel choice, can only be reported as a combination of life cycle and cohort effects.

Another limitation is assessing patterns by measuring household characteristics at the level of the reference person ("financially knowledgeable respondent"). Demographic discrepancies within the household, particularly in couples with large age or education differences, are therefore not captured. Although, this is standard to papers on HFCS, it should be kept in mind.

The original sample was reduced by roughly 5%, through filtering for households with strictly positive financial assets, and another 9%, excluding households with missing values on key reference-person controls. If systematic differences between this excluded sample and the analyzed one are of relevance in regard to portfolio choice cannot be examined within the present analysis.

Fourth, country-level sample sizes vary greatly across the participating countries. As a result, when looking into national sub samples and filtering for channel choice (RQ2) the number of observations falls below 100 for several countries, limiting the precision of country-level descriptive estimates reported in section 4.4. Regression results pool across countries and are not affected, but country-level point estimates from these small cells are flagged as such in the corresponding tables.

Further, the macro regions used in RQ3 follow the standard categories used in the HFCS literature but inevitably mask intra-region heterogeneity. Finland and Germany, for example, significantly differ in pension system design despite inclusion in the Western &

Nordic group. The regional aggregation is intended to provide statistical power for the age-by-region interaction tests rather than to claim institutional homogeneity within each group. Country dummies retained in all RQ3 specifications absorb the remaining cross-country variation within regions. However, they do not allow demand- and supply-side drivers of the regional differences to be separately identified.

Lastly, the dataset does not include a harmonized financial-literacy module and self-reported risk attitudes are available only for a subset. The role of these behavioral variables, which the literature identifies as quantitively important determinant of participation, can therefore not be examined directly. Education is used as a proxy for the financial-literacy channel, with the limitations of this proxy acknowledged in the discussion of results in chapter 4.

4. Results and Findings

In the following chapter the results and empirical findings of in-depth data analysis of Wave 4 will be presented. Each section follows the structure of documenting descriptive patterns, the corresponding regression estimates, and finally the findings are discussed in light of the theoretical and regulatory frameworks and the prior evidence outlined in chapter 2.

All point estimates combine information from the five imputation implicates using Rubin's (1987) combination requirements, and standard errors depend on the 1,000 bootstrap replicate weight given in the HFCS W-file, following the rescaling method of Rao and Wu (1988). Regression coefficients will be reported to three decimal places paired with the bootstrap standard errors in parentheses. Statistical significance at the 1%, 5%, 10% level is denoted by three, two, and one asterisks, respectively.

Two technical caveats apply throughout. First, because the HFCS provides only five imputed datasets, standard errors on some coefficients are slightly conservative. Second, the linear probability model used as the primary specification produces fitted probabilities that occasionally fall outside the unit interval. This is a known property of the model and is addressed by comparing the linear coefficients with the corresponding probit estimates, which yield substantively identical conclusions.

4.1. Descriptive Statistics

Cross-country participation patterns

Across the complete European sample, weighted risky asset participation equals 19 percent, hence fewer than one in every five households hold risky financial instruments. This aggregate number masks substantial cross-country heterogeneity within the dataset. Participation rates range from below 1.9% in Greece and 2.5% Latvia to more than 30% in Luxembourg (33.2%) and Finland (43.2%), with Belgium (29.2%) and Germany (27.9%) just below the 30% threshold (Figure 1 Panel A). Spreads of more than 40 percentage points between highest- and lowest participating countries are consistent with patterns documented in prior HFCS waves by Christelis, Georgarakos, and Haliassos (2013) and Guiso, Haliassos, and Jappelli (2003), confirming that the limited-participation puzzle manifests itself unevenly across European economies.

Distributing the observations into regional (defined in 3.2.3) and characteristically similar groups helps structure this heterogeneity, disclosing that Western and Nordic households

exhibit the highest participation rate of 23.7%. Southern European households participate at substantially lower rates of about 14.1%, while Eastern European countries also remain below the European average at 7.8% (Figure 1 Panel B). Figure 2 extends the view by decomposing the full financial portfolio into deposits, risky assets, bonds, and other components. Deposits consistently dominate the asset classes across all countries, accounting for more than 50% of financial wealth in most countries and exceed 65% in several Southern and Eastern households. These descriptive patterns set the stage for the regression analysis, seeking the identification of household-level determinants associated with participation conditional on country context.

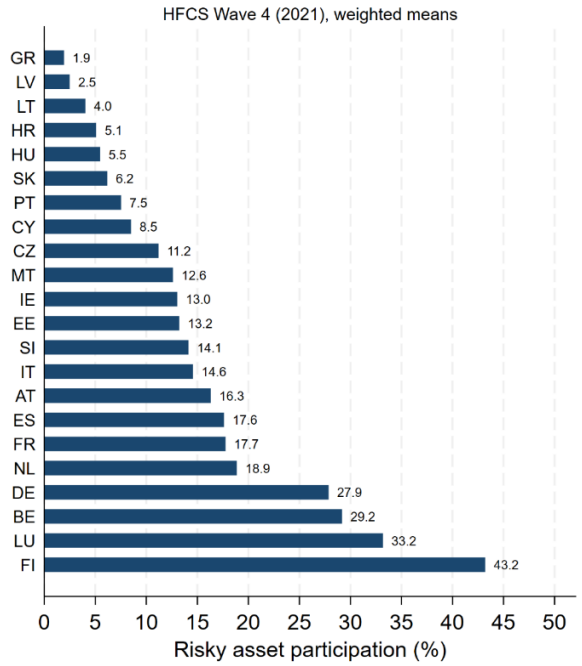

Figure 1 Panel A. Risky Asset Participation by Country

Note. Author's calculations based on wave 4

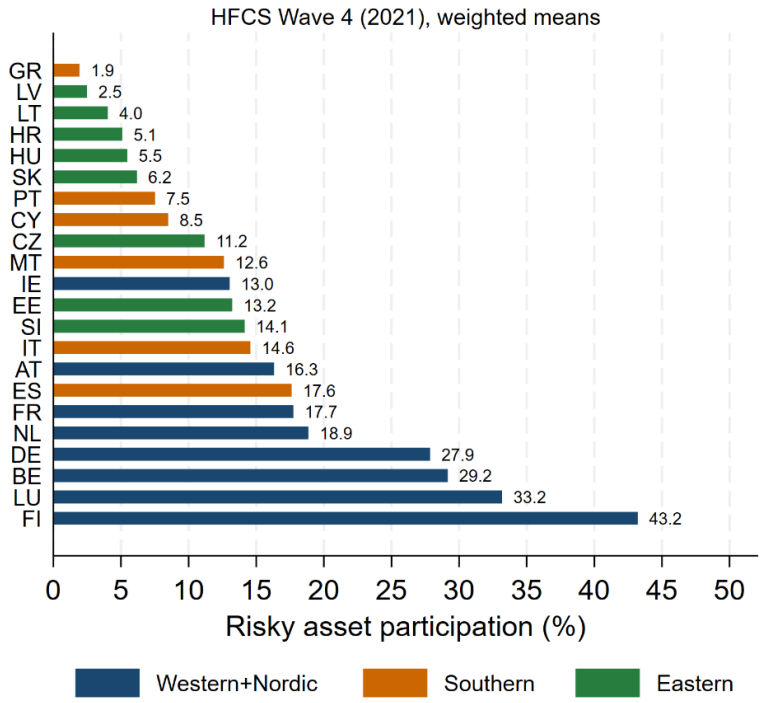

Figure 1 Panel B. Risky Asset Participation by Macro-Region

Note. Author's calculations based on wave 4

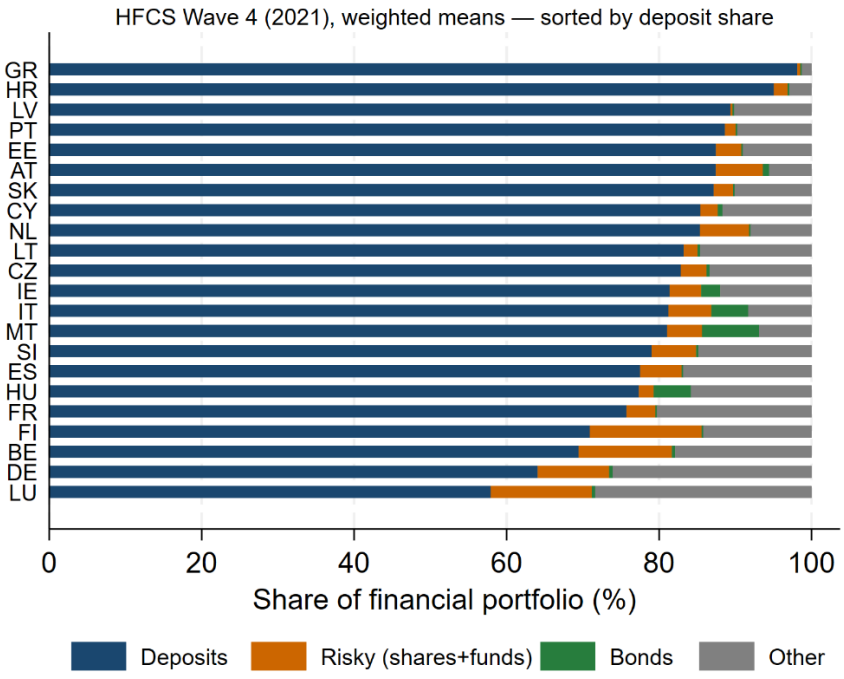

Figure 2. Composition of Financial Portfolios by Country

Note. Author's calculations based on wave 4

Direct versus indirect channel choice

Conditional on holding risky assets, the average European household splits its risky portfolio almost evenly between directly held shares and investment funds, with a weighted EU22 average direct share of approximately 42.7%. The cross-country variation, however, is dramatic. *Direct_share* values range from 23.4% in the Netherlands and 25 % in Belgium at the low end to 75.7% in Ireland, 90.7% percent in Croatia, and 92.4% in Cyprus at the high end (Figure 4 and Table 5). The countries with the lowest direct shares are characterised by mature mutual-fund industries and well-developed occupational and private pension schemes, while the countries with the highest direct shares are smaller economies in which fund-based intermediation is comparatively underdeveloped. Direct ownership is therefore not simply the choice of the most sophisticated investors, but in several European countries it is also the default channel by which households gain equity exposure when intermediated alternatives are scarce.

Two further descriptive features stand out. First, several countries with extreme *direct_share* values are based on small national sub-samples. Cyprus consists of 104 risky-asset holders, Croatia of 56, Greece of 60, Latvia of 48, and Lithuania of 74. The corresponding country-level point estimates are reported in Table 5 but should be interpreted with caution. Second, in the unconditional Panel B of Table 5, the direct share appears almost flat across age groups (43.3% for households under 35 and 43.6 for those aged 65 and above). The conditional regression analysis below tells a different story.

Life-cycle patterns

The descriptive cross-tabulation of risky-asset participation by age group and macro-region (Table 6, visualised in Figure 3) already gives away the substantial regional heterogeneity in life-cycle patterns. In Western and Nordic Europe, participation is hump-shaped, rising from 21.3% among households under 35 to 25.8% at ages 35 to 54 and 25.4% at ages 55 to 64, before declining to 21.7% among households aged 65 and above. In Southern Europe, the pattern is slightly increasing in age up to the 55-64 bracket: participation rises from 10.3% under 35 to 13.7% at ages 35-54 and 16.5% at ages 55-64, then drops slightly to 13.6% among the eldest. In Eastern Europe, participation is essentially flat across the life cycle at approximately seven to eight percent in all four age brackets, with a marginal decline among the oldest group. These three profiles are visually distinct and motivate the formal test of life-cycle heterogeneity below.

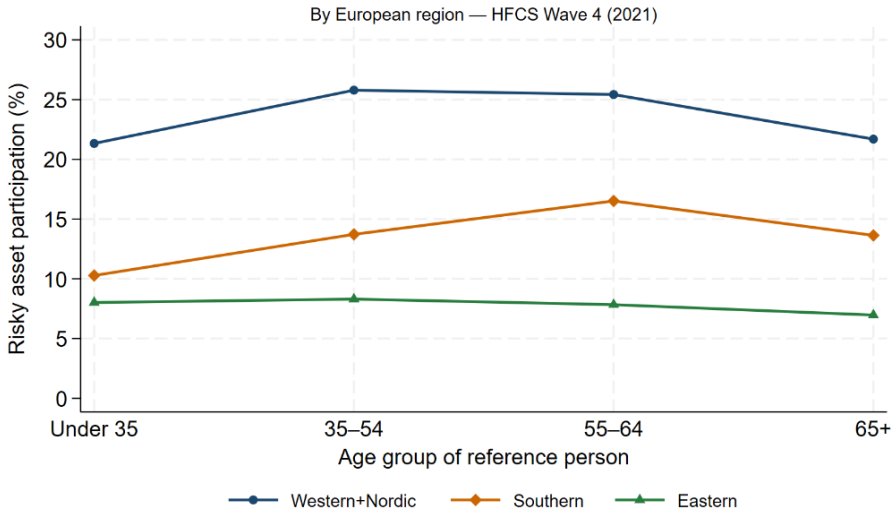

Figure 3. Life-Cycle Profile by Region

Note. Author's calculations based on wave 4

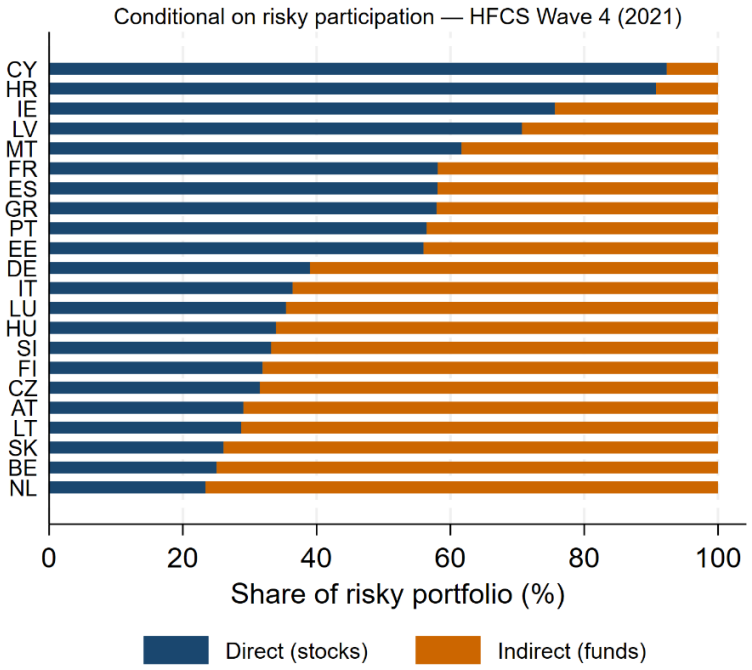

Figure 4. Direct vs Indirect Risky Asset Holdings by Country

Note. Author's calculations based on wave 4

4.2. Portfolio allocation across European countries and determinants of cross-country heterogeneity

RQ1: How does the allocation of financial wealth across major asset classes vary across European countries, and what household characteristics are associated with risky asset participation within and across these countries?

Table 1 reports the estimates of the three specifications introduced in section 3.3.2. The progression from Model 1a to Model 1c builds up the primary specification step by step and clarifies the role of country-level heterogeneity in the participation decision.

Model 1a estimates a pooled probit without country dummies, uncovering that household-level gradients in column 1 are economically meaningful. Log income, the inverse-hyperbolic-sine of net wealth, and tertiary education are all strongly positive, while the female indicator and household size enter with significantly negative coefficients. Two limitations, however, restrict the usefulness of these estimates. The probit coefficients are not directly interpretable as percentage-point effects, and the specification ignores the substantial cross-country differences in baseline participation documented above.

Model 1b addresses this second limitation by incorporating country dummies, with Austria as reference. The country dummies are jointly highly significant ($F(19)=149.8$, $p<0.001$), confirming that a substantial share of cross-country variation in participation stems from country-specific factors rather than from household characteristics. Once national effects are accounted for, several household coefficients change in magnitude. Most notably, the negative coefficient on retired employment status (-0.266 in Model 1b versus -0.203 in Model 1a) becomes more pronounced, suggesting that its unconditional gradient was partially absorbed by national differences.

Model 1c replaces the probit link with the linear probability model, retaining country dummies, serving as the primary specification throughout the following sub-chapters. The coefficients can be read directly as percentage-point differences in the probability of participation, which simplifies interpretation and ensures consistency with the comparable specifications used for RQ2 and RQ3. All three models are estimated on the sample of 72,515 households with strictly positive financial assets, representing a weighted population of approximately 152 million households across the 22 countries.

The household-level gradient in Model 1c is dominated by three variables: income, net wealth, and education. A doubling of household gross income (≙ log(income) increases

by 0.69) is associated with a 5.0 percentage point increase in the probability of participation ($\beta = 0.072$, $SE = 0.005$). The identical increase in net wealth is associated with a significant but smaller increase of 1.0 p.p. ($\beta = 0.010$, $SE = 0.001$). The largest household-level determinant for risky asset participation is education. Relative to households whose reference person has lower-secondary education, households with upper-secondary education are about 3.9 percentage points more likely to participate, and households with tertiary education are about 15.9 percentage points more likely to do so, both findings being strongly significant with $p < 0.01$. Female reference persons exhibit a lower participation probability of 3.4 percentage points, while homeowners are 5.4 percentage points more likely to hold risky assets, possibly reflecting their higher overall wealth and broader financial experience. Household size enters negatively, with each additional household member associated with a 2.2 percentage-point lower participation probability. Country dummies, employment status indicators, and household size are included in all specifications but, for readability, only the substantively relevant coefficients are shown in Table 1.

Two patterns deserve particular attention. First, although age groups are included as control variables in all three specifications, the four age-group dummies are individually small and statistically indistinguishable from zero in Model 1c, and a joint Wald test cannot reject the hypothesis that they are jointly zero ($F(3) = 0.37$, $p = 0.77$). Conditional on income, wealth, education, and country, age has no independent effect on participation in the linear probability model. The probit specification of Model 1b delivers small negative age effects of three to four percentage points, all statistically significant, but the discrepancy with the LPM is modest in magnitude and reflects the well-known sensitivity of probit coefficients to the location of the index function. The substantive implication is that age operates predominantly through correlated socio-economic variables (income, wealth, and education) rather than as an independent driver of participation. Heterogeneity in life-cycle patterns therefore needs to be sought in interactions with regional context, which is the focus of section 4.4.

Second, robustness check R4 reveals that the wealth gradient itself depends on education (Wald test on the education-by-wealth interaction: $F(2) = 34.47$, $p < 0.001$). The marginal effect of an additional unit of the inverse-hyperbolic-sine of net wealth on participation is approximately 0.003 for households with low education, 0.010 for households with medium-education, and 0.015 for households with tertiary education. The same increment of wealth therefore translates into participation about five times more

strongly for tertiary-educated households than for households with at most lower-secondary education. The interpretation of the interaction is supported by the small correlation between education and wealth in the analytical sample. The Pearson correlation between the three-level education indicator and the inverse hyperbolic sine of net wealth is 0.147, and the corresponding Spearman rank correlation is 0.268. These values indicate that education and net wealth capture distinct dimensions of household resources rather than redundant constructs. This complementarity between human and financial capital is consistent with the financial-literacy mechanism advanced by Jappelli and Padula (2015) and with the Calvet, Campbell, and Sodini (2007) finding that investment frictions are concentrated among the less educated.

Two diagnostics on the linearity assumption are reported. First, fitted values from Model 1c fall outside the unit interval for approximately 21.1% of observations (averaged across implicates: 11.1% below zero, 10.0% above one). This share is non-trivial and reflects a known property of the LPM in settings where the dependent variable's mean is far from 0.5; with weighted average participation at 19%, many households lie in regions where the linear approximation extrapolates beyond the bounded interval. Second, the parallel probit specification (Model 1b) is reported as a methodological cross-check: the sign, statistical significance, and approximate magnitude of all key coefficients on income, wealth, and education are substantively identical across the two estimators, indicating that the conclusions do not depend on the linearity assumption.

Taking together, these results confirm three well-established patterns in the household finance literature. First, wealth, income, and education are in their order the most influential drivers of risky asset participation, with education being the key gradient. The estimated gap of approximately 16 p.p. between tertiary- and lower-educated households is substantially larger than the income gradient over a comparable range, underscoring that education operates beyond income, potentially through the financial-literacy channel emphasised by Jappelli and Padula (2015).

Second, robustness check R4 provides insights into additionally nuances of the education gradient. The interaction coefficient between tertiary education and *ihs_netwealth* is 0.011 (*SE 0.002, $p < 0.01$*), while the baseline effect of *ihs_netwealth* for households with low education is only 0.003. Essentially, this means that the effect of additional wealth on participation is almost five times larger for tertiary-educated households than in the group of low-educated households. Once more this is in line with Jappelli and Padula's (2015)

framework, explaining how education reduces the effective cost of entering risky asset markets and thereby amplifies the translation of available wealth into actual participation.

Third, country dummies absorb a sizeable share of the cross-country variation in participation, confirming the central finding of Christelis et al. (2013) that the economic environment matters beyond household characteristics. At the same time, the individual-level coefficients remain essentially unchanged once country dummies are introduced, indicating that the intra-country gradients of participation follow comparable trends across the 22 European countries. While levels of participation differ substantially across countries, the household characteristics driving participation appear to operate through a shared logic. Subject to country-specific institutional constraints, this implies that policy interventions targeting household-level channels, such as financial education or tax incentives, can be expected to behave similarly across national contexts. This last point is of practical importance as it suggests that policy lessons travel reasonably well between European countries, allowing governments to learn from both successful and unproductive interventions in their peer economies, rather than having to assume that each national setting requires a fundamentally different approach.

4.3. Channel choice between direct equity and investment funds across European countries

RQ2: Conditional on holding risky financial assets, how do households allocate their risky portfolios between directly held shares and investment funds, and how does this allocation vary across countries and household characteristics?

Having documented who participates in risky assets, the analysis now turns to how participating households allocate their risky portfolios between directly held shares and investment funds. As discussed in chapter 2 (sections 2.4 and 2.5), the two channels imply substantially different cost structures, information requirements, and degrees of diversification.

Table 2 reports the estimates of the two specifications introduced in section 3.3.3. As in the participation analysis, the progression from Model 2a to Model 2b clarifies which conclusions follow from the household-level controls alone and which require the regional interaction terms. Model 2a is the base specification, an ordinary least squares regression of the direct share on the standard household controls and country dummies. It is estimated on

the 17,955 risky-asset-holding households in the regression sub-sample, using the same multiple-imputation and bootstrap variance framework as the participation models.

Two patterns dominate the results of the analysis about channel choice. First, in contrast to the unconditional Panel B of Table 5, age has a substantial and statistically significant effect on the direct share once income, wealth, education, and country are controlled for. Households whose reference person is aged 55 to 64 hold approximately 10 percentage points less of their risky portfolio in directly held shares than households under 35 ($\beta = -0.100,\ SE = 0.036$), and households aged 65 and above hold approximately 10.9 percentage points less ($\beta = -0.109,\ SE = 0.046$). Both effects are significant at the five percent level. The discrepancy with Panel B reflects a textbook suppression effect. Younger households tend to hold less wealth and lower incomes, which in turn correlate with higher fund holdings. Controlling for these covariates reveals a clear age gradient toward fund-based intermediation among older households.

Second, none of the other household covariates explains a substantial share of the variation in the direct share. Income, education, and employment status all have small and mostly insignificant coefficients, and average R^2 in Model 2a is only 0.076. Country dummies, by contrast, absorb a large share of the explained variation. A second specification, Model 2b (Table 2, column 2) replaces the country dummies with three macro region dummies and adds the age-by-region interaction. The average R^2 falls to 0.015, an order of magnitude lower than Model 2a. The drop in R^2 shows that most of the country-level variation in channel choice exists between countries within the same region, not between the regions themselves. The channel-choice decision is therefore driven by national structural features that cannot be captured by the macro region split. The Wald test on the six age-by-region interaction terms cannot reject the null of no interaction ($F(6) = 0{,}83,\ p = 0.55$), indicating that the age gradient in channel choice is uniform across European regions. This null result complements the participation findings of 4.4, where the same interaction is highly significant for the participation outcome but not for the channel-choice outcome. The substantive implication is that the life-cycle shift from direct stock holdings toward intermediated vehicles is a general feature of household portfolio behaviour in Europe, not a region-specific phenomenon.

Two observations merit attention regarding channel choice in Europe. Primary, the cross-country heterogeneity in the direct-versus-indirect channel choice is immense and is not explained by household demographics, income, wealth, or education. The contrast

between Cyprus (92.4% direct) and the Netherlands (23.4% direct) cannot be reconciled with the household-level controls available in the HFCS, and the regression framework attributes the remaining variation to country dummies. The further drop in R^2 when country dummies are replaced by three regional dummies confirms that the relevant variation operates at the country rather than the regional level. This is consistent with the literature on the role of financial-system architecture (Guiso et al., 2003; Bekhtiar et al., 2019), which emphasises the depth of the mutual-fund industry, the availability of occupational pension schemes, and the cost structure of brokerage services as primary determinants of channel choice. Supply-side factors, such as range of product availability and barriers to investing into foreign markets, are absorbed within country dummies rather than measured directly. This is a recognised limitation.

Second, the results identify a robust life-cycle pattern at the household level. Households aged 55 and above hold approximately ten percentage points less of their risky portfolio directly, conditional on income, wealth, and education, and this pattern is consistent across European regions. The shift toward intermediated investment vehicles in later life is consistent with both a behavioural delegation channel documented by Brown et al. (2023) and a portfolio-rebalancing interpretation in which older households exchange direct stock exposure for the diversification provided by mutual funds as their investment horizon shortens.

4.4. Life-cycle patterns across European regions

RQ3: How do risky asset participation rates and risky portfolio compositions differ across age groups, and are these life-cycle patterns consistent across European regions?

Table 3 reports three specifications that test the central hypothesis of RQ3, namely whether life-cycle patterns differ across European regions. Model 3b (Table 3, column 1) is the linear probability participation model with the age-by-region interaction and serves as the primary specification of this section. Model 3a is the probit version of the same specification, reported as a methodological cross-check on the linearity assumption. Model 3c (Table 3, column 2) applies the same interaction structure to the direct-share outcome of section 4.3, allowing a direct comparison of regional heterogeneity at the extensive margin of participation versus the intensive margin of channel choice.

Model 3b confirms that the age slope of risky-asset participation differs significantly across the three European macro-regions. The Wald test on the six age-by-region interaction coefficients yields $F(6) = 5.87,\ p < 0.001$, even after controlling for household income, wealth, education, and country. The probit version (Model 3a) confirms this conclusion ($F(6) = 4.50,\ p < 0.001$), establishing that the regional heterogeneity is not a product of the linearity assumption.

The pattern of the interaction coefficients in Model 3b is consistent with the descriptive evidence. Relative to the Western-and-Nordic baseline, households aged 65 and above in Southern Europe are 8.6 p.p. more likely to participate than the regional age slope alone would predict ($\beta=0.086,\ p<0.01$), and households aged 65 and above in Eastern Europe are 5.9 percentage points more likely to participate ($\beta=0.059,\ p<0.01$). The 55-64 interaction with Southern Europe yields a coefficient of 6.9 percentage points ($\beta=0.069,\ p<0.05$), while the corresponding interaction with Eastern Europe is small and insignificant ($\beta=-0.003$). These positive interactions imply that the negative age slope observed in Western and Nordic Europe flattens out, and in Southern Europe partially reverses, in the older brackets. Age-related declines in participation are therefore a Western and Nordic phenomenon rather than a universal European one.

Model 3c applies the same age-by-region interaction structure to the direct-share outcome, mirroring the channel-choice analysis of section 4.3. In contrast to the participation result, the Wald test on the six interaction coefficients is not significant ($F(6) = 0.83,\ p = 0.55$), indicating that the regional heterogeneity is concentrated at the extensive margin of participation rather than at the intensive margin of channel choice. The shift from direct to indirect risky-asset holding among older households documented in section 4.3 occurs uniformly across all three macro-regions.

The interpretation of the results center on one finding. Life-cycle patterns in risky-asset participation are not consistent across European regions. The hump-shaped age profile predicted by standard portfolio theory and documented for Western and Nordic households (Ameriks & Zeldes, 2004; Cocco et al., 2005) is absent in Southern Europe, where participation rises monotonically into the 55-64 bracket, and absent in Eastern Europe, where participation is essentially flat. The size of the regional differences (interaction coefficients of 6 to 9 percentage points among older households) is large relative to the regional main effects and statistically robust. Another interesting finding is that life-cycle patterns in the channel-choice decision are consistent across regions. As reported in 4.3, the

shift toward intermediated investment vehicles among households aged 55 and above is observed uniformly in all three macro-regions, suggesting that regional institutional differences shape who participates more strongly than they shape how participants allocate their risky portfolios.

Two interpretations of the participation finding deserve note. The first is a cohort interpretation, which could be described by households currently aged 65 and above in Southern and Eastern Europe have lived through markedly different macroeconomic histories from their Western and Nordic counterparts. Older Eastern European households experienced the post-1990 economic transition during their working lives, while older Southern European households have lived through the legacy of the sovereign debt crisis of the early 2010s. Both experiences plausibly shape risk attitudes in ways that the available controls cannot capture. The second is a structural interpretation about the institutional environments in which older Southern and Eastern European households participate may have evolved differently from those of younger households in the same regions. Later-life participation reflecting accumulated wealth in homeownership as well as small businesses being released into financial asset markets. Disentangling these two interpretations would require longitudinal data and explicit demand- versus supply-side control, neither of which is available in the HFCS as identified in section 3.5. The cross-sectional evidence alone is consistent with both readings and is reported descriptively.

The Pearson correlation between log households income and the age of the reference person vary significantly across the three macro-regions. It is almost zero in Western and Nordic Europe ($r = -0.003$) and in Southern Europe ($r = -0.032$), but distinctly negative in Eastern Europe ($r = -0.355$). These results reflect a negative gradient in Eastern Europe of lower post-employment incomes of older households relative to younger working-age households. In Western and Nordic Europe, the near-zero income-age profile in participation is not driven by life-cycle income variation, further supporting the structural and behavioral interpretation discussed above.

5. Conclusion

This dissertation explores cross-country heterogeneity in a sample of 72,515 households' financial portfolio composition across 22 European countries, using harmonized microdata from the fourth wave of the Household Finance and Consumption Survey. Examined households held strictly positive financial assets and were analyzed on patterns of risky asset participation and channel-choice, representing approximately 152 million households.

Research question one documents how the allocation of financial wealth across general asset classes varies across countries and identifies correlated household characteristics. Research question two examines participating households' distribution between directly held shares and investment funds. Research question 3 tests whether life-cycle patterns in participation and channel choice are consistent across European regions.

The empirical analysis produced three core findings. Insight number one is that household-level participation in risky assets is mainly associated with income, net wealth, and education. With education being the clearest indicator, tertiary-educated households are roughly 16 percentage points more likely to participate in risky financial assets than the lowest education group, in this case lower secondary. This observation was made holding income and wealth constant and the wealth gradient is around five times stronger for tertiary-educated households than for low-educated ones.

When researching the split of risky portfolios between direct equity holdings and indirect investment funds, the analysis documents dramatic variation across countries, ranging from 23% direct holdings in Netherlands to 92% in Cyprus. This variation is largely absorbed by country dummies rather than by household characteristics, indicating that channel choice is shaped primarily by structural features of national financial systems.

The final observation is about life-cycle patterns in participation not being consistent across European regions. The hump-shaped age profile found for Western and Nordic Europe is absent in Southern Europe, where participation rises monotonically into the 55-64 age bracket, while Eastern European participation stays mostly flat across all age groups. The corresponding shift in channel choice, by contrast, is uniform across regions.

Contribution to the Literature

In addressing the three gaps identified in Section 2.7, this dissertation makes three corresponding contributions. Above all, this dissertation provides an updated baseline across 22 European countries and the most recent HFCS wave. Additionally, this fourth wave was

conducted in 2021, a pandemic-era defined by an unusual macroeconomic backdrop due to high savings rates and exceptional fiscal and monetary support.

Besides, this dissertation synthesizes three fields that have typically been studied separately, namely cross-country participation, direct and indirect channel choice, and life-cycle heterogeneity. The joint analysis reveals both consistencies and inconsistencies revealing that national context, households' characteristics, and demographic components differ in degree of impact, depending on the observed dimension. This discovery is of importance when both nationally and internationally designing and adapting policies.

Lastly, the analysis confirms and quantifies the complementarity between human capital and financial capital. The interaction between education and net wealth, examined in robustness check R4, shows that the equivalent increases in wealth translate into participation about five times more strongly for certain compared education levels. This result gave direct empirical support for the financial-literacy mechanism in a comparative European setting.

Policy Implications

The findings can be of high value for policy-relevant stakeholder groups for various reasons which will be outlined. For institutions concerned with retirement security and the goals of the Capital Markets Union, the results suggest that interventions should be targeted at financial education, with mechanisms expected to operate broadly similarly across European countries. While levels of participation significantly differ across countries, the household determinants associated with participation, namely income, wealth, and especially education, follow a shared logic. Policy lessons on developing these characteristics should therefore travel reasonably well between European peer economies, allowing governments to learn from each other's interventions rather than designing each national programme from scratch.

This topic was refined by additionally learnings about the relationship between education and wealth. Wealth alone does not convert into risky asset participation when households lack the human capital to act on it. Initiatives aimed at broadening participation are likely to be most effective when financial literacy incentives are paired with measures to reduce the cost of converting savings into risky assets. This could come in the form of default options in occupational pension schemes, intermediated investment products, or encouraging competition among fintechs.

Financial service providers and intermediaries could benefit from identification of country variation in channel choice and the role of supply-side factors in shaping households'

behavior. Continued development of intermediated investment vehicles in markets currently dominated by direct holdings could lower entry barriers and improve diversification for participating households.

Limitations

The HFCS dataset is a single cross-section, therefore age-related variation cannot be cleanly decomposed into life cycle and cohort effects, nor can the regional life-cycle differences documented in section 4.4 be attributed to specific demand- or supply-side mechanisms. Household characteristics are usually measured at the reference person, masking intra household differences in age and education. To gain statistically significant results, households without banking relations or any financial assets (including deposits, securities, and other financial instruments) were excluded, representing a meaningful but separate group. With the number of per-country registered observations varying strongly and limited risky asset participation, specific country-level sample sizes in some methods fell below 100. Finally, the HFCS does not include a harmonised financial-literacy module in Wave 4, so education serves as a proxy for the literacy channel rather than as a direct measure.

Future Research

While several extendable research fields have been mentioned, the most valuable areas are briefly outlined here. Soon the fifth wave of the HFCS will be made available to the researchers, allowing insights into patterns on data as new as 2024, answering question related to investment behaviour shaped by the pandemic-era or shifts to normalization. A two-wave comparison limited to Waves 4 and 5 (restricted to countries with no methodical questionnaire change) would partially mitigate the methodological-discontinuity concern that affects pooling across the full HFCS history.

Another valuable extension could be to gather time series data, enabling the disentangling of life-cycle effects from cohort effects. Pursuable options could be to link HFCS-based descriptive evidence to longitudinal sources such as the Survey of Health, Ageing and Retirement in Europe (SHARE) or constructing synthetic cohorts across HFCS waves, with appropriate caveats.

Furthermore, behavioural and cognitive determinants such as financial literacy, trust in financial institutions, and risk preferences could not be examined directly with the harmonised HFCS variables. Future research that combines HFCS data with country-level literacy measures or that leverages national surveys including these modules would refine the financial-literacy interpretation advanced in this dissertation.

Finally, the rapid expansion of low-cost intermediated investment products, particularly exchange-traded funds and digital-broker platforms, may be reshaping the direct-versus-indirect channel choice in ways that the 2021 wave could not yet capture. Future HFCS waves should assess whether the democratisation of capital markets translates into measurable shifts in household portfolio composition, what role national reforms play and if there are generational differences in the use of these new options.

Tables

Table 1

Determinants of Risky Asset Participation

	(1)	(2)	(3)
Predictor	*Probit (no dummy)*	*Probit (Country Dummy)*	*LPM (Country Dummy)*
Age 35-54	−0.135***	−0.116**	−0.011
	(0.045)	(0.047)	(0.011)
Age 55-64	−0.176***	−0.143***	−0.011
	(0.051)	(0.053)	(0.012)
Age 65 and above	−0.178**	−0.165**	−0.012
	(0.077)	(0.079)	(0.018)
Log household income	0.459***	0.371***	0.072***
	(0.029)	(0.032)	(0.005)
Net wealth (asinh)	0.070***	0.072***	0.010***
	(0.010)	(0.010)	(0.001)
Medium education	0.283***	0.250***	0.039***
	(0.042)	(0.045)	(0.008)
Tertiary education	0.628***	0.623***	0.159***
	(0.043)	(0.045)	(0.009)
Female reference person	−0.120***	−0.127***	−0.034***
	(0.032)	(0.032)	(0.007)
Household size	−0.147***	−0.115***	−0.022***
	(0.013)	(0.014)	(0.003)
Homeowner	0.111***	0.181***	0.054***
	(0.041)	(0.043)	(0.008)
Observations	72,515	72,515	72,515

Note. Dependent variable is *risky_part*, a binary indicator equal to 1 if the household holds either listed shares or mutual fund shares. Coefficients in columns (1) and (2) are probit estimates; column (3) reports linear probability model estimates and is the primary specification. Standard errors in parentheses, obtained from 1,000 bootstrap replicate weights and pooled across five imputation implicates using Rubin's (1987) rules. Country dummies in columns (2) and (3) include 21 country dummies with Austria as reference; $F(19) = 149.8$, $p < .001$. Joint Wald test on age groups in column (3): $F(3) = 0.37$, $p = .77$.

*$p < .10$. **$p < .05$. ***$p < .01$.

Table 2

Direct vs. Indirect Allocation of the Risky Portfolio

	(1)	(2)
Predictor	*OLS (Country Dummy)*	*OLS (Age × Region)*
Age 35-54	−0.046	−0.046
	(0.036)	(0.040)
Age 55-64	−0.100***	−0.100**
	(0.036)	(0.040)
Age 65 and above	−0.109**	−0.127**
	(0.046)	(0.051)
Log household income	0.025	0.006
	(0.017)	(0.015)
Net wealth (asinh)	0.005*	0.006**
	(0.003)	(0.003)
Medium education	0.003	−0.010
	(0.031)	(0.033)
Tertiary education	0.003	0.005
	(0.030)	(0.032)
Female reference person	−0.043**	−0.042**
	(0.020)	(0.020)
Household size	−0.013	−0.002
	(0.009)	(0.009)
Homeowner	0.009	0.013
	(0.026)	(0.025)
Observations	17,955	17,955

Note. Dependent variable is *direct_share*, defined as listed shares divided by the sum of listed shares and mutual fund shares. Sample restricted to households with positive risky asset holdings. Standard errors in parentheses, obtained from 1,000 bootstrap replicate weights and pooled across five imputation implicates. Both columns include country dummies (not shown). Joint Wald test on the six age × region interaction terms in column (2): $F(6) = 0.83$, $p = .55$.

*$p < .10$. **$p < .05$. ***$p < .01$.

Table 3

Life-Cycle Patterns Across European Regions

	(1)	(2)
Predictor	*LPM: Participation*	*OLS: Direct share*
Age 35-54	−0.014	−0.046
	(0.014)	(0.040)
Age 55-64	−0.029*	−0.100**
	(0.017)	(0.040)
Age 65 and above	−0.035*	−0.127**
	(0.021)	(0.051)
Southern Europe	−0.110***	0.017
	(0.023)	(0.084)
Eastern Europe	−0.114***	−0.091
	(0.015)	(0.060)
Age 55-64 × Southern	0.069**	0.033
	(0.027)	(0.091)
Age 65+ × Southern	0.086***	0.032
	(0.025)	(0.090)
Age 65+ × Eastern	0.059***	0.059
	(0.018)	(0.068)
Log household income	0.081***	0.006
	(0.005)	(0.015)
Net wealth (asinh)	0.009***	0.006**
	(0.001)	(0.003)
Tertiary education	0.162***	0.005
	(0.009)	(0.032)
Female reference person	−0.034***	−0.042**
	(0.007)	(0.020)
Homeowner	0.052***	0.013
	(0.008)	(0.025)
Observations	72,515	17,955

Note. Column (1): dependent variable is *risky_part* (binary participation indicator); column (2): dependent variable is *direct_share*. Region categories: Western and Nordic Europe (reference), Southern Europe, Eastern Europe. Standard errors in parentheses, pooled across five imputation implicates. Country dummies included in both columns (not shown). Joint Wald test on the six age

× region interactions: column (1) *F(6) = 5.87, p < .001; column (2) F(6) = 0.83, p = .55.* Selected interaction coefficients are shown; full results available on request. * $p < .10$. ** $p < .05$. *** $p < .01$.

Table 4

Robustness Checks for the Linear Probability Participation Model

	(1)	(2)	(3)	(4)
Predictor	*R1: Age polynomial*	*R2: Broad risky*	*R3: IHS income*	*R4: Edu × Wealth*
Log household income	0.073***	0.078***	—	0.071***
	(0.005)	(0.005)		(0.005)
Household income (asinh)	—	—	0.034***	—
			(0.003)	
Net wealth (asinh)	0.010***	0.010***	0.011***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Medium education	0.038***	0.044***	0.047***	−0.032***
	(0.008)	(0.008)	(0.008)	(0.010)
Tertiary education	0.158***	0.164***	0.179***	0.032*
	(0.009)	(0.009)	(0.008)	(0.018)
Wealth × Medium edu	—	—	—	0.007***
				(0.001)
Wealth × Tertiary edu	—	—	—	0.011***
				(0.002)
Female reference person	−0.034***	−0.032***	−0.040***	−0.034***
	(0.007)	(0.008)	(0.007)	(0.007)
Household size	−0.023***	−0.024***	−0.014***	−0.022***
	(0.003)	(0.003)	(0.003)	(0.003)
Homeowner	0.054***	0.060***	0.060***	0.056***
	(0.008)	(0.008)	(0.008)	(0.008)
Observations	72,515	72,515	72,910	72,515

Note. All four columns are linear probability models with country dummies (not shown). R1 replaces the four age-group indicators with a continuous age polynomial (age and age-squared). R2 broadens the risky asset definition to include bonds. R3 substitutes the inverse hyperbolic sine of income for log income, increasing the sample by approximately 400 households. R4 adds the interaction between education and the inverse hyperbolic sine of net wealth; joint Wald test on the interaction terms*: F(2) = 34.5, p < .001*. Standard errors in parentheses.

** p < .10. ** p < .05. *** p < .01.*

Table 5

Direct vs. Indirect Share of the Risky Portfolio (Weighted %)

Panel A: By Country

Country	*Direct share (%)*	*Indirect share (%)*	*n*
CY †	92.4	7.6	104
HR †	90.7	9.3	56
IE	75.7	24.3	1,169
LV †	70.7	29.3	48
MT	61.7	38.3	137
FR	58.1	41.9	3,172
ES	58.1	41.9	1,974
GR †	57.9	42.1	60
PT	56.4	43.6	713
EE	56.0	44.0	343
DE	39.0	61.0	1,664
IT	36.4	63.6	1,149
LU	35.4	64.6	803
HU	33.9	66.1	389
SI	33.2	66.8	232
FI	31.9	68.1	5,258
CZ	31.5	68.5	308
AT	29.1	70.9	339
LT †	28.7	71.3	74
SK	26.1	73.9	109
BE	25.0	75.0	665
NL	23.4	76.6	534
EU22 Average	*42.7*	*57.3*	*19,300*

Panel B: By Age Group

Age group	*Direct share (%)*	*Indirect share (%)*	*n*
Under 35	43.3	56.7	1,496
35-54	43.1	56.9	6,170
55-64	39.4	60.6	4,318
65 and above	43.6	56.4	6,010

Note. Panel A reports the country-level mean direct share (listed shares as a percentage of total risky assets) for households holding at least one risky financial asset. Panel B reports the unconditional

age-group mean of the direct share for the pooled European sample. Cell entries are weighted percentages using HW0010 and pool five imputation implicates. n indicates the unweighted number of risky-asset-holding households used to estimate the country or age-group cell. Country-level estimates with n < 100 (CY, HR, GR, LV, LT) are based on small national sub-samples and should be interpreted with caution and are highlighted with †. Countries are sorted in Panel A by direct share in descending order.

Table 6

Risky Asset Participation Rate by Age Group and European Region

	Western and Nordic	Southern	Eastern	All regions
Age group	*%*	*%*	*%*	*%*
Under 35	21.3	10.3	8.0	18.4
35-54	25.8	13.7	8.3	19.8
55-64	25.4	16.5	7.8	20.7
65 and above	21.7	13.6	7.0	17.6

Note. Cell entries are weighted percentages of households holding any risky financial asset (listed shares or mutual fund shares) within the indicated age group and region. Estimates use the household sampling weight and pool five imputation implicates using Rubin's (1987) rules. Sample restricted to households with positive financial assets (n = 79,742). Western and Nordic Europe comprises AT, BE, DE, FI, FR, IE, LU, NL; Southern Europe comprises CY, ES, GR, IT, MT, PT; Eastern Europe comprises CZ, EE, HR, HU, LT, LV, SI, SK. All regions are weighted aggregates.

Appendix A - Multiple Imputation and Bootstrap Variance Estimation

Multiple Imputation and Rubin's Combination Rules

For every missing value of a wealth, consumption, or income variable, the HFCS provides M=5 imputed replacement values, generated through a chained-equation model with country-specific specifications. The five resulting datasets, called implicates, together represent the distribution of plausible values for the missing observations.

Inference based on multiple imputation follows Rubin's (1987) combination rules. Let $\hat{\theta}_m$ denote the point estimate of a parameter of interest (for example, a regression coefficient) computed on implicate m, and let U_m denote the corresponding within-imputation variance that is, the variance of $\hat{\theta}_m$ that would be obtained if implicate m were the only available dataset. The pooled point estimate is the simple average across implicates,

$$\bar{\theta} = \frac{1}{M} \sum_{m=1}^{M} \hat{\theta}_m, \tag{A.1}$$

and the total variance (T) combines two components,

$$T = \bar{U} + \left(1 + \frac{1}{M}\right) B, \tag{A.2}$$

where

$$\bar{U} = \frac{1}{M} \sum_{m=1}^{M} U_m \tag{A.3}$$

is the average of the within-imputation variances, and

$$B = \frac{1}{M-1} \sum_{m=1}^{M} (\hat{\theta}_m - \bar{\theta})^2 \tag{A.4}$$

is the between-imputation variance, capturing the dispersion of point estimates across implicates. The factor

$$(1 + 1/M) \tag{A.5}$$

is a finite-sample correction reflecting the limited number of imputations. With M = 5 implicates, this correction equals 1.2.

The within-imputation variance $\bar{U}$ would dominate the total variance if the imputed values agreed closely; the between-imputation variance B would dominate if they disagreed substantially. Reporting only $\bar{U}$ (as would happen if a single implicate were analysed) understates the true uncertainty of the estimate.

Rescaling Bootstrap for the Within-Imputation Variance

The HFCS sampling design includes stratification, multi-stage clustering, and deliberate oversampling of wealthy households. Standard variance formulas valid under simple random sampling do not apply. The HFCS therefore provides 1,000 bootstrap replicate weights generated through the rescaling bootstrap of Rao and Wu (1988), which preserves the complex sampling design.

The within-imputation variance for implicate $m$is computed from the replicate weights as

$$U_m = \frac{1}{R}\sum_{r=1}^{R}\left(\hat{\theta}_m^{(r)} - \hat{\theta}_m\right)^2, \tag{A.6}$$

where $\hat{\theta}_m^{(r)}$is the point estimate computed on implicate $m$using the r-th set of replicate weights, and $R = 1{,}000$is the number of bootstrap replicates. This variance accounts for stratification, clustering, and oversampling without requiring an explicit specification of the sampling design.

Implementation in Stata

The framework is implemented through Stata's mi estimate: svy: prefix combination. The survey design is declared with mi svyset [pw=hw0010], bsrweight(wr0001-wr1000) vce(bootstrap), which informs Stata of the household sampling weight HW0010 and the 1,000 bootstrap replicate weights stored in the W-file. Each regression is then estimated with mi estimate: svy: <model>, which applies the bootstrap to compute $U_m$on each implicate and pools the results via Rubin's combination rules.

Two implementation choices deserve note. First, subpopulation analysis, such as the restriction of RQ2 to risky-asset holders, is handled through the subpop() option rather than if conditions. The reason is that if removes primary sampling units that should still contribute to variance estimation, whereas subpop() retains them. Second, joint significance tests on multiple coefficients (such as the Wald F-tests on country dummies or age-by-region interactions in Chapter 4) are computed via mi test, which applies the Rubin combination rules to multivariate hypotheses.

The combined two-stage procedure produces standard errors that reflect both imputation and sampling uncertainty. All confidence intervals and p-values reported in Chapter 4 are derived from the total variance T.

References

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, *80*(1), 123-129. https://doi.org/10.1016/S0165-1765(03)00032-6

Ameriks, J., & Zeldes, S. P. (2004). *How do household portfolio shares vary with age?* (Working paper). Columbia Business School. https://business.columbia.edu/sites/default/files-efs/imce-uploads/szeldes/pdfs/Ameriks_Zeldes_age_Sept_2004d.pdf

Ampudia, M., & Ehrmann, M. (2017). Macroeconomic experiences and risk taking of euro area households. *European Economic Review*, *91*, 146-156. https://doi.org/10.1016/j.euroecorev.2016.09.012

Arrondel, L., Bartiloro, L., Fessler, P., Lindner, P., Mathä, T. Y., Rampazzi, C., Savignac, F., Schmidt, T., Schürz, M., & Vermeulen, P. (2014). *How do households allocate their assets? Stylised facts from the Eurosystem Household Finance and Consumption Survey* (ECB Working Paper No. 1722). European Central Bank. https://doi.org/10.2139/ssrn.2490969

Balogh, E., Kékesi, Z., & Sisak, B. (2019). Analysis of households' investment decisions based on international data. *Financial and Economic Review*, *18*(1), 60-87. https://doi.org/10.33893/FER.18.1.6087

Bekhtiar, K., Fessler, P., & Lindner, P. (2019). *Risky assets in Europe and the US: Risk vulnerability, risk aversion and economic environment* (ECB Working Paper No. 2270). European Central Bank. https://doi.org/10.2866/018078

Bertaut, C. C. (1998). Stockholding behavior of U.S. households: Evidence from the 1983-1989 Survey of Consumer Finances. *The Review of Economics and Statistics*, *80*(2), 263-275. https://doi.org/10.1162/003465398557500

Bertaut, C. C., & Starr-McCluer, M. (2002). Household portfolios in the United States. In L. Guiso, M. Haliassos, & T. Jappelli (Eds.), *Household portfolios* (pp. 181–217). MIT Press.

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. https://doi.org/10.1016/0165-1889(92)90044-F

Brown, S., Kontonikas, A., Montagnoli, A., Pickard, H., & Taylor, K. (2023). *Household portfolios and financial literacy: The flight to delegation* (SSRN Working Paper No. 4604763). https://doi.org/10.2139/ssrn.4604763

Calvet, L. E., Campbell, J. Y., & Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, *115*(5), 707-747. https://www.nber.org/system/files/working_papers/w12030/w12030.pdf

Campbell, J. Y. (2006). Household finance. *The Journal of Finance*, *61*(4), 1553-1604. https://doi.org/10.1111/j.1540-6261.2006.00883.x

Christelis, D., Georgarakos, D., & Haliassos, M. (2013). Differences in portfolios across countries: Economic environment versus household characteristics. *The Review of Economics and Statistics*, *95*(1), 220-236. https://doi.org/10.1162/REST_a_00260

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. https://doi.org/10.1093/rfs/hhi017

Curcuru, S., Heaton, J., Lucas, D., & Moore, D. (2010). Heterogeneity and portfolio choice: Theory and evidence. In Y. Aït-Sahalia & L. P. Hansen (Eds.), *Handbook of financial econometrics: Tools and techniques* (pp. 337-382). Elsevier. http://home.uchicago.edu/~lhansen/handbook12.pdf

D'Addio, A. C., Roger, M., & Savignac, F. (2020). *Pensions and household savings: Cross-country heterogeneity in Europe* (ECB Working Paper No. 2372). European Central Bank. https://www.econstor.eu/handle/10419/228986

European Central Bank. (2016a). *The Household Finance and Consumption Survey: Methodological report for the second wave* (Statistics Paper Series No. 17). https://www.ecb.europa.eu/pub/pdf/other/ecbsp17.en.pdf

European Central Bank. (2016b). *The Household Finance and Consumption Survey: Methodological report for the second wave - Updated tables*. https://www.ecb.europa.eu/home/pdf/research/hfcn/HFCS-Methodologicalreport-wave2updatedtables.pdf

European Central Bank. (2020). *The Household Finance and Consumption Survey: Methodological report for the 2017 wave* (Statistics Paper Series No. 35). https://www.ecb.europa.eu/pub/pdf/scpsps/ecb.sps35~b9b07dc66d.en.pdf

European Central Bank. (2023). *The Household Finance and Consumption Survey: Methodological report for the 2021 wave* (Statistics Paper Series No. 45). https://www.ecb.europa.eu/pub/pdf/scpsps/ecb.sps45~4bbcfb02eb.en.pdf

Gollier, C. (2001). *The economics of risk and time*. MIT Press. https://doi.org/10.7551/mitpress/2622.001.0001

Guiso, L., Haliassos, M., & Jappelli, T. (Eds.). (2002). *Household portfolios*. MIT Press.

Guiso, L., Haliassos, M., & Jappelli, T. (2003). Household stockholding in Europe: Where do we stand and where do we go? *Economic Policy*, *18*(36), 123-170. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=374500

Haliassos, M., & Bertaut, C. C. (1995). Why do so few hold stocks? *The Economic Journal*, *105*(432), 1110-1129. https://doi.org/10.2307/2235407

Jappelli, T., & Padula, M. (2015). Investment in financial literacy, social security, and portfolio choice. *Journal of Pension Economics and Finance*, *14*(4), 369-411. https://doi.org/10.1017/S1474747214000377

Kaustia, M., Conlin, A., & Luotonen, N. (2023). What drives stock market participation? The role of institutional, traditional, and behavioral factors. *Journal of Banking & Finance*, *148*, Article 106743. https://doi.org/10.1016/j.jbankfin.2022.106743

Kochaniak, K., & Ulman, P. (2020). Risk-intolerant but risk-taking - Towards a better understanding of inconsistent survey responses of the euro area households. *Sustainability*, *12*(17), Article 6912. https://doi.org/10.3390/su12176912

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77-91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, *51*(3), 247-257. https://doi.org/10.2307/1926560

Paxson, C. H. (1990). Borrowing constraints and portfolio choice. *The Quarterly Journal of Economics*, *105*(2), 535-543. https://doi.org/10.2307/2937799

Rao, J. N. K., & Wu, C. F. J. (1988). Resampling inference with complex survey data. *Journal of the American Statistical Association*, *83*(401), 231-241. https://doi.org/10.2307/2288945

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons. https://doi.org/10.1002/9780470316696

Samuelson, P. A. (1969). Lifetime portfolio selection by dynamic stochastic programming. *The Review of Economics and Statistics*, *51*(3), 239-246. https://doi.org/10.2307/1926559

Vissing-Jørgensen, A. (2002). *Towards an explanation of household portfolio choice heterogeneity: Nonfinancial income and participation cost structures* (NBER Working Paper No. 8884). National Bureau of Economic Research. https://doi.org/10.3386/w8884

Zhan, J. C. (2015). Who holds risky assets and how much? *Empirica*, *42*(2), 323-370. https://doi.org/10.1007/s10663-015-9295-1

Annex A - Variable Definitions and HFCS Dataset Codes

This annex lists all HFCS variables used in the empirical analysis, together with their dataset codes and definitions. The variables are drawn from the HFCS fourth wave.

Table A.1

HFCS Variables Used in the Empirical Analysis

Variable name in text	HFCS code	File	Description
Identifiers			
Country identifier	SA0100	All	Two-letter ISO code identifying the participating country.
Household identifier	SA0010	All	Unique identifier linking observations across files.
Survey weight			
Household sampling weight	HW0010	H-file	Cross-sectional weight scaling the sample to the national household population.
Wealth and income aggregates			
Total financial assets	DA2100	D-file	Sum of all financial assets held by the household.
Listed shares	DA2105	D-file	Direct holdings of publicly traded shares.
Mutual fund shares	DA2102	D-file	Holdings of mutual fund shares.
Bonds	DA2103	D-file	Holdings of bonds and debt securities.
Total household gross income	DI2000	D-file	Aggregate household gross income before taxes.
Net wealth	DN3001	D-file	Total assets minus total liabilities.
Reference-person characteristics			
Reference-person age	DHAGEH1	H-file	Age in years, top-coded at the country level.
Reference-person education	DHEDUH1	H-file	Highest educational attainment, ISCED 2011.
Reference-person gender	PA0100	P-file	Coded one if female, zero otherwise.
Other household variables			
Household size	HB0010	H-file	Number of persons in the household.
Constructed variables			
risky_part	derived	—	Equals one if the household holds DA2105 or DA2102; zero otherwise.
direct_share	derived	—	DA2105 / (DA2105 + DA2102); defined for households with positive risky asset holdings.
insample_fin	derived	—	Equals one for households with strictly positive total financial assets (DA2100 > 0).
insample_risky	derived	—	Equals one for households in insample_fin with positive risky asset holdings.

Log of household income	derived	—	Natural logarithm of DI2000; defined for strictly positive income.
Inverse hyperbolic sine of net wealth	derived	—	arsinh(DN3001) = ln(DN3001 + sqrt(DN3001^2 + 1)); defined for negative values.
Age group	derived	—	Four categories: under 35, 35-54, 55-64, 65 and above.
Education level	derived	—	Three categories (ISCED 2011): low (0-2), medium (3-4), high (5-8).
Employment status	derived	—	Five categories: employee, self-employed, retired, unemployed, other or not in labour force.
Female indicator	derived	—	Equals one if the reference person is female (PA0100 = 2).
Homeowner indicator	derived	—	Equals one if the household owns its main residence outright or with a mortgage.
Region	derived	—	Three macro-regions: Western and Nordic, Southern, Eastern Europe (see Section 3.2.3).

Note. Variable names in italics correspond to those used in the empirical analysis. Codes refer to the column identifiers in the HFCS User Database, fourth wave (2021). For complete variable definitions, see European Central Bank (2023), HFCS User Database Documentation.