Financial Literacy and Expectations of Asset Returns: Heterogeneity, Determinants, and Effects on Financial Decisions

Sarah Maria Kiesl-Reiter

Munich 2024

Financial Literacy and Expectations of Asset Returns: Heterogeneity, Determinants, and Effects on Financial Decisions

INAUGURAL-DISSERTATION
zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2024

vorgelegt von

Sarah Maria Kiesl-Reiter

Referent:	Prof. Panu Poutvaara, Ph.D.
Korreferent:	Prof. Dr. Joachim Winter
Mündliche Prüfung:	24. Juni 2024
Berichterstatter:	Prof. Panu Poutvaara, Ph.D.
	Prof. Dr. Joachim Winter
	Prof. Dr. Tabea Bucher-Koenen
Promotionsabschlussberatung:	10. Juli 2024

Acknowledgments

I thank my supervisor, Panu Poutvaara, for his support throughout my Ph.D. I am grateful for the freedom he gave me to pursue my research interests and for his encouragement to engage with the broader academic community; whether it was attending conferences or undertaking research stays abroad, he always supported these endeavors. Most importantly, I would like to thank him for his continued insightful feedback on my research.

Additionally, I want to express my gratitude to Joachim Winter for the guidance and mentorship he offered me in his role as my secondary supervisor. Joachim warmly welcomed me into his group, treating me as one of his own students, for which I am truly thankful. He connected me with other great researchers and introduced me to new research topics, enriching my Ph.D. in ways I had not anticipated. His optimism and willingness to brainstorm ideas were particularly encouraging, especially when I faced challenges. The guidance he provided was crucial in improving not only my research but also my writing and presentation skills.

Next, I want to thank Tabea Bucher-Koenen, who agreed to serve as a reviewer of this thesis, and who provided valuable feedback on the first chapter of my thesis. I met Tabea at a couple of scientific conferences, and talking to her—both about research and about personal topics—was always a great pleasure. I look up to her because she demonstrates that it is possible to be successful both as a researcher and as a mother.

This dissertation benefited significantly from the support and recognition of various awards and funding. I am grateful for the Rising Financial Literacy Scholar Award from The Global Financial Literacy Excellence Center and the Early Career Research Prize on Financial Literacy, presented by the Deutsche Bundesbank in collaboration with the Mannheim Institute for Financial Education. The Bernt Rohrer Foundation provided a scholarship for my research stay at Royal Holloway, University of London. My research on financial literacy was further supported by the Klaus Liebscher Economic Research

Scholarship of the Oesterreichische Nationalbank. Additionally, the research grant from the Think Forward Initiative was crucial for the work that contributed to the first chapter of this thesis. My sincere thanks go to these organizations for their support in facilitating my research.

The first chapter of this thesis is the result of a fruitful collaboration with Christa Hainz and Elisabeth Beckmann. Christa played a pivotal role during my Ph.D., providing crucial scientific guidance for our collaborative project and consistent support whenever required. Elisabeth is the one who ignited my curiosity in the topic of household financial decision-making during an internship at the Oesterreichische Nationalbank. Her encouragement was instrumental in my decision to pursue a Ph.D. It is fair to say that without her, this thesis would not have materialized.

Melanie Lührmann and Jonathan Shaw, alongside my secondary supervisor Joachim Winter, were my collaborators on the work leading to the second chapter of this thesis. Melanie taught me a lot about what it means to be both efficient and effective as a researcher. Additionally, she generously hosted me at the Royal Holloway University, which turned out an amazing experience. By taking time out of her busy schedule to show me around London and introducing me to fantastic local students, she made sure I felt comfortable from the beginning, leading me to perform at my best. Jonathan is one of the kindest people I know. Approachable as he is, he was always eager to provide valuable input, making sure we made continuous progress while keeping the bigger picture in perspective.

Apart from my direct collaborators, I want to express my heartfelt thanks to Helmut Elsinger, Helmut Stix, Pirmin Fessler, and Lukas Menkhoff, who spent significant amounts of their free time reading different chapters of my thesis and providing extremely valuable feedback.

My gratitude extends to my colleagues at the ifo Institute: Tanja, Yvonne, Clara, Mirely, Felicitas, Joop, Till, and Marcus. Our shared meals and the engaging discussions over coffee where something I always looked forward to. These simple joys significantly enriched my experience at the ifo Institute.

My friends from Royal Holloway University deserve special mention for their welcoming and supportive presence throughout my stay. Thank you, Dahab, Alex, Sidharth, Shahab, Arthur, Daim, Ren, and Constance, for integrating me into your community from the beginning. You showed me the real-life side of Ph.D. studies, balancing rigorous research with a genuine sense of belonging. By including me in both research seminars and private events, you have demonstrated the importance of a supportive academic and social network. I am really thankful we met and had the opportunity to share this part of our journeys together.

My gratitude goes out to my colleagues at the Munich Graduate School of Economics: Leonie, Luisa, Katharina, Victoria, and Maria. They were always there, ready to discuss the realities of Ph.D. life and its challenges. Their willingness to offer support and practical help made a significant difference in my academic experience. Most importantly, however, our relationships evolved into genuine friendships. These friendships stand out as a highlight of my time spent in academia.

I also want to sincerely thank my friends from outside academia: Anna and Florian, Isabella and Patrick, Thomas and Petra, Sandra, Theresa, Franziska, Manuel, Helena, Katrin, Maria, Julia, and Hima. Thank you for your continued support and for the wonderful and fun times we have had together. You helped me keep things in perspective, providing me with anecdotal evidence for what research has demonstrated repeatedly—that personal friendships are a strong predictor of one's happiness.

My family's support has been essential throughout my academic endeavors, from start to finish. I want to thank my parents-in-law, Gertrud and Gerhard, for their encouragement and positive presence in our lives. My brother David, along with his partner Valentina and their lovely daughters Alina and Livia, have been a source of joy and refreshment. David's empathy and understanding, especially regarding the balance between work and family life, have been invaluable. Our time together has been a cornerstone of my well-being during my Ph.D. studies. Above all, my parents, Elfriede and Johann, have been essential. They had to make financial sacrifices to enable my studies, and I truly appreciate that. Their constant acts of love, from visiting our home with home-cooked meals to helping with household chores and dedicating quality time to my son, who adores them, have been the foundation of my support system.

Finally, my greatest gratitude goes to Benjamin—my husband and best friend. Thank you for always having my back, for believing in me, for encouraging me to reach my full potential, and for your optimism and motivation during tough times. Without your support, completing this dissertation would not have been possible. My Ph.D. was also uniquely marked by a profound personal event: the birth of our son, Florentin. To my dear son, thank you for bringing so much joy into our lives. I love you.

Sarah Maria Kiesl-Reiter

Contents

List of Figures

List of Tables

Introduction

Throughout their entire lives, individuals are confronted with making financial decisions that impact their financial and overall well-being. For instance, individuals participating in the stock market and those financially planning for their long-term future tend to accumulate higher levels of wealth (van Rooij, Lusardi, and Alessie 2012; Ameriks, Caplin, and Leahy 2003). At the societal level, individuals' financial decisions also have important implications for wealth inequality (Lusardi, Michaud, and Mitchell 2017; Campbell 2016; Piketty 2014). Moreover, household financial decisions can drive dynamics at the aggregate level, including boom-bust cycles, putting at risk financial stability (Mian and Sufi 2010, 2018; International Monetary Fund 2017).

The literature identifies various factors that matter for individual financial decision-making. In this thesis, I focus on two key factors, (1) individual financial literacy and (2) subjective expectations about (uncertain) future financial and economic outcomes.

Individual financial literacy. A large literature studies the levels and determinants of individuals' financial literacy and its role in financial decision-making. As a measure of individual financial literacy, a set of three survey questions developed by Lusardi and Mitchell (2008)—covering knowledge of interest rates, inflation, and risk diversification—has become well-established in the literature. The measure has been used in large-scale projects to compare people's financial literacy across countries all over the world ("Financial Literacy around the World;" for an overview, see Lusardi and Mitchell 2011, 2014, 2023). While levels of financial literacy have been found to vary considerably across countries, patterns regarding age, gender, and education are relatively stable. In most countries, financial literacy is highest for the middle-age group and lowest for the very young and the very old; it is higher for men than for women, and positively associated with education (Lusardi and Mitchell 2014, 2023; Bucher-Koenen, Lusardi, Alessie, and van Rooij 2017).

In addition to this rather basic measure of financial literacy, there exists a range of alternative measures that aim at capturing specific aspects of financial literacy. For instance, van Rooij, Lusardi, and Alessie (2011) design a set of questions to measure individuals' understanding of the returns and risks associated with different financial assets, the workings of the stock market, and the concept of risk diversification. Lusardi and Tufano (2015) propose a novel set of questions on debt literacy. Gathergood and Weber (2017) introduce questions on mortgage literacy, and Beckmann and Stix (2015) develop a question to measure individuals' understanding of the exchange-rate risk of foreign-currency loans.

Ultimately, research in this area aims to understand to which extent variation in individuals' financial literacy can explain heterogeneity in financial decision-making. Concerned about endogeneity of financial literacy, studies in this area often resort to instrumental variable (IV) strategies as a methodological approach. With respect to savings and investment behavior, individuals with higher financial literacy have been found to be more likely to plan for retirement (e.g., Lusardi and Mitchell 2011a; Bucher-Koenen and Lusardi 2011), to build up savings (Beckmann and Kiesl-Reiter 2023), to participate in the stock market (van Rooij, Lusardi, and Alessie 2011), and to achieve better investment outcomes (von Gaudecker 2015; Guiso and Viviano 2015; Clark, Lusardi, and Mitchell 2017; Bianchi 2018). With respect to debt behavior, individuals with higher financial literacy are less likely to take out costly or risky loans (Disney and Gathergood 2013; Gathergood and Weber 2017; Beckmann and Stix 2015) or to carry high debt into retirement (Lusardi, Mitchell, and Oggero 2020).

The first chapter of this thesis, joint work with Elisabeth Beckmann and Christa Hainz, adds to the research on financial literacy. We introduce *contingent liabilities arising from third-party loan guarantees* as a novel aspect to the literature. We design a new question that measures how well individuals understand the consequences of granting a guarantee, and document considerable variation in the levels of guarantee literacy across countries and population subgroups. Most importantly, our results indicate that guarantee-literate individuals are significantly less likely to act as guarantors than illiterate individuals.

Given the important role of financial literacy for individual financial behavior and well-being, and more generally, for macro-financial stability, research on financial literacy has also become of increasing interest to policy makers. Indeed, many countries around the world are currently implementing national strategies with the aim of improving individuals' financial literacy. Measuring a population's level of financial literacy and identifying the most vulnerable population subgroups and their specific needs have been identified as important first steps towards reaching this goal (OECD 2015).

Subjective expectations. Individuals' subjective expectations about (uncertain) future events are another important factor in individual financial decision-making. Manski (2004) emphasizes the importance of explicitly measuring individuals' subjective expectations in surveys (rather than merely making assumptions about them), optimally using subjective probabilities. While researchers started to elicit so-called probabilistic expectations in surveys already in the 1990s (e.g., Dominitz and Manski 1997a,b; Dominitz 1998), there has been a noticeable increase in the recent past, with many large-scale surveys now including probabilistic questions on subjective expectations for different outcomes (for an overview, see Manski 2004; Hurd 2009; Manski 2018; Kuchler, Piazzesi, and Stroebel 2023).

In my thesis, I am interested in the subjective expectations and beliefs that individuals hold about the returns of different (financial and non-financial) assets. There is a large literature studying individuals' subjective expectations about *stock market returns* and, more recently, researchers also show increasing interest in individuals' subjective expectations about *house price changes*. Many studies in this research area document the importance of subjective asset return expectations for individual investment and debt behavior. For instance, individuals with more pessimistic expectations about future stock market returns are less likely to own stocks (Dominitz and Manski 2007; Hurd, van Rooij, and Winter 2011; Kézdi and Willis 2011). Further, stock market expectations play an important role not only for participation in the stock market but also for portfolio choice and trading behavior (Giglio, Maggiori, Stroebel, and Utkus 2021; Merkle and Weber 2014; Zimpelmann 2021; Amromin and Sharpe 2014).

In the context of housing markets, subjective expectations about house price changes have been found to affect individuals' decisions of whether to rent or to buy a home (Bailey, Cao, Kuchler, and Stroebel 2018; Adelino, Schoar, and Severino 2018), whether to sell a home (Bottan and Perez-Truglia 2020), and how to finance a home (Bailey, Dávila, Kuchler, and Stroebel 2019). In recent research, subjective house price expectations have also been linked to non-housing outcomes; conducting a field experiment, Chopra, Roth, and Wohlfart (2023) find that expectations about long-run house price growth play a role in individuals' consumption decisions. Finally, subjective house price expectations can also drive dynamics at the aggregate level, including housing booms and busts (Piazzesi and Schneider 2009; Case, Shiller, and Thompson 2012; Burnside, Eichenbaum, and Rebelo 2016; Landvoigt 2017; Kaplan, Mitman, and Violante 2020; Kindermann, Le Blanc, Piazzesi, and Schneider 2021).

Given the relevance of subjective asset return expectations, a large literature aims at developing a better understanding of how individuals form such expectations and how expectations are determined. Studies based on survey data consistently document large heterogeneity in subjective expectations across individuals—both for stock market returns (e.g., Heiss, Hurd, van Rooij, Rossmann, and Winter 2022; Giglio, Maggiori, Stroebel, and Utkus 2021; Hurd, van Rooij, and Winter 2011; Dominitz and Manski 2007, 2011) and for house price changes (e.g., Kuchler, Piazzesi, and Stroebel 2023; Armona, Fuster, and Zafar 2019).

Some of this heterogeneity can be explained by demographic characteristics: Females and individuals with lower socioeconomic status hold on average more pessimistic expectations about future stock market returns (e.g., Drerup, Enke, and von Gaudecker 2017; Hurd, van Rooij, and Winter 2011; Kuhnen and Miu 2017; Dominitz and Manski 2004). Interestingly, with respect to house price growth, Kuchler, Piazzesi, and Stroebel (2023) document more pessimistic expectations for individuals with high education and numeracy. However, as pointed out by Giglio, Maggiori, Stroebel, and Utkus (2021) and Kuchler, Piazzesi, and Stroebel (2023), demographic characteristics can explain only a small share of the heterogeneity in subjective asset return expectations.

Especially when it comes to subjective expectations about housing markets, the sources of heterogeneity are still poorly understood. To this end, Armona, Fuster, and Zafar (2019) point out that "home price expectations are believed to play an important role in housing dynamics, yet we have limited understanding of how they are formed" (p. 1371). Traditional prediction models for house prices, going back to Case and Shiller (1989), are based on past realized changes, establishing a natural starting point for modeling subjective house price expectations. Indeed, there is evidence that people extrapolate from past house price changes when forming expectations about future house price changes (Case, Shiller, and Thompson 2012; Armona, Fuster, and Zafar 2019). In addition, Armona, Fuster, and Zafar (2019) highlight that not only realizations but also perceptions of past (national) house price changes matter in the expectation formation process.

The second chapter of this thesis, joint work with Melanie Lührmann, Jonathan Shaw, and Joachim Winter, extends the existing literature on the formation of subjective house price expectations. Combining unique survey evidence with administrative data, we shed light on a new aspect and study how *local macroeconomic conditions* shape individuals' subjective expectations about house price growth in their local area of residence. Local macroeconomic conditions are part of an individual's salient experiences, and a growing literature shows that experiences, broadly defined, affect belief formation (Malmendier 2021b; Bailey, Cao, Kuchler, and Stroebel 2018; Kuchler and Zafar 2019).

Most of the existing work that uses survey data to study the formation and heterogeneity of subjective asset return expectations focuses on single assets, neglecting an important aspect of portfolio choice—*correlation of asset returns*. In the third chapter of this thesis, I fill this gap and extend the literature by studying individuals' subjective expectations about the *joint* return distribution for an investment in a "mixed asset," including both stocks *and* real estate. This task requires individuals to not only form expectations about the returns of assets but to also take into account the correlation between those returns. Exploring subjective expectations about joint return distributions can contribute to understanding the widely documented lack of diversification in household portfolios (see e.g., Blume and Friend 1975; Goetzmann and Kumar 2008; Gomes, Haliassos, and Ramadorai 2021; Badarinza, Campbell, and Ramadorai 2016).

In what follows, I provide a summary of the three chapters included in this thesis. Each chapter stands alone and permits independent reading. A consolidated bibliography is provided at the end of the thesis.

Chapter 1, joint work with Elisabeth Beckmann and Christa Hainz, studies individuals' financial literacy regarding third-party loan guarantees and analyzes its effect on the granting of such guarantees. Individuals who agree to grant a guarantee (so-called guarantors) take on contingent liabilities, i.e., while initially only agreeing to help the borrower gain access to credit, they are ultimately liable for the borrower's outstanding debt (including interest) if the borrower fails to repay. Since granting a third-party loan guarantee does not involve a financial transaction at the time of contracting, guarantors are often unaware of the associated risks and potential consequences.

We introduce contingent liabilities arising from third-party loan guarantees as a new aspect to the financial literacy literature. To measure individuals' understanding of the consequences of granting a guarantee, we develop a new survey question. The question was included in two waves of the OeNB Euro survey—a repeated cross-sectional survey on household finance conducted by the Austrian Central Bank in central, Eastern, and Southeastern Europe. In total, our analysis sample covers 18,000 individuals living in nine countries. Our results show that only 56 percent of individuals have an understanding of the consequences of granting a guarantee. We observe considerable variation across countries, with the percentage of guarantee-literate individuals ranging from 43 percent in Bosnia and Herzegovina to 70 percent in Croatia.

In line with previous literature, we document substantial heterogeneity of guarantee literacy in the population. Young people have lower levels of guarantee literacy than the middle-aged. Employed individuals and those with higher household income and higher levels of education are more literate. Notably, we do *not* find a gender gap in guarantee literacy, a finding that may not be too surprising given that the countries in our analysis used to be communist with comparatively equal gender roles. Other studies on former communist countries similarly find a low gender gap in financial literacy (Bucher-Koenen and Lusardi, 2011, Cupák, Fessler, Schneebaum, and Silgoner, 2018, Beckmann and Kiesl-Reiter, 2020).

Finally, we study how guarantee literacy affects the granting of third-party loan guarantees. To deal with endogeneity of guarantee literacy, we employ an instrumental variable strategy. Using *cohort-specific averages of financial literacy in the region in which a respondent lives* as an instrument for individual guarantee literacy (motivated by research of Agnew, Bateman, and Thorp, 2013, Bucher-Koenen and Lusardi, 2011, Klapper, Lusardi, and Panos, 2012, and van Rooij, Lusardi, and Alessie, 2011), we find that guarantee-literate individuals are 11 percentage points less likely to act as guarantors than guarantee-illiterate ones. Our results are robust to a placebo analysis and several sensitivity checks.

Chapter 2, joint work with Melanie Lührmann, Jonathan Shaw, and Joachim Winter, studies belief formation in survey data on subjective local house price expectations and perceptions of past house price changes from Great Britain—a country with high home ownership and transaction rates, and profound and persistent geographical variation in house price dynamics (Agrawal and Phillips 2020; Overman and Xu 2022). We focus on two predictors of subjective house price expectations, *past house price changes* and *local economic conditions*.

For our analysis, we use data from a newly designed survey module on subjective expectations, conducted by the Financial Conduct Authority as part of the Financial Lives survey between August 2019 and February 2020. Our comparatively large analysis sample covers almost 2,800 individuals living in 364 local housing markets. We elicit respondents' perceptions of house price changes over the past year in their local area of residence, and their subjective expectations of one-year-ahead local house price changes using probabilistic elicitation techniques (Manski 2004). Respondents' local area of residence is used to link the survey data with administrative data. More specifically, we link the survey data with (i) the UK House Price Index in a respondent's local area at the time of interview and (ii) locally experienced economic conditions (as measured by local unemployment rates at the time of interview, and alternatively, by local deprivation scores from 2019), to study the role of realized local house price changes and local economic conditions in the expectation formation process.

We derive four main results: First, individuals do not extrapolate from realized past one-year local house price changes (when forming subjective expectations about one-year-ahead local house price changes), but rather from their perceptions of past local house price changes. Second, locally experienced economic conditions are an additional important predictor of subjective local house price expectations: Individuals who live in local areas with higher unemployment rates and deprivation scores expect, on average, lower rates of house price growth. Third, the importance of such locally experienced economic conditions in individuals' beliefs varies across subgroups, and it matters in particular for those who are risk averse, less financially sophisticated, and residing in local housing markets where past prices display high volatility and no short-run momentum. Fourth, our results point to substantial heterogeneity in subjective expectations that are driven in part by large and heterogeneous gaps between perceived and realized price changes, echoing findings in Armona, Fuster, and Zafar (2019). While we find little evidence of systematically higher or lower levels of perceptions based on observables, perception gaps are driven both by local market factors such as past house price volatility, as well as individual characteristics; they are larger for women and also particularly pronounced for individuals with low financial sophistication.

We conclude that agents' belief formation process is not fully captured by models that only include (recent or more distant) past house price changes as predictors. Instead, agents' beliefs about future house prices react to salient local information as well.

Chapter 3 studies the formation of individuals' subjective expectations about the returns and risks of different investments, and their consistency with basic diversification properties. In my analysis, I use data from a unique survey module that asks individuals to report their subjective expectations about the joint return distribution of an investment in a portfolio which consists of two broad asset classes, housing and stock. This task requires individuals to not only form expectations about asset returns but to also incorporate expectations regarding the correlation between those returns. In addition, the survey elicits individuals' subjective expectations about the return distributions of the assets underlying the portfolio—separately for an investment in housing and for an investment in stocks.

The expectation questions included in the survey are designed in a way such that they elicit the whole return distribution using subjective probabilities, similar to Giglio, Maggiori, Stroebel, and Utkus (2021) and Laudenbach, Loos, Pirschel, and Wohlfart (2021). From these subjective return distributions, I obtain estimates of the mean and standard deviation to investigate how well individuals' subjective expectations regarding the returns (as measured by the mean of the subjective return distribution) and risks (as measured by the standard deviation of the subjective return distribution) of the three different investments align with two basic diversification properties: (1) The expected return of a two-asset portfolio lies within the range of the expected returns of its individual underlying assets. (2) The risk of a two-asset portfolio is lower than or equal to the maximum risk associated with its individual underlying assets.

My results show that a non-negligible share of respondents (every fourth) does *not* provide a response to the expectation questions about asset returns. Further, respondents who provide responses take into account basic diversification properties only partially or not at all in their expectation formation: 50% of respondents form subjective expectations in line with the first property (on asset returns), 79% form expectations in line with the second property (on asset risk), and only 41% form expectations in line with both properties. Respondents' socio-economic status and their overall financial literacy are strong predictors for participating in the expectation-elicitation task and forming expectations in line with basic diversification properties.

Individuals who underestimate the expected return of a portfolio investment and those who overestimate the risk of such, may refrain from diversifying their portfolios, which can result in sizeable return losses, particularly for individuals with low financial literacy (von Gaudecker 2015). My results suggest that understanding of probabilities and basic concepts of diversification are important topics to be covered in financial-education programs, which should specifically be targeted towards individuals with lower socio-economic status.

Chapter **1**

Third-Party Loan Guarantees: Measuring Literacy and Its Effect on Financial Decisions*

Abstract: The granting of a third-party guarantee for a loan does not directly involve a financial transaction. Therefore, guarantors might not understand that they are taking on a liability, albeit a contingent one. We introduce literacy about guarantees as a novel and distinct aspect of financial literacy. For nine Eastern European countries, we find that 44 percent of individuals lack this form of financial literacy. Using regional cohort-specific financial literacy as an instrument for individual guarantee literacy, we show that guarantee literacy significantly reduces the probability of acting as a guarantor. Our results are robust to a placebo analysis and several sensitivity checks.

*This chapter is based on joint work with Elisabeth Beckmann and Christa Hainz. A version of this chapter has been published in the Working Paper Series of the Oesterreichische Nationalbank (Beckmann, Hainz, and Kiesl-Reiter 2022).

Neither a borrower nor a lender be,
For loan oft loses both itself and friend.
—William Shakespeare (Hamlet, Act 1, Scene 3)

1.1 Introduction

For several centuries, it has been a truism that borrowing from a friend or lending money to a friend may put both the friendship and the money at risk. Compared to loans, third-party loan guarantees are often not treated with the same degree of caution. Since granting a third-party loan guarantee does not involve a financial transaction at the time of contracting, guarantors are often unaware of the associated risks and potential consequences.

Acting as a guarantor for a loan is common in both emerging and advanced economies. In North Macedonia, seven percent of the adult population acted as guarantors in 2019 (OeNB Euro Survey). In Poland, the share was four percent, with eleven percent of guaranteed loans in arrears (BIK 2018). In Germany, about three percent of over-indebted individuals identify guarantee-related issues as the main reason for their indebtedness (Creditreform Wirtschaftsforschung 2020). In the UK, nine percent of individuals have experience guaranteeing a loan (YouGov 2021). In recent years, guarantees have become widespread in the UK high-cost credit market—a development about which the Financial Conduct Authority has expressed concern (FCA 2017).

In this paper, we study individuals' financial literacy regarding third-party loan guarantees (short: *guarantees*) and analyze the effect of this literacy on the granting of guarantees. To measure how well individuals understand the consequences of acting as guarantors, we developed a new survey question on *guarantee literacy*. This question was included in the 2018 and 2019 waves of the OeNB Euro Survey—a survey on household finance conducted by the Austrian Central Bank in central, Eastern, and Southeastern European countries (short: *Eastern Europe*).

Our empirical analysis yields three main results that contribute to the understanding of individuals' financial decisions. First, 44 percent of individuals are not aware of the consequences associated with a guarantee. Second, our survey question on guarantee literacy captures a specific concept not covered by the well-known questions on financial literacy. Third, guarantee-literate people are 11 percentage points less likely to act as guarantors than those who are guarantee illiterate.

To address endogeneity concerns in estimating the effect of guarantee literacy on the probability of granting a guarantee, we develop an instrumental-variables strategy using regional cohort-specific general financial literacy as an instrument for individual guarantee literacy. In addition, we conduct a placebo analysis in which the information on whether someone currently grants an informal loan to family or friends is the dependent variable. We find that guarantee literacy has no effect on granting informal loans, which demonstrates that our results are not driven by unobserved characteristics, such as social norms or trust. This result and various robustness checks corroborate our finding that being guarantee literate lowers the probability that someone acts as a guarantor.

To conclude a third-party guarantee, three parties are required: the bank, the borrower, and the guarantor. For the borrower, providing a guarantor as security leads to lower interest rates and facilitates access to credit (De Blasio, De Mitri, D'Ignazio, Finaldi Russo, and Stoppani 2018; Bachas, Kim, and Yannelis 2021).[1] Guarantees grant the bank access to the guarantor's assets up to the outstanding amount, including interest, and unlike collateral, not only to the pledged assets (De Haas and Millone 2020). The guarantor, while initially only agreeing to help the borrower gain access to credit, has to step in if the borrower defaults.

By introducing the concept of guarantee literacy, our paper adds a new aspect to the research on financial literacy. There is a large body of research documenting the levels of financial literacy and its impact on savings and investment behavior.[2] By contrast, much less attention has been paid to the household liability side—even though a lack of financial literacy may result in poor borrowing decisions that ultimately have a highly negative impact on individuals' financial well-being, especially in times of crisis. With regard to financial literacy, the aspect of contingent liabilities that individuals assume when granting a guarantee has been neglected so far.[3]

Regarding household liabilities, using the "big three" financial literacy questions on interest rates, inflation, and risk diversification (Lusardi and Mitchell 2008, 2011b), it has been shown that individuals with higher financial literacy borrow less (Stango

[1] So far, this research has focused on access to credit for firms where guarantees are usually granted by the government.

[2] For an overview of the respective literature before 2014, see Lusardi and Mitchell (2014); for more recent studies, see for example, von Gaudecker (2015), Badarinza, Campbell, and Ramadorai (2016), Anderson, Baker, and Robinson (2017), Clark, Lusardi, and Mitchell (2017), Bianchi (2018), or Hastings and Mitchell (2020).

[3] The contingency aspect also plays a role in insurance decisions. In our case, the guarantor is not the policy holder but insures the bank against the default risk of the borrower. Measures for insurance literacy are used, for example, by Cole et al. (2013).

and Zinman 2009), are less likely to take out a costly loan (Disney and Gathergood 2013; Lusardi and Bassa Scheresberg 2013), and are less likely to default on a sub-prime mortgage (Gerardi, Goette, and Meier 2013). Those with high financial literacy less often borrow informally, but more often formally (Klapper, Lusardi, and Panos 2013). Moreover, individuals with high financial literacy are less likely to carry high debt into retirement (Lusardi, Mitchell, and Oggero 2020).

In addition, research has developed measures to capture specific liability aspects of financial literacy. Proposing a novel set of questions on debt literacy, Lusardi and Tufano (2015) show that people who are more literate with respect to the debt-specific questions are less likely to have high-cost debt products or excessive debt. Almenberg, Lusardi, Säve-Söderbergh, and Vestman (2020) add questions about attitudes towards debt and find that those who are uncomfortable with debt have lower debt ratios. Gathergood and Weber (2017) introduce questions on mortgage products and demonstrate that individuals with better mortgage literacy are less likely to choose expensive interest-only mortgages. Also focusing on mortgages, van Ooijen and van Rooij (2016) show that debt literacy is lower than financial literacy in general and that individuals who seek financial advice are particularly likely to take out riskier mortgages if they have a low level of debt literacy. Individuals with a better understanding of the exchange-rate risk of foreign-currency loans are less likely to take out such loans (Beckmann and Stix 2015).

The main contributions of our paper are the following: First, we conceptually introduce contingent liabilities arising from a guarantee as a new aspect of the financial literacy literature. For this purpose, we develop a measure of how well individuals understand the consequences of a guarantee. Second, we present novel evidence, which is harmonized and comparable across countries, on how widespread both third-party guarantees and guarantee literacy are, and how guarantee literacy is associated with individuals' characteristics.[4] Third, we analyze the effect of guarantee literacy on the granting of guarantees using an instrumental-variables approach. Financial literacy in the peer group serves as an instrument which we measure using the average regional cohort-specific financial literacy. For our instrument, we calculate leave-out means drawing on unique financial literacy data collected over the course of several survey waves of the OeNB Euro Survey.

The remainder of this paper is organized as follows. In Section 1.2, we describe our data and introduce our new survey question on guarantee literacy. In Section 1.3, we

[4]We are the first to provide evidence on third-party guarantees that is comparable across countries.

demonstrate the validity and specificity of our new question and present descriptive evidence on the correlates of guarantee literacy. In Section 1.4, we explain our empirical framework and introduce regional cohort-specific financial literacy as an instrument for individual guarantee literacy. In Section 1.5, we present our main findings from OLS and IV estimations and a placebo analysis, and perform several robustness checks. Finally, we summarize and discuss our findings in Section 1.6.

1.2 Data and Background

The main data source for our analysis is the *OeNB Euro Survey*, a survey of private individuals on household finance. It has been conducted by the Austrian Central Bank since 2007 as a repeated cross-sectional face-to-face survey in ten Eastern European countries: six EU member states that are not part of the euro area (Bulgaria, Croatia, the Czech Republic, Hungary, Poland, and Romania) and four EU candidates and potential candidates (Albania, Bosnia and Herzegovina, North Macedonia, and Serbia). In each country and in each survey wave, around 1,000 individuals are interviewed based on multistage random sampling procedures. The samples reflect a country's population characteristics in terms of age, gender, region, and ethnicity. The weights are calibrated using census population statistics for each country and each wave separately. When several countries are pooled, the weights also take into account the relative size of each country's population.[5] We use data from the survey waves conducted in the fall of 2018 and 2019. In these waves, we introduce a new survey question that is central to our analysis of guarantee literacy.[6]

The law of guarantees, which is based on contract law, stipulates that the guarantor is liable for the borrower's outstanding debt including interest, if the borrower fails to repay. Although there might be slight differences in the laws in the individual countries, the core of the guarantee, i.e., the associated legal obligation, is comparable in all nine countries. Table A.15 in the Appendix presents the relevant legislation for each country in our sample. By signing the guarantee, the guarantor assumes a contingent risk—a fact and the extent of which the guarantor may not be aware of. With our new question,

[5]For the remainder of the paper, we employ individual weights when reporting statistics for individual countries. We use the combined individual-population weights when presenting statistics that include multiple countries. We do not weight survey data when conducting regression analyses.

[6]We do not include Albania in our analysis as the data for the waves 2018 and 2019 do not cover North Albania and are, therefore, not representative of the population.

Table 1.1. Survey question on guarantee literacy

Concept	Survey question
Third-party guarantee	Suppose your friend has taken out a consumer loan from a bank to finance his/her new car and you acted as a guarantor for this consumer loan. Then your friend loses his/her job and therefore is no longer able to repay the loan. What is your legal obligation as a guarantor? As a guarantor, I am obliged to (1) immediately inform the bank about any financial difficulties my friend may run into, but I have no financial obligations. (2) financially support my friend but I do not have any financial obligations towards the bank where he/she took out the loan. (3) repay the outstanding amount of the loan excluding interest to the bank. (4) repay the outstanding amount of the loan including interest to the bank. (5) None of the statements is correct. (6) Do not know (7) No answer

Notes: The table shows the survey question on guarantee literacy included in the OeNB Euro Survey. The correct answer is (4).

shown in Table 1.1, we measure individuals' literacy about the consequences of granting a guarantee.

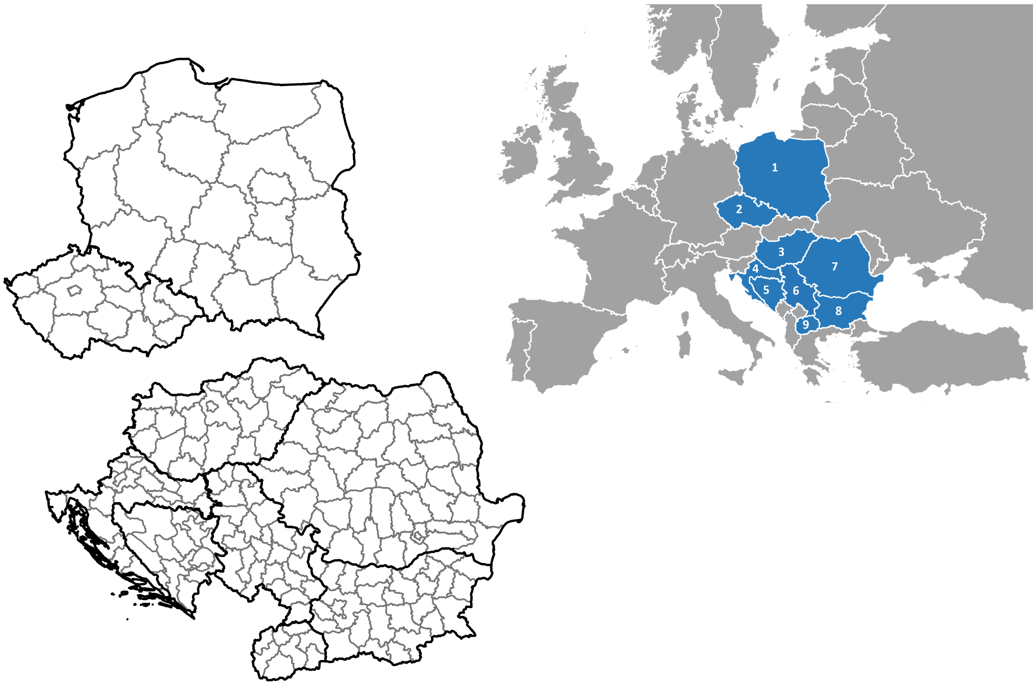
Respondents who choose answer (4) are fully aware of the risk involved in granting a guarantee; we classify them as *guarantee literate*. Respondents selecting an answer other than (4) are classified as *guarantee illiterate*. Respondents selecting answer (3) comprehend the contingent nature, but underestimate the amount for which they are liable. In a robustness check, we show that classifying respondents who choose answer (3) or (4) as guarantee literate does not change our results qualitatively.[7]

We also ask respondents whether they have helped a family member or a friend in the past twelve months by (i) granting a loan or (ii) acting as a guarantor for a loan. Given the known structure of loans, we can assume that both forms of financial assistance would still be ongoing at the time of the interview. In addition to information about current informal loans and guarantees, we collect information on whether individuals have ever granted an informal loan or a guarantee (available for 2018 only).

The OeNB Euro Survey data include a rich set of information, such as socio-demographic characteristics, individual beliefs and attitudes, and proxies for wealth and the use of financial products. The data also contains the addresses of the interviewer starting points for the random route sampling, which means that we know that a respondent's residence is within walking distance of that starting point. This allows us to geographically merge

[7]Respondents who do not provide an answer to the survey question on guarantee literacy (one percent of the overall sample) are excluded from our analysis.

Figure 1.1. Regional subdivisions for instrumental-variables calculation



Notes: The figure shows the regional subdivisions on which the calculation of our instrument is based. We distinguish the regional subdivisions according to the Eurostat NUTS 2016 classification. In general, our definition of regional subdivisions corresponds to the regions at the NUTS 3 level. In Poland, our definition of regional subdivisions is equivalent to the NUTS 2-level regions (due to the small number of observations at the NUTS 3 level). Countries under study: 1-Poland (PL), 2-Czech Republic (CZ), 3-Hungary (HU), 4-Croatia (HR), 5-Bosnia and Herzegovina (BA), 6-Serbia (RS), 7-Romania (RO), 8-Bulgaria (BG), 9-North Macedonia (MK).

the survey data with (two) indicators of the area in which the respondent lives: (i) an indicator of regional economic activity measured by nightlight data (following Henderson, Storeygard, and Weil, 2012), and (ii) an indicator of the regional banking environment (as in Beckmann, Kiesl-Reiter, and Stix, 2018).[8] All variables used in our empirical analysis are described in Table A.2 in the Appendix.

We further take advantage of the fact that the OeNB Euro Survey (i) has been conducted over a long period of time and (ii) contains the big three financial literacy questions (Lusardi and Mitchell 2008, 2011b) (see Table A.1 in the Appendix for the exact wording).[9]

[8] Matching the survey data with bank branch data based on geolocation information is a unique feature of the data. Unfortunately, the data does not allow matching guarantors and banks at the individual level; it also does not allow matching guarantors and borrowers.

[9] We use the terms *financial literacy* and *financial knowledge* as synonyms, i.e., we use a narrow definition of the financial literacy concept (see World Bank, 2014).

The data, which stems from a total of seven survey waves (2012–2016, 2018, and 2019), provides us with sufficient observations (around 63,000) to compute *regional cohort-specific financial literacy*, which we use as an instrument for individual guarantee literacy. Figure 1.1 illustrates the regional subdivisions we use, most of which correspond to the smallest regions of the NUTS-2016 classification developed by Eurostat.

1.3 Descriptive Evidence

In this section, we address our first research question: *How well do individuals understand the consequences of granting a guarantee*? We provide descriptive statistics on guarantee literacy and compare it to the big three questions on financial literacy. We also investigate how guarantee literacy is associated with individuals' socio-demographic and socio-economic characteristics.

1.3.1 Guarantee Literacy Versus General Financial Literacy

Our results in Table 1.2 show that 55.8 percent of the individuals answer the survey question on guarantee literacy correctly (by selecting answer 4) and can thus be considered guarantee literate.

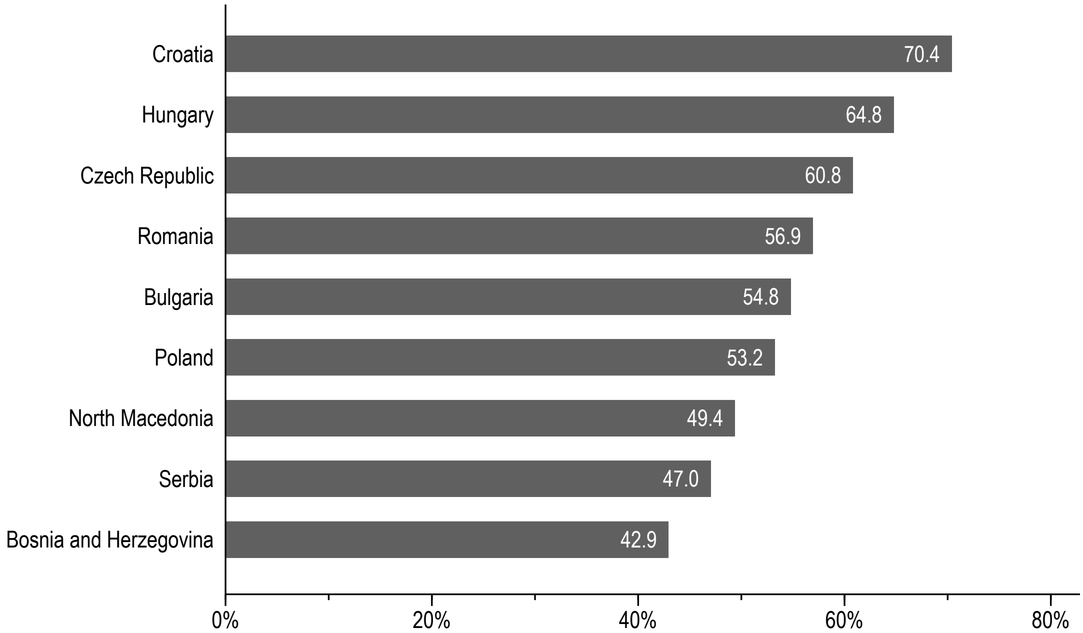
Table 1.2. Answers to guarantee literacy question

As a guarantor, I am obliged to …	% of individuals
(1) Immediately inform the bank (but no financial obligations)	6.2
(2) Financially support my friend (but no financial obligations towards bank)	6.4
(3) Repay the outstanding amount of the loan excluding interest to the bank	9.2
(4) Repay the outstanding amount of the loan including interest to the bank	55.8
(5) None of the statements is correct	6.1
(6) Do not know	16.3

Notes: The table shows the distribution of the responses to the survey question on guarantee literacy. The statistics are based on weighted data from the 2018 and 2019 waves of the OeNB Euro Survey, including nine Eastern European countries. N=17,985.

Figure 1.2 shows that the level of guarantee literacy varies considerably across countries. In Croatia, 70.4 percent of the individuals choose the correct answer. In Hungary and the Czech Republic, guarantee literacy is above 60 percent. In Romania, Bulgaria, and Poland, more than half of the individuals are guarantee literate. In North Macedonia, Serbia, and Bosnia and Herzegovina, figures are below 50 percent. Bosnia and Herzegovina ranks last, with only 42.9 percent of individuals answering correctly.

Figure 1.2. Variation in guarantee literacy across countries



Notes: The figure shows the country-specific percentage of individuals with correct answers to the survey question on guarantee literacy. Statistics are based on weighted data from the 2018 and 2019 waves of the OeNB Euro Survey. N=17,985.

To put our new survey measure into perspective, we compare the answers on guarantee literacy with the big three financial literacy questions on interest rates, inflation, and risk diversification. Table 1.3 shows that guarantee literacy is positively correlated with literacy about interest rates, inflation, and risk diversification. The correlation is most pronounced for inflation literacy, where two thirds of individuals with the correct answer on guarantees also provide the correct answer on inflation. At the same time, 58 percent of those who are guarantee illiterate also give an incorrect answer to the inflation question. For risk diversification, the positive correlation is smaller, which is not surprising given that literacy about risk diversification is much lower than about guarantees. While the association is positive, these results also indicate that guarantee literacy is a specific aspect of financial literacy that is not captured by the frequently used big three questions.

1.3.2 Heterogeneity in Guarantee Literacy

To investigate which groups are more likely to be guarantee literate, we perform a multivariate regression analysis. We present results from estimating a linear probability

Table 1.3. Cross-question consistency of guarantee literacy and financial literacy

	Interest-rate literate		Inflation literate		Risk-diversification literate	
	Yes (%)	No (%)	Yes (%)	No (%)	Yes (%)	No (%)
All individuals	54.6	45.4	57.6	42.4	44.5	55.5
Only individuals …						
Guarantee literate	62.8	37.2	69.5	30.5	50.8	49.2
Guarantee illiterate	44.2	55.8	42.4	57.6	36.3	63.7

Notes: The table shows the percentage of individuals with (in)correct answers to the survey questions on guarantees, interest rates, inflation, and risk diversification (detailed in Tables 1.1 and A.1). Statistics are based on weighted data from the 2018 and 2019 waves of the OeNB Euro Survey, including nine Eastern European countries. N=17,508.

model in Table 1.4. In the first specification, we study how individuals' guarantee literacy correlates with their socio-demographic characteristics. In the second specification, we add the three standard financial literacy questions. In the third specification, we control for interviewer characteristics as suggested by Crossley, Schmidt, Tzamourani, and Winter (2020), who show that interviewers introduce measurement error, especially when it comes to questions evaluating individuals' levels of financial literacy.

Our results show that younger individuals (18–35) are less likely to select a correct answer. Married individuals and those with higher levels of education are more literate. Guarantee literacy is also more prevalent among individuals with jobs and those with higher incomes. Our results are in line with findings of previous studies with respect to age and education (Lusardi and Mitchell 2011b) as well as income (Brown and Graf 2013). The lack of a gender gap in the nine Eastern European countries may not be too surprising, given that they used to be communist and had comparatively equal gender roles. Other studies on former communist countries similarly find a low gender gap (Bucher-Koenen and Lusardi, 2011; Cupák, Fessler, Schneebaum, and Silgoner, 2018). For the countries in our dataset, there are also no or only small gender differences in interest-rate, inflation and risk-diversification literacy (Beckmann and Kiesl-Reiter 2020).

Regarding the three standard financial literacy questions, our results are in line with those from our analysis on cross-question consistency (Table 1.3). The positive coefficient is highest for inflation and lowest for risk diversification. When adding interview duration and interviewer characteristics, the results for socio-demographic characteristics and financial literacy do not change. Of these additional control variables, the interviewer's age is positive and statistically significant, but the size of the coefficient is small.[10]

[10]We address interviewer effects in Section 1.4.1 and Section 1.5.3.

Table 1.4. Multivariate analysis of guarantee literacy

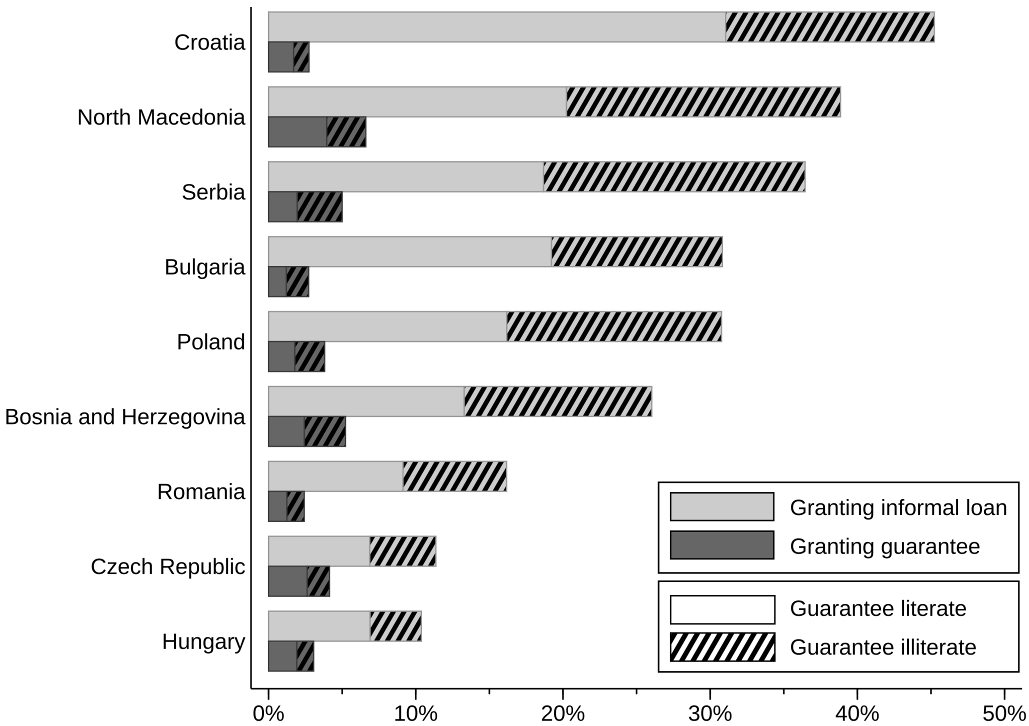
Dependent variable	Guarantee literate		
	(1)	(2)	(3)
Socio-demographic characteristics			
Female	−0.005	0.007	0.005
	(0.007)	(0.007)	(0.007)
Age (ref: 36–50)			
18–35	−0.063***	−0.056***	−0.051***
	(0.011)	(0.010)	(0.010)
51–65	0.027***	0.016	0.015
	(0.010)	(0.010)	(0.010)
65 or older	0.016	0.008	0.004
	(0.015)	(0.015)	(0.015)
Education (ref: Secondary)			
Primary	−0.105***	−0.070***	−0.072***
	(0.014)	(0.014)	(0.014)
Tertiary	0.048***	0.021**	0.023**
	(0.011)	(0.010)	(0.010)
Married	0.017*	0.018**	0.018**
	(0.009)	(0.009)	(0.009)
Working	0.039***	0.042***	0.045***
	(0.010)	(0.010)	(0.010)
Household income (ref: Low)			
Medium	0.046***	0.032***	0.030**
	(0.012)	(0.012)	(0.012)
High	0.104***	0.070***	0.070***
	(0.014)	(0.013)	(0.013)
Missing information	−0.007	−0.006	−0.005
	(0.015)	(0.015)	(0.014)
Size of town (log)	0.004	0.003	0.003
	(0.003)	(0.003)	(0.003)
Financial literacy (*Big Three*)			
Interest-rate literate		0.136***	0.134***
		(0.011)	(0.011)
Inflation literate		0.191***	0.185***
		(0.011)	(0.011)
Risk-diversification literate		0.041***	0.041***
		(0.010)	(0.010)
Interview(*er*) *characteristics*			
Interviewer female			0.026*
			(0.015)
Interviewer age			0.002***
			(0.001)
Interviewer education (ref: Secondary)			
Primary			0.146
			(0.135)
Tertiary			−0.002
			(0.014)
Interviewer experienced			−0.023
			(0.015)
Interview duration			−0.001
			(0.001)
Mean DepVar	0.56	0.57	0.57
Adj. R-squared	0.05	0.12	0.12
N	17,961	17,484	17,484
Country FE	✓	✓	✓
Wave FE	✓	✓	✓

Notes: The table shows estimates from a linear probability model. The dependent variable is equal to 1 if an individual is guarantee literate, i.e., correctly answering the survey question on guarantee literacy (as detailed in Table 1.1), and 0 otherwise. Standard errors in parentheses are adjusted for clustering at the *primary-sampling-unit* and *time* level. 'ref.' indicates the omitted category. The drop in the number of observations is due to item non-response on covariates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results are similar when estimating a probit model (see Table A.5 in the Appendix). *Data Source*: OeNB Euro Survey.

1.3.3 Granting of Guarantees

For each country, Figure 1.3 shows the percentage of individuals who are currently granting a guarantee (dark gray) or an informal loan (light gray). While individuals are more likely to provide informal loans, there is also a non-negligible share of individuals granting guarantees. In North Macedonia, for instance, the share of individuals granting a guarantee is as high as seven percent. For those currently granting a guarantee or an informal loan, the figure further distinguishes between individuals who are illiterate (striped) or literate (solid) about guarantees. In some countries, the majority of individuals who currently act as guarantors are unaware of the potential legal and financial consequences of guarantees.

Figure 1.3. Granting informal loans and guarantees



Notes: The figure shows the percentage of individuals currently granting an informal loan or a guarantee to another person. A solid (striped) pattern indicates being guarantee literate (illiterate). Statistics are based on weighted data from the 2018 and 2019 waves of the OeNB Euro Survey. For *granting an informal loan*, N=17,911; for *granting a guarantee*, N=17,847.

1.4 Empirical Methodology

In this section, we address our second research question: *Does guarantee literacy reduce the probability of granting a guarantee*? We describe our model, discuss identification challenges, and explain our identification strategy.

1.4.1 Model

First, we estimate a linear probability model of the following form:

$$\mathbb{1}(Guarantor)_i = \alpha + \beta \mathbb{1}(GuaranteeLit)_i + X_i'\gamma + X_r'\delta + CountryFE + WaveFE + \epsilon_i \tag{1.1}$$

The dependent variable, $\mathbb{1}(Guarantor)_i$, is an indicator of whether individual i is currently granting a guarantee. The main variable of interest, $\mathbb{1}(GuaranteeLit)_i$, indicates whether individual i is considered guarantee literate in the sense that they know that guarantors must repay outstanding loan amounts, including interest, if the main borrower defaults. X_i' is a vector of control variables for a person's socio-demographic characteristics (such as gender, age, education, and marital status) and socio-economic characteristics (such as labor-market status, income, wealth, and personal attitudes and beliefs). X_r' is a vector of control variables at the regional level r, including proxies for economic and financial development (such as night-light intensity and bank density). All regressions include country-fixed and wave-fixed effects.

Second, to isolate the effect of guarantee literacy from other factors and to address potential endogeneity issues, we propose an instrumental-variable strategy. To estimate Equation 1.1, we use two-stage least-squares. In the first stage, we estimate the effect of regional cohort-specific average financial literacy ($RCFLit_i$) on guarantee literacy.

$$\mathbb{1}(GuaranteeLit)_i = \alpha + \beta RCFLit_i + X_i'\gamma + X_r'\delta + CountryFE + WaveFE + u_i \tag{1.2}$$

1.4.2 Identification Challenges

Estimating Equation 1.1 using ordinary least squares (OLS) likely causes our point estimates for β to be biased. Our list of control variables may well exclude factors that are correlated with guarantee literacy and that might also drive the decision to grant a guarantee. *Cognitive ability* is one example of an omitted variable in the financial literacy research (Agarwal and Mazumder 2013; Lusardi and Mitchell 2014). While it is plausible to assume that a person's cognitive ability is positively correlated with guarantee literacy, it is not clear ex-ante whether individuals with higher cognitive ability are more or less likely to act as guarantors.

Reverse causality may be another issue as individuals who have granted a guarantee might have better literacy due to their experience as a guarantor. In particular, guarantees might have been called on and, as a result, a guarantor would have been obliged to make loan repayments on behalf of the main borrower, which in turn would improve the guarantor's understanding of the potential consequences of granting a guarantee. Guarantors may also be more literate simply because they have gone through the process of granting a guarantee.

In the literature on the effect of financial literacy on financial behavior, reverse causality usually leads to an upward bias of OLS estimates. For example, higher literacy increases the propensity to be financially included, and financial inclusion increases literacy—both effects are mutually reinforcing. In our case, however, the OLS estimates are attenuated because the effect is positive in one direction while negative in the other. Better guarantee literacy lowers the propensity to grant a guarantee, i.e., the expected coefficient is negative. Experience in granting a guarantee, however, increases guarantee literacy, i.e., the expected coefficient is positive. The OLS estimate would capture the combined effect, while the true effect of guarantee literacy on behavior would be a stronger negative one.

Another concern is that the responses to our survey question on guarantee literacy are a noisy measure of a person's true guarantee literacy, which can lead to measurement error. Such measurement error could arise, for example, from respondents guessing the answer. If a respondent guesses the correct answer, we would incorrectly classify this person as guarantee literate. As both the dependent variable and the main regressor are binary, the measurement error takes the form of misclassification. A positive probability of misclassification would lead to an attenuation bias in our estimates of β (Aigner 1973). Assuming that β is negative, this would imply a positive bias.

Lusardi and Mitchell (2017) and van Rooij, Lusardi, and Alessie (2011) provide evidence that guessing is indeed prevalent in financial literacy questions. To reduce the likelihood that a respondent would guess the correct answer, we included six different response options for our survey question on guarantee literacy. The standard financial literacy questions usually offer only up to four different response options (Lusardi and Mitchell

2014). Taken together, there is still a $\frac{1}{6}$ chance that a respondent will guess the right answer. As discussed in Section 1.3, measurement error could also arise from interviewer effects. Crossley, Schmidt, Tzamourani, and Winter (2020) show that such interviewer-induced measurement error is particularly pronounced for financial literacy questions. We address concerns regarding interviewer-induced measurement error by including interviewer-level control variables in our robustness analyses.[11]

1.4.3 Estimation Strategy

To address the concerns related to endogeneity, we perform instrumental-variables estimations. Agnew, Bateman, and Thorp (2013) and van Rooij, Lusardi, and Alessie (2011) use the financial literacy of siblings and parents as instruments for an individual's financial literacy. However, one may question whether the financial literacy of parents or siblings is beyond the control of the individual. Bucher-Koenen and Lusardi (2011) and Klapper, Lusardi, and Panos (2012) use regional financial literacy as an instrument for an individual's financial literacy. These papers employ proxies for regional financial literacy, such as the share of votes for liberal parties, the number of universities, or the newspapers in circulation.

We combine these two types of instruments and introduce a new instrument to the literature: We use *cohort-specific averages of financial literacy in the region in which the respondent lives* as an instrument for individual guarantee literacy. The instrument is based on data from seven survey waves of the OeNB Euro Survey (2012–2019),[12] which includes the big three questions on financial literacy (Lusardi and Mitchell 2008). These three questions serve to calculate a financial literacy score (for each respondent) which corresponds to the number of correctly answered financial literacy questions, ranging from 0 to 3. Our instrument is calculated as the average financial literacy score for all unique combinations of *region* and *cohort*. Regions are defined in line with the EU Nomenclature of Territorial Units (NUTS) at level 3 (see Figure 1.1 for an illustration).

Cohorts are defined according to whether the individuals experienced communism in their adult lives: The first cohort consists of individuals who experienced communism (*communist cohort*), i.e., individuals aged 18 or older in 1989. The second cohort consists of individuals who were younger than 18 in 1989, or who were not yet born (*post-communist cohort*). We define cohorts in this manner for two reasons. Firstly, the banking sector was

[11]The number of interviews per interviewer is too low for fixed-effects estimation.
[12]Unfortunately, the 2017 wave does not include the big three financial literacy questions.

merely used for transactional purposes in communist regimes. Financial markets that require consumers to make informed and more complex financial decisions developed only after the transition from a planned to a market economy. For the younger cohort, the formative years fall in this period, which is not the case for the older cohort. Secondly, during the transition from a planned to a market economy, most countries experienced banking, currency, or other economic crises. It is likely that such crisis experience will also affect literacy, e.g., in the form of an improved understanding of inflation after having experienced hyperinflation.

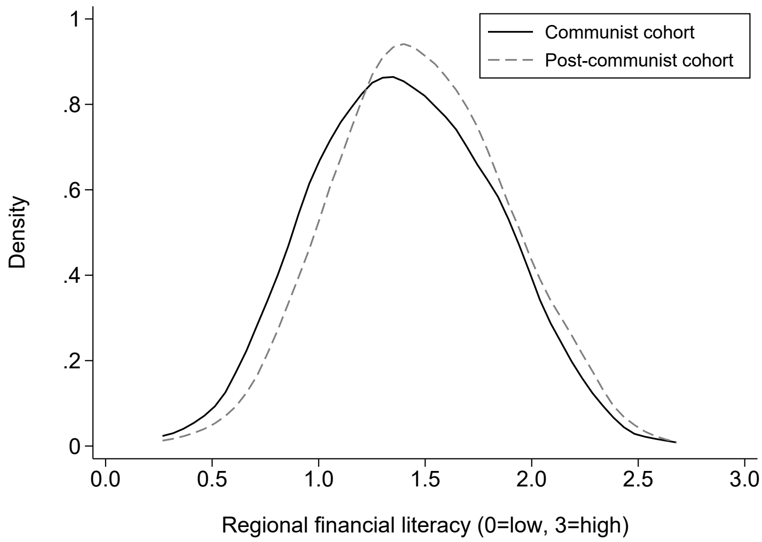
In terms of possible collinearities of our instrument and control variables, especially age, the following points should be noted: To construct our instrument of regional cohort-specific average financial literacy, we compute leave-out means (Townsend 1994), i.e., we take into account the responses of all respondents living in the respective region and belonging to the respective cohort, but exclude the financial literacy score of the respondent whose guarantee literacy we instrument. This means that our instrument varies at the individual-respondent level (and not at the regional level). It is also important to note that depending on the survey wave, some age groups may fall into different cohorts: For example, a 41-year old respondent in the 2012 wave would belong to the "communist" cohort. In contrast, a 41-year old respondent in the 2019 wave would belong to the "post-communist" cohort.

By using cohort-specific averages of financial literacy in the region in which the respondent lives as an instrument for guarantee literacy, we contend that exposure to more financially-literate individuals increases guarantee literacy.[13] Here, we draw on the empirical evidence that individuals' financial choices are influenced by those of their peers (Brown, Ivković, Smith, and Weisbenner 2008; Kaustia and Knüpfer 2012). It is further reasonable to assume that the financial literacy of an entire cohort is beyond the control of any single member of that cohort. Figure 1.4 shows the kernel densities of average regional financial literacy separately for the two cohorts. For the post-communist

[13]Individuals are influenced by their geographically distant friends when buying a house (Bailey, Cao, Kuchler, and Stroebel 2018) or a flood insurance (Hu 2022), which is strong evidence that social networks and the extent of "social connectedness" have an impact on economic activity. This would suggest that geographical exposure may only cover one aspect of exposure to financially-literate individuals. However, the countries we study have a relatively low indicator of geographically-distant social connectedness (Bailey, Cao, Kuchler, Stroebel, and Wong 2018). Moreover, in the countries that we study, internet penetration and access varies widely, from 65% of individuals with internet access at home in Romania to 84% in Poland. In the countries where internet penetration is low, the social connectedness indicator likely overstates the importance of geographically-distant social linkage, because the sample of individuals using the internet and social media is not representative of the population.

cohort (dashed line), the regional financial literacy score is slightly higher on average than for the communist cohort (solid line).

Figure 1.4. Kernel density plot of regional cohort-specific financial literacy



Notes: The figure shows kernel density estimates of the leave-out-mean regional financial literacy score (ranging between 0 and 3) for the communist cohort (solid line) and the post-communist cohort (dashed line). The expected financial literacy score would be 0.75 if the response options were chosen randomly. N=186 regions.

The identifying assumption underlying our estimation strategy is that, conditional on the observable characteristics of the individual and other controls, the instrument—regional cohort-specific financial literacy—is uncorrelated with the error term. The following two concerns may arise: First, regional financial literacy is likely to be correlated with economic prosperity or other characteristics of the region that may directly drive the prevalence of guaranteed loans. However, such regional factors are unlikely to be correlated with *cohort-specific* regional financial literacy. Second, it might be that the cohort-specific regional reference group which we use to calculate our instrument has similar social norms as the respondent, especially since the cohorts are defined according to experience with communism. Some of our control variables, particularly religion, may partially capture social norms. The fact that we are not able to fully control for social norms might weaken the validity of the exclusion restriction associated with our instrument. We address this concern by conducting a placebo analysis in which the dependent variable is an indicator of whether individuals currently lend money to family members or friends. The outcome for the main borrower (receiving a loan) and the risk of losing money for the person helping the main borrower are comparable. Of course, the two concepts differ in that not everyone has the necessary liquidity to lend money directly, which we take into account by controlling for income and wealth. But the decisions to financially support family members or friends directly (by lending money) or indirectly (by granting a guarantee) are correlated with similar social norms. Guarantee literacy, however, should only affect the granting of a guarantee. If we observe an effect of guarantee literacy on granting informal loans in the instrumental-variables estimation, this would suggest that the instrument captured omitted variables, such as social norms. If we do not find an effect of guarantee literacy on granting informal loans in the IV estimation, we are confident that the instrument does not pick up omitted variables, such as social norms.

1.5 Main Results

In this section, we study the effect of guarantee literacy on the granting of guarantees. After our baseline results, we present results from a placebo analysis, and additional robustness checks.

1.5.1 Baseline Analysis

In Table 1.5, we report the results of the OLS and IV estimation. In regression (1), we control for basic socio-demographic characteristics, in regression (2), we add control variables for income and wealth, and in regression (3), we additionally control for economic and financial development at the regional level.

OLS estimates (Panel A) show a negative and significant association between guarantee literacy and the probability of granting a guarantee. In the IV estimation, the results of the first stage (reported in Panel C) show a positive and highly significant relationship between regional cohort-specific financial literacy and an individual's guarantee literacy. The Kleibergen-Paap F-statistic varies between 153.8 and 179.1 (for the different specifications in columns 1–3),[14] indicating that the instrument of regional cohort-specific financial literacy is a strong predictor of individual guarantee literacy. The estimates

[14] According to Lee, McCrary, Moreira, and Porter (2022), two-stage-least-squares inference requires a correction if the first-stage F-statistic is below 104.7. In our analyses (see Table 1.5), the obtained F-statistics are above this threshold.

Table 1.5. Baseline and placebo analysis

	Baseline analysis: Granting guarantee			Placebo analysis: Granting informal loan		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
Guarantee literate	-0.007^{**}	-0.008^{***}	-0.008^{***}	0.010	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.007)	(0.007)	(0.007)
Mean DepVar	0.04	0.04	0.04	0.28	0.28	0.28
N	17,635	17,635	17,635	17,700	17,700	17,700
Panel B: 2SLS (second stage)						
Guarantee literate	-0.084^{***}	-0.106^{***}	-0.105^{***}	0.084	0.039	0.040
	(0.032)	(0.035)	(0.035)	(0.067)	(0.071)	(0.070)
Mean DepVar	0.04	0.04	0.04	0.28	0.28	0.28
N	17,635	17,635	17,635	17,700	17,700	17,700
Panel C: 2SLS (first stage) – Guarantee literate						
Regional cohort-specific financial literacy	0.220^{***}	0.204^{***}	0.205^{***}	0.220^{***}	0.204^{***}	0.206^{***}
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Kleibergen-Paap F-stat.	179.1	153.8	156.1	180.7	155.6	158.2
Panel D: Reduced form (OLS)						
Regional cohort-specific financial literacy	-0.018^{***}	-0.022^{***}	-0.022^{***}	0.018	0.008	0.008
	(0.007)	(0.007)	(0.007)	(0.015)	(0.015)	(0.015)
Mean DepVar	0.04	0.04	0.04	0.28	0.28	0.28
N	17,635	17,635	17,635	17,700	17,700	17,700
Country FE	✓	✓	✓	✓	✓	✓
Wave FE	✓	✓	✓	✓	✓	✓
Socio-demographic controls	✓	✓	✓	✓	✓	✓
Socio-economic controls		✓	✓		✓	✓
Regional controls			✓			✓

Notes: The table shows estimation results for granting a guarantee (columns 1 to 3), or granting an informal loan (columns 4 to 6). Socio-demographic controls include gender, age, education, marital status, employment status, religion, risk aversion, and size of town. Socio-economic controls include household income, savings, and secondary residence. Regional controls include local nightlight and local number of banks. For full results, see Appendix, Tables A.6, A.7, A.8, and A.9. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: OeNB Euro Survey.

of the reduced form (reported in Panel D) show a negative and significant association between the instrument and the probability of granting a guarantee, further supporting the validity of our instrument.

Panel B reports the results of the second stage. Across all specifications, we find that guarantee literacy has a negative effect on the granting of a guarantee: Being guarantee literate decreases the probability of granting a guarantee by 8 to 11 percentage points. This result is statistically significant and also economically relevant, as about 4 percent of the individuals in our sample are guarantors. The significance level and size of the coefficient do not change when we add controls for regional economic and banking market development (compare specifications 2 and 3), which reassures us that our instrument does not capture regional differences that drive our results. The IV estimates are larger (in absolute values) than the OLS estimates in all specifications, which we would expect based on our discussion of endogeneity concerns in Section 1.4.2.

1.5.2 Placebo Analysis

In Table 1.5, columns 4–6, we present our placebo analysis estimating the effect of guarantee literacy on granting an informal loan. As discussed in Section 1.4.3, guarantee literacy should not influence the decision to grant an informal loan to family and friends unless it is correlated with some unobservable characteristics, such as social norms. Indeed, we do not find a significant effect of guarantee literacy on lending to family and friends in any of the regression specifications. In the OLS estimation and the second stage of the IV estimation (Panels A and B), the coefficient of guarantee literacy is not significant, nor is the coefficient of regional cohort-specific financial literacy in the reduced-form estimation (Panel D).

1.5.3 Robustness

In our baseline analysis, we distinguish between individuals who currently grant a guarantee and those who do not. The group of individuals who do *not* currently grant a guarantee is likely to be heterogeneous in terms of their experiences with guarantees. In robustness checks, we restrict the sample so that we can compare individuals who currently grant a guarantee with individuals who have not yet had any experience with guarantees. First, we drop those who currently have a loan (not necessarily secured by a guarantee). Second, we drop the individuals who currently have a loan secured by a guarantee. Third, we drop the individuals who either have a loan secured by a guarantee or have ever granted a guarantee (see Table A.10). The results are similar to our baseline findings. Taken together, these results suggest that our main result is not driven by individuals who do not currently grant a guarantee but are more literate because they have had (poor) experience with guarantees in the past and are therefore less likely to grant a guarantee.

We follow Haliassos, Jansson, and Karabulut (2020) and vary the likelihood of interaction in the construction of our instrument. For example, an individual's literacy may be influenced by others who do not belong to the same cohort, e.g., parents. We calculate general financial literacy only at the regional level (and do *not* take into account potential differences in financial literacy between cohorts) (see Table A.11, column 1).

For a few regions, our instrument is based on a relatively small number of observations (see Table A.4). We repeat our estimations dropping these regions from our analysis (column 2). Finally, we present the estimation results with standard errors clustered at the time and primary-sampling-unit level (column 3). The results are stable for the specifications in columns (1) and (3), and slightly smaller in magnitude in column (2).

As discussed, it is unlikely that regional factors are correlated with *cohort-specific* regional financial literacy. In our baseline analysis, we control for local nightlight and the local number of banks. In addition, digital access could affect both, the credit market and financial literacy. Table A.12, column (1) shows the estimation results when we control for mobile coverage. Furthermore, in column (2), we control for an index of social connectedness at the NUTS 3 level (developed by Bailey, Cao, Kuchler, Stroebel, and Wong 2018). The results are very close to our baseline results in both magnitude and significance.

Even though there is only one correct answer to our question on guarantee literacy, it could be argued that an individual who thinks that the obligations associated with the granting of a guarantee consist in the repayment of the outstanding loan amount to the bank *excluding* interest understands the contingent nature of a guarantee and can therefore be considered literate. By classifying as *guarantee literate* those individuals who either state that the obligation of a guarantor is to repay the outstanding loan *including* interest or *excluding* interest, we show that the estimates obtained with the alternative measure are similar to those obtained with the original measure (see Appendix, column 1 of Table A.13).

Another concern is that IV estimation may not correct for interviewer-induced measurement error (Crossley, Schmidt, Tzamourani, and Winter 2020). We repeat our baseline analysis and control for interviewer age (see Table A.13, column 2), which we found to be correlated with a person's guarantee literacy in our regression analysis (see Table 1.4). As an alternative, we drop respondents from our analysis who were interviewed by older interviewers (column 3). The results are unchanged.

Finally, estimating linear IV, we cannot rule out a heterogeneous treatment effect—the effect of guarantee literacy may not be the same for all adults. The linear IV estimates show the effect of guarantee literacy on the probability of granting a guarantee for those who are guarantee literate because their cohort in the region where they live has a high level of financial literacy, i.e., the local average treatment effect. Instead of estimating a linear probability model using IV, we estimate a bivariate probit model and report marginal effects (Table A.14). We find that guarantee literate individuals are 5.8 to 6.7 percentage points less likely to grant a guarantee, which could be taken as an indication that the average treatment effect is smaller than the local average treatment effect (Chiburis, Das, and Lokshin 2012).

1.6 Conclusion

In this paper, we study guarantee literacy and its effect on financial decision-making. We develop a novel survey question to capture how well individuals understand the potential consequences of granting a guarantee. We find that almost half of the individuals lack literacy about guarantees. Similar to other financial literacy measures, guarantee literacy is positively associated with education, income, and employment status. In an IV estimation using regional cohort-specific financial literacy as an instrument for individual guarantee literacy, we show that literate individuals are 11 percentage points less likely to grant a guarantee than illiterate individuals.

Guarantors will be increasingly called upon to repay loans secured by guarantees when a recession or a sudden rise in interest rates leads to a surge in loan defaults. As a result, guarantors themselves may experience financial distress and lose a large portion of their wealth, potentially facing economic and social problems. This could lead to calls to severely restrict loan guarantees in the future. Before responding to these demands, policy makers should carefully consider the costs and benefits of guarantees for society.

On the benefit side, guarantees are an effective means of promoting access to loans which can be limited for two reasons. First, the bank may require additional security due to the characteristics of the borrower or the loan. Granting a guarantee may be far less costly than using an asset as collateral in terms of transaction costs. Second, in countries in which the institutional underpinnings of the market are less developed, guarantees are an important alternative to collateralization with immovable or movable property. Our results are based on nine countries that differ significantly in their economic and financial market performance and level of development—guarantees are likely to be used for both reasons, and contribute to making financial markets more efficient.

On the cost side, guarantors are primarily affected as they bear the risks associated with the contingent liability. Our research shows that individuals who are guarantee literate are less likely to grant a guarantee; they will consider the consequences of their decisions more carefully. The aim of any policy intervention should therefore be to enable individuals to make informed decisions by enhancing guarantee literacy. There is ample evidence that financial-education programs have positive effects on financial literacy and financial decision-making (Kaiser, Lusardi, Menkhoff, and Urban 2022). Including the topic of guarantees in these programs would therefore contribute to further improving financial decision-making.

Appendix A1 Survey Questions and Survey Data

Appendix A1 reports the exact wording of the big three financial literacy questions included in the OeNB Euro Survey (Table A.1), a description of all variables (Table A.2), sample summary statistics (Table A.3), and summary statistics of our instrument, *regional cohort-specific financial literacy* (Table A.4 and Figure A.1).

Table A.1. The big three financial literacy questions in the OeNB Euro Survey

Concept	Survey question
Interest rates	Suppose you had 100 [local currency] in a savings account and the interest rate was 2% per year. Disregarding any bank fees, how much do you think you would have in the account after 5 years if you left the money to grow: more than 102, exactly 102, less than 102 [local currency]? (i) More than 102 [local currency]* (ii) Exactly 102 [local currency] (iii) Less than 102 [local currency] (iv) Do not know (v) No answer
Inflation	Suppose that the interest rate on your savings account was 4% per year and inflation was 5% per year. Again disregarding any bank fees – after 1 year, would you be able to buy more than, exactly the same as, or less than today with the money in this account? (i) More (ii) Exactly the same (iii) Less* (iv) Do not know (v) No answer
Risk diversification	When an investor spreads his money among different assets, does the risk of losing money (i) Increase (ii) Decrease* (iii) Stay the same (iv) Do not know (v) No answer

Notes: The table shows the three standard financial literacy questions on interest rates, inflation, and risk diversification included in the OeNB Euro Survey. The correct answer is marked with an asterisk.

Table A.2. Description of variables

Label	Description
(*a*) *Respondents*	
	Dummy equal to 1 if . . .
Granting guarantee	Acting as guarantor for someone else's loan during the past 12 months prior interview.
Granting informal loan	Granting a loan to family or friends over the past 12 months prior interview.
Guarantee literate	Providing a correct answer to survey question on guarantees (see Table 1.1).
Interest-rate literate	Providing a correct answer to survey question on interest rates (see Table A.1).
Inflation literate	Providing a correct answer to survey question on inflation (see Table A.1).
Risk-diversification literate	Providing a correct answer to survey question on risk diversification (see Table A.1).
Female	Female.
Age 18–35	Aged between 18 and 35 years.
Age 36–50	Aged between 36 and 50 years.
Age 51–65	Aged between 51 and 65 years.
Age 65 or older	Aged 65 or older.
Education primary	Having primary education.
Education secondary	Having lower secondary, upper secondary, or post-secondary non-tertiary education.
Education tertiary	Having first or second stage of tertiary education.
Married	Being married or living with a partner.
Working	Being employed, self-employed, a contributing family worker, or an own account worker.
Religious	Being religious (e.g., Christian, Muslim, Jew, Buddhist, etc.).
Risk averse	Selecting response option "4" to the question "In managing your financial investments, would you say you have a preference for investments that offer (1) Very high returns, but with a high risk of losing part of the capital, (2) A good return, but also a fair degree of protection for the investment capital, (3) A fair return, with a good degree of protection for the invested capital, or (4) Low returns, with no risk of losing the invested capital.
Household income low	Net household income included in the first tercile; sample values are used to construct terciles.
Household income medium	Net household income included in the second tercile; sample values are used to construct terciles.
Household income high	Net household income included in the last tercile; sample values are used to construct terciles.
Household income info missing	Providing no answer to the household income question.
Savings	Having any of the following forms of savings: cash, bank accounts, life insurance, mutual funds, stocks, pension funds, bonds, or current account.
Secondary residence	Respondent or someone else in the household owns a secondary residence.
(*b*) *Primary sampling unit*	
Size of town (log)	Logarithm of the number of inhabitants living in the town/village in which the respondent lives.
Local nightlight (asinh)	Inverse hyperbolic sine of VIIRS nightlight within a radius of 20km around the respondent's place of residence.
Local number of banks	Number of banks within a radius of 20km around the respondent's place of residence.
(*c*) *Interviewers*	
Interviewer female	=1 if interviewer is female, and 0 otherwise.
Interviewer age	Age of the interviewer; integer value ranging from 18 upwards.
Interviewer education primary	=1 if the interviewer has primary education, and 0 otherwise.
Interviewer education secondary	=1 if the interviewer has lower secondary, upper secondary, or post-secondary non-tertiary education, and 0 otherwise.
Interviewer education tertiary	=1 if the interviewer has first or second stage of tertiary education, and 0 otherwise.
Interviewer experienced	=1 if the interviewer has conducted interviews on behalf of the OeNB Euro Survey during the two survey waves prior the current interview.
Interview duration	Duration of the total interview in minutes.

Notes: The table shows a detailed description of all variables used in our analyses.

Table A.3. Summary statistics

	Min	Max	*N*	BA	BG	CZ	HR	HU	MK	PL	RO	RS	Total
(*a*) *Respondents*													
Granting guarantee	0	1	17,847	0.06	0.03	0.04	0.03	0.03	0.07	0.03	0.02	0.06	0.04
				(0.23)	(0.17)	(0.2)	(0.16)	(0.17)	(0.25)	(0.18)	(0.15)	(0.24)	(0.20)
Granting informal loan	0	1	17,911	0.27	0.32	0.11	0.45	0.1	0.37	0.31	0.16	0.4	0.28
				(0.44)	(0.47)	(0.32)	(0.5)	(0.31)	(0.48)	(0.46)	(0.37)	(0.49)	(0.45)
Guarantee literate	0	1	17,985	0.44	0.56	0.61	0.71	0.65	0.49	0.53	0.57	0.5	0.56
				(0.5)	(0.5)	(0.49)	(0.45)	(0.48)	(0.5)	(0.5)	(0.5)	(0.5)	(0.50)
Interest-rate literate	0	1	17,955	0.38	0.5	0.65	0.73	0.5	0.54	0.6	0.37	0.68	0.55
				(0.49)	(0.5)	(0.48)	(0.44)	(0.5)	(0.5)	(0.49)	(0.48)	(0.47)	(0.50)
Inflation literate	0	1	17,838	0.4	0.77	0.68	0.63	0.61	0.46	0.5	0.56	0.63	0.58
				(0.49)	(0.42)	(0.47)	(0.48)	(0.49)	(0.5)	(0.5)	(0.5)	(0.48)	(0.49)
Risk-diversification literate	0	1	17,948	0.35	0.29	0.62	0.42	0.45	0.29	0.53	0.29	0.38	0.40
				(0.48)	(0.45)	(0.48)	(0.49)	(0.5)	(0.45)	(0.5)	(0.45)	(0.49)	(0.49)
Female	0	1	18,189	0.51	0.55	0.5	0.57	0.57	0.58	0.51	0.53	0.51	0.54
				(0.5)	(0.5)	(0.5)	(0.49)	(0.5)	(0.49)	(0.5)	(0.5)	(0.5)	(0.50)
Age 18–35	0	1	18,182	0.29	0.2	0.27	0.33	0.22	0.28	0.34	0.29	0.28	0.28
				(0.45)	(0.4)	(0.44)	(0.47)	(0.42)	(0.45)	(0.47)	(0.45)	(0.45)	(0.45)
Age 36–50	0	1	18,182	0.26	0.32	0.32	0.31	0.36	0.27	0.25	0.3	0.33	0.30
				(0.44)	(0.47)	(0.47)	(0.46)	(0.48)	(0.44)	(0.44)	(0.46)	(0.47)	(0.46)
Age 51–65	0	1	18,182	0.3	0.33	0.24	0.27	0.3	0.27	0.23	0.25	0.28	0.28
				(0.46)	(0.47)	(0.43)	(0.45)	(0.46)	(0.45)	(0.42)	(0.44)	(0.45)	(0.45)
Age 65 or older	0	1	18,182	0.15	0.15	0.17	0.09	0.12	0.18	0.17	0.16	0.1	0.14
				(0.36)	(0.36)	(0.37)	(0.29)	(0.33)	(0.38)	(0.38)	(0.36)	(0.3)	(0.35)
Education primary	0	1	18,170	0.2	0.1	0.06	0.08	0.11	0.23	0.23	0.02	0.18	0.13
				(0.4)	(0.29)	(0.24)	(0.27)	(0.32)	(0.42)	(0.42)	(0.15)	(0.38)	(0.34)
Education secondary	0	1	18,170	0.68	0.66	0.81	0.73	0.74	0.57	0.6	0.79	0.57	0.68
				(0.47)	(0.47)	(0.39)	(0.44)	(0.44)	(0.49)	(0.49)	(0.4)	(0.5)	(0.46)
Education tertiary	0	1	18,170	0.12	0.24	0.13	0.19	0.14	0.2	0.17	0.18	0.25	0.18
				(0.32)	(0.43)	(0.34)	(0.39)	(0.35)	(0.4)	(0.38)	(0.39)	(0.43)	(0.38)
Married	0	1	18,189	0.58	0.7	0.69	0.59	0.65	0.68	0.66	0.66	0.63	0.65
				(0.49)	(0.46)	(0.46)	(0.49)	(0.48)	(0.47)	(0.48)	(0.47)	(0.48)	(0.48)
Working	0	1	18,189	0.39	0.64	0.7	0.61	0.74	0.42	0.57	0.57	0.61	0.58
				(0.49)	(0.48)	(0.46)	(0.49)	(0.44)	(0.49)	(0.49)	(0.5)	(0.49)	(0.49)
Religious	0	1	18,189	0.99	0.93	0.3	0.89	0.77	0.99	0.86	0.98	0.99	0.86
				(0.11)	(0.26)	(0.46)	(0.32)	(0.42)	(0.09)	(0.34)	(0.13)	(0.12)	(0.35)
Risk averse	0	1	18,189	0.24	0.26	0.22	0.31	0.36	0.34	0.26	0.27	0.13	0.27
				(0.43)	(0.44)	(0.42)	(0.46)	(0.48)	(0.47)	(0.44)	(0.45)	(0.34)	(0.44)
Household income low	0	1	18,189	0.2	0.19	0.31	0.29	0.21	0.3	0.27	0.25	0.21	0.25
				(0.4)	(0.4)	(0.46)	(0.45)	(0.41)	(0.46)	(0.44)	(0.43)	(0.41)	(0.43)
Household income medium	0	1	18,189	0.2	0.23	0.33	0.29	0.24	0.28	0.26	0.25	0.24	0.26
				(0.4)	(0.42)	(0.47)	(0.45)	(0.43)	(0.45)	(0.44)	(0.43)	(0.42)	(0.44)
Household income high	0	1	18,189	0.19	0.22	0.32	0.31	0.21	0.24	0.24	0.25	0.27	0.25
				(0.39)	(0.41)	(0.47)	(0.46)	(0.41)	(0.43)	(0.43)	(0.43)	(0.44)	(0.43)
Household income info missing	0	1	18,189	0.42	0.36	0.04	0.12	0.34	0.18	0.23	0.26	0.28	0.25
				(0.49)	(0.48)	(0.19)	(0.32)	(0.47)	(0.39)	(0.42)	(0.44)	(0.45)	(0.43)
Savings	0	1	18,189	0.21	0.35	0.81	0.52	0.4	0.37	0.48	0.28	0.26	0.41
				(0.41)	(0.48)	(0.39)	(0.5)	(0.49)	(0.48)	(0.5)	(0.45)	(0.44)	(0.49)
Secondary residence	0	1	18,189	0.08	0.12	0.05	0.09	0.04	0.1	0.11	0.05	0.14	0.09
				(0.27)	(0.33)	(0.23)	(0.29)	(0.19)	(0.29)	(0.32)	(0.22)	(0.34)	(0.28)
(*b*) *Primary sampling unit*													
Size of town (log)	4	14	2,593	8.53	10.05	9.79	9.19	10.21	9.90	9.69	10.09	10.07	9.67
				(2.33)	(2.65)	(2.37)	(2.50)	(2.45)	(2.39)	(2.55)	(2.25)	(2.50)	(2.50)
Local nightlight (asinh)	0	4	2,593	1.02	1.13	1.79	1.64	1.45	1.16	1.74	1.29	1.71	1.43
				(0.43)	(0.78)	(0.71)	(0.84)	(0.98)	(0.70)	(0.88)	(0.82)	(0.87)	(0.84)
Local number of banks	0	31	2,593	11.75	14.99	15.20	17.25	8.24	11.60	16.12	16.29	22.78	14.79
				(5.01)	(6.31)	(4.10)	(7.38)	(2.39)	(3.30)	(6.63)	(8.34)	(7.22)	(7.18)
(*c*) *Interviewers*													
Number of interviewers both waves				138	207	101	136	214	149	188	158	153	1,444
Number of interviewers 2018 wave				70	104	51	65	100	80	94	85	78	727
Number of interviewers 2019 wave				68	103	50	71	114	69	94	73	75	717
Interviewer female	0	1	1,444	0.63	0.88	0.71	0.79	0.81	0.79	0.81	0.80	0.82	0.79
				(0.48)	(0.33)	(0.45)	(0.41)	(0.39)	(0.41)	(0.39)	(0.40)	(0.38)	(0.40)
Interviewer age	18	78	1,444	34.49	52.71	50.22	42.57	48.74	39.07	43.94	42.42	42.03	44.44
				(11.47)	(11.47)	(13.00)	(13.56)	(11.25)	(12.09)	(10.47)	(14.03)	(11.01)	(13.08)
Interviewer education primary	0	1	1,444	0	0	0	0	0	0	0	0	0.01	0.00
				0	0	0	0	0	0	0	0	(0.11)	(0.04)
Interviewer education secondary	0	1	1,444	0.65	0.40	0.79	0.85	0.85	0.46	0.73	0.41	0.39	0.61
				(0.48)	(0.49)	(0.41)	(0.36)	(0.36)	(0.50)	(0.44)	(0.49)	(0.49)	(0.49)
Interviewer education tertiary	0	1	1,444	0.35	0.60	0.21	0.15	0.15	0.54	0.27	0.59	0.59	0.39
				(0.48)	(0.49)	(0.41)	(0.36)	(0.36)	(0.50)	(0.44)	(0.49)	(0.49)	(0.49)
Interviewer experienced	0	1	1,444	0.46	0.28	0.12	0.48	0.11	0.50	0.11	0.55	0.11	0.34
				(0.50)	(0.45)	(0.33)	(0.50)	(0.31)	(0.50)	(0.31)	(0.50)	(0.32)	(0.47)
Interview duration	10	152	18,189	25.24	23.11	33.03	25.19	29.7	28.73	34.67	19.99	22.9	26.94
				(7.93)	(8.28)	(9.24)	(7.39)	(8.23)	(10.48)	(8.43)	(5.88)	(10.23)	(9.73)

Notes: The table shows the (unweighted) sample means and standard deviations (in parentheses) of the respective variables. *Total* refers to the entire sample of observations without adjusting for country size. Panel (a) shows summary statistics for variables measured at the respondent level (varying number of observations due to item-nonresponse), panel (b) shows summary statistics for variables measured at level of primary sampling units, and panel (c) shows summary statistics for interviewers. Countries under study: Bosnia and Herzegovina (BA), Bulgaria (BG), Czech Republic (CZ), Croatia (HR), Hungary (HU), North Macedonia (MK), Poland (PL), Romania (RO), and Serbia (RS). *Data Source*: OeNB Euro Survey.

Table A.4. Summary statistics of regional cohort-specific financial literacy

		Financial literacy score	
NUTS 3 Region	*N*	Communist cohort	Post-communist cohort
Bulgaria			
BG311	50	1.65	1.43
BG312	255	1.37	1.43
BG313	116	1.50	1.45
BG314	375	1.68	1.43
BG315	130	1.75	1.66
BG321	220	1.59	1.79
BG322	43	1.96	1.48
BG323	281	1.47	1.57
BG324	210	1.59	1.94
BG325	55	1.51	1.59
BG331	541	1.87	1.96
BG332	190	1.72	1.88
BG333	86	0.94	1.24
BG334	42	1.69	1.19
BG341	446	1.75	1.76
BG342	85	1.20	1.30
BG343	181	1.63	1.78
BG344	491	1.86	1.95
BG411	1,242	1.62	1.52
BG412	158	1.47	1.47
BG413	374	1.62	1.61
BG414	44	1.96	1.92
BG415	224	2.01	2.07
BG421	773	1.46	1.53
BG422	143	1.34	1.36
BG423	201	2.06	2.21
BG424	59	1.15	1.02
BG425	91	1.58	1.60
Bosnia and Herzegovina			
BH011	1,126	1.28	1.41
BH012	351	1.29	1.42
BH020	456	1.11	1.26
BH021	787	1.12	1.16
BH022	925	1.02	1.14
BH023	775	0.98	1.13
BH024	458	1.01	1.08
BH025	130	1.13	1.08
BH026	152	1.05	1.24
BH027	509	0.84	1.02
BH028	105	0.60	0.86
BH029	110	0.94	1.02
BH031	196	1.31	1.56
BH041	195	0.98	0.88
BH042	321	1.19	1.40
BH043	111	0.64	0.64
BH044	122	1.14	1.07
BH045	258	1.25	1.33
Czech Republic			
CZ010	842	2.15	2.00
CZ020	895	2.04	2.00
CZ031	521	1.32	1.69
CZ032	315	2.26	2.32

Continued on next page

Table A.4 (*Continued*)

NUTS 3 Region	*N*	Financial literacy score: Communist cohort	Financial literacy score: Post-communist cohort
CZ041	138	2.07	2.09
CZ042	630	1.75	2.12
CZ051	175	1.56	1.45
CZ052	482	1.93	2.20
CZ053	377	2.00	1.97
CZ063	515	1.64	1.57
CZ064	645	1.85	2.11
CZ071	129	2.02	2.09
CZ072	708	1.70	1.94
CZ080	852	1.75	1.98
Croatia			
HR031	432	1.71	1.56
HR032	88	2.03	2.04
HR033	131	1.91	1.89
HR034	289	1.73	1.78
HR035	695	1.20	1.40
HR036	422	1.64	1.65
HR037	not covered in 2018 and 2019 survey waves		
HR041	1,327	1.74	1.75
HR042	468	1.58	1.77
HR043	not covered in 2018 and 2019 survey waves		
HR044	563	1.54	1.56
HR045	172	1.36	1.57
HR046	162	2.23	2.01
HR047	154	1.60	1.80
HR048	not covered in 2018 and 2019 survey waves		
HR049	141	1.58	1.77
HR04A	406	1.87	1.86
HR04B	579	1.45	1.62
HR04C	179	1.56	1.48
HR04D	161	1.51	1.45
HR04E	434	1.48	1.57
Hungary			
HU110	1,259	1.66	1.60
HU120	824	1.64	1.65
HU211	294	1.81	2.03
HU212	214	1.80	1.67
HU213	258	1.60	1.43
HU221	314	1.52	1.45
HU222	179	1.13	1.14
HU223	195	1.77	1.64
HU231	286	1.81	1.92
HU232	243	1.58	1.56
HU233	170	2.04	1.94
HU311	460	1.88	1.95
HU312	227	2.00	1.67
HU313	132	1.96	1.96
HU321	380	1.45	1.53
HU322	265	1.32	1.07
HU323	379	1.66	1.83
HU331	360	1.60	1.63
HU332	261	1.83	1.91
HU333	307	1.92	2.15

Continued on next page

Table A.4 (*Continued*)

NUTS 3 Region	*N*	Financial literacy score: Communist cohort	Financial literacy score: Post-communist cohort
North Macedonia			
MK001	567	1.64	1.74
MK002	748	1.19	1.44
MK003	700	1.30	1.25
MK004	519	1.27	1.32
MK005	911	1.43	1.46
MK006	1,062	1.00	1.07
MK007	597	1.23	1.18
MK008	2,016	1.28	1.30
Poland			
PL21	555	1.55	1.62
PL22	869	1.13	1.47
PL41	592	1.80	1.86
PL42	342	0.83	0.90
PL43	122	1.60	1.55
PL51	604	1.36	1.39
PL52	180	1.21	1.31
PL61	437	1.34	1.36
PL62	169	1.23	1.26
PL63	422	1.32	1.43
PL71	570	1.10	1.23
PL72	230	1.25	1.05
PL81	393	1.44	1.63
PL82	411	1.29	1.45
PL84	271	1.81	1.68
PL91	491	1.63	1.56
PL92	361	1.28	1.40
Romania			
RO111	242	0.76	1.14
RO112	45	0.60	0.38
RO113	247	0.88	1.14
RO114	166	1.10	0.96
RO115	133	1.38	1.55
RO116	103	1.62	1.90
RO121	141	1.08	1.13
RO122	248	0.97	1.03
RO123	83	0.84	0.81
RO124	74	0.54	0.93
RO125	163	1.33	1.60
RO126	155	1.10	1.18
RO211	183	0.89	1.11
RO212	202	1.03	1.25
RO213	250	1.22	1.62
RO214	144	1.03	1.27
RO215	219	1.32	1.35
RO216	166	0.47	0.67
RO221	144	1.01	1.14
RO222	204	1.22	1.34
RO223	253	1.00	0.93
RO224	234	1.00	1.18
RO225	48	0.94	1.27
RO226	not covered in 2018 and 2019 survey waves		
RO311	244	1.31	1.18
RO312	not covered in 2018 and 2019 survey waves		
RO313	198	0.76	0.95

Continued on next page

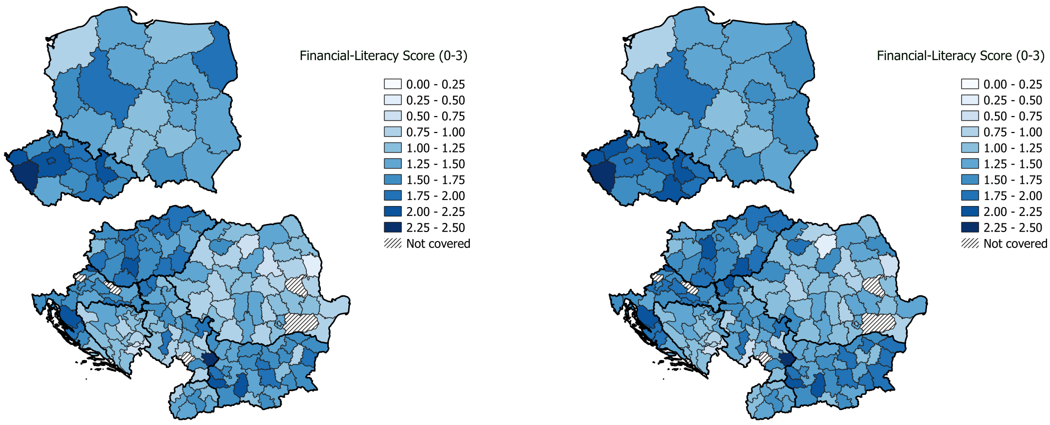
Table A.4 (*Continued*)

		Financial literacy score	
NUTS 3 Region	*N*	Communist cohort	Post-communist cohort
RO314	183	1.20	1.14
RO315	not covered in 2018 and 2019 survey waves		
RO316	298	1.04	1.40
RO317	126	0.81	1.06
RO321	665	1.12	1.23
RO322	69	1.26	1.37
RO411	239	0.90	1.11
RO412	142	0.98	1.19
RO413	108	1.09	1.56
RO414	166	1.30	1.36
RO415	104	1.26	1.33
RO421	187	1.31	1.51
RO422	136	0.99	1.15
RO423	176	0.96	0.79
RO424	222	1.03	1.19
Serbia			
RS110	1,679	1.51	1.52
RS121	212	1.84	1.85
RS122	284	1.06	1.28
RS123	543	1.32	1.65
RS124	134	1.15	1.61
RS125	166	1.99	2.22
RS126	163	1.26	1.24
RS127	425	1.51	1.68
RS211	231	1.24	1.34
RS212	141	1.13	1.42
RS213	356	1.23	1.00
RS214	297	1.95	1.92
RS215	274	1.42	1.37
RS216	294	1.00	1.06
RS217	291	0.77	0.75
RS218	238	1.00	1.21
RS221	167	1.76	1.87
RS222	149	1.56	1.55
RS223	183	0.93	0.97
RS224	196	1.59	1.66
RS225	273	1.10	1.42
RS226	140	2.48	2.44
RS227	270	1.27	1.40
RS228	202	0.91	1.15
RS229	not covered in 2018 and 2019 survey waves		

Notes: The table shows the (unweighted) sample means of the cohort-specific financial literacy score at the *NUTS 3* regional level, and the underlying number of observations. For Poland, the table shows the cohort-specific financial literacy score on the *NUTS 2* regional level (due to small numbers of observations on the NUTS 3 regional level). For the calculation of the cohort-specific financial literacy scores, we use data from seven survey waves of the OeNB Euro Survey (survey waves 2012, 2013, 2014, 2015, 2016, 2018, and 2019); for each NUTS region and each cohort, the average number of correctly-answered financial literacy questions (ranging between 0 and 3) – excluding the respondent her/himself – is calculated. The expected financial literacy score would be 0.75 if response options were chosen randomly. *Communist cohort* refers to the group of individuals aged 18 or older in 1989; *post-communist cohort* refers to the group of individuals aged 17 or younger, or not yet born in 1989.

Figure A.1. Mapping of regional cohort-specific financial literacy

(a) Communist cohort (b) Post-communist cohort



Notes: The figure maps the summary statistics of regional cohort-specific financial literacy from Table A.4.

Appendix A2 Additional Analyses of Guarantee Literacy

Appendix A2 includes additional analyses of individual guarantee literacy. In Table A.5, we report results from a probit model with "guarantee literate" as the dependent variable. Results are qualitatively similar to estimating a linear probability model (see Table 1.4).

Table A.5. Multivariate analysis of guarantee literacy: Probit model

Dependent variable	Guarantee literate		
	(1)	(2)	(3)
Socio-demographic characteristics			
Female	−0.005	0.006	0.005
	(0.007)	(0.007)	(0.007)
Age (ref: 36–50)			
18–35	−0.063***	−0.056***	−0.051***
	(0.011)	(0.010)	(0.010)
51–65	0.026***	0.015	0.015
	(0.010)	(0.010)	(0.010)
65 or older	0.015	0.007	0.004
	(0.015)	(0.015)	(0.015)
Education (ref: Secondary)			
Primary	−0.104***	−0.069***	−0.071***
	(0.014)	(0.014)	(0.014)
Tertiary	0.049***	0.021**	0.023**
	(0.011)	(0.011)	(0.011)
Married	0.017*	0.017**	0.018**
	(0.009)	(0.009)	(0.009)
Working	0.038***	0.042***	0.045***
	(0.010)	(0.010)	(0.010)
Household income (ref: Low)			
Medium	0.045***	0.031***	0.029**
	(0.012)	(0.012)	(0.011)
High	0.104***	0.069***	0.069***
	(0.014)	(0.013)	(0.013)
Missing information	−0.006	−0.006	−0.004
	(0.015)	(0.014)	(0.014)
Size of town (log)	0.004	0.003	0.003
	(0.003)	(0.003)	(0.003)
Financial literacy (*Big Three*)			
Interest-rate literate		0.132***	0.129***
		(0.010)	(0.010)
Inflation literate		0.181***	0.176***
		(0.009)	(0.009)
Risk-diversification literate		0.041***	0.041***
		(0.010)	(0.010)
Interview(*er*) *characteristics*			
Interviewer female			0.025
			(0.015)
Interviewer age			0.002***
			(0.001)
Interviewer education (ref: Secondary)			
Primary			0.144
			(0.134)
Tertiary			−0.001
			(0.014)
Interviewer experienced			−0.022
			(0.014)
Interview duration			−0.001
			(0.001)
Mean DepVar	0.56	0.57	0.57
Log-L	−11,817	−10,889	−10,848
N	17,961	17,484	17,484
Country FE	✓	✓	✓
Wave FE	✓	✓	✓

Notes: The table shows marginal effects from a probit model. The dependent variable is equal to 1 if an individual is guarantee literate, i.e., correctly answering the survey question on guarantee literacy (as detailed in Table 1.1), and 0 otherwise. Standard errors in parentheses are adjusted for clustering at the *primary-sampling-unit* and *time* level. 'ref.' indicates the omitted category. The drop in the number of observations is due to item non-response on covariates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: OeNB Euro Survey.

Appendix A3 Guarantee Literacy and Granting of Guarantees: Full Results

Appendix A3 includes detailed regression results underlying the OLS and IV estimation in Section 1.5. Detailed regression results of the OLS estimation are provided in Table A.6. Detailed regression results of the IV estimation are provided in Table A.7 (second-stage full results) and Table A.8 (first-stage full results). Detailed reduced-form regression results are provided in Table A.9. The results of the robustness checks are provided in Tables A.10, A.11, A.12, A.13, and A.14.

Table A.6. Baseline and placebo analysis: Full OLS regression results

	Baseline analysis: Granting guarantee			Placebo analysis: Granting informal loan		
	(1)	(2)	(3)	(4)	(5)	(6)
Guarantee literate	−0.007**	−0.008***	−0.008***	0.010	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.007)	(0.007)	(0.007)
Female	−0.002	−0.002	−0.002	−0.025***	−0.023***	−0.023***
	(0.003)	(0.003)	(0.003)	(0.007)	(0.006)	(0.006)
Age (ref: 36–50)						
18–35	−0.012***	−0.013***	−0.012***	0.015	0.016*	0.016*
	(0.004)	(0.004)	(0.004)	(0.009)	(0.009)	(0.009)
51–65	0.013***	0.013***	0.013***	0.012	0.009	0.009
	(0.004)	(0.004)	(0.004)	(0.009)	(0.009)	(0.009)
65 or older	0.001	0.001	0.001	−0.004	−0.018	−0.017
	(0.005)	(0.005)	(0.005)	(0.012)	(0.012)	(0.012)
Education (ref: Secondary)						
Primary	−0.003	−0.001	−0.000	−0.040***	−0.024**	−0.023**
	(0.004)	(0.004)	(0.005)	(0.010)	(0.010)	(0.010)
Tertiary	0.018***	0.011**	0.012**	0.068***	0.038***	0.038***
	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)	(0.009)
Married	0.002	−0.001	−0.001	0.011	−0.001	−0.001
	(0.003)	(0.003)	(0.003)	(0.007)	(0.007)	(0.007)
Working	0.025***	0.022***	0.022***	0.063***	0.046***	0.046***
	(0.004)	(0.004)	(0.004)	(0.008)	(0.008)	(0.008)
Religious	0.014***	0.014***	0.014***	0.021**	0.023**	0.023**
	(0.005)	(0.005)	(0.005)	(0.011)	(0.010)	(0.010)
Risk averse	−0.016***	−0.015***	−0.015***	−0.053***	−0.045***	−0.046***
	(0.003)	(0.003)	(0.003)	(0.007)	(0.007)	(0.007)
Size of town (log)	0.000	0.000	0.001	−0.000	−0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Household income (ref: Low)						
Medium		0.005	0.005		−0.004	−0.004
		(0.004)	(0.004)		(0.009)	(0.009)
High		0.011**	0.011**		0.018*	0.020*
		(0.005)	(0.005)		(0.010)	(0.010)
Missing information		−0.009**	−0.009**		−0.023**	−0.022**
		(0.004)	(0.004)		(0.010)	(0.010)
Savings		0.015***	0.015***		0.155***	0.155***
		(0.003)	(0.003)		(0.007)	(0.007)
Secondary residence		0.042***	0.042***		0.087***	0.087***
		(0.007)	(0.007)		(0.013)	(0.013)
Local nightlight (asinh)			−0.000			−0.009
			(0.003)			(0.007)
Local number of banks			−0.000			−0.001
			(0.000)			(0.001)
Mean DepVar	0.04	0.04	0.04	0.28	0.28	0.28
Adj. R-squared	0.01	0.02	0.02	0.09	0.12	0.12
N	17,635	17,635	17,635	17,700	17,700	17,700
Country FE	✓	✓	✓	✓	✓	✓
Wave FE	✓	✓	✓	✓	✓	✓

Notes: The table shows detailed regression estimation results underlying Table 1.5, Panel A, columns 1–3 and 4–6. ‘ref.’ indicates the omitted category. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: OeNB Euro Survey.

Table A.7. Baseline and placebo analysis: Full second-stage regression results

	Baseline analysis: Granting guarantee			Placebo analysis: Granting informal loan		
	(1)	(2)	(3)	(4)	(5)	(6)
Guarantee literate	−0.084***	−0.106***	−0.105***	0.084	0.039	0.040
	(0.032)	(0.035)	(0.035)	(0.067)	(0.071)	(0.070)
Female	−0.003	−0.002	−0.002	−0.025***	−0.023***	−0.023***
	(0.003)	(0.003)	(0.003)	(0.007)	(0.006)	(0.006)
Age (ref: 36–50)						
18–35	−0.017***	−0.018***	−0.018***	0.018*	0.018*	0.019*
	(0.005)	(0.005)	(0.005)	(0.010)	(0.010)	(0.010)
51–65	0.015***	0.015***	0.015***	0.010	0.008	0.008
	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)	(0.009)
65 or older	0.001	−0.000	−0.000	−0.004	−0.018	−0.017
	(0.005)	(0.005)	(0.005)	(0.012)	(0.012)	(0.012)
Education (ref: Secondary)						
Primary	−0.012**	−0.011*	−0.010*	−0.032**	−0.020	−0.019
	(0.006)	(0.006)	(0.006)	(0.013)	(0.012)	(0.012)
Tertiary	0.023***	0.016***	0.016***	0.064***	0.036***	0.036***
	(0.005)	(0.005)	(0.005)	(0.010)	(0.010)	(0.010)
Married	0.005	0.001	0.001	0.009	−0.002	−0.002
	(0.003)	(0.003)	(0.003)	(0.007)	(0.007)	(0.007)
Working	0.030***	0.025***	0.025***	0.059***	0.045***	0.044***
	(0.004)	(0.004)	(0.004)	(0.009)	(0.009)	(0.009)
Religious	0.014***	0.014***	0.014***	0.021**	0.023**	0.023**
	(0.005)	(0.005)	(0.005)	(0.011)	(0.010)	(0.010)
Risk averse	−0.009**	−0.006	−0.006	−0.060***	−0.049***	−0.050***
	(0.004)	(0.004)	(0.004)	(0.010)	(0.010)	(0.010)
Size of town (log)	0.001	0.001	0.001	−0.001	−0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Household income (ref: Low)						
Medium		0.009*	0.009*		−0.005	−0.005
		(0.005)	(0.005)		(0.010)	(0.010)
High		0.020***	0.021***		0.014	0.016
		(0.006)	(0.006)		(0.012)	(0.012)
Missing information		−0.009**	−0.009**		−0.023**	−0.022**
		(0.004)	(0.004)		(0.010)	(0.010)
Savings		0.020***	0.020***		0.153***	0.153***
		(0.004)	(0.004)		(0.008)	(0.008)
Secondary residence		0.036***	0.036***		0.089***	0.090***
		(0.008)	(0.008)		(0.014)	(0.014)
Local nightlight (asinh)			−0.003			−0.007
			(0.004)			(0.007)
Local number of banks			−0.000			−0.001
			(0.000)			(0.001)
Mean DepVar	0.04	0.04	0.04	0.28	0.28	0.28
Kl.-Paap F-stat. first stage	179.1	153.8	156.1	180.7	155.6	158.2
N	17,635	17,635	17,635	17,700	17,700	17,700
Country FE	✓	✓	✓	✓	✓	✓
Wave FE	✓	✓	✓	✓	✓	✓

Notes: The table shows detailed second-stage regression estimation results underlying Table 1.5, Panel B, columns 1–3 and 4–6. 'ref.' indicates the omitted category. First-stage-regression results are shown in Table A.8. Robust standard errors in parentheses. 'ref.' indicates the omitted category. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: OeNB Euro Survey.

Table A.8. Baseline and placebo analysis: Full first-stage regression results

Dependent variable	Guarantee literate					
	Baseline analysis			Placebo analysis		
	ad (1)	ad (2)	ad (3)	ad (4)	ad (5)	ad (6)
Regional cohort-specific financial literacy	0.220***	0.204***	0.205***	0.220***	0.204***	0.206***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Female	−0.009	−0.007	−0.007	−0.010	−0.008	−0.008
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Age (ref: 36–50)						
18–35	−0.056***	−0.060***	−0.060***	−0.056***	−0.060***	−0.060***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
51–65	0.027***	0.029***	0.029***	0.027***	0.030***	0.030***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
65 or older	−0.002	−0.002	−0.002	0.001	0.001	0.001
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Education (ref: Secondary)						
Primary	−0.112***	−0.103***	−0.102***	−0.111***	−0.102***	−0.102***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Tertiary	0.059***	0.048***	0.048***	0.060***	0.049***	0.048***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Married	0.031***	0.017**	0.017**	0.030***	0.015*	0.016*
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Working	0.054***	0.036***	0.036***	0.054***	0.036***	0.036***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Religious	0.006	0.005	0.003	0.005	0.004	0.003
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Risk averse	0.090***	0.092***	0.092***	0.093***	0.095***	0.095***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Size of town (log)	0.004**	0.003**	0.005***	0.004**	0.003**	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Household income (ref: Low)						
Medium		0.037***	0.037***		0.038***	0.038***
		(0.011)	(0.011)		(0.011)	(0.011)
High		0.088***	0.090***		0.089***	0.091***
		(0.012)	(0.012)		(0.012)	(0.012)
Missing information		−0.001	0.000		0.001	0.002
		(0.011)	(0.011)		(0.011)	(0.011)
Savings		0.054***	0.053***		0.053***	0.052***
		(0.008)	(0.008)		(0.008)	(0.008)
Secondary residence		−0.063***	−0.064***		−0.065***	−0.065***
		(0.013)	(0.013)		(0.013)	(0.013)
Local nightlight (asinh)			−0.035***			−0.036***
			(0.008)			(0.008)
Local number of banks			0.004***			0.004***
			(0.001)			(0.001)
Kl.-Paap F-stat.	179.1	153.8	156.1	180.7	155.6	158.2
N	17,635	17,635	17,635	17,700	17,700	17,700
Country FE	✓	✓	✓	✓	✓	✓
Wave FE	✓	✓	✓	✓	✓	✓

Notes: The table shows detailed first-stage regression estimation results underlying Table 1.5, Panel C, columns 1–3 and 4–6. ‘ref.’ indicates the omitted category. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: OeNB Euro Survey.

Table A.9. Baseline and placebo analysis: Full reduced form regression results

	Baseline analysis: Granting guarantee			Placebo analysis: Granting informal loan		
	(1)	(2)	(3)	(4)	(5)	(6)
Regional cohort-specific financial literacy	−0.018***	−0.022***	−0.022***	0.018	0.008	0.008
	(0.007)	(0.007)	(0.007)	(0.015)	(0.015)	(0.015)
Female	−0.002	−0.002	−0.002	−0.026***	−0.023***	−0.023***
	(0.003)	(0.003)	(0.003)	(0.007)	(0.006)	(0.006)
Age (ref: 36–50)						
18–35	−0.012***	−0.012***	−0.012***	0.014	0.016*	0.016*
	(0.004)	(0.004)	(0.004)	(0.009)	(0.009)	(0.009)
51–65	0.012***	0.012***	0.012***	0.013	0.009	0.009
	(0.004)	(0.004)	(0.004)	(0.009)	(0.009)	(0.009)
65 or older	0.001	0.000	0.000	−0.004	−0.018	−0.017
	(0.005)	(0.005)	(0.005)	(0.012)	(0.012)	(0.012)
Education (ref: Secondary)						
Primary	−0.002	0.000	0.000	−0.041***	−0.024**	−0.023**
	(0.004)	(0.005)	(0.005)	(0.010)	(0.010)	(0.010)
Tertiary	0.018***	0.011**	0.011**	0.069***	0.038***	0.038***
	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)	(0.009)
Married	0.002	−0.001	−0.001	0.012*	−0.001	−0.001
	(0.003)	(0.003)	(0.003)	(0.007)	(0.007)	(0.007)
Working	0.025***	0.022***	0.022***	0.063***	0.046***	0.046***
	(0.004)	(0.004)	(0.004)	(0.008)	(0.008)	(0.008)
Religious	0.014***	0.014***	0.014***	0.021**	0.023**	0.023**
	(0.005)	(0.005)	(0.005)	(0.011)	(0.010)	(0.010)
Risk averse	−0.016***	−0.016***	−0.016***	−0.052***	−0.045***	−0.046***
	(0.003)	(0.003)	(0.003)	(0.007)	(0.007)	(0.007)
Size of town (log)	0.001	0.000	0.001	−0.000	−0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Household income (ref: Low)						
Medium		0.005	0.005		−0.004	−0.004
		(0.004)	(0.004)		(0.009)	(0.009)
High		0.011**	0.011**		0.018*	0.019*
		(0.005)	(0.005)		(0.010)	(0.010)
Missing information		−0.009**	−0.009**		−0.023**	−0.022**
		(0.004)	(0.004)		(0.010)	(0.010)
Savings		0.015***	0.015***		0.155***	0.155***
		(0.003)	(0.003)		(0.007)	(0.007)
Secondary residence		0.043***	0.043***		0.087***	0.087***
		(0.007)	(0.007)		(0.013)	(0.013)
Local nightlight (asinh)			0.000			−0.009
			(0.003)			(0.007)
Local number of banks			−0.000			−0.001
			(0.000)			(0.001)
Mean DepVar	0.04	0.04	0.04	0.28	0.28	0.28
Adj. R-squared	0.01	0.02	0.02	0.09	0.12	0.12
N	17,635	17,635	17,635	17,700	17,700	17,700
Country FE	✓	✓	✓	✓	✓	✓
Wave FE	✓	✓	✓	✓	✓	✓

Notes: The table shows detailed regression estimation results underlying Table 1.5, Panel D, columns 1–3 and 4–6. 'ref.' indicates the omitted category. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: OeNB Euro Survey.

Table A.10. Robustness: Variation in past guarantee exposure

	(1)	(2)	(3)
Panel A: OLS			
Guarantee literate	-0.005	-0.008^{**}	-0.001
	(0.004)	(0.003)	(0.005)
Mean DepVar	0.05	0.04	0.05
N	13,471	16,790	7,354
Panel B: 2SLS (second stage)			
Guarantee literate	-0.131^{***}	-0.110^{***}	-0.149^{**}
	(0.044)	(0.035)	(0.062)
Mean DepVar	0.05	0.04	0.05
N	13,471	16,790	7,354
Panel C: 2SLS (first stage)			
Regional cohort-specific financial literacy	0.204^{***}	0.210^{***}	0.200^{***}
	(0.019)	(0.017)	(0.026)
Kleibergen-Paap F-stat.	119.0	154.9	60.8
Panel D: Reduced form (OLS)			
Regional cohort-specific financial literacy	-0.027^{***}	-0.023^{***}	-0.030^{**}
	(0.009)	(0.007)	(0.012)
Mean DepVar	0.05	0.04	0.05
N	13,471	16,790	7,354
Country FE	✓	✓	✓
Wave FE	✓	✓	✓
Socio-demographic controls	✓	✓	✓
Socio-economic controls	✓	✓	✓
Regional controls	✓	✓	✓

Notes: The table shows estimation results for granting a guarantee. Column (1) excludes individuals who are currently *not* acting as guarantor, but having a loan. Column (2) excludes individuals who are currently *not* acting as guarantor, but having a loan secured with a guarantee. Column (3) excludes individuals who are currently *not* acting as guarantor, but having a loan secured with a guarantee or having ever granted a guarantee. Socio-demographic controls include gender, age, education, marital status, employment status, religion, risk aversion, and size of town. Socio-economic controls include household income, savings, and secondary residence. Regional controls include local nightlight and local number of banks. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: OeNB Euro Survey.

Table A.11. Robustness: Instrument calculation and clustering

	(1)	(2)	(3)
Panel A: OLS			
Guarantee literate	-0.008^{***}	-0.008^{**}	-0.008^{**}
	(0.003)	(0.003)	(0.004)
Mean DepVar	0.04	0.04	0.04
N	17,635	16,112	17,635
Panel B: 2SLS (second stage)			
Guarantee literate	-0.100^{***}	-0.070^{**}	-0.105^{**}
	(0.035)	(0.032)	(0.046)
Mean DepVar	0.04	0.04	0.04
N	17,635	16,112	17,635
Panel C: 2SLS (first stage)			
Regional financial literacy	0.221^{***}		
	(0.017)		
Regional cohort-specific financial literacy		0.242^{***}	0.205^{***}
		(0.018)	(0.030)
Kleibergen-Paap F-stat.	159.8	173.8	47.4
Panel D: Reduced form (OLS)			
Regional financial literacy	-0.022^{***}		
	(0.008)		
Regional cohort-specific financial literacy		-0.017^{**}	-0.022^{**}
		(0.008)	(0.009)
Mean DepVar	0.04	0.04	0.04
N	17,635	16,112	17,635
Country FE	✓	✓	✓
Wave FE	✓	✓	✓
Socio-demographic controls	✓	✓	✓
Socio-economic controls	✓	✓	✓
Regional controls	✓	✓	✓

Notes: The table shows estimation results for granting a guarantee. In column (1), we use an alternative instrument, *regional financial literacy*. In column (2), we keep the original instrument, *regional cohort-specific financial literacy*, but exclude observations where the sample size for estimating regional cohort-specific financial literacy yields a power of less than 80%, assuming z=1.96. Robust standard errors in parentheses. In column (3), we repeat our baseline analysis and account for clustering standard errors at the *time* and *primary-sampling-unit* level. Socio-demographic controls include gender, age, education, marital status, employment status, religion, risk aversion, and size of town. Socio-economic controls include household income, savings, and secondary residence. Regional controls include local nightlight and local number of banks. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: OeNB Euro Survey.

Table A.12. Robustness: Mobile coverage and social connectedness

	(1)	(2)
Panel A: OLS		
Guarantee literate	-0.008^{***}	-0.008^{**}
	(0.003)	(0.003)
Mean DepVar	0.04	0.04
N	17,635	15,680
Panel B: 2SLS (second stage)		
Guarantee literate	-0.104^{***}	-0.097^{***}
	(0.035)	(0.035)
Mean DepVar	0.04	0.04
N	17,635	15,680
Panel C: 2SLS (first stage)		
Regional cohort-specific financial literacy	0.205^{***}	0.206^{***}
	(0.016)	(0.017)
Kleibergen-Paap F-stat.	155.4	145.3
Panel D: Reduced form (OLS)		
Regional cohort-specific financial literacy	-0.021^{***}	-0.020^{***}
	(0.007)	(0.007)
Mean DepVar	0.04	0.04
N	17,635	15,680
Country FE	✓	✓
Wave FE	✓	✓
Socio-demographic controls	✓	✓
Socio-economic controls	✓	✓
Regional controls	✓	✓

Notes: The table shows estimation results for granting a guarantee. Socio-demographic controls include gender, age, education, marital status, employment status, religion, risk aversion, and size of town. Socio-economic controls include household income, savings, and secondary residence. Regional controls include local nightlight and local number of banks, and in column (1), a proxy for local mobile coverage (indicator ranging from 0, no mobile coverage, to 1, 4G coverage since 2012, based on annual maps from 2011 to 2018 by Collins Bartholomew's Mobile Coverage Explorer), and in column (2), an index for social connectedness (based on Bailey, Cao, Kuchler, Stroebel, and Wong, 2018, *gadm1_nuts3*, maximum value of social connectedness outside the region of individuals' residence). Note that the social connectedness index is not available for Bosnia and Herzegovina. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: OeNB Euro Survey.

Table A.13. Robustness: Measurement of guarantee literacy and interviewer effects

	(1)	(2)	(3)
Panel A: OLS			
Guarantee literate redefined	-0.006^{*}		
	(0.003)		
Guarantee literate		-0.008^{**}	-0.006^{**}
		(0.003)	(0.003)
Mean DepVar	0.04	0.04	0.04
N	17,635	17,635	16,091
Panel B: 2SLS (second stage)			
Guarantee literate redefined	-0.121^{***}		
	(0.040)		
Guarantee literate		-0.112^{***}	-0.104^{***}
		(0.037)	(0.032)
Mean DepVar	0.04	0.04	0.04
N	17,635	17,635	16,091
Panel C: 2SLS (first stage)			
Regional cohort-specific financial literacy	0.179^{***}	0.192^{***}	0.224^{***}
	(0.015)	(0.016)	(0.017)
Kleibergen-Paap F-stat.	134.1	135.9	167.1
Panel D: Reduced form (OLS)			
Regional cohort-specific financial literacy	-0.022^{***}	-0.022^{***}	-0.023^{***}
	(0.007)	(0.007)	(0.007)
Mean DepVar	0.04	0.04	0.04
N	17,635	17,635	16,091
Country FE	✓	✓	✓
Wave FE	✓	✓	✓
Socio-demographic controls	✓	✓	✓
Socio-economic controls	✓	✓	✓
Regional controls	✓	✓	✓
Interviewer age		✓	

Notes: The table shows estimation results for granting a guarantee. In column (1), we use an alternative measure of guarantee literacy, equal to 1 if a respondent answers (3) or (4) in the survey question in Table 1.1, and 0 otherwise. In column (2), we repeat the baseline analysis and additionally control for interviewer age. In column (3), we winsorize interviewer age by country excluding all observations collected by interviewers whose age is above the 90th percentile of each country's interviewer age distribution. Socio-demographic controls include gender, age, education, marital status, employment status, religion, risk aversion, and size of town. Socio-economic controls include household income, savings, and secondary residence. Regional controls include local nightlight and local number of banks. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: OeNB Euro Survey.

Table A.14. Robustness: Probit and bivariate-probit models

	(1)	(2)	(3)
Panel A: Probit			
Guarantee literate	-0.007^{**}	-0.008^{***}	-0.008^{***}
	(0.003)	(0.003)	(0.003)
Mean DepVar	0.04	0.04	0.04
N	17,635	17,635	17,635
Panel B: Outcome equation			
Guarantee literate	-0.058^{**}	-0.067^{**}	-0.065^{**}
	(0.025)	(0.027)	(0.026)
Mean DepVar	0.04	0.04	0.04
N	17,635	17,635	17,635
Panel C: Selection equation – Guarantee literate			
Regional cohort-specific financial literacy	0.219^{***}	0.203^{***}	0.205^{***}
	(0.016)	(0.016)	(0.016)
Panel D: Reduced form (Probit)			
Regional cohort-specific financial literacy	-0.018^{***}	-0.021^{***}	-0.021^{***}
	(0.007)	(0.007)	(0.007)
Mean DepVar	0.04	0.04	0.04
N	17,635	17,635	17,635
Country FE	✓	✓	✓
Wave FE	✓	✓	✓
Socio-demographic controls	✓	✓	✓
Socio-economic controls		✓	✓
Regional controls			✓

Notes: The table shows marginal effects from probit models (Panel A and D) and bivariate probit models (Panel B and C). The dependent variable is equal to 1 for individuals currently granting a guarantee, and 0 otherwise. Socio-demographic controls include gender, age, education, marital status, employment status, religion, risk aversion, and size of town. Socio-economic controls include household income, savings, and secondary residence. Regional controls include local nightlight and local number of banks Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: OeNB Euro Survey.

Appendix A4 Regulation of Guarantees

Appendix A4 provides background information on the relevant legislation of guarantees in the nine central, Eastern, and Southeastern European countries under study.

Table A.15. Regulation of guarantees in central, Eastern, and Southeastern Europe

Country	Legislation	Nature of guarantee	Extent of guarantor's liability	Source
Bulgaria	Law of obligations and contracts: General Part, Title VII. Security on claims, Chapter 3. Guarantee, Articles 138–148.	**Art. 138 (1)** Under a guarantee contract the guarantor undertakes an obligation before another person's creditor to be liable for the performance of the other person's obligation. This contract must be in writing.	*Explicitly* **Art. 139** The guarantee may also be undertaken for a part of the debtor's obligations or under alleviated terms. If the guarantor has undertaken an obligation exceeding what the debtor owes or under more aggravated terms, his obligation shall be reduced to the amount of the principal obligation. **Art. 140** The guarantee shall cover all consequences of the non-performance of the principal obligation, including costs on collection of the claim. **Art. 141 (1)** The guarantor shall be liable jointly and severally with the principal debtor. *Implicitly* **Art. 143 (1)** A guarantor who has performed the obligation is entitled to claim from the debtor the principal, the interest, and the expenses he has made after notifying the debtor of the action he has brought against him. He shall also be entitled to the interest on the amounts paid from the date of payment.	Unofficial translation made available by law firm Ruskov & Colleagues.
Czech Republic	Civil Code (Občanský zákoník): Part Four – Relative Property Rights, Title 1. General Provisions on Obligations, Chapter 8. Securing and corroboration of debts, Section 2. Security of a debt – Suretyship, Articles 2018–2028.	**Art. 2018 (1)** A person who declares in relation to a creditor that he will satisfy him if the creditor's debtor fails to discharge his debt becomes the creditor's surety. A creditor who does not accept a surety may require nothing from him. **(2)** A suretyship declaration must be in writing.	*Implicitly* **Art. 2019 (1)** Suretyship is contingent on the debtor's debt being valid; suretyship may also be provided for future or conditional debts, as well as for a set of certain kind of debts incurred by the debtor at a particular time or a set of various debts arising from the same legal cause. **Art. 2020** If suretyship secures only part of a debt, the scope of the suretyship is not reduced by partial performance as long as the debt is not discharged to the extent to which it is secured by the suretyship. **Art. 2027** If the same debt is secured by several sureties, each of them is liable as a surety for the entire debt in respect of the creditor. A surety has the same rights with respect to the other sureties as a co-debtor.	Unofficial translation made available by the Ministry of Justice of the Czech Republic.

Continued on next page

Table A.15 (*Continued*)

Country	Legislation	Nature of guarantee	Extent of guarantor's liability	Source
Hungary	Act V of 2013 on the Civil Code (2013. évi V. törvény a Polgári Törvénykönyvről): Book 6. Law of Obligations, Part 3. Express Contracts, Title XXI. Guarantee Agreements, Chapter LX. Contracts of Suretyship, Articles 6:416–6:430.	**Art. 6:416 (1)** Under a contract of suretyship the surety undertakes the obligation of performance to the creditor in the event of nonperformance by the principal debtor. **(2)** Suretyship may be provided to guarantee one or more existing or future, conditional or unconditional pecuniary claims of a specific amount or the amount of which can be determined, or any other claims of monetary value. **(3)** The contract shall be executed in writing.	**Art. 6:417 (1)** The obligation of a surety shall be adjusted to the obligation for which he has promised to answer. The obligation of a surety shall not and cannot subsequently exceed the original obligation; however, it shall cover the consequences of the debtor's non-performance and shall include the collateral claims that fall due after the suretyship is undertaken.	Unofficial translation made available by the Ministry of Justice of Hungary.
Poland	Civil Code (Kodeks cywilny): Third Book. Obligations, Title XXXII. Guarantee, Articles 876–887.	**Art. 876 (1)** Under a guarantee contract, the guarantor undertakes to the creditor to fulfill a debt in the event that the debtor fails to do so. **(2)** For a guarantee declaration to be valid, it must be in writing.	**Art. 879 (1)** The scope of the guarantor's obligation is determined by the respective scope of the debtor's debt. **(2)** However, the guarantor's obligation may not be extended by a legal transaction entered into by the debtor with the creditor after the guarantee has been assumed. **Art. 881** In the absence of an agreement to the contrary, the guarantor shall be jointly and severally liable, similar to a co-debtor.	Translation made available by the *Institut für Ostrecht* München, Regensburg.

Continued on next page

Table A.15 (*Continued*)

Country	Legislation	Nature of guarantee	Extent of guarantor's liability	Source
Romania	Civil Code (Codul civil): Book V. On Obligations, Title X. Personal Security, Articles 2279–2322.	**Art. 2280** A guarantee is a contract by which one party, the guarantor, agrees to the other party, who is the creditor in another binding contractual relationship, to fulfill the debtor's obligation free of charge or for a fee if the latter fails to do so. **Art. 2282** A guarantee contract must be in writing.	*Explicitly* **Art. 2290 (1)** Unless otherwise specified, the guarantee of a principal obligation extends to all ancillary services, including the costs after the notarization of the guarantor and the costs associated with his summons to appear in court. **(2)** The guarantor shall be liable for the costs of litigation and enforcement brought by the creditor in proceedings against the principal debtor only if the creditor has notified him in advance. *Implicitly* **Art. 2306 (1)** The guarantor who has entered into the debtor's agreement may claim from the debtor what he has paid, namely principal, interest, and costs, as well as compensation for damages he has suffered as a result of the guarantee. He may also claim interest on any amount he has had to pay to the creditor, even if the principal debt had not produced interest.	Unofficial translation (special thanks to Codruta Rusu).

Continued on next page

Table A.15 (*Continued*)

Country	Legislation	Nature of guarantee	Extent of guarantor's liability	Source
Bosnia and Herzegovina, Croatia, North Macedonia, Serbia*	**B&H**: Federation of Bosnia and Herzegovina: Law on obligations (Zakon o obligacionim odnosima) Republika Srpska: Law on obligations (Zakon o obligacionim odnosima): Part One, General Part, Chapter 29. Guarantee, Articles 997–1019. **Croatia**: Civil Obligations Act (Zakon o obveznim odnosima): Part One. General Part, Title VI. Alterations in Obligations, Chapter 1. Alterations on the Side of the Subject, Section 6. Guarantee, Articles 104–126. **North Macedonia**: Law on obligations (Zakon za obligaczionite odnosi): Chapter XXXIII Guarantee, Articles 1036–1058. **Serbia**: Law on obligations (Zakon o obligacionim odnosima): Part One, General Part, Chapter 29. Guarantee, Articles 997–1019.	**Art. A** By entering into a contract of guarantee the guarantor undertakes to the creditor to perform a valid and due obligation of the debtor if the latter fails to do. **Art. B** A contract of guarantee shall be binding for the guarantor only if the guarantor has given a statement of guarantee in writing.	**Art. C (1)** A guarantor's obligation shall not exceed the obligation of the principal debtor and if agreed otherwise, it shall be reduced to the scope of the debtor's obligation. **(2)** A guarantor shall be liable for the performance of the whole obligation for which he provided guarantee unless his liability is limited to one of its parts or is subject to less strict terms and conditions. **(3)** He shall be liable for reimbursement of all necessary expenses incurred by the creditor for the purpose of collecting the debt from the principal debtor. **(4)** The guarantor shall also be liable for every increase in the obligation arising from a delay or fault on the part of the debtor, unless agreed otherwise. **(5)** He shall be liable only for contractual interest due after entering into a contract of guarantee.	Information on the legal texts was made available by the *Institut für Ostrecht* München, Regensburg. Unofficial translation made available by the Supreme Court of the Republic of Croatia.

*Note: In former Yugoslavia, the Law on Obligations (Zakon o obligacionim odnosima) – which also regulated guarantees – was adopted in 1978. The law was adopted in all the successor states after the break-up of Yugoslavia. **Art. A/B/C** refers to Art. 997/998/1002 in B&H and Serbia, to Art. 104/105/109 in Croatia, and to Art. 1036/1037/1041 in North Macedonia.

Notes: The table provides information on the relevant legislation of guarantees separately for each of the nine countries in our study. Note that translations are tentative and do *not* represent official documents of the respective countries.

Chapter 2

The Formation of Subjective House Price Expectations*

Abstract: Subjective house price expectations are an important driver of individual housing choices and market dynamics. We study the formation of subjective expectations about local house prices using novel survey data from Britain, a country with high homeownership rates and widely varying local housing dynamics. There is a substantial and heterogeneous perception gap and individuals extrapolate strongly from perceived but not from realized past price changes. In addition, expectations are predicted by wider, easily observable measures of local economic conditions, especially among individuals with low financial sophistication. Individuals residing in local housing markets where past prices are less informative or less observable rely more strongly on local economic conditions in their belief formation. Our results emphasize the role of heterogeneity in expectations formation processes.

*This chapter is based on joint work with Melanie Lührmann, Jonathan Shaw, and Joachim Winter. A version of this chapter has been published in the Rationality and Competition Discussion Paper Series (Kiesl-Reiter, Lührmann, Shaw, and Winter 2024).

2.1 Introduction

Subjective expectations about economic and financial outcomes crucially affect individual economic choices.[1] On housing markets, subjective expectations play an important role in shaping individuals' investment and debt behavior (Armona, Fuster, and Zafar 2019; Bailey, Dávila, Kuchler, and Stroebel 2019; Bailey, Cao, Kuchler, and Stroebel 2018; Bottan and Perez-Truglia 2020; Chopra, Roth, and Wohlfart 2023). They can also drive dynamics at the aggregate level, including housing booms and busts (Piazzesi and Schneider 2009; Case, Shiller, and Thompson 2012; Burnside, Eichenbaum, and Rebelo 2016; Landvoigt 2017; Kaplan, Mitman, and Violante 2020; Kindermann, Le Blanc, Piazzesi, and Schneider 2021). The literature on the formation of subjective house price expectations shows that there is substantial heterogeneity but its sources are still poorly understood (Koşar and O'Dea 2023; Kuchler, Piazzesi, and Stroebel 2023).

In this paper, we study belief formation in survey data on subjective local house price expectations and perceptions of past house price changes from Britain, a country with high home ownership and transaction rates, and profound and persistent geographical variation in house price dynamics (Agrawal and Phillips 2020; Overman and Xu 2022). We focus on two predictors of subjective house price expectations, past house price changes and local economic conditions.

Traditional prediction models for house prices, going back to Case and Shiller (1989), are based on past realized changes, establishing a natural starting point for modeling subjective house price expectations. More recently, perceptions of past price changes have been shown to matter for expectations regarding house price changes and inflation (Armona, Fuster, and Zafar 2019; Cavallo, Cruces, and Perez-Truglia 2017; Fuster, Perez-Truglia, Wiederholt, and Zafar 2022; Kuchler and Zafar 2019). Moreover, Armona, Fuster, and Zafar (2019) argue that individuals may "take into account information other than past home price growth, and we do not know their 'mental model' nor their information set." However, there is little theoretical guidance as to which variables individuals might use in their mental models. In the context of house price expectations, local macroeconomic conditions might matter as they are part of an individual's salient experiences. A growing

[1] A large literature shows that subjective expectations matter for individual decision-making. For example, subjective expectations about future equity returns and risks predict individual portfolio choices (e.g., Dominitz and Manski 2007; Hurd, van Rooij, and Winter 2011; Merkle and Weber 2014; Ameriks, Kézdi, Lee, and Shapiro 2020; Giglio, Maggiori, Stroebel, and Utkus 2021), and subjective inflation expectations predict individual consumption-savings decisions (e.g., Armantier et al. 2013; Coibion, Georgarakos, Gorodnichenko, and van Rooij 2023; Vellekoop and Wiederholt 2019).

literature shows that experiences, broadly defined, affect belief formation.[2] Our findings are consistent with these views: Perceived house price changes *and* local economic conditions matter and there are important differences between individuals.

We present four important observations in this paper. First, individuals do *not* extrapolate from realized past house price changes, but rather from their *perceptions* of past local house price changes. Second, locally experienced economic conditions also matter in the formation of subjective expectations about local house price changes. The importance of such locally experienced economic conditions in individuals' beliefs varies across subgroups, and it matters in particular for respondents who are less financially sophisticated, risk averse, and reside in local housing markets where past prices display high volatility and no short-run momentum. These results are consistent with locally experienced economic conditions being an easily observable predictor of subjective expectations, particularly in settings where past house prices may be less informative or individuals may be less informed.

Finally, our results point to large heterogeneity in subjective expectations that are driven in part by large and heterogeneous gaps between perceived and realized price changes, echoing findings in Armona, Fuster, and Zafar (2019). While we find little evidence of systematically higher or lower levels of perceptions based on observables, perception gaps are driven both by local market factors such as past house price volatility, as well as individual characteristics. They are larger for women and particularly for individuals with low financial sophistication.

The data we analyze come from a newly designed survey module on subjective expectations, conducted by the Financial Conduct Authority as part of the Financial Lives survey between August 2019 and February 2020, i.e., shortly before the onset of the COVID-19 pandemic. Our comparatively large analysis sample covers almost 2,800 individuals living in 364 local housing markets. We elicit perceptions of house price changes over the past year in their local area of residence, and their subjective expectations of one-year-ahead local house price changes using probabilistic elicitation techniques (Manski 2004). We link this survey data with the UK House Price Index, in a respondent's local area at the time of interview, and with locally experienced economic conditions to study the role of realized local house price changes and local economic

[2]For instance, Malmendier (2021b) discusses the role of individual long-run experiences in the formation of subjective inflation expectations. Bailey, Cao, Kuchler, and Stroebel (2018) find social interactions, through out-of-town friends' experiences of housing investment, to be an important influence in belief formation.

conditions in the expectation formation process. More specifically, we consider local unemployment rates at the time of interview and, as an alternative measure of local economic conditions, local deprivation scores from 2019.

Our paper contributes to the literature on the formation of subjective expectations about housing markets. There is evidence that people extrapolate from past house price changes when forming expectations about future house price changes (Case, Shiller, and Thompson 2012; Armona, Fuster, and Zafar 2019). In addition, Armona, Fuster, and Zafar (2019) point out that not only realizations but also perceptions of past (national) house price changes matter in the expectation formation process. Our results emphasize the importance of perceptions in belief formation. Estimating a reduced-form model of local house price expectations in Britain, we find that realized local house price changes over the past year do not predict subjective expectations of local house price changes over the next 12 months.[3] Rather, individuals form house price expectations by extrapolating strongly from their perceptions: A one percentage-point increase in the perceived past one-year local house price change is associated with a 0.13 percentage-point increase in individuals' house price expectations (which is 0.99 of a standard deviation).

Individuals form house-price beliefs in a manner and magnitude that echoes the well-established short-run momentum in house-price fundamentals (e.g., Case and Shiller 1989; Guren 2018). Using monthly-level data from the UK HPI, we estimate dependencies in local housing markets in Britain. For the time period between 2010 and 2019, we find an average short-run momentum of 0.156.[4] Yet, individual perceptions are biased, hence realizations deviate substantially from individuals' beliefs of past house price changes (by around 5 percentage points), creating a sizeable perception gap. While individuals extrapolate in a manner that is consistent with the average short run-momentum in house price fundamentals, they overestimate the short-run momentum due to inflated perceptions. This result rationalizes a frequently stated stylized fact about house-price beliefs; namely that individuals overestimate the short-run momentum in *realized* house prices (Glaeser and Nathanson 2017).

[3]For ease of exposition, we use "house price expectations" as a short hand for "subjective expectations about local house price changes over the next year."

[4]A closer look shows that local house price dynamics are heterogeneous: for 44.2% of the local authorities, there is a positive and significant relation between past one-year and future one-year house price growth, while the remainder of localities display negative or no momentum. Since at the time of the survey, the consequences of the COVID-19 pandemic – which only unfolded after – could not have been anticipated by survey respondents, we cannot compare the extent of extrapolation in the expectations data with the extent of autocorrelation in the realized house price data for the time after the survey was conducted.

Building on this, a series of papers looks at various definitions of "personalized" past house price changes and their role in explaining subjective house price expectations (Malmendier 2021b), taking into account personal background characteristics such as an individual's place of residence (Kuchler and Zafar 2019) or an individual's social network (Bailey, Cao, Kuchler, and Stroebel 2018; Bailey, Dávila, Kuchler, and Stroebel 2019). For instance, Kuchler and Zafar (2019) use past house price changes in an individual's place of residence as measure of personal housing experiences, and identify a positive relationship with individuals' expectations about nationwide house price changes. Bailey, Cao, Kuchler, and Stroebel (2018) and Bailey, Dávila, Kuchler, and Stroebel (2019) emphasize the role of (geographically distant) friends' housing-market experiences in shaping individuals' subjective expectations about local house price growth. In this paper, we consider whether local macroeconomic conditions matter for subjective expectations as they are part of an individual's salient experiences. We find that individuals expect lower house price growth when local unemployment rates are higher. A one standard-deviation change in local economic conditions leads to a change in subjective beliefs of around -0.17 percentage points—about one sixth the magnitude of a one standard-deviation change in perceived local house price changes. Taken together, our findings suggest that individuals use a wider set of local factors in their belief formation models, and provide new support for the rising body of evidence that personal experiences matter in the formation of subjective expectations.

Finally, we consider heterogeneity in individuals' extrapolation models depending on features of local housing markets, and individuals' financial sophistication. We find that extrapolation from local economic conditions is stronger in local housing markets where past price changes are less informative or less observable, i.e., in markets characterized by high house price volatility over the past five years or those that did not display short-run momentum in prices.[5]

Similarly, individuals with low financial sophistication additionally use easily observable local economic conditions as a heuristic in their formation of house price expectations. Our survey data allows us to distinguish between more and less financially-sophisticated individuals (using different measures such as general financial literacy about interest compounding, inflation, and risk diversification following Lusardi and Mitchell (2008), or understanding of the risk-and-return-profile of savings accounts). Both those with

[5]A closer look at dependencies in local housing markets in Britain revealed that a positive and significant relation between past one-year and future one-year house price growth was present in 44.2% of the local authorities, while the remainder of localities displayed negative or no momentum.

high and low financial sophistication extrapolate from perceived rather than realized past house price changes, but those with high sophistication rely more heavily on their perceptions.

The paper proceeds as follows. Section 2.2 outlines the survey data set and introduces our measures of local house price changes and local economic conditions. In Section 2.3, we describe the empirical framework and report reduced-form empirical evidence on perceived and realized local house price changes and local economic conditions as predictors of subjective expectations regarding local house price changes, and their heterogeneity across local markets and individuals, and we study dependence in realized price changes in the local housing markets in Great Britain. Section 2.4 concludes.

2.2 Data

In a newly designed survey module, we measure people's perceptions of recent local house price changes, and subjective expectations of future house price dynamics. Respondents were also asked about subjective expected and perceived past stock market returns.[6] It was pretested by the Financial Conduct Authority (FCA) and implemented in the 2020 wave of the *Financial Lives survey*—a comprehensive, large, and nationally representative survey of 16,000 adults aged 18 and older living in the UK.[7] The module was presented to a randomized subset of just under 4,000 participants who were interviewed between August 2019 and February 2020, i.e., fieldwork completed shortly before the onset of the COVID-19 pandemic. COVID-related online searches increased only after the end of the survey in March 2020 (see Appendix B3 for a detailed analysis). Social distancing measures were not introduced until mid March 2020. It is therefore unlikely that the beliefs and expectations of survey respondents about future house price changes were distorted by the upcoming COVID-19 pandemic. Importantly, given the potential impact of the pandemic on *actual* house price changes, we refrain from contrasting individuals' expectations with actual realizations post-survey. However, we do compare individuals' perceptions of past house price changes (elicited in the survey) with realized ones.

[6]It also elicited savings account returns as well as the relative riskiness of broad asset classes.

[7]The wording of the survey questions on perceptions and expectations of house price changes was carefully pre-tested and piloted by the Financial Conduct Authority (FCA) in cooperation with an independent survey institute. Pre-testing included cognitive testing in different local authorities at different points in time.

The Financial Lives survey includes rich information on socio-economic characteristics and attitudes, individuals' use of financial products, and their experiences in dealing with financial products and services. In addition, it elicits measures of financial sophistication. We restrict our sample to respondents residing in Great Britain, for whom information on the place of residence is available, and who provide answers to the survey questions about perceived past and expected future house price changes, resulting in a final sample of 2,799 respondents living in 364 local authorities (LAs).[8] Respondents' local authority of residence[9] is used to link the survey data with administrative data to obtain relevant local measures of realized house prices and economic conditions.

The survey data reflects the high home ownership rates in Great Britain and the prime role housing plays in households' portfolios: Almost three quarters of respondents in our sample report owning a home (either outright or with a mortgage); 42% hold other assets (e.g., stocks, bonds, or investment funds). Table B.2 in the Appendix reports summary statistics.

2.2.1 Subjective House Price Expectations

We measure subjective expectations about local house price changes using probabilistic elicitation techniques (see Dominitz and Manski 1997b; Hurd and McGarry 2002; Manski 2004; Giglio, Maggiori, Stroebel, and Utkus 2021). More specifically, we ask respondents to assign probabilities to a range of possible future house price changes, requiring them to add up to 100%.[10] We tie their beliefs to house prices in their local area, and not to their own homes. Respondents were prompted to imagine that they received an unexpected inheritance of £100,000 which they put towards buying a house in their local area, and to subjectively assess the percentage chances that—in 12 months' time—the house will have *decreased in value* by (i) 10% or more, (ii) 9.9% to 5%, or (iii) 4.9% to 0%; or *increased in value* by (iv) 0.1% to 5%, (v) 5.1% to 10%, (vi) 10.1% to 15%, or (vii) 15.1% or more.[11]

[8] The randomized sample comprises 3,662 individuals of which we drop 3 observations with a missing local authority and 860 observations who did not understand the probabilistic question format of the risk-and-return questions, and thus refused to answer. These were more likely to be female, younger, less educated and less financially literate. Our sample does not include respondents from seven local authorities (Barrow-in-Furness, City of London, Derbyshire Dales, Hertsmere, Isles of Scilly, Blaenau Gwent, and Merthyr Tydfil).

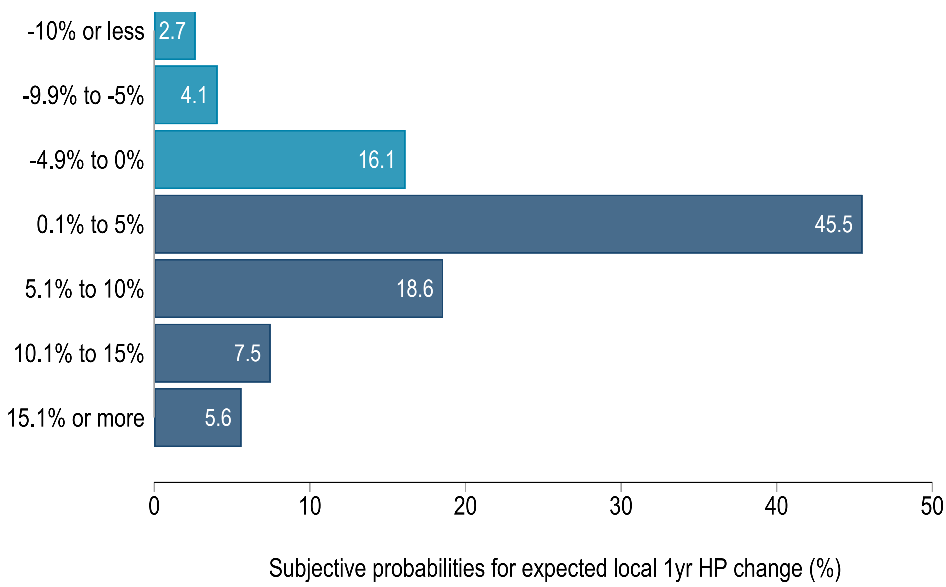
[9] We use the terms *local authority* and *local area* as synonyms.

[10] If reported probabilities did not add up to a 100%, respondents were shown a message reminding them of this requirement to ensure consistency (see also Giglio, Maggiori, Stroebel, and Utkus 2021).

[11] Brackets were chosen to reflect the distribution of past returns. For the exact wording of the survey questions, see Appendix B1.

Figure 2.1 depicts the distribution of subjective expectations about future changes in local house prices. The realization of gains in local housing markets is assessed as considerably more likely than the realization of losses. Modest positive changes (between 0.1% and 5%) were deemed most likely. The average probability of large gains (of 15.1% or more) is—at around 6%—twice as large as that of large losses (of at least −10%).

Figure 2.1. Subjective expected one-year local house price changes



Notes: The figure shows the distribution of one-year-ahead expected local house price (HP) changes, N=2,799. For detailed summary statistics, see Table B.3 in the Appendix. *Data Source*: Financial Lives 2020 survey.

Using these responses, we construct a measure of subjective expected one-year house price changes, adopting the estimation approach suggested in Hurd, van Rooij, and Winter (2011). They construct non-parametric estimates of the mean of the expected rate of return distribution for stock market investments. The model is given by

$$E(\pi) = \sum_j P(\pi \in B_j) E(\pi | \pi \in B_j) \tag{2.1}$$

where, $P(\pi \in B_j)$, is an individual's subjective probability assigned to bracket j, and, $E(\pi|\pi \in B_j)$, is the historical average of one-year rates of return conditional on the return being in bracket j.

We use the same methodology to construct a measure of the subjective expected house price change from the probabilities corresponding to the seven brackets B_j described above. We compute historical UK-wide year-on-year house price changes r_t for each month using the quality-adjusted UK House Price Index for the time between January 2002 and July 2019.[12] We then assign these historical returns to the return brackets B_j to get bracket-specific average returns $E(r|r \in B_j)$. Panel A in Table 2.1 shows the non-parametric estimates: in 12 months' time, respondents expect a mean change in local house prices of 3.71%, with a standard deviation of 4.45 percentage points. For more detail on the non-parametric estimation, see Appendix B2.

Table 2.1. Summary statistics: Expectations, perceptions, and realizations

	Mean	Std Dev	P10	P50	P90
Panel A: Expectations					
Expected 1yr HP change (%)	3.71	4.45	0.13	2.69	8.52
Panel B: Perceptions					
Perceived 1yr HP change (%)	3.79	7.52	0.00	3.00	10.00
Panel C: Realizations					
Realized 1yr HP change (%)	1.11	2.68	−2.16	1.20	4.20
Panel D: Perception gap					
Absolute perception gap (%-points)	5.04	6.73	0.63	3.51	10.24
N	2,799				

Notes: The table shows summary statistics. The absolute perception gap denotes the difference between realized and perceived past one-year (1yr) local house price (HP) changes in absolute terms. Computation of expectations is based on the non-parametric estimation approach by Hurd, van Rooij, and Winter 2011; for details, see Appendix B2. *Data Source*: Financial Lives 2020 survey, and historical values from the UK HPI.

2.2.2 Perceptions of Past Local House Price Changes

We also elicited respondents' perceptions of house price changes in their local area in the past year. Following the *Survey of Consumer Expectations Housing Survey* (fielded by the Federal Reserve Bank of New York), we use a two-step format in the survey where

[12]Figure B.4 in the Appendix shows that estimates of subjective expected house price changes are qualitatively unchanged when we instead use house price changes that are vary by geographic localities (the individual's government office region or local authority instead of the whole UK) or use alternative time horizons (1969–2019 instead of 2002–2019) in the computation of $E(r|r \in B_j)$.

respondents first report their beliefs regarding the direction of house price changes in the last 12 months and are then requested to give a point estimate (see Appendix B.1). On average, respondents' perceptions were that house prices increased by 3.79% over the previous 12 months (Panel B in Table 2.1). Only 5.8% of respondents believed house price had fallen. The standard deviation of 7.52 percentage points is high, pointing to large variation in respondents' perceptions of past local house price growth.

2.2.3 Past Local House Prices

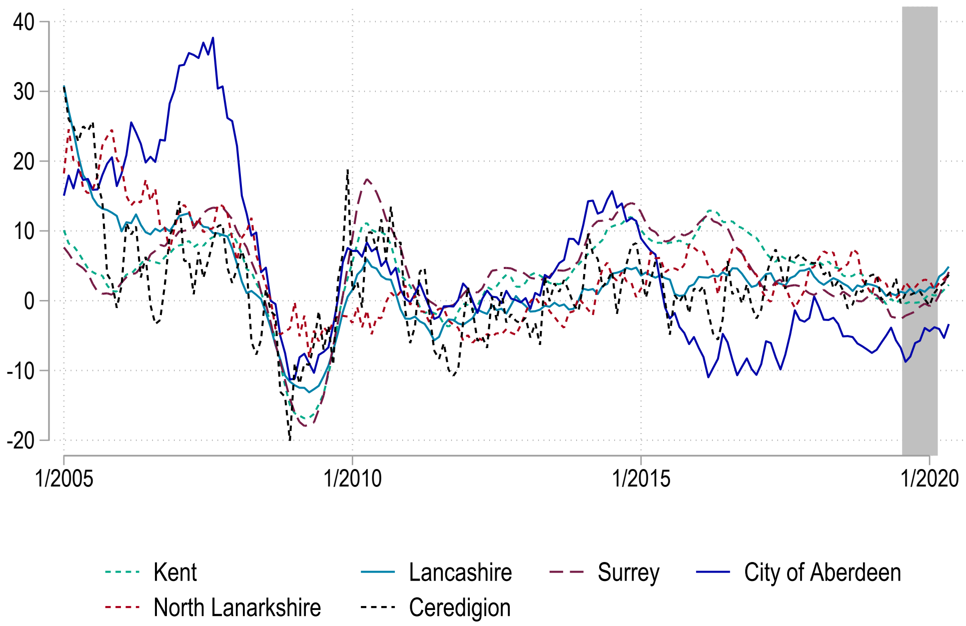
Next, we construct measures of the realized past house price change individuals experienced in their local area of residence. We compute past house price changes from the *UK House Price Index* (*UK HPI*), published by *HM Land Registry*. The local authority-specific index is updated monthly, and is mix-adjusted to account for changes in housing quality and composition (HM Land Registry 2021). It produces highly accurate price trends as residential property sales occur frequently in the majority of local authorities. In 2019 (2018), more than 1,000 residential properties were sold annually in 95% (96%) of the 339 local authorities in England and Wales, and above 2,000 in 50% (58%) of them (Office for National Statistics 2023).

To compute past experienced local house price changes, we follow the approach in Kuchler and Zafar (2019).[13] We link local house prices into the Financial Lives survey data by respondents' area of residence and interview month, and use the most recent annual percentage change in local house prices, relative to the month a respondents' interview was conducted.

British house price dynamics vary considerably across time and place, as Figure 2.2 exemplifies for six local authorities (see also Appendix B4). Panel C in Table 2.1 shows that local house prices rose on average by a modest 1.11% with a standard deviation of 2.68.

[13]In the literature on the effects of past experiences on subjective expectations, more sophisticated aggregation functions for historic experiences have been proposed. As a robustness check, we followed the approach by Malmendier and Nagel (2011) and constructed a weighted average of annual local house price changes, selecting a combination of lookback period and weights that yielded the best goodness-of-fit. We achieve the best fit for the 5-year lookback period with slowly decreasing weights. Our main results are robust to using this alternative measure of historic return experiences. See Appendix B6 for details, and robustness of our estimates.

Figure 2.2. Realized local house price dynamics



Notes: The figure shows (monthly computed) annual house price changes for the time from 2005 to 2020 for an arbitrary selection of six different local authorities. The shaded area indicates the survey period. *Data Source*: Historical values from the UK HPI.

2.2.4 Local Economic Conditions

As suggested by Armona, Fuster, and Zafar (2019), individuals may take into account 'information other than past home price growth, [...] including local macroeconomic conditions', when they form beliefs about future house prices. We consider two salient, widely used measures of local macroeconomic conditions, (i) the unemployment rate and (ii) an index of social and economic deprivation.

Local unemployment rate. The unemployment rate is frequently reported in the news specifically for local areas, and is one of the most commonly used measures of local economic conditions. We measure local unemployment rates as the total number of people (i) claiming Jobseeker's Allowance, plus those (ii) claiming Universal Credit and being out of work, divided by the resident local population aged 16–64. These administrative "claimant counts," are lower than officially reported unemployment rates.[14] We consider unemployment rates in the respondent's local authority of residence in the month in which the survey interview was conducted. Table 2.2 reports average local unemployment rates for the survey period. Even in a period of low unemployment, we see large variation across the local authorities in our sample, visualized in Figure 2.3. The mean (median) unemployment rate is 2.6% (2.4%), with a standard deviation of 1.13 percentage points. Older industrial areas, some seaside towns, and some London boroughs are among the places with the highest unemployment rates.

Table 2.2. Summary statistics: Local economic conditions

	N	Mean	Std Dev	P10	P50	P90
Local unemployment rate	364	2.59	1.13	1.33	2.41	4.23
Local deprivation score	312	19.66	7.99	10.24	18.58	30.72

Notes: The table shows summary statistics for the local authorities covered in our sample. Monthly local unemployment rates averaged for the survey period between August 2019 and February 2020. Local deprivation score as of 2019 (only available for English local authorities). *Data Source*: Office for National Statistics (data from Nomis) and Ministry of Housing, Communities & Local Government.

Local deprivation score. Our alternative measure of local economic conditions, the social and economic deprivation score, is taken from the 2019 Index of Multiple Deprivation, provided by the *Ministry of Housing, Communities & Local Government.* It aggregates indicators of local conditions across a broad spectrum of social and economic dimensions, and is frequently considered in the design of local and national policies, such as the UK government's 'Levelling Up' agenda.

The score aggregates indicators from seven domains, including (i) income, (ii) employment, (iii) education, skills, and training, (iv) health and disability, (v) crime, (vi) barriers to housing and services, and (vii) living environment. Due to methodological differences in their measurement and underlying indicators across the constituent countries of Great Britain, estimates based on this measure of local economic conditions are produced for England only. Appendix B5 shows that both measures are positively correlated with each other and with other measures of local conditions.

[14] Officially reported unemployment rates are based on self-reports in the Labour Force survey and Annual population survey, whose sample sizes are not sufficient for granular spatial analysis at monthly level. Claimant counts yield systematically lower unemployment rates as they exclude, for example, those searching for a job who have not claimed unemployment or other benefits. E.g., the unemployment rate based on claimant counts across local authorities was on average 2.4% in 2019; in contrast, the model-based estimate of the officially reported unemployment rate across local authorities was 3.6% in 2019 (Office for National Statistics 2022).

Figure 2.3. Local unemployment rates

Notes: The map shows the average monthly local unemployment rates for the survey period between August 2019 and February 2020, measured as claimants per resident population aged 16–64. *Claimants* denote those claiming Jobseeker's Allowance and out-of-work claimants of Universal Credit. *Data Source*: Office for National Statistics (data from Nomis).

2.2.5 Financial Sophistication

In Section 2.3.3, we will consider heterogeneity in belief formation about future house prices by respondents' financial sophistication, using the canonical four-item construct of financial literacy covering knowledge of interest rates, interest compounding, inflation, and risk diversification (for measurement details, see Lusardi and Mitchell, 2008). In our sample, 48% of the respondents answer all four questions correctly. We define these as having *high* financial literacy; the remaining 52% of the sample are defined as having *low* financial literacy.[15] In Table B.9 in the Appendix, we show that our results are robust to using an alternative measure of sophistication based on knowledge about interest rates for savings accounts.

2.3 Empirical Analysis

The primary goal of this paper is to characterize the role of local economic conditions and past (perceived and realized) local house price changes in shaping the formation of individuals' subjective expectations about future changes in local house prices. Our baseline regression specification is as follows:

$$E(\Delta HPI^{t+12})_{ilt} = \beta LEC_{lt} + \delta \Delta HPI_{ilt}^{t-12} + \gamma X_{ilt} + \eta_t + \epsilon_{ilt} \tag{2.2}$$

The dependent variable $E(\Delta HPI^{t+12})_{ilt}$ is the subjective expected rate of one-year change in the house price index in individual i's local area l, relative to the interview month t.[16] LEC_{lt} is a measure of the local economic condition in locality l, e.g., in our baseline specification the monthly unemployment rate in the individual's residential local authority. ΔHPI_{ilt}^{t-12} refers to the perceived past one-year rate of change in the local house price index in local authority l relative to the interview month t. In some regression specifications, we consider *realized* rather than *perceived* past local house price changes; in these cases, the variable simplifies to ΔHPI_{lt}^{t-12}, as it varies only by local authority and interview month. X_{ilt} is a vector of individual-specific controls, and η_t are interview-month fixed effects.[17] Standard errors are clustered at the level of local authorities. The number of observations in the full regression sample drops from 2,799 to 2,731 due to item non-response on covariates.

[15] In line with previous research (Lusardi and Mitchell 2008), we find that women, individuals with low education, and those with low household income are less likely to have high financial literacy. Also, individuals living in regions with higher rates of unemployment are less likely to have high financial literacy (see Table B.5 in the Appendix).

[16] The subscript refers to the point in time at which expectations are formed and the superscript refers to the period of time over which they are formed.

[17] We do not include local-authority fixed effects because there is little within-LA variation in local unemployment rates over the short interview period of the survey.

2.3.1 Predictors of Subjective House Price Expectations

We estimate Equation 2.2 to study the role of past house price changes and of locally experienced economic conditions in shaping the formation of subjective expectations about one-year-ahead changes in local house prices. In column (1) of Table 2.3, we include local unemployment rates and *realized* local one-year house price growth as explanatory variables; in column (2), we replace realized local one-year house price growth with *perceived* local one-year house price growth; in column (3), we include all three variables.

Table 2.3. Predictors of subjective expected house price changes

	Expected 1yr HP change (%)		
	(1)	(2)	(3)
Local unemployment rate	-0.159^{**}	-0.155^{**}	-0.153^{**}
	(0.074)	(0.067)	(0.067)
Realized 1yr HP change	-0.001		-0.003
	(0.032)		(0.030)
Perceived 1yr HP change		0.133^{***}	0.133^{***}
		(0.029)	(0.029)
Month Fixed Effects	✓	✓	✓
Socio-demographics	✓	✓	✓
Effect of 1 std in Local unemployment rate	-0.18	-0.17	-0.17
(in %)	(-4.78)	(-4.65)	(-4.60)
Effect of 1 std in Realized 1yr HP change	-0.00		-0.01
(in %)	(-0.07)		(-0.24)
Effect of 1 std in Perceived 1yr HP change		0.99	0.99
(in %)		(26.82)	(26.83)
Mean DepVar	3.70	3.70	3.70
R-squared	0.03	0.08	0.08
N	2,731	2,731	2,731

Notes: The table shows regression estimates, with *expected 1yr house price (HP) change* as the dependent variable. Realized, perceived, and expected 1yr changes in prices refer to the *local* housing market. Local unemployment rates refer to the month of interview. *Socio-demographics* include indicators for age categories, education categories, household-income categories, whether respondents are female, married, or working, and whether they finished the interview during one day. Standard errors in parentheses are adjusted for clustering at the *local-authority* level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

The estimation results show that individuals do not extrapolate from past realized one-year local house price changes when forming subjective expectations about one-year-ahead local house price changes. Instead, they heavily extrapolate from perceived past local house price changes. The estimate of the coefficient of perceived local one-year house price growth is positive and significant at the 1% level. Individuals who perceive the past one-year house price growth in their local area to be higher are also more likely to expect higher local house price growth in the future.

A strong positive association between past perceived and future expected house price changes was also documented in Armona, Fuster, and Zafar (2019). Yet, our estimates reveal that economic conditions in the local area are an additional important predictor of subjective house price expectations. Individuals who live in local areas with higher unemployment rates expect, on average, lower rates of house price growth: A one standard-deviation increase in local unemployment rates is associated with a decrease in individuals' expected local house price changes by 0.17 percentage points (which corresponds to 5% of one-year-ahead expected local house price changes).

Comparing the magnitude of the effects on expected local house price growth between local economic conditions and past perceived local house price changes, we find that a one standard-deviation change in local economic conditions is associated with a decrease in subjective price-change beliefs of around 0.17 percentage points—about one sixth the magnitude of a one standard-deviation change in past perceived local house price changes.

2.3.2 House Price Perception Gaps

The empirical analysis so far revealed that individuals base their subjective house price expectations on perceived rather than realized past house price changes, suggesting a deviation between them, in short: a "perception gap." Panel D in Table 2.1 indeed displays a large absolute difference between realized and perceived one-year local house price changes. On average, the absolute perception gap is 5.04 percentage points. Further, it is above 3.5 percentage points for half of the sample; for around 10% of the sample, the absolute perception gap is above 10 percentage points.

In Table 2.4 (columns 3 and 4), we investigate the correlates of individuals' perception gaps. First, we find that the gap is larger for women and those living in local authorities that experienced high variability in house price changes in the past, and smaller among homeowners. At the same time, results in columns (1) and (2) show that they do not predict the *level* of perceived house price changes. Hence, while, for example, women's perceptions tend to be less accurate, they do not systematically over- or underestimate past house price changes. We return to these dimensions of heterogeneity in Section 2.3.4 below.

Table 2.4. Correlates of perceived past house price changes and the perception gap

Dependent variable	Perceived 1yr HP change (%)		Absolute perception gap (%-points)	
	(1)	(2)	(3)	(4)
Local area characteristics				
Realized 1yr HP change	−0.065	−0.068		
	(0.056)	(0.057)		
Realized 1yr HP change (absolute value)			0.148	0.096
			(0.117)	(0.128)
Local unemployment rate	−0.089	−0.088	0.083	−0.027
	(0.177)	(0.181)	(0.183)	(0.194)
5yr local house price volatility	−0.181	−0.225	0.362**	0.385***
	(0.171)	(0.172)	(0.149)	(0.143)
Log property sales	0.524	0.161	0.730	0.539
	(1.065)	(1.082)	(1.081)	(1.071)
Household (or head) characteristics				
Female		0.399		0.580**
		(0.310)		(0.274)
Age (ref: 18–44)				
45–64		−0.121		−0.047
		(0.461)		(0.386)
65 or older		−0.238		−0.610
		(0.558)		(0.485)
Partner in household		−0.213		0.299
		(0.348)		(0.318)
Education (ref: Higher)				
Lower or medium		0.621*		0.270
		(0.322)		(0.300)
None		−0.874		0.006
		(0.633)		(0.500)
Info missing		−6.350*		6.418**
		(3.472)		(2.893)
Working		0.484		−0.120
		(0.503)		(0.430)
Annual HH income (ref: £70k or more)				
Less than £20k		−0.184		0.227
		(0.504)		(0.422)
£20k - <£40k		−0.226		0.390
		(0.350)		(0.314)
£40k - <£70k		0.861**		0.673**
		(0.333)		(0.296)
Info missing		0.119		1.030**
		(0.520)		(0.479)
Risk averse		0.281		0.361
		(0.325)		(0.287)
High financial literacy		−0.447		−1.056***
		(0.308)		(0.283)
Recent mover		−0.504		−0.604
		(0.430)		(0.426)
Homeowner		−0.038		−0.955**
		(0.456)		(0.402)
Survey finished on same day		0.282		0.375
		(0.335)		(0.269)
Constant	1.874	2.810	2.823	3.751
	(4.029)	(4.100)	(3.933)	(3.845)
Mean DepVar	3.82	3.84	5.05	5.04
Adj. R-squared	0.01	0.02	0.01	0.05
N	2,516	2,456	2,516	2,456
Month Fixed Effects	✓	✓	✓	✓

Notes: The table shows regression estimates, with *perceived 1yr HP change* (columns 1 and 2) and *absolute perception gaps* (columns 3 and 4) as dependent variables. 'ref.' indicates the omitted category. "Log property sales" denotes the logarithm of local residential property sales per 1,000 inhabitants (note that using instead the "logarithm of local residential property sales per 1,000 dwellings," results remain qualitatively unchanged). Standard errors in parentheses are adjusted for clustering at the *local-authority* level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

Secondly, individuals' financial sophistication is a key predictor for the perception gap. Table 2.4 shows that high financial sophistication is associated with lower perception gaps (by about one-fifth compared to the overall average), and is strongly statistically significant. That is, financially sophisticated individuals assess realized past local one-year changes in house prices more accurately than individuals with low financial literacy.[18] In summary, we find that individuals' perceptions about past house prices are far from accurate. We document large heterogeneity in perception gaps which vary with the uncertainty in the local housing market, individuals' financial sophistication, and other observables.

2.3.3 Subjective Expectations and Financial Sophistication

A large literature documents considerable individual heterogeneity in subjective expectations of aggregate and individual-level economic outcomes. One important dimension of this heterogeneity is related to sophistication, as measured for instance by numeracy or educational attainment, cognitive ability and intelligence (e.g., see D'Acunto, Hoang, Paloviita, and Weber, 2019 for inflation expectations and Kuchler and Zafar, 2019 and Kuchler, Piazzesi, and Stroebel, 2023 for house price expectations). Our data allow us to study the importance of domain-specific skills, in particular financial sophistication. Table 2.4 already suggested that the measure of financial literacy, whose construction we described in Section 2.2.5 above, is a strong predictor of the gap between perceived and realized past house price changes. Indeed, those with high financial literacy exhibit an absolute perception gap of, on average, 4.23 percentage points. The gap is significantly higher (5.77 percentage points) among those with low financial sophistication. Compared to individuals with high financial literacy, those with low financial literacy also expect larger increases in house prices over the next 12 months (4.35% vs. 3%); in addition, we observe substantial heterogeneity in expectations among individuals with low financial literacy. Related, Kuchler and Zafar (2019) show that sophistication lowers the extent to which individuals naively extrapolate from local house price changes when asked about their beliefs regarding future national house price changes.

Next, we study whether the formation of subjective expectations varies by financial sophistication, measured by financial literacy. Results are qualitatively similar (see Appendix B6) when we use the secondary measure capturing interest rate knowledge.

[18] Replacing the main financial sophistication measure with our secondary measure, *interest rate knowledge*, the regression coefficient remains negative and significant at the 0.01 level.

Table 2.5. Heterogeneity of subjective expectations: Financial literacy

	Expected 1yr HP change (%)					
			Financial literacy			
	Pooled		High		Low	
	(1)	(2)	(3)	(4)	(5)	(6)
Local unemployment rate	-0.159^{**}	-0.155^{**}	-0.014	0.009	-0.347^{***}	-0.347^{***}
	(0.074)	(0.067)	(0.077)	(0.059)	(0.120)	(0.112)
Realized 1yr HP change	-0.001		0.022		-0.008	
	(0.032)		(0.032)		(0.052)	
Perceived 1yr HP change		0.133^{***}		0.342^{***}		0.086^{***}
		(0.029)		(0.043)		(0.028)
Effect of 1 std in Local unemployment rate	-0.18	-0.17	-0.01	0.01	-0.39	-0.39
(in %)	(-4.78)	(-4.65)	(-0.48)	(0.32)	(-9.12)	(-9.14)
Effect of 1 std in Realized 1yr HP change	-0.00		0.06		-0.02	
(in %)	(-0.07)		(1.95)		(-0.47)	
Effect of 1 std in Perceived 1yr HP change		0.99		1.57		0.81
(in %)		(26.82)		(51.49)		(18.73)
Mean DepVar	3.70	3.70	3.04	3.04	4.31	4.31
R-squared	0.03	0.08	0.03	0.27	0.03	0.05
N	2,731	2,731	1,309	1,309	1,422	1,422
Month Fixed Effects	✓	✓	✓	✓	✓	✓
Socio-demographics	✓	✓	✓	✓	✓	✓

Notes: The table shows regression estimates, with *expected 1yr house price (HP) change* as the dependent variable. Realized, perceived, and expected 1yr changes in prices refer to the *local* housing market. Local unemployment rates refer to the month of interview; local deprivation scores from 2019. Financial literacy is *high* if all four standard financial literacy questions are answered correctly, and *low* otherwise. *Socio-demographics* include indicators for age categories, education categories, household-income categories, whether respondents are female, married, or working, and whether they finished the interview during one day. Standard errors in parentheses are adjusted for clustering at the *local-authority* level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

Table 2.5 shows estimates for the empirical model specified in Equation 2.2, separately for individuals with high and low financial literacy. Past house price changes matter for the formation of expectations in both groups; but as before and irrespective of the level of financial literacy, individuals extrapolate from *perceived* rather than from *realized* one-year house price changes. Yet, individuals with low financial literacy rely much less on perceived past house price changes. Instead, individuals use local economic conditions when forming expectations about local house price changes (see columns 5 and 6).[19] These results are intuitive once information costs are taken into account. For less financially sophisticated individuals, it may be cognitively less costly to learn about past local economic conditions than about past local house price changes, as local economic conditions are more persistent and less volatile over time, and arguably, also more salient.

[19] Estimates in Table 2.5, columns (2), (4), and (6), are virtually unchanged if we include realized one-year house price changes as a control variable.

2.3.4 Subjective Expectations, Information Acquisition, and Market Uncertainty

Information acquisition costs and benefits regarding past house price changes may depend on individual incentives to be informed and on characteristics of local housing markets. First, we consider several individual experiences that may shift the incentives to be informed, such as a recent move and homeownership status, and present estimates of heterogeneity in belief formation with respect to these factors. Second, we explore a set of such local market characteristics: past price volatility and the salience of local price information.

First, we consider whether risk averse individuals are more likely to diversify their information sources and rely on local economic conditions *and* own perceptions in their belief formation. We split our sample into those with higher self-reported risk aversion, measured through a widely used and validated survey question on risk attitudes.[20] We find that risk averse individuals extrapolate indeed from both perceived past house price changes and local conditions, while individuals who are more tolerant of risk do not rely on local economic conditions (see panel A of Table 2.6).

Whether home owners have a stronger incentive to monitor house prices in their local area (and other price dynamics) than renters is still subject to debate – with mixed evidence.[21] Intuitively, housing is by far the largest asset in their portfolio for the majority of home owners, warranting attention to the dynamic evolution of their wealth. Yet, renters who may aspire to 'climb onto the housing ladder'[22] or who wish to form expectations of future rents may also have an incentive to monitor local house prices.

[20] This question has been shown to be a good predictor of risk taking behavior across different domains (e.g., Bonin, Dohmen, Falk, Huffman, and Sunde, 2007; Jaeger et al., 2010; Dohmen et al., 2011).

[21] Adelino, Schoar, and Severino (2018) focus on measuring perceptions of house price risk and show that renters view housing as riskier than owners. In a study of regional house prices in Germany, Kindermann, Le Blanc, Piazzesi, and Schneider (2021) find that on average, households underpredict local price growth. Yet, renters make on average higher and hence more accurate forecasts than owners, although their forecasts are more dispersed and their mean squared forecast errors are higher. Ahn and Yang (2022) find that homeowners are attentive to news on interest rates – driven by changes in mortgage-rates, and adjust their inflation expectations accordingly, but do not investigate subjective house price expectations. Gohl, Haan, Michelsen, and Weinhardt (2024) find no evidence of house price expectation biases related to individual housing tenure decisions.

[22] *The Economist* describes the notion of the housing ladder as follows: "The ladder is deeply embedded into British thinking. On its most narrow definition, it is usually taken to mean the idea of first-time buyers purchasing a modest dwelling (a flat, say) and then trading up to something larger as their incomes grow and their housing equity increases. More broadly, the metaphor reflects Britons' general aspiration to residential-property ownership." (*The Economist*, 13 January 2024, p. 23)

While we find no evidence that home ownership leads to systematic shifts in perceptions of past house prices, home owners possess more accurate information regarding past house price changes, as evidenced by an about a 1 percentage-point lower perception gap (see Table 2.4). They also rely more strongly on perceived price changes when forming beliefs about future local house prices, and on local economic conditions, than renters (see panel B of Table 2.6).

While it is not clear a priori whether home owners or renters have a stronger incentive to be informed about local house prices, individuals who recently moved house may have better information about past house prices than non-movers. Yet, we do not find evidence that a recent moving experience shifts the individuals' level of house price perceptions, nor that their perceptions are more accurate (see Table 2.4). The heterogeneity estimates in panel C of Table 2.6, however, suggest that recent movers rely more strongly on their perceptions of past house prices than non-movers, and the latter rely on both when forming beliefs about future house price changes. A plausible interpretation of this finding is that recent movers overestimate how well they are informed about local prices.

Studies of subjective expectations often find systematic gender differences and more distorted beliefs among women (D'Acunto, Malmendier, and Weber 2021). In our data, men have lower perception gaps (see Table 2.4), and rely more strongly on their perceptions of past changes, but we do not find consistent differences in the reliance on local economic conditions between men and women (see Appendix B6, Table B.10).

Next, we turn to investigating the role of local market characteristics, i.e., past price volatility and the salience of local price information, in belief formation. First, perception gaps are higher in local housing markets that display higher price uncertainty where forming accurate beliefs about past house price changes is difficult. Some local authorities experienced highly volatile house prices in the past five years. While individuals' perception gaps are higher in local housing markets with more volatile prices (see columns 3 and 4 in Table 2.4), they do not lead to systematically higher perceptions (columns 1 and 2). A one standard-deviation rise in local house prices in the last 5 years increases the perception gap by 0.394 percentage points or 8%. Panel A of Table 2.7 shows that market uncertainty also matters for the belief formation about future house prices. In local authorities where price uncertainty is high, individual perceptions of past price changes remain an important predictor of subjective house price expectations. However, individuals also rely strongly on local economic fundamentals. A one percentage-point increase in the local unemployment rate reduces expected future house price growth

Table 2.6. Heterogeneity of subjective expectations: Risk aversion, homeownership, and recent movers

	Dependent variable: Expected 1yr HP change (%)			
	(1)	(2)	(3)	(4)
	A. Risk averse			
	Yes		No	
Local unemployment rate	−0.295**	−0.299**	−0.081	−0.072
	(0.124)	(0.117)	(0.090)	(0.084)
Realized 1yr HP change	−0.034		0.021	
	(0.048)		(0.041)	
Perceived 1yr HP change		0.121***		0.140***
		(0.044)		(0.035)
Mean DepVar	3.72	3.72	3.70	3.70
R-squared	0.03	0.08	0.05	0.10
N	1,142	1,142	1,583	1,583
	B. Homeowner			
	Yes		No	
Local unemployment rate	−0.199**	−0.200***	−0.121	−0.116
	(0.077)	(0.069)	(0.162)	(0.157)
Realized 1yr HP change	0.033		−0.066	
	(0.034)		(0.079)	
Perceived 1yr HP change		0.173***		0.087***
		(0.048)		(0.032)
Mean DepVar	3.54	3.54	4.11	4.11
R-squared	0.02	0.10	0.06	0.08
N	1,966	1,966	765	765
	C. Recent mover			
	Yes		No	
Local unemployment rate	−0.213	−0.310	−0.148*	−0.138**
	(0.237)	(0.223)	(0.076)	(0.070)
Realized 1yr HP change	0.008		0.001	
	(0.080)		(0.033)	
Perceived 1yr HP change		0.216***		0.125***
		(0.054)		(0.029)
Mean DepVar	3.61	3.61	3.71	3.71
R-squared	0.19	0.27	0.03	0.07
N	252	252	2,479	2,479

Notes: The table shows regression estimates, with *expected 1yr house price (HP) change* as the dependent variable. Realized, perceived, and expected 1yr changes in prices refer to the *local* housing market. Local unemployment rates refer to the month of interview. In all specifications, we control for socio-demographics and interview-month fixed effects. *Socio-demographics* include indicators for age categories, education categories, household-income categories, whether respondents are female, married, or working, and whether they finished the interview during one day. Note that estimates in columns (2) and (4) are virtually unchanged if we include realized one-year house price changes as a control variable. Standard errors in parentheses are adjusted for clustering at the *local-authority* level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

Table 2.7. Heterogeneity of subjective expectations: Local market uncertainty and local information

	Dependent variable: Expected 1yr HP change (%)			
	(1)	(2)	(3)	(4)
	A. Volatility of local house prices (past 5 years)			
	Low		High	
Local unemployment rate	−0.149	−0.142	-0.275^{**}	-0.275^{**}
	(0.098)	(0.092)	(0.130)	(0.117)
Realized 1yr HP change	−0.039		0.008	
	(0.049)		(0.042)	
Perceived 1yr HP change		0.097^{***}		0.176^{***}
		(0.030)		(0.049)
Mean DepVar	3.79	3.79	3.61	3.61
R-squared	0.02	0.05	0.05	0.13
N	1,365	1,365	1,366	1,366
	B. Local media coverage of house prices			
	Low		High	
Local unemployment rate	−0.213	-0.202^{*}	−0.136	−0.130
	(0.130)	(0.118)	(0.090)	(0.082)
Realized 1yr HP change	−0.017		0.017	
	(0.038)		(0.050)	
Perceived 1yr HP change		0.135^{**}		0.130^{***}
		(0.056)		(0.029)
Mean DepVar	3.70	3.70	3.71	3.71
R-squared	0.06	0.10	0.03	0.08
N	1,108	1,108	1,623	1,623

Notes: The table shows regression estimates, with *expected 1yr house price (HP) change* as the dependent variable. Realized, perceived, and expected 1yr changes in prices refer to the *local* housing market. Local unemployment rates refer to the month of interview. In all specifications, we control for socio-demographics and interview-month fixed effects. *Socio-demographics* include indicators for age categories, education categories, household-income categories, whether respondents are female, married, or working, and whether they finished the interview during one day. Note that estimates in columns (2) and (4) are virtually unchanged if we include realized one-year house price changes as a control variable. Standard errors in parentheses are adjusted for clustering at the *local-authority* level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

by 0.275 percentage points. This is in contrast to individuals who live in areas with less volatile prices where we find no evidence that individuals rely on local economic conditions in their belief formation.

Second, local house price reporting in the media may lower the cost of information acquisition, reduce individuals' perceptions gaps, and change the weight of perceptions in individuals' belief formation. To explore this possibility, we develop a measure of how frequently the topic 'house prices' or 'property prices' is reported in the media in the different local authorities. We compile a list of local newspapers (i.e., newspapers covering only certain local authorities), and focus on the time period from August 2018 to February 2020, which covers the 12-months period before the first survey interview was conducted and the months during which the survey was conducted. Using Google search results extraction, we count, for each local newspaper, the number of articles that contain these search terms. We use this measure as a proxy for the local media coverage of the topic of house prices.

We find that the relation between individuals' perceptions and local house price expectations does not vary with the level of local media salience regarding house prices (see panel B in Table 2.7). However, there is some indication that local unemployment rates are a stronger predictor of local house price expectations in areas with little coverage of house prices in the media. Yet, our measure of media reporting likely underestimates the variation in salient information on house prices across local areas, as it does not include widely used commercial property search engines.[23]

In summary, we find that the extent to which individuals rely on their perceptions of past house prices and particularly the weight they attribute to easily observable local economic conditions and perceptions in their belief formation varies with the volatility of local house prices, the intensity of local reporting of price information, and with individual financial sophistication, risk preferences, and incentives to be informed about recent price dynamics.

2.3.5 Subjective Expectations, Short-Run Momentum, and Local House Price Fundamentals

There is widespread evidence that subjective house price expectations may be a source of house price bubbles, and affect outcomes in the housing market. The pandemic shock hit shortly after the completion of our survey, creating idiosyncratic market dynamics such as a sudden high demand for larger properties and those with outside space, and a temporary slump in demand for property in well-connected locations. This precludes us from linking the elicited subjective house price expectations to housing market outcomes in Great Britain in 2020. Instead, we use (historic) panel information on local house price dynamics between 2010 and 2019 to study short-run momentum in the housing market, i.e., provide evidence on the extent to which information on local house prices at t predict house prices a year ahead. Based on Case and Shiller (1989) and Guren (2018),

[23]Unfortunately, more granular localized information on the search intensity for house price information was not available from these providers due to their business sensitive nature.

the existence of short-run momentum is seen as a stylized fact of house price dynamics. Yet, this evidence is mostly based on data from housing markets in the US, hence we estimate an AR(1) model[24] of short-run house price evolution in Great Britain. We follow Armona, Fuster, and Zafar (2019) to estimate local AR(1) models:

$$\Delta HPI_{l,t+12} = \alpha_l + \sigma_l \Delta HPI_{l,t} + \epsilon_{l,t} \tag{2.3}$$

where $\Delta HPI_{l,t}$ is the rate of one-year change in the UK House Price Index in local authority l and month t. The rate of change in house prices is calculated over the horizon of 12 months. σ_l shows whether the local house price change in the past is a useful predictor of the current local house price change.

We estimate the autoregressive coefficients σ_l for each of the 364 local authorities by an OLS model with Newey–West standard errors for the time horizon from 7/2010 to 7/2019 (i.e., after the financial crisis occurred, and before the Financial Lives survey was conducted). In Table 2.8, we report the mean and the standard deviation (in parentheses) of these estimates (column 1), as well as the share of local authorities for which σ_l is positive (negative) and significant at the 5% level (column 2 respectively 3).

Table 2.8. Dependence in realized local house price changes

	7/2010–7/2019		
	Mean (Std) (1)	Percent positive[a] (2)	Percent negative[b] (3)
One-year local HP growth on lagged one-year local HP growth	0.156 (0.226)	44.2	6.9

Notes: The table shows regression estimates of local house price (HP) growth dependence on previous local HP growth. Mean coefficient across local authorities shown in column "mean," and standard deviation (std) across local authorities shown in parentheses. We exclusively consider local authorities that are covered in our survey sample, $N = 364$. Number of monthly-computed annual house price growth per local authority that estimates are based on: 97 observations. [a] [[b]] Indicates percent of estimates statistically significantly positive [negative] at the 5% level, based on Newey–West standard errors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: UK HPI.

For the majority of localities in our sample, the autoregressive coefficient is positive and statistically significant. This short-term momentum in house prices is a frequent phenomenon in local housing markets in Great Britain. The average estimate across the local authorities in our sample is 0.156, indicating that across local areas in our sample, a one percentage-point higher house price growth in month t is followed, on average, by

[24]We base this on monthly-computed one-year local changes in house prices and their lagged values.

0.156 percentage-point higher rates of house price growth in the consecutive 12 months. This estimate is very similar to the estimated percentage-point change in subjective house price expectations upon a one percentage-point change in *perceived* local house prices (see our main estimates in Table 2.3). While individuals do not perfectly recall the realized past local house price change, they appear to extrapolate according to the average short-run momentum in realized house prices.

However, the average momentum in Table 2.8 masks strong heterogeneity: we find statistically significant short-run momentum in 44.2% of all local authorities, but also reversal in 6.9% of areas, and no evidence of autocorrelation in prices for the remainder of areas.[25] Local markets with short-run momentum can be found across all regions of Great Britain; they do include most local authorities in London (91%) and more than half of local authorities in the East of England and the West Midlands but also more than a third of local authorities in the South East, the North West, and Yorkshire and Humberside.

Do individuals' subjective expectations reflect the varying informativeness of past local house prices, i.e., do individuals rely on different underlying information depending on the momentum in their area of residence? Our findings imply that individuals' belief formation models vary with local housing market fundamentals. Table 2.9 shows that those living in areas with short-run momentum extrapolate strongly from perceived past house price change—and in a manner that is quantitatively similar to our main findings in Table 2.3. However, those who reside in areas where past prices are not predictive for future prices extrapolate very differently: they rely strongly on local unemployment rates in their belief formation; in fact, similarly to those living in areas with high price volatility. A one percentage-point increase in the local unemployment rate is associated with an 0.28 percentage-point decrease in expected local house price growth. This suggests that individuals may adapt the model which they use to form beliefs to the fundamentals of the local housing markets they find themselves in.

2.3.6 Stock Market Return Expectations

The analysis so far revealed that individuals use local economic conditions in their mental models of subjective house price expectations. Our interpretation of this finding is that

[25] We also consider a five-year lag in prices (using *annualized* house price changes) to investigate long run mean reversion, and find predictive power for about 40% of local authorities – but only weak evidence of mean reversion for about half of them. Results available from the authors upon request.

Table 2.9. Heterogeneity of subjective expectations: Local short-run momentum

	Dependent variable: Expected 1yr HP change (%)			
	(1)	(2)	(3)	(4)
	Local housing market with short-run momentum			
	Yes		No	
Local unemployment rate	−0.093	−0.125	-0.280^{**}	-0.222^{*}
	(0.096)	(0.081)	(0.121)	(0.116)
Realized 1yr HP change	−0.037		0.052	
	(0.046)		(0.043)	
Perceived 1yr HP change		0.129^{***}		0.148^{***}
		(0.039)		(0.040)
Mean DepVar	3.82	3.82	3.58	3.58
R-squared	0.04	0.09	0.03	0.09
N	1,431	1,431	1,300	1,300

Notes: The table shows regression estimates, with *expected 1yr house price (HP) change* as the dependent variable. Realized, perceived, and expected 1yr changes in prices refer to the *local* housing market. Local unemployment rates refer to the month of interview. In all specifications, we control for socio-demographics and interview-month fixed effects. *Socio-demographics* include indicators for age categories, education categories, household-income categories, whether respondents are female, married, or working, and whether they finished the interview during one day. Note that estimates in columns (2) and (4) are virtually unchanged if we include realized one-year house price changes as a control variable. Areas with 'No' momentum comprise those exhibiting no autocorrelation or reversal. Results are qualitatively unchanged when we exclude areas exhibiting reversal. Standard errors in parentheses are adjusted for clustering at the *local-authority* level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

for many individuals, in particular those who are less knowledgeable about past local house prices and those with lower financial sophistication, this makes sense: measures like the local unemployment rate are, arguably, more easily accessible than past house price changes. Similarly sensibly, we find that where local housing markets are more volatile and for which there is no short-run momentum, individuals rely more heavily on a wider set of local fundamentals.

More generally, our findings so far are also in line with a major theme of recent research on subjective expectations: Salient experiences matter. A natural next question to ask is then: Do individuals also use local economic conditions when they form beliefs about national outcomes? For instance, Kuchler and Zafar (2019) study whether individuals extrapolate from local (i.e., zip code-level) experiences in their formation of beliefs about national outcomes, including stock market returns.

The Financial Lives survey data allow us to explore this issue. In addition to the questions on house price changes, we elicited subjective expectations, perceptions of

Table 2.10. Predictors of subjective expected stock market returns

	Expected 1yr stock market return (%)		
	(1)	(2)	(3)
Local unemployment rate	−0.058	−0.035	−0.038
	(0.078)	(0.074)	(0.075)
Realized 1yr stock market return	0.098***		0.087***
	(0.029)		(0.029)
Perceived 1yr stock market return		0.128***	0.126***
		(0.021)	(0.021)
Socio-demographics	✓	✓	✓
Effect of 1 std in Local unemployment rate	−0.06	−0.04	−0.04
(in %)	(−3.02)	(−1.83)	(−1.98)
Effect of 1 std in Realized 1yr stock market return	0.34		0.30
(in %)	(16.16)		(14.27)
Effect of 1 std in Perceived 1yr stock market return		1.07	1.06
(in %)		(50.22)	(49.66)
Mean DepVar	2.13	2.13	2.13
R-squared	0.02	0.06	0.06
N	2,731	2,731	2,731

Notes: The table shows regression estimates, with *expected 1yr stock market return* as the dependent variable. Realized, perceived, and expected 1yr returns refer to the *national* stock market. Local unemployment rates refer to the month of interview. *Socio-demographics* include indicators for age categories, education categories, household-income categories, whether respondents are female, married, or working, and whether they finished the interview during one day. Standard errors in parentheses are adjusted for clustering at the *local-authority* level. Note that interview-month fixed effects are captured in *realized 1yr stock market returns*, which only vary across interview months, but not across local authorities. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

price changes over the last 12 months, and compiled data on realized price changes for an *aggregate* rather than *local* asset – the FTSE-100 UK stock market index.[26]

In this final part of our empirical analysis, we present estimation results from models that are specified analogously to those reported for local house price expectations. We find that both realized and perceived past stock market returns predict subjective expectations about national stock market returns (see Table 2.10). However, similar to the analysis on expected house price changes, perceived (rather than realized) past returns have more predictive power in explaining expected stock market returns. In contrast, we do not find evidence of local economic conditions predicting subjective expectations about expected returns of the FTSE-100 UK stock market index—an outcome at the aggregate level. Taken together, our results suggest that individuals use local economic conditions as salient characteristic when forming expectations about a local outcome only.

[26] Again, we follow the non-parametric approach in Hurd, van Rooij, and Winter (2011) to construct a measure of subjective expected stock market returns; for more detail on the summary statistics, see Appendix B7.

2.4 Conclusion

In this paper, we study the role of local economic conditions in shaping the formation of subjective expectations about one-year-ahead local house prices. Using survey data from the UK Financial Lives survey, and exploiting considerable variation in house price changes and economic conditions across local authorities in Great Britain, we find that in addition to perceived past local house price changes, local economic conditions have significant power in predicting expectations about local house prices. Interestingly, there is substantial heterogeneity in belief formation by financial sophistication. Individuals with low financial sophistication complement perceptions of past house price growth with a wider set of local economic indicators in their formation of subjective expectations about future house prices. We conclude that agents' belief formation process is not fully captured by models that only include (recent or more distant) past house price changes as predictors. Instead, agents' beliefs about future house prices react to salient local information as well.

We further find that the weight individuals give to their perceptions and to local economic conditions in their belief formation also depends on features of the local housing market they reside in. In local markets with high past price volatility and where there is no short-run momentum in realized prices, i.e., where past prices are not very informative, individuals rely more on local economic conditions than on their perception of past price changes.

Our findings have novel implications for models of the expectation formation process. Survey-based measures of subjective expectations have been repeatedly found to deviate from full-information rational expectations (e.g., Adam, Pfäuti, and Reinelt 2022; Mankiw, Reis, and Wolfers 2003). Recent studies highlight the benefits of developing empirically-founded models of beliefs that go beyond rational expectations and take into account empirical evidence on subjective expectations (e.g., Coibion, Gorodnichenko, and Kamdar 2018; Barberis, Greenwood, Jin, and Shleifer 2015; Fuster, Laibson, and Mendel 2010; Gelain and Lansing 2014; Kuchler, Piazzesi, and Stroebel 2023). Andre, Pizzinelli, Roth, and Wohlfart (2022) provide evidence that even when individuals have access to similar information about macroeconomic fundamentals, there is substantial heterogeneity in their subjective models resulting in different expectations. It is therefore crucial to empirically measure expectations and to better understand their formation. The evidence provided in the present paper highlights the importance of heterogeneity in belief formation: Individuals rely on perceptions and a wider set of economic indicators, and by how much they rely on various predictors varies with local market dynamics, with information that is likely salient to them, and with financial sophistication.

Appendix B1 Survey Questions and Survey Data

Appendix B1 reports the exact wording of the survey questions on subjective expectations and perceptions about local house price changes and aggregate stock market returns (Figure B.1), a description of all variables (Table B.1), and sample summary statistics (Tables B.2 and B.3). It shows that perceived past one-year house price changes differ from realized ones (Figure B.2). Table B.4 shows the exact wording of the financial literacy questions; financial literacy is *high* if all four standard financial literacy questions are answered correctly, and *low* otherwise. Table B.5 shows that the probability of having high financial literacy is lower among women, those with low education and low household income.

Figure B.1. Measurement of subjective expectations and perceptions about local house price changes and aggregate stock market returns in the Financial Lives survey

RISK4_INTRO [STATE TO ALL]

Now you've seen the examples of the chances of rain next July in Edinburgh and Barcelona, we want to ask you about the chances of different investments making money.

For the next few questions, imagine you receive an unexpected inheritance of £100,000.

RISK5

Imagine you put the £100,000 towards buying **a house in your local area**.

What do you think are the percentage chances that the house will have gone up or down in value by the amounts given below in 12 months' time?

Write a percentage chance in each box-to reflect how likely you think different outcomes are. Make sure your percentages add up to 100%.

	0% *Still to use*	
Rise in house value		% chance of 15.1% or more **rise** in house value
		% chance of 10.1% to 15% **rise** in house value
		% chance of 5.1% to 10% **rise** in house value
		% chance of 0.1% to 5% **rise** in house value
Fall in house value		% chance of 0% to 4.9% **fall** in house value
		% chance of 5% to 9.9% **fall** in house value
		% chance of 10% or more **fall** in house value
Total		**% Note all answers must total to 100%.**

RISK6

Imagine instead you invest the £100,000 in **the FTSE 100, which is the main UK stock market index**.

What do you think are the percentage chances that your stock market investment will have gone up or down in value by the amounts given below in 12 months' time?

Write a percentage chance in each box to reflect how likely you think different outcomes are. Make sure your percentages add up to 100%.

	0% *Still to use*	
Rise in stock market value		% chance of 15.1% or more **rise** in stock market value
		% chance of 10.1% to 15% **rise** in stock market value
		% chance of 5.1% to 10% **rise** in stock market value
		% chance of 0.1% to 5% **rise** in stock market value
Fall in stock market value		% chance of 0% to 4.9% **fall** in stock market value
		% chance of 5% to 9.9% **fall** in stock market value
		% chance of 10% or more **fall** in stock market value
Total		**% Note all answers must total to 100%.**

Figure B.1. Measurement of subjective expectations and perceptions about local house price changes and aggregate stock market returns in the Financial Lives survey (cont.)

RISK9b Now let's imagine you actually received the unexpected £100,000 inheritance **12 months ago**.

If you put your £100,000 towards a house **in your local area**, do you think the value of your £100,000 investment would have increased, decreased or stayed the same over the last 12 months?

1. Increased
2. Stayed the same
3. Decreased

IF CODE 1 SELECTED SHOW THE FOLLOWING

By how much do you think it would have increased? ______ %

IF CODE 3 SELECTED SHOW THE FOLLOWING

By how much do you think it would have decreased? ______ %

RISK9c If instead you invested your £100,000 in the main UK stock market index, called the FTSE 100 index, do you think the value of your £100,000 investment would have increased, decreased or stayed the same over the last 12 months?

1. Increased
2. Stayed the same
3. Decreased

IF CODE 1 SELECTED SHOW THE FOLLOWING

By how much do you think it would have increased? ______ %

IF CODE 3 SELECTED SHOW THE FOLLOWING

By how much do you think it would have decreased? ______ %

Data Source: Financial Lives 2020 survey.

Table B.1. Description of variables

Label	Description
(*a*) *Socio-demographic characteristics*	
	Dummy equal to 1 if …
Female	Female, and 0 if male.
Age 44 or younger	Aged between 18 and 44.
Age 45 to 64	Aged between 45 and 64.
Age 65 or older	Aged 65 or older.
Higher education	One of the following qualifications: (1) Higher degree, or (2) Degree or degree equivalent, or (3) Other Higher Education below degree level.
Lower or medium education	One of the following qualifications: (1) A level, vocational level 3 and equivalents, or (2) Trade Apprenticeships, or (3) O level/ GCSE Grades 4-9/A*-C, vocational level 2 and equivalents, or (4) Qualifications at level 1 and below, or (5) Other qualifications including overseas.
No education	No qualifications, or question about qualifications answered with "don't know."
Education info missing	No information on respondent's education.
Partner in household	Married, in a registered civil partnership, or living with someone in the household as a couple.
Working	Employed or self-employed.
	Total annual household income from all sources (including benefits) before taxes and other deductions …
Annual HH income: less than £20k	…less than £20,000 a year.
Annual HH income: £20k–<£40k	…£20,000 or more but less than £40,000 a year.
Annual HH income: £40k–<£70k	…£40,000 or more but less than £70,000 a year.
Annual HH income: £70k or more	…£70,000 or more a year.
Annual HH income: info missing	No information on respondents' total annual household income.
Holding risky asset	Currently having one of the following investments (either in own name or in joint names): (1) Investment fund, e.g. unit trust, OEIC, ETF, or endowment, or (2) Shares/equities, or (3) Corporate bond or gilt / government bond, or (4) Investment-based crowdfunding, or (5) Peer-to-peer lending, or (6) Structured deposit (sometimes referred to as a savings bond) or structured investment, or (7) Stocks and shares ISA, or (8) Lifetime ISA that is invested, or (9) Insurance bonds (investment bonds), or (10) Innovative Finance ISA (IFISA).
Homeowner	Owning the property currently living in (1) outright, or (2) with a mortgage (or a different kind of loan).
Recent mover	Event of moving house experienced in the last 12 months.
Risk averse	Answer to the question "Are you a person who is generally willing to take risks?" (on a scale from 0 to 10, where 0 is "Not at all willing to take risks" and 10 is "Very willing to take risks") $\leq$ 5.
High financial literacy	All four standard financial-literacy questions are answered correctly.
High interest rate knowledge	Understanding three out of three concepts related to past financial developments and future financial expectations.
Interview finished on same day	Survey questionnaire finished on the same day.
(*b*) *Expectations, perceptions, realizations, and perception gaps*	
Expected 1yr HP change	Non-parametric estimates of the subjective expected rate of one-year change in the house price (HP) index in the individual's local authority, relative to the interview month.
Perceived 1yr HP change	Perceived past one-year house price (HP) change in the individual's local authority, relative to the interview month.
Realized 1yr HP change	Realized past one-year house price (HP) change in the individual's local authority, relative to the interview month.

Continued on next page

Table B.1 (*Continued*)

Label	Description
Absolute perception gap	Difference in absolute terms between realized and perceived past one-year local house price (HP) changes.
Expected 1yr stock market return	Non-parametric estimates of the subjective expectations about one-year ahead returns in the FTSE-100 UK stock market index, relative to the interview month.
Perceived 1yr stock market return	Perceived past one-year return in the FTSE-100 UK stock market index, relative to the interview month.
Realized 1yr stock market return	Realized past one-year return in the FTSE-100 UK stock market index, relative to the interview month.
(*c*) *Local conditions*	
Local unemployment rate	Unemployment rate in the individual's local authority in the interview month; proxied by the number of claimants as a proportion of the resident population of a local area aged 16–64. *Claimants* are defined as the number of people claiming Jobseeker's Allowance, plus those claiming Universal Credit and being out of work. Data from Nomis.
Local deprivation score	Deprivation score in the individual's local authority, taken from the 2019 Index of Multiple Deprivation (comparable data available for England only).
Log property sales	Logarithm of residential property sales per 1,000 inhabitants in the individual's local authority (comparable data available for England and Wales only). Data from: ONS.
5yr local house price volatility	Standard deviation of monthly-computed one-year local changes in house prices in the individual's local authority over the five years before the interview month; definition of "low" and "high" based on sample median split. Data from: UK HPI / HM Land Registry.
High local media coverage of house prices	Dummy equal to 1 if self-constructed measure of the media coverage of the topic of "house prices" in the individual's local authority above median (computed at local-authority level).

Notes: The table shows a detailed description of the variables used in the analyses.

Table B.2. Summary statistics

	N	Mean	Std Dev
(*a*) *Socio-demographic characteristics*			
Female	2,774	0.47	0.50
Age 44 or younger	2,799	0.43	0.49
Age 45 to 64	2,799	0.34	0.48
Age 65 or older	2,799	0.23	0.42
Higher education	2,799	0.63	0.48
Lower or medium education	2,799	0.30	0.46
No education	2,799	0.05	0.22
Education info missing	2,799	0.03	0.16
Partner in household	2,744	0.71	0.45
Working	2,799	0.62	0.49
Annual HH income: less than £20k	2,799	0.15	0.36
Annual HH income: £20k - <£40k	2,799	0.23	0.42
Annual HH income: £40k - <£70k	2,799	0.23	0.42
Annual HH income: £70k+	2,799	0.20	0.40
Annual HH income: info missing	2,799	0.18	0.39
Holding risky asset	2,799	0.42	0.49
Homeowner	2,799	0.72	0.45
Recent mover	2,799	0.09	0.29
Risk averse	2,789	0.42	0.49
High financial literacy	2,799	0.48	0.50
High interest rate knowledge	2,799	0.40	0.49
Interview finished on same day	2,799	0.91	0.28
(*b*) *Expectations, perceptions, realizations, and perception gaps*			
Expected 1yr HP change (%)	2,799	3.71	4.45
Perceived 1yr HP change (%)	2,799	3.79	7.52
Realized 1yr HP change (%)	2,799	1.11	2.68
Realized 1yr HP change (absolute value)	2,799	2.28	1.80
Absolute perception gap (%-points)	2,799	5.04	6.73
Expected 1yr stock market return (%)	2,799	2.13	4.94
Perceived 1yr stock market return	2,799	2.58	8.34
Realized 1yr stock market return	2,799	3.30	3.52
(*c*) *Local conditions*			
Local unemployment rate	2,799	2.73	1.11
Local deprivation score	2,375	20.55	8.12
Log property sales	2,516	2.63	0.23
5yr local house price volatility	2,799	3.53	1.65
High local media coverage of house prices	2,799	0.59	0.49

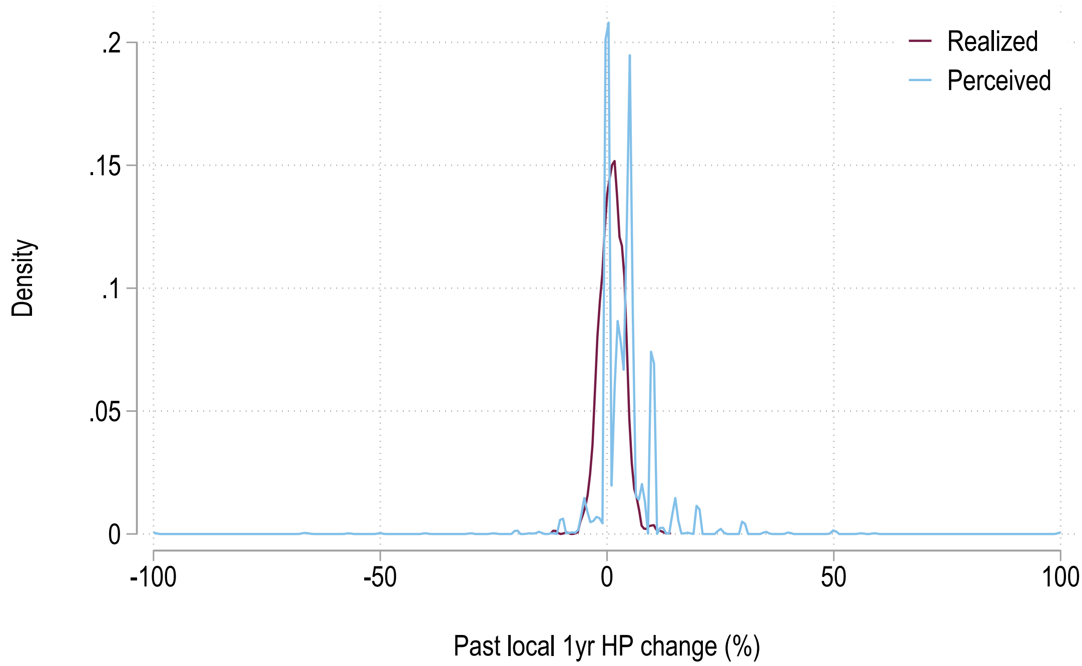
Notes: The table shows summary statistics. For a detailed explanation of the variables, see Table B.1. *Data Source*: Financial Lives 2020 survey, Office for National Statistics, Ministry of Housing, Communities & Local Government, and historical values from the UK HPI and the FTSE-100 Index.

Table B.3. Summary statistics of subjective expected one-year local house price changes

	N	Mean	Std Dev	P10	P25	P50	P75	P90
B1: -10% or less	2,799	2.65	8.87	0	0	0	1	10
B2: -9.9% to -5%	2,799	4.07	8.29	0	0	0	5	10
B3: -4.9% to 0%	2,799	16.13	19.67	0	0	10	25	50
B4: 0.1% to 5%	2,799	45.51	31.69	0	20	49	70	95
B5: 5.1% to 10%	2,799	18.57	22.54	0	0	10	25	50
B6: 10.1% to 15%	2,799	7.48	14.42	0	0	0	10	20
B7: 15.1% or more	2,799	5.59	15.64	0	0	0	5	12

Notes: The table shows summary statistics of the raw subjective probabilities for one-year-ahead expected local house price changes. *Data Source*: Financial Lives 2020 survey.

Figure B.2. Realized and perceived past local one-year house price changes



Notes: The figure shows kernel densities of realized and perceived past local one-year house price (HP) changes in our survey sample. *Data Source*: Financial Lives 2020 survey and UK HPI. *N*=2,799.

Table B.4. Measurement of financial literacy in the Financial Lives survey

Concept	Survey question
Interest rate	Suppose you put £100 into a savings account with a guaranteed interest rate of 2% per year. There are no fees or tax to pay. You don't make any further payments into this account and you don't withdraw any money. How much would be in the account at the end of the first year, once the interest payment is made? Please type in your answer to the nearest pound.
Interest compound	And how much would be in the account at the end of five years (remembering that there are no fees or tax deductions)? 1. More than £110 2. Exactly £110 3. Less than £110 4. It is impossible to tell from the information given 5. Do not know
Inflation	If the inflation rate is 5% and the interest rate you get on your savings is 3%, will your savings have more, less or the same amount of buying power in a year's time? 1. More 2. The same 3. Less 4. Do not know
Risk diversification	Is the following statement true or false? Buying shares in a single company usually provides a safer return than buying shares in a range of companies. 1. True 2. False 3. Do not know

Notes: The table shows the financial literacy questions on interest rates, interest compound, inflation, and risk diversification included in the Financial Lives 2020 survey.

Table B.5. Correlates of financial literacy

Dependent variable	High financial literacy
Female	-0.188^{***}
	(0.018)
Age (ref: 18–44)	
45–64	0.207^{***}
	(0.020)
65 or older	0.206^{***}
	(0.028)
Partner in household	0.025
	(0.021)
Education (ref: Higher)	
Lower or medium	-0.156^{***}
	(0.018)
None	-0.211^{***}
	(0.041)
Info missing	-0.296^{***}
	(0.054)
Working	0.001
	(0.024)
Annual HH income (ref: £70k or more)	
Less than £20k	-0.248^{***}
	(0.032)
£20k - <£40k	-0.137^{***}
	(0.028)
£40k - <£70k	-0.087^{***}
	(0.025)
Info missing	-0.201^{***}
	(0.031)
Local unemployment rate	-0.027^{***}
	(0.008)
Constant	0.599^{***}
	(0.070)
Mean DepVar	0.48
R-squared	0.18
N	2,731
Month Fixed Effects	✓

Notes: The table shows estimates from a linear probability model. The dependent variable is equal to 1 if an individual has high financial literacy (i.e., answers all four standard financial literacy questions, as shown in Table B.4, correctly), and 0 otherwise. 'ref.' indicates the omitted category. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

Appendix B2 Non-Parametric Estimation

Appendix B2 provides a detailed description of how we construct our measure of subjective expected one-year house price changes. We adopt the non-parametric estimation approach suggested in Hurd, van Rooij, and Winter (2011). They construct non-parametric estimates of the mean of the expected rate of return distribution for stock market investments. The model is given by

$$E(\pi) = \sum_j P(\pi \in B_j)E(\pi|\pi \in B_j) \tag{B.1}$$

where, $P(\pi \in B_j)$, is an individual's subjective probability assigned to bracket j, and, $E(\pi|\pi \in B_j)$, is the historical average of one-year rates of return conditional on the return being in bracket j. The number and thresholds of the brackets are pre-determined by the respective survey questions.

Table B.6. Conditional averages of historical house price changes

	Average of historical UK-wide house price changes (1/2002–7/2019)
Brackets:	
B1: -10% or less	−14.21
B2: -9.9% to -5%	−8.36
B3: -4.9% to 0%	−1.54
B4: 0.1% to 5%	2.64
B5: 5.1% to 10%	7.33
B6: 10.1% to 15%	11.95
B7: 15.1% or more	21.77
N	211

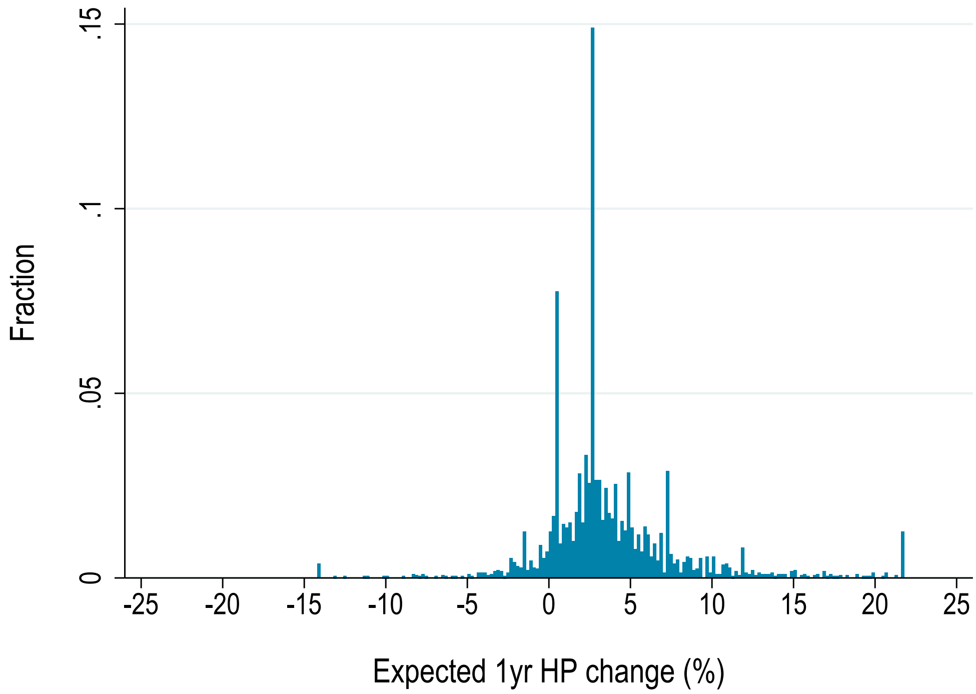
Notes: The table shows bracket-specific averages of monthly computed historical year-on-year house price changes computed at the UK level, $E(r|r \in B_j)$. *Data Source*: UK HPI.

We use the same methodology to construct a measure of the subjective expected house price change from the probabilities corresponding to the seven brackets B_j in our survey question on expected house price changes, where $B_1 = (-\infty, -10\%]$, $B_2 = (-10\%, -5\%]$, $B_3 = (-5\%, 0\%]$, $B_4 = (0\%, 5\%]$, $B_5 = (5\%, 10\%]$, $B_6 = (10\%, 15\%]$, and $B_7 = (15\%, \infty)$.

We compute $E(r|r \in B_j)$ based on historical UK-wide year-on-year house price changes r for each month from the UK House Price Index for the time between January 2002 and July 2019 (the period corresponds to the period before the survey was conducted, and for which annual historical house price changes are available on a monthly basis). We then assign these historical house price changes to the brackets B_j to get bracket-specific average house price changes $E(r|r \in B_j)$. Table B.6 shows the resulting bracket-specific averages of the historical year-on-year house price changes.

Following Equation B.1, we use the bracket-specific averages of the historical year-on-year house price changes, $E(r|r \in B_j)$, and weight them with the respondents' probabilities assigned to the seven different brackets, $P(r \in B_j)$, to estimate the subjective expected house price change. In Figure B.3, we show the corresponding sample distribution.[27]

Figure B.3. Distribution of the expected one-year house price change

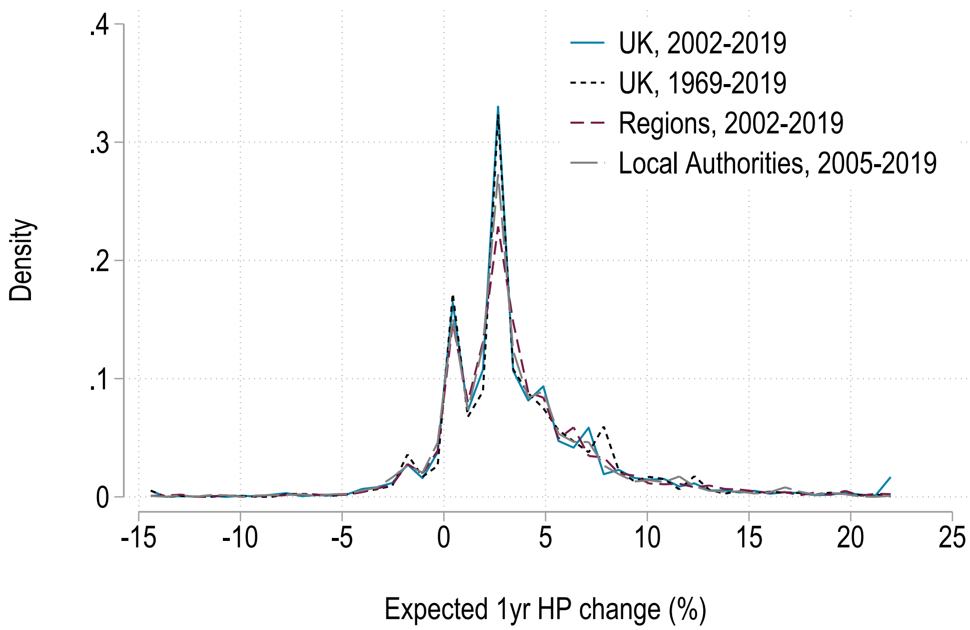


Notes: Computation of expectations, $E(r)$, is based on the non-parametric estimation approach by Hurd, van Rooij, and Winter 2011. *Data Source*: UK HPI and Financial Lives 2020 survey, N=2,799.

[27]There are two outliers; the first outlier is driven by respondents assigning a probability of 50% to the bracket "between −4.9% and 0%," and another 50% to the bracket "between 0.1% and 5%;" the second outlier is driven by respondents assigning a probability of 100% to the bracket "between 0.1% and 5%."

Figure B.4 shows that relying on different specifications of geographic localities (the individual's government office region or local authority instead of the whole UK) and time horizons (1969–2019 instead of 2002–2019) in the computation of $E(r|r \in B_j)$, resulting estimates of subjective expected house price changes are qualitatively unchanged.

Figure B.4. Kernel-density plot of the expected one-year house price change



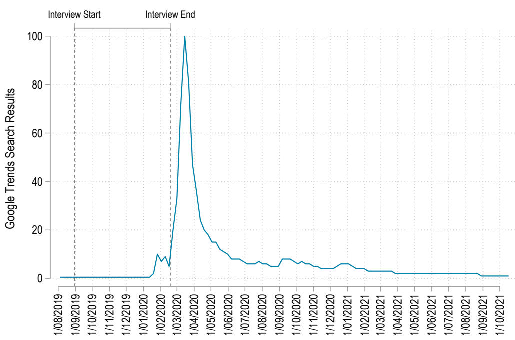
Notes: The figure shows kernel densities for the expected 1yr house price (HP) change, $E(r)$, resulting from different specifications of geographic localities and time horizons in the non-parametric estimation approach by Hurd, van Rooij, and Winter (2011). *Data Source*: UK HPI and Financial Lives 2020 survey, N=2,799.

Appendix B3 Survey Fielding and Pandemic Onset

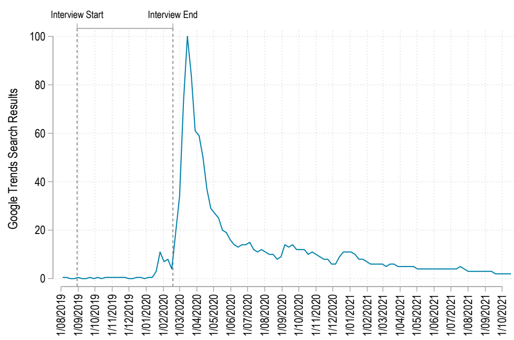
Appendix B3 contains additional information on the survey fielding and pandemic onset. As of February 2020, which constitutes the end of our survey period, there were no more than 47 registered COVID-19 cases in the UK; while the first case in the UK was registered in January 2020, the first COVID-19 related death was registered only on March 2, 2020. The first lockdown in the UK started on March 23, 2020 (https://coronavirus.data.gov.uk/details). In Figure B.5, we report *Google Trends* search results for various search terms related to COVID for the time period during and after which the survey was conducted. Searches of related terms increased rapidly only in March 2020, i.e., *after* the survey was conducted. In contrast, search intensity was close to zero during the time of the survey. It is therefore unlikely that respondents' expectations about future house price changes were distorted by the upcoming pandemic.

Figure B.5. Google Trends search results for COVID-related search terms

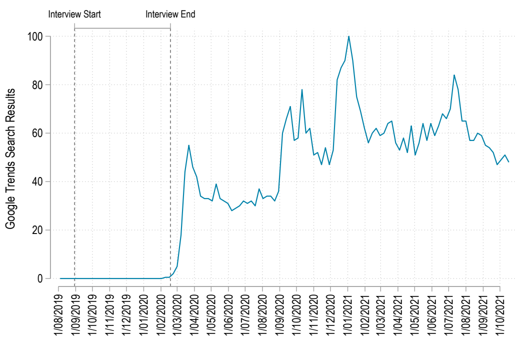
(a) Search term: *Corona*



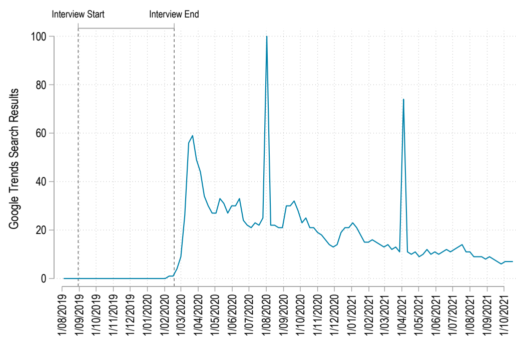
(b) Search term: *Coronavirus*



(c) Search term: *COVID*



(d) Search term: *COVID-19*



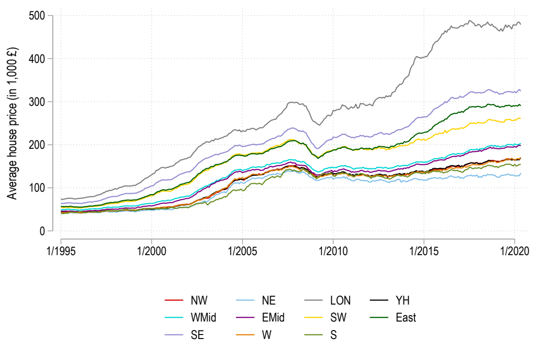
Data Source: https://trends.google.com/trends/?geo=GB, data retrieved at October 20, 2021.

Appendix B4 House Price Dynamics in Great Britain

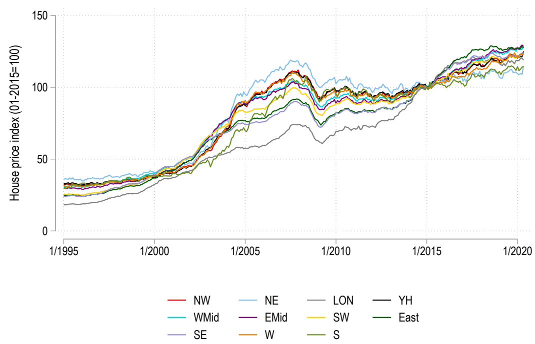
Appendix B4 reports house price dynamics in Great Britain. Figure B.6 reports at the government-office-regional level the development of average house prices and the annual changes in the HPI since 1995. Figure B.7a shows the spatial variation in the annual change in house prices across local authorities as of January 2020, the survey month in which the majority of interviews were conducted; annual house price changes vary between -13.6% in the *City of London* and +12.3% in *Burnley*. Given that respondents of the Financial Lives survey (who live in different areas) were interviewed at different points in time over a period of seven months, our measure of past local house price changes varies not only across different local authorities but also over time. Figure B.7b shows that for the survey period, the average local-authority-level annual house price changes ranges between 0.6% in August 2019 and 1.4% in February 2020.

Figure B.6. Regional house price dynamics in Great Britain

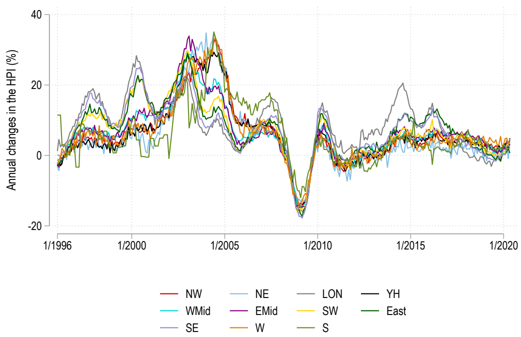
(a) Average house prices



(b) House price index (HPI)



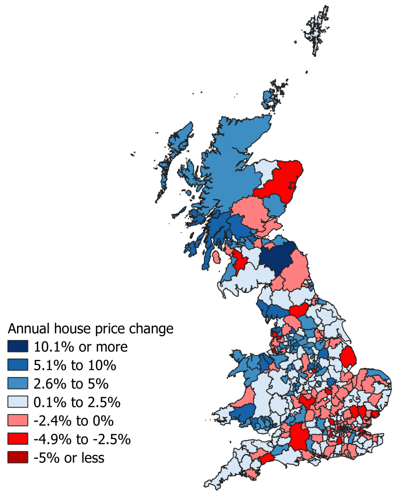
(c) Percentage change in HPI



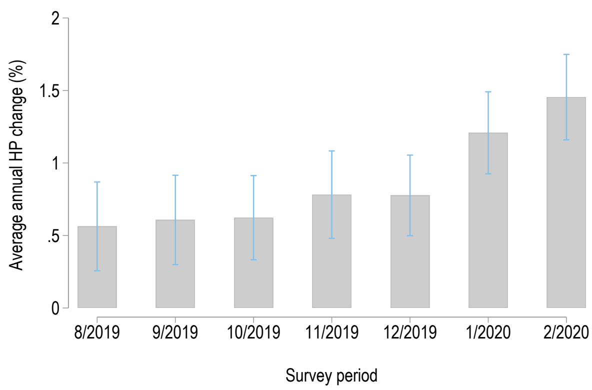
Notes: The figure at the top left shows the development of the average house prices (including all property types) in Wales, Scotland, and the regions in England from 1995. The figure at the top right shows the (quality and composition adjusted) *UK House Price Index* (*HPI*) (January 2015=100). The figure at the bottom left shows the percentage change in the HPI compared to the same period twelve months earlier. *Data Source*: UK HPI.

Figure B.7. Variation in annual house price changes across locality and survey period

(a) Spatial variation (in January 2020)



(b) Time variation (across local authorities)



Notes: The figure at the left shows the spatial variation in annual (year-over-year) house price changes in January 2020 across the local authorities in Great Britain . The figure at the right shows the average and the 95% confidence interval of the annual house price changes across local authorities for the survey period (between August 2019 and February 2020); we treat all local authorities identical (i.e., we compute *unweighted* averages) and we exclusively consider local authorities that are covered in our survey sample, $N = 364$. *Data Source*: UK HPI.

Appendix B5 Local Economic Conditions

Appendix B5 shows that different measures of local economic conditions are strongly positively correlated. Figure B.8a shows that local unemployment rate, our main measure of local economic conditions, is highly positively correlated with the local deprivation score measure. In addition, we investigate the relationship with a different measure—local monetary losses from welfare reforms. After the General Election 2010, the Conservative-Liberal-Democrat coalition government implemented far-reaching welfare cuts via the Welfare Reform Act 2012 and followup reforms. We draw on estimates from Beatty and Fothergill (2016) who provide local-authority-level estimates of the total annual financial benefit loss per working-age adult, separately for pre-2015 and post-2015 welfare reforms. The authors base their estimates on official statistics from the HM Treasury, the Department for Work and Pensions, and the HM Revenue and Customs department, including data on benefit claimants, statistics on welfare-related financial savings, and government's Impact Assessment. Figure B.8b shows that these losses from welfare reforms are strongly positively associated with the unemployment rates in a given local authority.

Figure B.8. Relationship between different measures of local economic conditions

(a) Local unemployment and local deprivation in England

(b) Local unemployment and local monetary loss from welfare cuts in Great Britain

Notes: The figures show the relationship between different measures of local economic conditions. *Data Source*: Office for National Statistics (data from Nomis), Ministry of Housing, Communities & Local Government, and Beatty and Fothergill 2016. Panel (a), N=312 local authorities (information on the index-of-multiple-deprivation score is restricted to LAs in England); panel (b), N=364 local authorities.

Appendix B6 Sensitivity Analyses

Appendix B6 reports a series of sensitivity analyses, using (i) alternative measures of past house price changes, (ii) alternative measures of local economic conditions, and (iii) alternative measures of financial sophistication.

Alternative Measures of Past House Price Changes

In robustness checks, we follow the approach by Malmendier and Nagel (2011) considering different time horizons in the past and allowing different returns in the past to carry different weights. Malmendier and Nagel (2011) define the concept of *weighted average annual returns* of individual i in year t, A_{it}, as a weighted sum of individually experienced annual returns, $R_{i,t-s}$, at time, $t-s$, over an individual's lifetime, S_i:

$$A_{it} = \sum_{s=0}^{S_i-1} w_{i,s}(\lambda) R_{i,t-s}, \tag{B.2}$$

where

$$w_{i,s}(\lambda) = \frac{(S_i - s)^{\lambda}}{\sum_{s=0}^{S_i-1}(S_i - s)^{\lambda}}. \tag{B.3}$$

The weights, $w_{i,s}$, depend on (i) the time horizon of the individual's experience, S_i (short: *lookback period*), (ii) how much time ago, s, the return was realized, and (iii) the weighting parameter, λ. The weighting parameter, λ, allows experiences that have been made at different points in the past to carry different weights. If $\lambda < 0$, more distant returns get higher weights than more recent returns. Conversely, if $\lambda > 0$, more recent returns get higher weights than more distant returns (the weighting function is linear if $\lambda = 1$, concave if $\lambda < 1$, and convex if $\lambda > 1$). If $\lambda = 0$, weights are constant, and, A_{it}, refers to the simple (unweighted) average annual return.

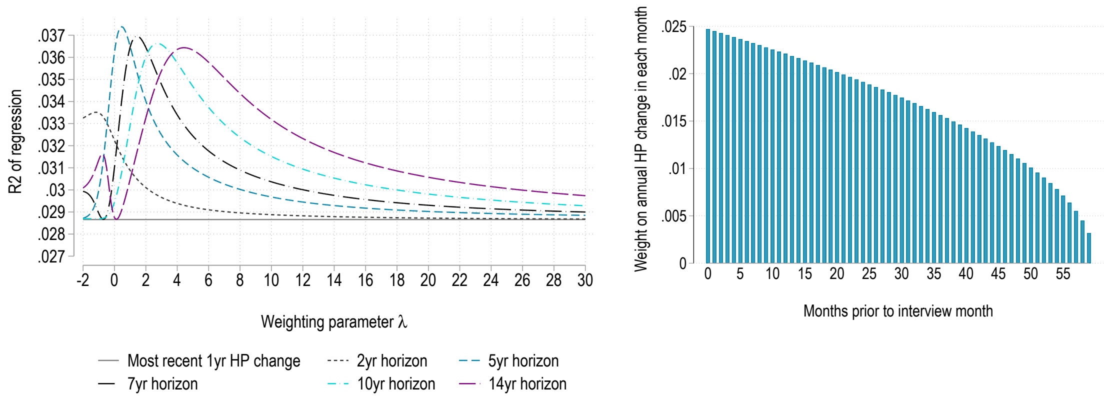
Kuchler and Zafar (2019) apply this concept to house price changes but instead assume a constant lookback period, S, leading to common weights, w_s, across individuals. They determine, S, and, λ, via simulation to maximize their goodness-of-fit in predicting subjective house price expectations. They use zipcode-level annual house price changes to construct a measure of personally experienced returns. We aim to capture local experience and thus construct weighted average annual local house price changes based on past returns at the local-authority level.

We compute the weighted average annual local housing return (where annual returns are computed monthly) as outlined in Equation B.2 for different time horizons S (12, 24, 36, 48, 60, 72, 84, 96, 108, 120, 132, 144, 156, and 168 months);[28] we perform return calculations for a series of different weights λ, ranging from -2 to 30, considering steps of 0.1 (i.e., $-2, -1.9, \ldots, 29.9, 30$). We then select the weighted average annual local housing return (and the underlying combination of S and λ) that yields the highest fit of our model (see Equation 2.2), as measured by the R-squared.

Figure B.9a illustrates the R-squared for each combination of time horizon S and weighting parameter λ when predicting subjective house price expectations. The longer the time horizon, the higher the weighting parameter λ at which R-squared is maximized, indicating that respondents' recent experiences matter more for their future local house price expectations than experiences that lie further in the past. A similar pattern has been identified by Kuchler and Zafar (2019) for the US local housing market. We achieve the best goodness-of-fit for $S = 60$ (months, i.e., a 5-year lookback period) and $\lambda = 0.5$, implying slowly decreasing return weights across the lookback period (see Figure B.9b).

Figure B.9. Model selection and weights on house price changes

(a) Model selection (b) Weights



Notes: Panel (a) shows the R-squared of the regression estimates for local house price changes for each combination of time horizon S and weighting parameter λ . Panel (b) shows weights on the monthly-calculated annual local house price changes for the combination of S and λ that achieves the best goodness-of-fit in predicting subjective house price expectations, that is $S = 60$ (5yr horizon) and $\lambda = 0.5$. *Data Source*: Financial Lives 2020 survey.

[28]Housing return data from the UK HPI are available since 2005 for all local authorities, i.e., for 14 years (168 months) prior interview. We can therefore not consider longer time horizons, e.g., the time horizon since a respondent's birth. $S = 12$ refers to the *most recent available* annual local housing return.

For this specification, the weighted average annual local housing return in the sample is equal to 3.7% (with a standard deviation of 1.5 percentage points).[29] In a different regression specification, we consider realized (annualized) local house price changes over the past five years, similar to Armona, Fuster, and Zafar (2019).[30]

Table B.7 shows that both the coefficients on the weighted average of local house price changes over the past 60 months (column 1) and on the annualized house price change over the past five years (column 2) are positive and statistically significant at the 1% level. It is not implausible that respondents' one-year local house price expectations are positively correlated with realized house price changes over the past five years, but not with realized house price changes over the past one year. Information on how house prices have changed over the past 12 months might not be publicly available that easily and therefore less salient than information on how house prices have evolved on average over the past five years.[31]

Table B.7 also shows that for the different specifications of past local house price changes, the effect of the locally experienced unemployment rates on expected local house price changes remains negative and statistically significant. Repeating the regression analyses with the alternative measures of past local house price experiences for the subgroups of individuals with high and low financial literacy, regression results remain qualitatively unchanged (not shown). While individuals with high financial literacy rely more heavily on past returns (here, over the past 5 years) in the expectation formation process, those with low financial literacy rely more heavily on locally experienced economic conditions.

[29] We also construct weighted average annual local house price changes taking into account only those past annual returns that refer to the month of a year in which a respondent was interviewed. Here, the best fit is achieved for $S = 5$ years and $\lambda = 0.1$. Using this measure of past local house price experiences, results are qualitatively similar.

[30] Note that in the survey, we elicit respondents' *perception* of local house price changes over the past one-year horizon, but not over the past five-year horizon.

[31] Also, since the UK HPI is based on completed sales, it is publicly available only with a delay of three months. For instance, if a respondent is interviewed in January 2020, the last housing return available to the public is from October 2019.

Table B.7. Alternative measures of past house price changes

	Expected 1yr HP change (%)	
	(1)	(2)
Local unemployment rate	−0.190***	−0.185***
	(0.069)	(0.070)
Weighted average of local HP change[a]	0.276***	
	(0.050)	
Realized 5yr HP change		0.233***
		(0.046)
Month Fixed Effects	✓	✓
Socio-demographics	✓	✓
Effect of 1 std in Local unemployment rate	−0.21	−0.21
(in %)	(−5.72)	(−5.58)
Effect of 1 std in Weighted average of local HP change[a]	0.42	
(in %)	(11.31)	
Effect of 1 std in Realized 5yr HP change		0.40
(in %)		(10.77)
Mean DepVar	3.70	3.70
R-squared	0.04	0.04
N	2,731	2,731

Notes: The table shows regression estimates, with *expected 1yr house price (HP) change* as the dependent variable. Realized and expected changes in prices refer to the *local* housing market. Realized 5yr HP change is annualized. [a]Calculation of weighted average of past local house price changes with best fit (see Malmendier and Nagel 2011) for $S = 60$ and $\lambda = 0.5$. *Socio-demographics* include indicators for age categories, education categories, household-income categories, whether respondents are female, married, or working, and whether they finished the interview during one day. Standard errors in parentheses are adjusted for clustering at the *local-authority* level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

Alternative Measures of Local Economic Conditions

Our results are robust to using local deprivation scores as an alternative measure of local economic conditions (see Table B.8). Individuals use local economic conditions as salient characteristics when forming expectations about local house price changes, in particular those with low financial literacy. Since deprivation can only be consistently measured for the subsample residing in England, estimates are less precisely estimated. Similar to our baseline results, individuals extrapolate from past perceived one-year house price changes, and not from realized ones.

Table B.8. Deprivation score as alternative measure of local economic conditions

	Expected 1yr HP change (%)					
			Financial literacy			
	Pooled		High		Low	
	(1)	(2)	(3)	(4)	(5)	(6)
Local deprivation score	−0.015	−0.018*	0.011	0.009	−0.045**	−0.047**
	(0.011)	(0.010)	(0.011)	(0.009)	(0.019)	(0.018)
Realized 1yr HP change	−0.019		−0.022		0.001	
	(0.038)		(0.036)		(0.066)	
Perceived 1yr HP change		0.139***		0.345***		0.092***
		(0.031)		(0.048)		(0.030)
Effect of 1 std in Local deprivation score (in %)	−0.12	−0.14	0.09	0.07	−0.37	−0.38
	(−3.19)	(−3.85)	(2.85)	(2.27)	(−8.45)	(−8.71)
Effect of 1 std in Realized 1yr HP change (in %)	−0.05		−0.05		0.00	
	(−1.21)		(−1.74)		(0.03)	
Effect of 1 std in Perceived 1yr HP change (in %)		1.05		1.60		0.87
		(27.98)		(52.37)		(19.77)
Mean DepVar	3.75	3.75	3.06	3.06	4.41	4.41
R-squared	0.03	0.09	0.05	0.29	0.03	0.06
N	2,324	2,324	1,122	1,122	1,202	1,202
Month Fixed Effects	✓	✓	✓	✓	✓	✓
Socio-demographics	✓	✓	✓	✓	✓	✓

Notes: The table shows regression estimates, with *expected 1yr house price (HP) change* as the dependent variable. Realized, perceived, and expected 1yr changes in prices refer to the *local* housing market. Local unemployment rates refer to the month of interview. *Socio-demographics* include indicators for age categories, education categories, household-income categories, whether respondents are female, married, or working, and whether they finished the interview during one day. Standard errors in parentheses are adjusted for clustering at the *local-authority* level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

Alternative Measures of Financial Sophistication

In robustness checks, we split the sample along a different dimension of financial sophistication—interest rate knowledge. A person's interest rate knowledge is considered *high* if three concepts related to the riskiness of savings accounts, and ranges of past and expected future interest rates are understood, and *low* otherwise. More precisely, if they (i) know that the interest rate on a savings account was not higher than 2% in the year before interview, (ii) think there is no chance of earning an interest rate of 4.1% or more on money kept in a savings account in the year after interview, and (iii) believe that keeping their money in a savings account over the next 12 months will be less risky than investing in the stock or local housing market. Compared to our baseline measure, financial sophistication using this measure is slightly lower with 37% of respondents with high interest rate knowledge.

Using this alternative measure for financial sophistication, we find that results are qualitatively similar to our baseline results (see Table B.9). While individuals with high interest rate knowledge draw more heavily on past house price changes when forming subjective expectations about year-ahead local house prices, individuals with low interest rate knowledge draw less heavily on past house price changes, but instead also take into account local economic conditions.

Table B.9. Interest rate knowledge as alternative measure of financial sophistication

	Expected 1yr HP change (%)					
			IR Knowledge			
	Pooled		High		Low	
	(1)	(2)	(3)	(4)	(5)	(6)
Local unemployment rate	-0.159^{**}	-0.155^{**}	-0.156	-0.147^{*}	-0.214^{**}	-0.211^{**}
	(0.074)	(0.067)	(0.096)	(0.084)	(0.107)	(0.100)
Realized 1yr HP change	-0.001		0.011		-0.004	
	(0.032)		(0.046)		(0.047)	
Perceived 1yr HP change		0.133^{***}		0.379^{***}		0.092^{***}
		(0.029)		(0.029)		(0.027)
Month Fixed Effects	✓	✓	✓	✓	✓	✓
Socio-demographics	✓	✓	✓	✓	✓	✓
Effect of 1 std in Local unemployment rate	−0.18	−0.17	−0.17	−0.16	−0.24	−0.24
(in %)	(−4.78)	(−4.65)	(−5.50)	(−5.19)	(−5.86)	(−5.76)
Effect of 1 std in Realized 1yr HP change	−0.00		0.03		−0.01	
(in %)	(−0.07)		(0.99)		(−0.26)	
Effect of 1 std in Perceived 1yr HP change		0.99		1.67		0.83
(in %)		(26.82)		(53.92)		(20.17)
Mean DepVar	3.70	3.70	3.10	3.10	4.11	4.11
R-squared	0.03	0.08	0.03	0.24	0.04	0.06
N	2,731	2,731	1,097	1,097	1,634	1,634

Notes: The table shows regression estimates, with *expected 1yr house price (HP) change* as the dependent variable. Realized, perceived, and expected 1yr changes in prices refer to the *local* housing market. Local unemployment rates refer to the month of interview. Interest rate (IR) knowledge is *high* if all three concepts related to past financial developments and future financial expectations are answered correctly, and *low* otherwise. *Socio-demographics* include indicators for age categories, education categories, household-income categories, whether respondents are female, married, or working, and whether they finished the interview during one day. Standard errors in parentheses are adjusted for clustering at the *local-authority* level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

Heterogeneity of Subjective Expectations: Gender

We also test for systematic gender differences in beliefs. We find that men have lower perception gaps (see Table B.10) and rely more strongly on their perceptions of past changes, but we do not find consistently that men and women differ in their reliance on local economic conditions.

Table B.10. Heterogeneous estimates of subjective expectations: Gender

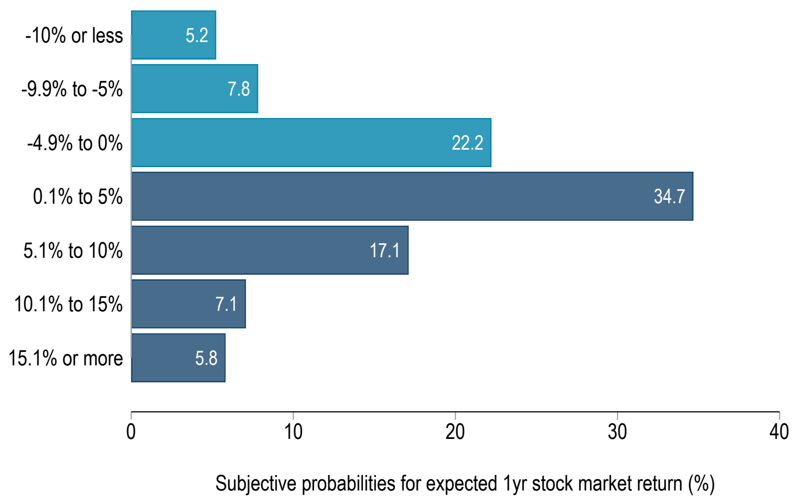
	Dependent variable: Expected 1yr HP change (%)			
	(1)	(2)	(3)	(4)
	Female		Male	
Local unemployment rate	−0.175	-0.218^{**}	-0.163^{*}	−0.107
	(0.111)	(0.110)	(0.095)	(0.085)
Realized 1yr HP change	−0.056		0.052	
	(0.051)		(0.046)	
Perceived 1yr HP change		0.107^{***}		0.186^{***}
		(0.033)		(0.051)
Mean DepVar	3.81	3.81	3.61	3.61
R-squared	0.02	0.06	0.06	0.12
N	1,274	1,274	1,457	1,457

Notes: The table shows regression estimates, with *expected 1yr house price (HP) change* as the dependent variable. Realized, perceived, and expected 1yr changes in prices refer to the *local* housing market. Local unemployment rates refer to the month of interview. In all specifications, we control for socio-demographics and interview-month fixed effects. *Socio-demographics* include indicators for age categories, education categories, household-income categories, whether respondents are female, married, or working, and whether they finished the interview during one day. Note that estimates in columns (2) and (4) are virtually unchanged if we include realized one-year house price changes as a control variable. Standard errors in parentheses are adjusted for clustering at the *local-authority* level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

Appendix B7 Stock Market Returns

Appendix B7 reports summary statistics for stock market return expectations, perceptions, realizations, and perception gaps.

Figure B.10. Subjective expected one-year stock market returns



Notes: The figure shows the distribution of one-year-ahead expected stock market returns, N=2,799. For detailed summary statistics, see Table B.11. *Data Source*: Financial Lives 2020 survey.

Table B.11. Summary statistics: Stock market return expectations, perceptions, and realizations

	Mean	Std Dev	P10	P50	P90
Panel A: Expectations					
Expected 1yr stock market return (%)	2.13	4.94	−1.99	1.76	7.50
Panel B: Perceptions					
Perceived 1yr stock market return (%)	2.58	8.34	−3.00	2.50	10.00
Panel C: Realizations					
Realized 1yr stock market return (%)	3.30	3.52	−1.36	4.55	5.25
Panel D: Perception gap					
Absolute perception gap (%-points)	5.39	7.12	0.45	4.55	10.98
N	2, 799				

Notes: The table shows summary statistics. The absolute perception gap denotes the difference between realized and perceived past one-year stock market returns in absolute terms. Computation of expectations is based on the non-parametric estimation approach by Hurd, van Rooij, and Winter (2011); for details, see Appendix B2. *Data Source*: Financial Lives 2020 survey and historical values from the FTSE-100 Index.

Chapter 3

Subjective Expectations About Joint Return Distributions

Abstract: A large literature studies the formation and heterogeneity of subjective expectations of asset returns and their role in individual financial decision-making. Yet, most of the existing work focuses on subjective return expectations of single assets. In this paper, I use data from a unique survey module to study the formation of subjective expectations with respect to the joint return distribution of a mixed asset, consisting of housing and stock. This task requires individuals not only to form expectations about asset returns but also to incorporate expectations regarding the correlation between those returns—a key aspect of portfolio choice. I find that a non-negligible share of individuals takes into account basic diversification properties only partially or not at all in their expectation formation. More generally, I find that assessing outcomes in terms of probabilities is challenging for many individuals, particularly for those with low financial sophistication and low socio-economic status.

3.1 Introduction

Individuals' financial decisions have crucial implications for their wealth accumulation (van Rooij, Lusardi, and Alessie 2012; Ameriks, Caplin, and Leahy 2003). At the societal and economic level, they also have important consequences for wealth inequality and financial stability (Lusardi, Michaud, and Mitchell 2017; Campbell 2016; Mian and Sufi 2010, 2018; International Monetary Fund 2017). Subjective expectations, such as those regarding future asset returns, have been found to significantly shape individuals' financial decisions (e.g., Hurd, van Rooij, and Winter 2011; Giglio, Maggiori, Stroebel, and Utkus 2021; Kuchler, Piazzesi, and Stroebel 2023), and as a result, a large literature aims at better understanding how individuals form expectations (Manski 2004, 2018; Bachmann, Topa, and van der Klaauw 2023).

Most of the existing work focuses on subjective expectations about *univariate* outcomes, such as returns associated with a single asset. In the context of investment behavior, however, it is essential for better diversification and risk management to consider investment options simultaneously rather than in isolation, taking into account correlations of different investments. A natural next step is therefore to extend the literature by studying subjective expectations about *joint* outcomes. Exploring joint return distributions can contribute to understanding the widely documented lack of diversification in household portfolios (e.g., Blume and Friend 1975; Goetzmann and Kumar 2008; Gomes, Haliassos, and Ramadorai 2021; Badarinza, Campbell, and Ramadorai 2016).

In this paper, I analyze data from a unique survey module that asks individuals to report their subjective expectations about the joint return distribution of an investment in a portfolio consisting of two broad asset classes—housing and stock. This task requires individuals not only to form expectations about asset returns but also to incorporate expectations regarding the correlation between those returns. In addition, the survey elicits individuals' subjective expectations about the return distributions of the assets underlying the portfolio—separately for an investment in housing and for an investment in stocks. The expectation questions included in the survey are designed in a way such that they elicit the whole return distribution using subjective probabilities, similar to Giglio, Maggiori, Stroebel, and Utkus (2021) or Laudenbach, Loos, Pirschel, and Wohlfart (2021). From the subjective return distributions, I can obtain estimates of the mean and standard deviation for each of the three investments (housing, stock, and a two-asset portfolio including both), which is crucial for my analysis.

Drawing on this survey data, I investigate how well individuals' subjective expectations regarding the returns (as measured by the mean of the subjective return distribution) and risks (as measured by the standard deviation of the subjective return distribution) of the three investments align with the following two basic diversification properties: (1) The expected return of a two-asset portfolio lies within the range of the expected returns of its individual underlying assets. (2) The risk of a two-asset portfolio is lower than or equal to the maximum risk associated with its individual underlying assets. I refer to individuals forming subjective expectations in line with the first property as *satisfying the return constraint*. Those aligning with the second property are referred to as *satisfying the risk constraint*. Individuals may form subjective expectations satisfying either one, both, or neither of these constraints.

My results show that a non-negligible share of respondents (one out of four) does *not* provide a response to the probabilistic expectation questions about asset returns.[1] Further, respondents who do provide responses take into account basic diversification properties only partially or not at all in their expectation formation: 50% of respondents satisfy the return constraint, 79% satisfy the risk constraint, and only 41% satisfy both constraints. Respondents' socio-economic status and their overall financial literacy are strong predictors for participating in the expectation-elicitation task and satisfying the return and risk constraint.

I analyze data from a survey module integrated in the 2020 wave of the Financial Lives survey—a nationally representative survey of the UK adult population, conducted by the Financial Conduct Authority (FCA) between August 2019 and February 2020. The survey module was presented to a randomly selected subset of 3,843 individuals, among which 2,926 completed the whole module. In addition to measuring individuals' subjective expected return distributions over the 12-month horizon for three different investments, the survey module elicits individuals' general understanding of the risk and growth potential of different financial investments.

My paper makes two main contributions. First, it contributes to the literature on individual financial literacy by studying how well individuals take into account basic diversification properties in their expectation formation. Although studying individuals' understanding of the concept of risk diversification has been the focus of many research papers (see e.g., Lusardi and Mitchell 2014; van Rooij, Lusardi, and Alessie 2011), most of this research relies on knowledge-based questions integrated in surveys to test people's

[1] Possible reasons for item non-response are discussed in Section 3.4.1.

understanding of the concept of risk diversification. My findings show that individuals with higher education, those with higher household income, and men are more likely to participate in the expectation-elicitation task and to take into account basic diversification properties in their expectation formation. These findings are consistent with well-established patterns in the literature on the correlates of individual financial literacy (Lusardi and Mitchell 2014; Klapper and Lusardi 2020; Lusardi and Mitchell 2023).

Second, my findings contribute to the literature on the elicitation and formation of subjective expectations about asset returns. It does so by studying subjective expectations about the *joint* return distribution for an investment in a "mixed asset," including both stock and real estate. When forming expectations about joint return distributions, survey participants not only have to take into account the return and risk of one single asset but also the *correlation of asset returns*. I use survey data from the UK, a country in which stocks and real estate make up a significant proportion (39.4%) of the total wealth held by households, with real estate being by far the most important component in UK households' portfolios (Badarinza, Campbell, and Ramadorai 2016).[2]

There is a large literature studying individuals' subjective expectations about *stock market returns* (see e.g., Hurd, van Rooij, and Winter 2011; Heiss, Hurd, van Rooij, Rossmann, and Winter 2022; Dominitz and Manski 2007; Kézdi and Willis 2011; Giglio, Maggiori, Stroebel, and Utkus 2021; Merkle and Weber 2014; Drerup, Enke, and von Gaudecker 2017; Amromin and Sharpe 2014; Dominitz and Manski 2011; Hudomiet, Kézdi, and Willis 2011; Vissing-Jorgensen 2003). More recently, researchers have also shown increasing interest in individuals' subjective expectations about *house price changes* (see e.g., Kuchler and Zafar 2019; Armona, Fuster, and Zafar 2019; De Stefani 2021; Kindermann, Le Blanc, Piazzesi, and Schneider 2021; Bailey, Cao, Kuchler, and Stroebel 2018; Bottan and Perez-Truglia 2020; Chopra, Roth, and Wohlfart 2023; Kuchler, Piazzesi, and Stroebel 2023; Kiesl-Reiter, Lührmann, Shaw, and Winter 2024). While the methodology of eliciting subjective expectations about return distributions has been applied to single assets in several studies, I am not aware of survey-based research that measures expectations with respect to the joint return distribution for more than one asset.[3]

The paper proceeds as follows. Section 3.2 provides theoretical background. Section 3.3 introduces the data and expectation measures, and the estimation approach used

[2]Figures (as of 2012) include the main residence, other real estate, directly held stocks, mutual funds, and bonds.

[3]There is related work by Drerup (2019), which is not survey-based but investigates the elicitation of subjective expectations about joint return distributions in a laboratory experiment.

in deriving the moments of the subjective return distributions. Section 3.4 presents summary statistics and information on non-participation in the expectation-elicitation task. Section 3.5 presents my main results of how well individuals' subjective expectations regarding the returns and risks of three different investments align with basic diversification properties, and Section 3.6 examines the robustness of the results. Section 3.7 discusses limitations and implications for future survey design. Section 3.8 concludes.

3.2 Theoretical Background

I study whether individuals' subjective expectations align with the following two basic diversification properties: (1) The expected return of a two-asset portfolio lies within the range of the expected returns of its individual underlying assets. (2) The risk of a two-asset portfolio is lower than or equal to the maximum risk associated with its individual underlying assets. In this section, I summarize the theory showing that these properties hold in the following setting, foundational for the remainder of this paper.

Let us consider an investment of a sum I in two risky assets S (for stock) and H (for housing), forming a portfolio P. Let w_S and w_H be the proportions ("weights") of I invested in assets S and H, respectively. Weights are assumed (i) to take on values that are between 0 and 1, i.e., w_S and $w_H \in [0, 1]$ and (ii) to sum up to 1, i.e., $w_S + w_H = 1$. The investment horizon is one period. The investor only cares about the expected return and risk (as measured by the variance) of the portfolio.

3.2.1 Expected Return and Risk

Let R_S and R_H denote the returns on assets S and H, respectively, and assume that the expected returns, variances, covariance, and correlation of R_S and R_H are as follows:[4]

$$\mu_S = E[R_S], \sigma_S^2 = var(R_S), \mu_H = E[R_H], \sigma_H^2 = var(R_H)$$

$$\sigma_{SH} = cov(R_S, R_H), \rho_{SH} = corr(R_S, R_H)$$

The portfolio's return, R_P, is given by the weighted average of the returns of the individual assets:

$$R_P = w_S R_S + w_H R_H$$

[4]Returns can follow any continuous distribution with a well-defined mean and variance. Notation and calculation of the portfolio's expected return and variance as in Zivot (2021).

The portfolio's expected return, μ_P, and its variance, σ_P^2, are given by:

$$\mu_P = E[R_P] = w_S\mu_S + w_H\mu_H$$

$$\sigma_P^2 = var(R_P) = w_S^2\sigma_S^2 + w_H^2\sigma_H^2 + 2w_Sw_H\sigma_{SH}$$

3.2.2 Implication for Expected Portfolio Return

Since μ_P is a linear combination of μ_S and μ_H, and from the assumptions on the weights w_S and w_H, it follows that[5]

$$min(\mu_S, \mu_H) \leq \mu_P \leq max(\mu_S, \mu_H) \tag{3.1}$$

In other words, Equation 3.1 establishes that the expected return of a portfolio is greater than or equal to the minimum expected return of its individual underlying assets and lower than or equal to the maximum expected return of its individual underlying assets.

3.2.3 Implication for Portfolio Risk

Recall that the risk of the portfolio is measured by its variance:

$$\sigma_P^2 = var(R_P) = w_S^2\sigma_S^2 + w_H^2\sigma_H^2 + 2w_Sw_H\sigma_{SH} \tag{3.2}$$

where σ_P^2 depends on the variance of the returns of the two individual underlying assets, σ_S^2 and σ_H^2, and on the covariance between the returns of the two assets, σ_{SH}.

We can rewrite Equation 3.2 in terms of standard deviation and correlation:

$$\sigma_P^2 = var(R_P) = w_S^2\sigma_S^2 + w_H^2\sigma_H^2 + 2w_Sw_H\rho_{SH}\sigma_S\sigma_H \tag{3.3}$$

where ρ_{SH} is the correlation between the returns of the two individual assets, with $\rho_{SH} \in [-1, 1]$. The smaller the correlation between the returns of assets S and H, the higher the benefits from diversification, and consequently, the lower the risk of the portfolio, cet.par.

[5] By definition, $\mu_P = w_S\mu_S + w_H\mu_H = w_S\mu_S + (1 - w_S)\mu_H$. Since w_S and $w_H \in [0, 1]$, $\mu_P \leq w_S max(\mu_S, \mu_H)(1 - w_S)max(\mu_S, \mu_H)$, and thus $\mu_P \leq max(\mu_S, \mu_H)$. Likewise for $min(\mu_S, \mu_H) \leq \mu_P$.

To show that $\sigma_P^2 \leq max(\sigma_S^2, \sigma_H^2)$, assume without loss of generality that $\sigma_H \geq \sigma_S$. Then,

$$\begin{aligned}
\sigma_P^2 &= w_S^2\sigma_S^2 + w_H^2\sigma_H^2 + 2w_S w_H \rho_{SH}\sigma_S\sigma_H \\
&\leq w_S^2\sigma_S^2 + w_H^2\sigma_H^2 + 2w_S w_H \sigma_S\sigma_H \\
&= (w_S\sigma_S + w_H\sigma_H)^2 \\
&\leq (w_S\sigma_H + w_H\sigma_H)^2 \\
&= (w_S\sigma_H + (1 - w_S)\sigma_H)^2 \\
&= \sigma_H^2.
\end{aligned}$$

It follows that:

$$\sigma_P^2 \leq max(\sigma_S^2, \sigma_H^2) \tag{3.4}$$

In other words, Equation 3.4 establishes that the risk of a portfolio is always lower than or equal to the maximum risk associated with its individual underlying assets.

3.3 Data and Estimation

In this section, I describe the survey data and introduce the survey measures of subjective asset return expectations. Further, I provide a detailed explanation of the estimation approach used in deriving the first two moments of the subjective return distributions.

3.3.1 Financial Lives Survey

For my analysis, I use data from a unique survey module on asset return and risk. In the module, individuals are asked questions about their subjective expectations of the one year-ahead returns associated with an investment in the housing market, and another one in the stock market. What is unique about the module is that it asks a similar question about a joint (pair-wise) outcome, i.e., it elicits individuals' subjective expectations of the return associated with a portfolio investment in the housing *and* stock market. Additionally, the module includes questions measuring individuals' general understanding of the risk and growth potential of different financial investments.

The survey module on asset return and risk was implemented in the 2020 wave of the *Financial Lives survey*—a nationally representative survey conducted by the Financial Conduct Authority (FCA), covering 16,000 adults aged 18 and older living in the UK. The module was presented to a randomized subset of 3,843 participants, with the vast majority (94%) conducting the survey online.[6] The interview process spanned from August 2019 to February 2020, with over half of the participants being interviewed in January 2020.[7]

The Financial Lives survey includes a broad range of information on individuals' socio-demographic and economic background characteristics and attitudes. It also elicits rich information on individuals' use of financial products, their investment behavior, their experiences in dealing with financial products and services, and their financial sophistication.

3.3.2 Measurement of Subjective Return Expectations

The Financial Lives survey measures individuals' subjective expectations of the year-ahead returns associated with three types of investments:

1. Investment of £100,000 in a house in the respondent's local area (H),
2. Investment of £100,000 in the FTSE-100 UK stock-market index (S), and
3. Investment of £100,000 in a portfolio (P), with 50% invested in a house in the respondent's local area (H) and 50% invested in the FTSE-100 UK stock-market index (S).

For all three investments, the survey elicits the subjective return expectations in probabilistic form. This means that respondents are asked to assign subjective probabilities to multiple possible future return brackets in a way such that the probabilities add up to 100%.[8] In the literature, expectations elicited in the form of subjective probabilities are commonly referred to as *probabilistic expectations* (e.g., Manski 2004; Heiss, Hurd, van Rooij, Rossmann, and Winter 2022).

In a first survey question, respondents are asked to imagine that they received an unexpected inheritance of £100,000 which they put towards buying a house in their local area, and to subjectively assess the percentage chances that—in 12 months' time—the house will have *decreased in value* by (i) 10% or more, (ii) 9.9% to 5%, or (iii) 4.9% to 0%; or *increased in value* by (iv) 0.1% to 5%, (v) 5.1% to 10%, (vi) 10.1% to 15%, or (vii) 15.1%

[6] The Financial Lives 2020 survey used a mixed-mode approach in its data collection, combining online and face-to-face interviews. This approach was specifically designed to include individuals with no or infrequent internet access, and those aged 70 or older (Financial Conduct Authority 2021).

[7] For details on the design of the survey module and the survey fielding, see Chapter 2.

[8] While the majority of respondents assigned integer values to the different brackets, they could generally input decimal values up to two decimal places. Respondents could only move on with the next survey question if the probabilities assigned to the different brackets added up to 100%.

or more.[9] A similar expectation question (with the same return brackets and investment amount) was asked for investing in the FTSE-100 UK stock-market index.

Most importantly, the survey additionally elicits the subjective distribution of expected returns for an investment in a portfolio, where respondents are asked to imagine to split the investment and put £50,000 (i.e., half of the total investment amount) towards buying a house in their local area, and another £50,000 in the FTSE-100 UK stock-market index. The return brackets for the portfolio investment are identical to those for the housing and stock market investment.

The probabilistic expectation questions for the three types of investments (H, S, and P) were displayed on different screens, but respondents could revisit previous survey questions to review and potentially modify their responses. Before respondents were asked the expectation questions on asset return, they were shown two examples to familiarize them with the probabilistic question format. The examples were about the *number of days of rain in July* in (i) Edinburgh and in (ii) Barcelona (where it is usually way less rainy than in Edinburgh). For the exact wording of the survey questions and the examples introducing the probabilistic expectation questions, see Figure C.1 in the Appendix.[10]

3.3.3 Non-Parametric Estimation of Moments of Subjective Return Distributions

To analyze my research question of how well individuals' subjective expectations regarding the returns and risks of three different investments align with basic diversification properties, I derive estimates of the first two moments of the subjective return distribution (the mean and the standard deviation) for each of the three investments. As pointed out by Armona, Fuster, and Zafar (2019), we do not know individuals' "mental model" when forming their return expectations. To account for the diverse ways individuals might form their expectations, I consider different approaches in the estimation of the mean and standard deviation of the subjective return distributions. Here, the literature

[9] The design of the survey questions is similar to the one by Giglio, Maggiori, Stroebel, and Utkus (2021) or Laudenbach, Loos, Pirschel, and Wohlfart (2021) who study subjective expectations of stock market returns. The *SCE Housing Survey* fielded by the Federal Reserve Bank of New York adopts a similar question design to elicit subjective expectations of future house price changes (see e.g., Armona, Fuster, and Zafar 2019; Kuchler and Zafar 2019).

[10] For the whole survey questionnaire, see https://www.fca.org.uk/publication/research/financial-lives-survey-2020-questionnaire.pdf.

distinguishes between non-parametric and parametric estimation approaches. In my main specification, I adopt the non-parametric estimation approach suggested in Hurd, van Rooij, and Winter (2011), which relies on incorporating historical realizations of asset returns in the computation of the mean and standard deviation of the subjective return distribution. The approach has been heavily used in the recent literature (see e.g., Giglio, Maggiori, Stroebel, and Utkus, 2021 or Laudenbach, Loos, Pirschel, and Wohlfart, 2021).

In robustness analyses in Section 3.6, I review a range of alternative estimation approaches. Among others, I apply an alternative non-parametric estimation approach where I use midpoints of the return brackets in the survey (following Bailey, Cao, Kuchler, Stroebel, and Wong 2018; Kuchler and Zafar 2019). Further, I consider a parametric estimation approach, fitting a log-normal distribution to the cumulative distribution function of expectations (following Drerup, Enke, and von Gaudecker 2017; Zimpelmann 2021). Given the variety of assumptions underlying the different estimation approaches, my results are fairly robust.

In the following, I explain in more detail the estimation approach by Hurd, van Rooij, and Winter (2011) that I use in my main specification. They construct non-parametric estimates of the mean of the expected rate-of-return distribution for stock market investments. The model is given by

$$E(\pi) = \sum_j P(\pi \in B_j) E(\pi | \pi \in B_j) \tag{3.5}$$

where, $P(\pi \in B_j)$, is the subjective probability assigned to return bracket j, and, $E(\pi|\pi \in B_j)$, is the historical average of one-year rates of return conditional on the return being in bracket j (in the rest of the paper, I refer to these conditional expectations as *bracket points*).

The non-parametric estimate of the standard deviation of the expected rates of return is given by

$$SD(\pi) = \sqrt{E(\pi^2) - E(\pi)^2} \tag{3.6}$$

where

$$E(\pi^2) = \sum_j P(\pi \in B_j) E(\pi^2 | \pi \in B_j) \tag{3.7}$$

I use this methodology to obtain estimates of the mean of the subjective expected return distribution for each of the three investments. In the computation of the standard deviation, I follow Bailey, Cao, Kuchler, Stroebel, and Wong (2018), and assign a standard deviation of zero to respondents who enter probabilities in one return bracket only.

In my analysis, $P(r_x \in B_j)$ denotes a respondent's subjective probability assigned to return bracket j for investment x with $x \in \{H, S, P\}$. Data on $P(r_x \in B_j)$ comes from the Financial Lives survey; for each of the three investments in the survey, there are seven return brackets B_j.

Bracket points are computed as follows: For investment H, I compute $E(r_H|r_H \in B_j)$ based on historical UK-wide year-on-year house price changes r_H for each month from the *UK House Price Index* (*UK HPI*).[11] For investment S, I compute $E(r_S|r_S \in B_j)$ based on historical year-on-year returns for each month from the *Financial Times Stock Exchange* (*FTSE*) *100 Index* using *yahoo!Finance* data on adjusted close prices.[12] For investment P (50/50 investment in H and S), I first compute historical year-on-year portfolio returns, $r_P = 0.5r_H + 0.5r_S$, for each month, which I then use to compute $E(r_P|r_P \in B_j)$. For all three investments, I consider historical returns from May 2003 to July 2019, i.e., I consider the period before the survey was conducted, and I look back in time just long enough to ensure that for each investment type and each return bracket, there is at least one historical return belonging to the bracket.[13]

Table 3.1 reports the bracket points, $E(r|r \in B_j)$, for each of the three investments, derived from the historical returns data.[14] Historically, the FTSE-100 UK stock market index took on more extreme positive and negative return values than the UK HPI, which explains that the derived points for the open-ended brackets are higher (in absolute terms) for the stock market investment than for the housing market investment.

Note that for a given return bracket, derived bracket points for the portfolio investment are not necessarily between the derived points for the stock and housing market investment.[15] To demonstrate, let us examine a simplified example: Consider historical returns over two periods $(t1, t2)$, where the returns were $(-16, 6)$ for a housing

[11] For more information, see https://landregistry.data.gov.uk/app/ukhpi.

[12] Close prices are adjusted for dividends and splits.

[13] The choice to stick to the shortest possible lookback period is motivated by the literature that highlights that individuals' investment decisions, beliefs, and risk attitudes are particularly influenced by *recent* experiences of returns ("recency bias") (Malmendier and Nagel 2011; Malmendier 2021a,b). In robustness analyses in Section 3.6, I also consider alternative time spans.

[14] For more information on the historical returns used in the computation of $E(r|r \in B_j)$, see Figure C.2 in the Appendix.

[15] If in the historical returns data, for each month, the annual return of the FTSE-100 UK stock market index and the UK HPI were in the same return bracket, the derived bracket points for the portfolio would be between the derived points for the stock and housing market investment in each bracket; however, historical returns data shows that this is not the case.

Table 3.1. Conditional averages of historical asset returns

Bracket	Housing	Stock	Portfolio
B1: -10% or less	−14.21	−20.91	−21.05
B2: -9.9% to -5%	−8.36	−7.28	−7.97
B3: -4.9% to 0%	−1.54	−2.54	−1.78
B4: 0.1% to 5%	2.61	2.89	2.84
B5: 5.1% to 10%	7.28	7.50	7.51
B6: 10.1% to 15%	11.81	12.45	11.54
B7: 15.1% or more	17.84	20.67	19.35

Notes: The table shows historical averages of one-year rates of return conditional on the returns being in the respective brackets, $E(r|r \in B_j)$, for three different investments (following the non-parametric estimation approach by Hurd, van Rooij, and Winter 2011). Historical returns data from 5/2003–7/2019. *Data Source*: Historical values from the UK HPI and the FTSE-100 UK stock-market index.

investment and $(-34, -10)$ for a stock-market investment. By calculating the returns for the portfolio investment for each time period using the formula $r_P = 0.5r_H + 0.5r_S$, we find the portfolio returns to be $(-25,-2)$. Calculating bracket points for bracket B1, i.e., averages of the historical returns conditional on the return being -10% or less, we find the outcomes to be -16 for the housing investment, -22 for the stock-market investment, and -25 for the portfolio investment.

3.3.4 Additional Survey Data

Risk aversion. The Financial Lives survey also contains information on individuals' risk aversion. More specifically, it asks individuals to assess their willingness to take risks on a scale of 0 to 10, where 0 is "not at all willing to take risks" and 10 is "very willing to take risks." I define individuals to be "risk averse" if they answer 4 or less.

Financial sophistication. The survey elicits a range of different measures of individuals' financial sophistication. First, it includes the standard financial literacy questions following Lusardi and Mitchell (2011b, 2014), testing individuals' knowledge about (i) interest rates, (ii) interest compounding, (iii) inflation, and (iv) risk diversification. I consider individuals to have "high financial literacy" if they have an understanding of all four financial concepts. Second, the survey measures individuals' knowledge about the relative risk of different assets; I consider individuals to "understand relative asset risk" if they believe that keeping their money in a savings account is less risky than investing in the stock market or in their local housing market. Third, the survey elicits individuals' confidence working with numbers.[16]

3.4 Descriptive Analysis

In this section, I analyze non-response to the probabilistic expectation questions about asset return. Further, I report summary statistics of the elicited subjective asset return distributions, along with summary statistics of the estimated moments of the subjective return distributions.

3.4.1 Non-Response to Subjective Expectation Questions

The survey module on asset return and risk includes 3,843 respondents in total.[17] However, not all of the respondents in the survey module provided a response to the probabilistic expectation questions.[18] Table 3.2 shows that only 76% were able to do so. Even though respondents were shown two introductory examples (about the numbers of days of rain in July in Edinburgh and in Barcelona) that should help clarify the probabilistic question format, one fourth of respondents dropped out of the survey module altogether.[19] There can be a variety of reasons leading to non-response in probabilistic expectation questions in surveys. Some respondents may not provide a response due to unwillingness, while others may lack the cognitive ability to do so.

In Table 3.2, I report regression results from a linear probability model with *providing a response to the probabilistic expectation questions* as the dependent variable. I find that women are less likely than men to provide answers to the probabilistic expectation questions. Compared to individuals aged 35–50, those relatively younger are also less

[16]For the exact wording of the underlying survey questions, see Table C.1 in the Appendix; for the description of the variables used in the analyses, see Table C.2.

[17]Table C.3 in the Appendix reports summary statistics.

[18]In the survey module on asset return and risk, individuals were asked to report expectations in probabilistic form for different investments (a savings account, the local housing market, the FTSE-100 UK stock-market index, and a portfolio). If a respondent indicated an inability to answer the probabilistic expectation question for one investment, they dropped out from the survey module, i.e., they were not asked to report expectations for any of the other investments.

[19]Hudomiet, Kézdi, and Willis (2011)/Binswanger and Salm (2017) study survey data from three/five waves of the Health and Retirement Study and report a similar item non-response rate of 19%/18.1% to the main probabilistic expectation question about stock market returns. Hurd, van Rooij, and Winter (2011) use household survey data to study stock market expectations eliciting subjective probabilities for different return thresholds, and find non-response rates varying between 13% and 21% (across the different waves and return thresholds).

likely to answer the expectation questions. Consistent with findings from Kleinjans and van Soest (2014) and Binswanger and Salm (2017), who find individuals with higher education and cognitive ability to be less likely to refuse answering the probabilistic expectation questions, I also find a negative association between education and non-response. Further, I find that self-employed individuals (as compared to employed ones) are more likely to select themselves into answering the expectation questions. The chance of providing a response also increases with a person's household income. These findings consistently align with well-established patterns in the literature on the correlates of individuals' financial literacy (see Lusardi and Mitchell 2014; Klapper and Lusardi 2020).

I consider two alternative regression specifications, including indicator variables reflecting individuals' financial sophistication (column 2) and their financial behavior (column 3). Individuals' financial sophistication is highly predictive of providing a response to the probabilistic expectation questions: Individuals who are knowledgeable about the concepts of interest compounding, inflation, risk diversification, and relative asset risk, as well as those who consider themselves confident working with numbers, are significantly more likely to provide answers to the probabilistic expectation questions. This is also true for individuals owning their homes or participating in the stock market.

In summary, I find that a non-negligible share of respondents has difficulty dealing with the assessment of different outcomes in terms of probabilities. The sample of respondents that ultimately provide answers to the probabilistic expectation questions is a selected one, characterized by higher socio-economic status (as defined by higher educational attainment, higher income and wealth, and higher financial sophistication).

3.4.2 Summary Statistics of Subjective Return Expectations

Survey responses. In Figure 3.1, I report summary statistics of the survey responses on return expectations for the investments in the local housing market, the FTSE-100 UK stock-market index, and the portfolio containing both (see Panels a, c, and e).[20] For each of the three different investments, I also report historical return frequencies (see Panels b, d, and f). I do so purely for informational purposes; I do not suggest that subjective expected return distributions should mirror historical return distributions.

At the time of interview (end of 2019 and early 2020), respondents were more likely to expect *positive* returns on an investment in the local housing market than in the national

[20]The figure shows subjective probabilities assigned by the average respondent; for detailed summary statistics, see Table C.4 in the Appendix.

Table 3.2. Correlates of responding to probabilistic expectation questions

Dependent variable	Response to probabilistic expectation questions		
	(1)	(2)	(3)
Gender (ref: Male)			
Female	-0.083^{***}	-0.050^{***}	-0.050^{***}
	(0.013)	(0.013)	(0.013)
Other	-0.066	-0.069	-0.062
	(0.060)	(0.066)	(0.065)
Age (ref: 35–50)			
18 to 34	-0.094^{***}	-0.075^{***}	-0.058^{***}
	(0.017)	(0.017)	(0.017)
51 to 64	0.036^{*}	0.008	-0.006
	(0.019)	(0.018)	(0.019)
65 or older	0.005	-0.014	-0.026
	(0.030)	(0.030)	(0.030)
Education (ref: Higher)			
Lower or medium	-0.092^{***}	-0.059^{***}	-0.054^{***}
	(0.014)	(0.014)	(0.014)
None	-0.331^{***}	-0.271^{***}	-0.256^{***}
	(0.025)	(0.025)	(0.025)
Info missing	-0.332^{***}	-0.263^{***}	-0.257^{***}
	(0.031)	(0.032)	(0.032)
Employment status (ref: Employed)			
Self-employed	0.079^{***}	0.070^{***}	0.063^{**}
	(0.026)	(0.026)	(0.026)
Retired	0.061^{**}	0.053^{**}	0.043
	(0.027)	(0.027)	(0.027)
Other	0.001	0.009	0.023
	(0.020)	(0.020)	(0.020)
Annual HH income (ref: £20k - <£40k)			
Less than £20k	-0.056^{***}	-0.036^{*}	-0.023
	(0.021)	(0.021)	(0.021)
£40k - <£70k	0.058^{***}	0.037^{*}	0.032^{*}
	(0.019)	(0.019)	(0.019)
£70k or more	0.080^{***}	0.045^{**}	0.038^{*}
	(0.021)	(0.020)	(0.020)
Info missing	-0.144^{***}	-0.120^{***}	-0.117^{***}
	(0.019)	(0.019)	(0.019)
Risk averse		-0.010	-0.008
		(0.013)	(0.013)
High financial literacy		0.163^{***}	0.153^{***}
		(0.014)	(0.015)
Understands relative asset risk		0.062^{***}	0.060^{***}
		(0.014)	(0.014)
Confident working with numbers		0.032^{**}	0.030^{**}
		(0.013)	(0.013)
Holding shares/equities			0.046^{***}
			(0.016)
Homeowner			0.058^{***}
			(0.015)
Constant	0.906^{***}	0.768^{***}	0.726^{***}
	(0.048)	(0.049)	(0.050)
Mean DepVar	0.76	0.77	0.77
Adj. R-squared	0.18	0.21	0.21
N	3,843	3,806	3,806
Month Fixed Effects	✓	✓	✓

Notes: The table shows estimates from a linear probability model. The dependent variable is equal to 1 if an individual provides a response to the probabilistic expectation questions about asset return, and 0 otherwise. 'ref.' indicates the omitted category. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

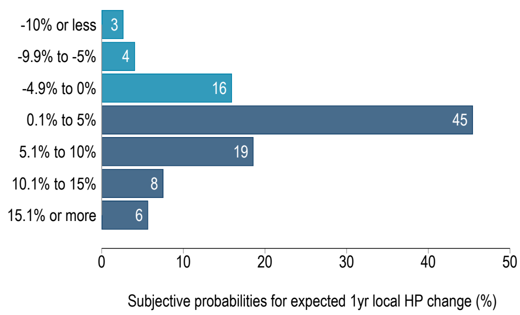
stock market (77.3% vs. 64.7%, when adding up the dark blue bars). For both, modest positive returns (between 0.1% and 5%) were deemed most likely. For the stock market, extreme gains (of 15.1% or more) or extreme losses (of −10% or less) were stated as almost equally likely, with average probabilities around 5–6%. In contrast, for the local housing market, the average probability of extreme gains was (at around 6%) considered twice as large as that of extreme losses. For the portfolio investment it holds true that for each return bracket, the assigned probabilities lie between the probabilities assigned to the respective brackets of the housing and stock market investment. Table C.4 in the Appendix shows that this pattern is also present for the median respondent.

Historically, annual house price changes in the UK have predominantly been positive, with the most frequent occurrences falling within the two middle brackets (0.1% to 5%, and 5.1% to 10%). This pattern is also reflected in people's house price expectations. Historical annual returns of the FTSE-100 UK stock market index in each month are more evenly distributed across the different return brackets. Here, the subjective return distribution differs considerably from the historical return distribution. Compared to historical patterns, respondents assign much more probability mass to the middle brackets and much less to the tails (particularly to the one representing high positive returns) when forming subjective expectations about the return distribution of the FTSE-100. Historical annual returns of the portfolio investment mostly take on positive values. This pattern is also reflected in the expectation formation where respondents assign most of the probability mass to the positive return brackets; however, the distribution of subjective probabilities across the return brackets is quite different from the distribution observed in the historical return data.

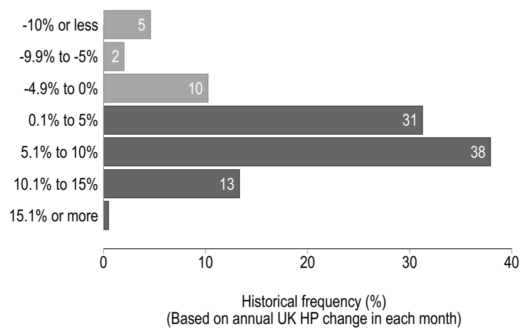
Estimated moments of the subjective return distributions. In Table 3.3, I report summary statistics of the mean and standard deviation of the subjective return distributions that I obtained (for the three different investments) adopting the non-parametric estimation approach by Hurd, van Rooij, and Winter (2011) (as outlined in Section 3.3.3). Respondents expect that, 12 months from the time of their response, there will be a higher mean change for the housing market investment than for the stock market investment, with estimated standard deviations being markedly lower for the housing market investment. The mean and standard deviation of the subjective return distribution of the portfolio investment are between the corresponding values of the single investments on average, but not across all moments (for the distribution of the estimated mean and standard deviation for each of the three investments, see Figure C.3 in the Appendix).

Figure 3.1. Subjective and historical return distributions

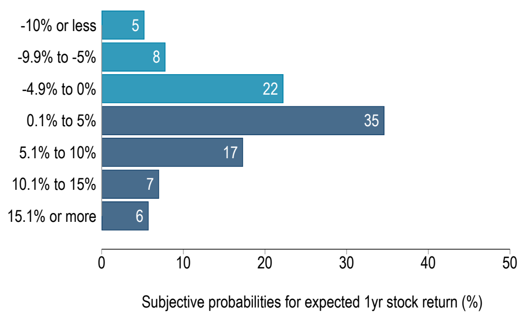
(a) Probability 1yr expected HP change



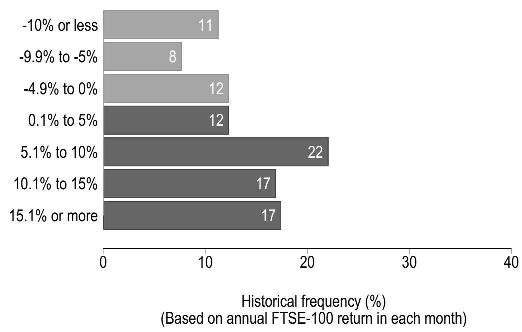
(b) Distribution of historical UK HP change



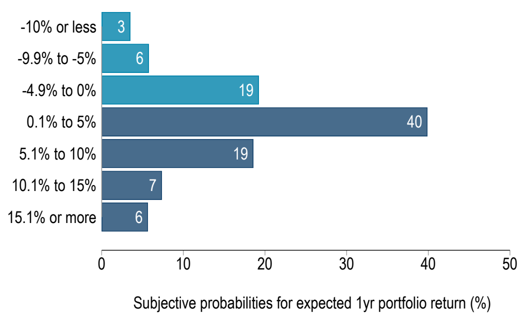
(c) Probability 1yr expected FTSE-100 return



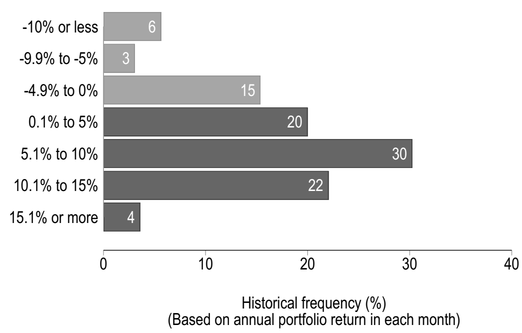
(d) Distribution of historical FTSE-100 return



(e) Probability 1yr expected portfolio return



(f) Distribution of historical portfolio return



Notes: Panels (a/c/e) show the subjective probability distributions of expected 1-year local house-price changes/ FTSE-100 stock-market returns/ portfolio returns. Statistics are based on data from the Financial Lives 2020 survey. N=2,926. Panels (b/d/f) show the historical return frequencies based on monthly-computed annual UK HP changes/ FTSE-100 returns/ portfolio returns. Historical values from the UK HPI and the FTSE-100 UK stock-market index for the period from May 2003 to July 2019. Historical annual portfolio returns calculated for each month as $r_P = 0.5r_H + 0.5r_S$.

Table 3.3. Summary statistics: Non-parametric estimates of the mean and standard deviation of the subjective return distributions

	N	Mean	Std Dev	P10	P50	P90
$E(r_H)$	2,926	3.48	4.04	0.08	2.63	8.15
$E(r_S)$	2,926	2.14	4.91	−1.99	1.76	7.50
$E(r_P)$	2,926	2.94	4.36	−0.15	2.71	7.51
$SD(r_H)$	2,926	3.96	2.79	0.00	3.35	8.07
$SD(r_S)$	2,926	6.08	4.23	0.00	5.15	11.85
$SD(r_P)$	2,926	5.06	3.75	0.00	4.01	10.77

Notes: The table shows summary statistics of the mean, $E(r)$, and standard deviation, $SD(r)$, of the subjective return distributions (obtained using the non-parametric estimation approach by Hurd, van Rooij, and Winter 2011) for three different investments: Housing (H), stock (S), and a two-asset portfolio including both (P). *Data Source*: UK HPI, FTSE-100 UK stock-market index, and Financial Lives 2020 survey, N=2,926.

3.5 Subjective Expectations and Basic Diversification Properties

The primary goal of this paper is to analyze the formation of individuals' subjective expectations about the returns and risks of different investments and the consistency of these expectations with basic diversification properties. First, I study whether individuals form expectations about univariate and joint return distributions in a manner such that the expected return of a two-asset portfolio lies within the range of the expected returns of its individual underlying assets (see Equation 3.1)—I refer to these individuals as "satisfying the return constraint." Second, I study whether individuals form expectations about univariate and joint return distributions in a manner such that the risk of a two-asset portfolio is always lower than or equal to the maximum risk associated with its individual underlying assets (see Equation 3.4)—I refer to these individuals as "satisfying the risk constraint." Ultimately, I am interested in whether individuals form expectations in a manner such that they satisfy both the return *and* the risk constraint.

I first investigate, for each respondent, if the return and risk constraints are generally satisfiable when following the non-parametric estimation approach by Hurd, van Rooij, and Winter (2011). More specifically, I investigate whether there exists (at least) one set of probabilities that can be assigned to the brackets of the portfolio investment such that both the return and the risk constraint are satisfied, considering as given (i) the respondents' set of probabilities assigned to the housing market and stock market investment (taken from the Financial Lives survey), as well as (ii) the bracket points for all three investments (see Table 3.1; obtained from historical returns data following the non-parametric estimation approach by Hurd, van Rooij, and Winter 2011). I do so for all respondents in my sample using the Z3 satisfiability-modulo-theories (SMT) solver (De Moura and Bjørner 2008).[21] For 2,888 of the total 2,926 respondents who provide an answer to the probabilistic expectation questions, there exists a set of probabilities for the portfolio investment such that both constraints can be satisfied; however, for 38 observations, no such set of probabilities exists.[22] Hence, when adopting the non-parametric estimation approach by Hurd, van Rooij, and Winter (2011), it is not possible for both constraints to be satisfied in the case of these 38 observations;[23] as a consequence, I exclude them from my analysis.

Table 3.4 (columns 1–3) reports regression results from a linear probability model with *satisfying the return constraint* as the dependent variable. It shows that only half of the respondents satisfy the return constraint. Even though the sample of individuals who provide a response to the probabilistic expectation questions is already a selected one—characterized by higher income, education, and financial sophistication (see Section 3.4.1), the regression results show that among this selected sample, there are still certain characteristics that predict whether or not an individual satisfies the return constraint: The higher an individuals' household income, the higher their chances of satisfying the return constraint. Similarly, home owners are more likely to satisfy the return constraint. Also, individuals with high financial literacy and those who consider themselves confident working with numbers are more likely to provide subjective probabilities that are consistent with the return constraint. Regarding the employment status, individuals in the category "other" (comprising unemployed and sick ones, those looking after the home, but also students) are significantly less likely to satisfy the return constraint than employed ones. From these results it follows that among the group of individuals that is already characterized by higher socio-economic status there is yet

[21] The Z3 SMT solver allows me to efficiently check satisfiability of both the return and the risk constraint, for all 2,926 observations in my sample.

[22] The majority of these cases refer to respondents who show bunching in their response behavior around the value of 100. If a respondent assigns 100 to the same return bracket for the housing and stock market investment, the respondent would also have to assign 100 to the return bracket for the portfolio investment in order to satisfy the risk constraint. However, since the derived bracket points of the portfolio investment are not necessarily between the respective bracket points of the housing and stock market investment, assigning 100 to one bracket for the portfolio investment can result in not satisfying the return constraint.

[23] In Section 3.6, I discuss the robustness of my results considering alternative approaches in estimating the mean and standard deviation of the subjective return distributions.

another subgroup of individuals that particularly stands out in terms of financial knowledge about asset-related concepts—those owning real estate and having high household income.

Still, the adjusted R-squared reported in Table 3.4 is small, indicating a low overall fit of the model. One likely reason for that is the lack of variability in the regressors (as individuals with certain background characteristics have been more likely to provide answers to the probabilistic expectation questions in the first place, see Section 3.4.1). Due to the resulting sample being very homogeneous in its composition, it is not surprising that the regression performs rather poorly at explaining variation in satisfying the return constraint. Measurement error in the elicitation of expectations and in the computation of the mean and standard deviation of the subjective return distributions might also contribute to explaining the low model fit.

The return constraint is satisfied by 50% of respondents, indicating violation by the remaining half. Violation of the return constraint can occur because respondents either underestimate or overestimate the expected return of the portfolio investment. My results show that 21% of respondents underestimate the expected portfolio return (i.e., they expect the return of the two-asset portfolio to be lower than the minimum expected return of the two individual underlying assets), whereas 29% of respondents overestimate the expected portfolio return (i.e., they expect the return of the two-asset portfolio to be greater than the maximum expected return of the two individual underlying assets). It is reasonable to assume that respondents who expect the return of a portfolio to be below the return of its individual underlying assets have no incentive to spread their investments across different assets (other things equal)—potentially resulting in portfolio underdiversification. The tendency to overestimate the expected portfolio return can equally result in investment mistakes. When individuals expect higher returns for a portfolio than are realistically attainable, it can have severe consequences, especially in the context of long-term investments and retirement planning. Furthermore, when realized portfolio returns fail to meet (unrealistic) expectations, investors may start to question the benefits of diversification.

Table 3.4 (columns 4–6) reports regression results with *satisfying the risk constraint* as the dependent variable. It shows that three out of four respondents satisfy the risk constraint, i.e., they form expectations about three different investments in a way such that Equation 3.4 about asset risk holds. Self-employed individuals (as compared to employed ones) and home owners are more likely to satisfy the risk constraint in their

Table 3.4. Correlates of satisfying the return/risk constraint

Dependent variable	Satisfying the return constraint			Satisfying the risk constraint		
	(1)	(2)	(3)	(4)	(5)	(6)
Gender (ref: Male)						
Female	0.001	0.016	0.012	−0.003	0.009	0.008
	(0.019)	(0.020)	(0.020)	(0.016)	(0.016)	(0.016)
Other	−0.069	−0.125	−0.125	0.108	0.088	0.085
	(0.104)	(0.113)	(0.113)	(0.085)	(0.092)	(0.092)
Age (ref: 35–50)						
18 to 34	0.024	0.028	0.040	−0.023	−0.021	−0.004
	(0.026)	(0.026)	(0.027)	(0.021)	(0.021)	(0.022)
51 to 64	0.051^{*}	0.041	0.038	0.040^{*}	0.030	0.019
	(0.027)	(0.027)	(0.028)	(0.022)	(0.022)	(0.023)
65 or older	0.014	0.011	0.009	0.009	0.002	−0.008
	(0.043)	(0.043)	(0.043)	(0.035)	(0.035)	(0.035)
Education (ref: Higher)						
Lower or medium	0.002	0.012	0.013	−0.010	0.001	0.004
	(0.021)	(0.021)	(0.021)	(0.017)	(0.017)	(0.017)
None	0.021	0.039	0.042	0.024	0.046	0.056
	(0.045)	(0.045)	(0.045)	(0.036)	(0.037)	(0.037)
Info missing	−0.017	−0.017	−0.017	−0.028	−0.024	−0.020
	(0.063)	(0.066)	(0.066)	(0.052)	(0.054)	(0.054)
Employment status (ref: Employed)						
Self-employed	0.062^{*}	0.053	0.054	0.072^{**}	0.067^{**}	0.065^{**}
	(0.037)	(0.037)	(0.037)	(0.030)	(0.030)	(0.030)
Retired	0.045	0.034	0.031	0.004	0.002	−0.004
	(0.038)	(0.038)	(0.038)	(0.031)	(0.031)	(0.031)
Other	-0.078^{**}	-0.078^{**}	-0.067^{**}	−0.004	−0.003	0.011
	(0.033)	(0.033)	(0.033)	(0.027)	(0.027)	(0.027)
Annual HH income (ref: £20k - <£40k)						
Less than £20k	0.013	0.022	0.032	−0.030	−0.022	−0.010
	(0.032)	(0.032)	(0.032)	(0.026)	(0.026)	(0.026)
£40k - <£70k	0.058^{**}	0.053^{*}	0.052^{*}	0.019	0.015	0.011
	(0.028)	(0.028)	(0.028)	(0.022)	(0.022)	(0.022)
£70k or more	0.073^{**}	0.057^{*}	0.056^{*}	0.036	0.025	0.019
	(0.029)	(0.029)	(0.029)	(0.024)	(0.024)	(0.024)
Info missing	0.051^{*}	0.056^{*}	0.061^{**}	-0.043^{*}	−0.038	−0.034
	(0.030)	(0.030)	(0.030)	(0.024)	(0.024)	(0.024)
Risk averse		−0.023	−0.026		−0.012	−0.013
		(0.019)	(0.019)		(0.016)	(0.016)
High financial literacy		0.045^{**}	0.044^{**}		0.045^{***}	0.039^{**}
		(0.021)	(0.021)		(0.017)	(0.017)
Understands relative asset risk		0.004	0.005		0.030^{*}	0.030
		(0.022)	(0.022)		(0.018)	(0.018)
Confident working with numbers		0.045^{**}	0.044^{**}		0.018	0.017
		(0.020)	(0.020)		(0.016)	(0.016)
Holding shares/equities			−0.027			0.016
			(0.023)			(0.019)
Homeowner			0.055^{**}			0.064^{***}
			(0.024)			(0.020)
Constant	0.353^{***}	0.309^{***}	0.277^{***}	0.758^{***}	0.707^{***}	0.662^{***}
	(0.069)	(0.073)	(0.075)	(0.057)	(0.060)	(0.061)
Mean DepVar	0.50	0.50	0.50	0.79	0.79	0.79
Adj. R-squared	0.01	0.01	0.01	0.01	0.01	0.01
N	2,888	2,878	2,878	2,888	2,878	2,878
Month Fixed Effects	✓	✓	✓	✓	✓	✓

Notes: The table shows estimates from a linear probability model. In columns 1–3, the dependent variable is equal to 1 if an individual is "satisfying the return constraint," i.e., forming subjective expectations about three different investments in a way consistent with Equation 3.1. In columns 4–6, the dependent variable is equal to 1 if an individual is "satisfying the risk constraint," i.e., forming subjective expectations about three different investments in a way consistent with Equation 3.4. 'ref.' indicates the omitted category. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

expectations formation. Again, we see that individuals with a good grasp of different financial concepts are also more likely to form subjective expectations in line with the risk constraint.[24]

In total, 21% of respondents do not satisfy the risk constraint, i.e., they expect the risk of a portfolio to exceed that of the individual underlying assets. These individuals clearly lack an understanding of the benefits of diversification for reducing investment risk, potentially translating into investment mistakes such as underdiversification.

Interestingly, the share of respondents who meet the risk constraint is greater than the share of respondents who meet the return constraint. This might be due to the fact that Equation 3.4 only tests a rather loose condition—the risk of a portfolio being always *lower than or equal* to the maximum risk associated with its individual underlying assets. In fact, diversification can considerably reduce the risk of an investment (with the risk of a portfolio being *strictly lower* than the maximum risk of its underlying assets), in particular when the returns of the underlying assets are negatively or only poorly positively correlated. Since respondents' beliefs about the correlation of returns of the two individual assets underlying the portfolio are not explicitly elicited in the Financial Lives survey, Equation 3.4 is the only condition that I can reliably test.

To summarize, I find that 41% of respondents satisfy both constraints (see Table 3.5), 9% only satisfy the return constraint, 38% only satisfy the risk constraint, and 12% satisfy neither of the two constraints. There is substantial heterogeneity by socio-economic status. Notably, among the group of high socio-economic-status individuals who self-select into answering the probabilistic expectation questions, there is yet another subgroup of individuals that stands out when it comes to forming subjective expectations about the return and risk of different investments consistent with basic diversification properties—individuals with higher household income and wealth (acquired through home ownership), those who are self-employed, and those who possess a broad understanding of financial matters.

From the literature we know that individuals with low financial sophistication are more likely to make poor financial decisions (e.g., van Rooij, Lusardi, and Alessie 2011; Disney and Gathergood 2013; Gathergood and Weber 2017; Bianchi 2018), which in turn can adversely affect their financial well-being. In light of this evidence, my analysis reveals a concerning trend: a significant share of individuals struggles to take into account basic diversification properties in their expectation formation. This issue is

[24]The arguments regarding the low model fit raised above also apply here.

Table 3.5. Correlates of satisfying the return and risk constraint

Dependent variable	Satisfying the return and risk constraint		
	(1)	(2)	(3)
Gender (ref: Male)			
Female	−0.006	0.006	0.003
	(0.019)	(0.020)	(0.020)
Other	0.030	−0.031	−0.033
	(0.102)	(0.111)	(0.111)
Age (ref: 35–50)			
18 to 34	−0.007	−0.005	0.011
	(0.026)	(0.026)	(0.026)
51 to 64	0.063^{**}	0.056^{**}	0.048^{*}
	(0.027)	(0.027)	(0.027)
65 or older	0.033	0.032	0.026
	(0.042)	(0.042)	(0.043)
Education (ref: Higher)			
Lower or medium	−0.004	0.003	0.006
	(0.021)	(0.021)	(0.021)
None	0.039	0.052	0.059
	(0.044)	(0.045)	(0.045)
Info missing	−0.037	−0.047	−0.045
	(0.062)	(0.065)	(0.065)
Employment status (ref: Employed)			
Self-employed	0.075^{**}	0.069^{*}	0.068^{*}
	(0.036)	(0.037)	(0.037)
Retired	0.015	0.005	0.000
	(0.038)	(0.038)	(0.038)
Other	-0.064^{**}	-0.063^{**}	−0.049
	(0.032)	(0.032)	(0.033)
Annual HH income (ref: £20k - <£40k)			
Less than £20k	0.010	0.016	0.028
	(0.031)	(0.032)	(0.032)
£40k - <£70k	0.063^{**}	0.059^{**}	0.056^{**}
	(0.027)	(0.027)	(0.027)
£70k or more	0.072^{**}	0.060^{**}	0.056^{**}
	(0.028)	(0.029)	(0.029)
Info missing	0.021	0.024	0.030
	(0.029)	(0.029)	(0.029)
Risk averse		−0.023	−0.025
		(0.019)	(0.019)
High financial literacy		0.037^{*}	0.033
		(0.020)	(0.021)
Understands relative asset risk		0.003	0.003
		(0.022)	(0.022)
Confident working with numbers		0.027	0.026
		(0.019)	(0.019)
Holding shares/equities			−0.009
			(0.022)
Homeowner			0.068^{***}
			(0.024)
Constant	0.269^{***}	0.239^{***}	0.194^{***}
	(0.068)	(0.072)	(0.074)
Mean DepVar	0.41	0.41	0.41
Adj. R-squared	0.01	0.01	0.01
N	2,888	2,878	2,878
Month Fixed Effects	✓	✓	✓

Notes: The table shows estimates from a linear probability model. The dependent variable is equal to 1 if an individual is "satisfying the return and risk constraint," i.e., forming subjective expectations about three different investments in a way consistent with Equation 3.1 and Equation 3.4. 'ref.' indicates the omitted category. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

especially pronounced among individuals with lower socio-economic status. It would be interesting to investigate further whether individuals who do not satisfy the return and risk constraint—in particular those who underestimate the expected return of a portfolio investment and those who overestimate the risk of such—are more likely to hold underdiversified portfolios, which can cause them to incur sizeable return losses (von Gaudecker 2015).[25]

3.6 Robustness

In this section, I study the robustness of my results. First, I take into account whether respondents finished the survey on the same day, and whether they participated in the survey online or face-to-face.[26] Second, I take a closer look at individuals' response behavior in the probabilistic expectation questions: I study how straightlining (i.e., respondents assigning identical probability distributions to the different investments) affects the chances of satisfying the return and/or risk constraint. In the literature on the elicitation of expectations using subjective probabilities, it has also been found that individuals show "bunching"[27] in their answers at the values of 0, 50, and 100 (Kleinjans and van Soest 2014; Hurd 2009; Binswanger and Salm 2017; Bruine de Bruin and Carman 2012). In additional robustness checks, I therefore also study how bunching in the probabilistic expectation questions is associated with satisfying the two constraints.[28] Finally, I study the sensitivity of my results with respect to using different approaches (other non-parametric ones as well as parametric ones) in the estimation of the mean and standard deviation of the subjective return distributions.

[25] For a related discussion on the role of individuals' subjective expectations about the correlation of asset returns on diversification decisions, see Drerup (2019).

[26] In the Financial Lives survey, it was not recorded how long respondents spent on each survey question. Thus, it is unfortunately not possible to conduct robustness checks distinguishing between respondents who took more or less time to answer the probabilistic expectation questions.

[27] Sometimes also referred to as "heaping" or "providing focal-point answers."

[28] It would be interesting to assess respondents' chances of satisfying the constraints by just randomly assigning probabilities to the return brackets of the portfolio investment. To do so, one would have to compute the number of different sets of probabilities that can be assigned to the return brackets of the portfolio investment satisfying the constraints (considering as given the probabilities assigned to the housing and stock investment, and the derived bracket points) as a share of the total number of all possible sets of probabilities. In total, there exist more than 1.7 billion possible different sets of probabilities that can be assigned to the return brackets of the portfolio investment (following the Stars-and-Bars theorem, assuming seven return brackets and respondents assigning integer values only). I refrain from this exercise as it is computationally very demanding.

3.6.1 Survey Duration and Survey Mode

In Table 3.6, I report results from a regression analysis with *satisfying the return and risk constraint* as dependent variable, including additional control variables. I find that completing the survey within one day or over the course of various days plays no role in satisfying the return and risk constraint (see column 1). Also, conducting the survey online or face-to-face does not seem to matter for satisfying the constraints (see column 2). Excluding respondents from the sample who did not finish the survey on the same day, Table 3.7 shows that the percentage of respondents satisfying the constraints is virtually unchanged (as compared to my baseline results); the same is true when excluding respondents who conducted the survey face-to-face. Including control variables for the survey duration and survey mode in the regression analysis does neither cause the sign, the significance levels, nor the size of the previously included regressors to change.

3.6.2 Response Behavior

Straightlining. As outlined in Section 3.3.2, the Financial Lives 2020 survey elicits expectations of the subjective return distribution associated with three different investments using three separate, yet consecutive survey questions with identical return brackets. The literature on survey design highlights an increased risk of respondents providing similar or identical answers when being confronted with a block of survey questions that use the same response scale, potentially reducing data quality (Krosnick 1991; Kim, Dykema, Stevenson, Black, and Moberg 2019). This response strategy, which has come to be known as straightlining or non-differentiation, might also be employed by respondents in the Financial Lives survey when providing answers to the probabilistic expectation questions. I find that 6% of the survey participants report identical probability distributions across all three investments (“straightlining fully”). For 23%, the probability distribution associated with the portfolio investment is identical to the probability distribution of the stock market investment or to the probability distribution of the housing investment (“straightlining partially”) (see Table C.3 in the Appendix). Because the bracket points obtained following the non-parametric estimation approach by Hurd, van Rooij, and Winter (2011) are not identical across the three different investments (see Table 3.1), it does not necessarily follow that respondents exhibiting straightlining behavior automatically satisfy the return and risk constraint.[29]

I control for this type of response behavior in the regression analysis and find that straightlining, both fully and partially, is positively associated with satisfying the return and risk constraint (see columns 3 and 4 in Table 3.6). Given the existing empirical evidence that straightlining is usually applied by lower-educated individuals (Kim, Dykema, Stevenson, Black, and Moberg 2019),[30] it is not implausible to assume that individuals employing this strategy in the Financial Lives survey are classified as satisfying the constraints but are actually lacking a proper understanding of the underlying asset-related concepts. Table 3.7 shows that excluding straightliners from the sample, the percentage of respondents satisfying both constraints drops by 5 percentage points.

Bunching. From the existing literature we know that respondents have a tendency to bunch responses at 0, 50, or 100 when being asked about their expectations using subjective probabilities (Kleinjans and van Soest 2014; Hurd 2009; Binswanger and Salm 2017; Bruine de Bruin and Carman 2012). In a next step, I therefore study the role of bunching in the probabilistic expectation questions when it comes to satisfying the return and risk constraint. More specifically, I identify respondents who assign a percentage of (i) 100 (to any bracket, and 0 to the remaining ones) for all three investments (7% of respondents), or (ii) 50/50 (to any two brackets, and 0 to the remaining ones) for all three investments (3.5% of respondents).[31] Controlling for bunching in the regression analysis (see column 5 in Table 3.6), I find that "bunchers" (irrespective of the type) are significantly more likely to be classified as satisfying the return and risk constraint.[32]

While for some respondents, reported probabilities at 0, 50, and 100 might represent meaningful expectations, for others they can be an expression of uncertainty (Bruine de Bruin, Fischhoff, Millstein, and Halpern-Felsher 2000; Hudomiet and Willis 2013; Hurd

[29] Consider the following example: A respondent employing full straightlining assigns 45% to bracket B3 (-4.9% to 0%) and 55% to bracket B4 (0.1% to 5%), similarly for all three investments. Then, computing the sum of weighted averages as outlined in Equation 3.5 for the housing, stock-market, and portfolio investment (H, S, P) using the obtained bracket points $(-1.54, -2.54, -1.78)$ for B3 and $(2.61, 2.89, 2.84)$ for B4, results in the following expected returns: $(0.74, 0.45, 0.76)$. Since the expected return of the portfolio investment is not within the range of the expected returns of its individual underlying assets, the return constraint is not satisfied.

[30] Similarly, in my survey data I find that straightlining is negatively associated with overall financial literacy (see Table C.6 in the Appendix.)

[31] Across all three investments, bunching of 100 occurred most frequently in the bracket "0.1% to 5%," and bunching of 50/50 occurred most frequently in the brackets "-4.9% to 0%" and "0.1% to 5%."

[32] This is also true when looking at satisfying the return and risk constraint separately (see Table C.5 in the Appendix).

2009; Kleinjans and van Soest 2014), with the latter one being particularly pronounced among individuals with lower numeracy and lower education (Binswanger and Salm 2017; Bruine de Bruin and Carman 2012).[33] Unfortunately, the Financial Lives 2020 survey does not include any questions measuring respondents' (un)certainty when reporting their probabilistic return expectations. Hence, it is difficult to assess whether or not bunchers who satisfy the return and risk constraint have a genuine understanding of the underlying asset-related concepts. Excluding observations characterized by bunching, the percentage of respondents satisfying both constraints drops by 3.4 percentage points to 37.6% (see Table 3.7).

Including control variables for straightlining and bunching in the regression analysis leads to an increase in the R-squared. Also, it becomes evident that controlling for response behavior is essential in order to avoid the introduction of omitted-variable bias: Once I control for straightlining and bunching in the regression analysis, the size of the effect of financial literacy on satisfying the return and risk constraint considerably increases, with significance levels changing from 10% to 1% (see columns 3–5 in Table 3.6).[34]

3.6.3 Alternative Estimation of Moments of Subjective Return Distributions

I further investigate the sensitivity of my results with respect to using alternative established approaches in the literature in the estimation of the mean and standard deviation of the subjective return distributions for the three different investments. In my baseline analysis, I use the non-parametric estimation approach proposed by Hurd, van Rooij, and Winter (2011), drawing on historical returns data from 5/2003 to 7/2019 in the derivation of the bracket points, which are then to be weighted with the respondents' probabilities assigned to the different brackets. To test the sensitivity of my results, I first repeat the approach by Hurd, van Rooij, and Winter (2011) assuming a different time horizon, taking into account all available historical returns data from 3/1985 to 7/2019.[35] Next, I

[33]Similarly, in my survey data I find that bunching is negatively associated with education and overall financial literacy (see Table C.6 in the Appendix).

[34]This is due to financial literacy and straightlining/bunching being both positively associated with satisfying the return and risk constraints, but financial literacy and straightlining/bunching being negatively associated. Note that the patterns of an increased R-squared and an increase in size and improvement in significance of the financial-literacy coefficient are also present when looking at satisfying the return and risk constraint separately (see Table C.5 in the Appendix).

[35]The corresponding bracket points are reported in Table C.7 in the Appendix.

Table 3.6. Satisfying the return and risk constraint: Role of survey duration, survey mode, and response behavior

Dependent variable	Satisfying the return and risk constraint				
	(1)	(2)	(3)	(4)	(5)
Gender (ref: Male)					
Female	0.006	0.007	0.011	0.004	0.009
	(0.020)	(0.020)	(0.019)	(0.019)	(0.019)
Other	−0.032	−0.029	0.003	−0.019	−0.002
	(0.111)	(0.111)	(0.109)	(0.109)	(0.108)
Age (ref: 35–50)					
18 to 34	−0.006	−0.005	−0.003	0.001	−0.003
	(0.026)	(0.026)	(0.025)	(0.025)	(0.025)
51 to 64	0.056^{**}	0.056^{**}	0.041	0.044^{*}	0.049^{*}
	(0.027)	(0.027)	(0.026)	(0.026)	(0.026)
65 or older	0.031	0.023	0.008	0.013	0.018
	(0.043)	(0.043)	(0.042)	(0.042)	(0.042)
Education (ref: Higher)					
Lower or medium	0.004	0.003	0.006	0.001	0.005
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
None	0.053	0.043	0.042	0.050	0.029
	(0.045)	(0.045)	(0.044)	(0.044)	(0.044)
Info missing	−0.047	−0.044	−0.067	−0.065	−0.068
	(0.065)	(0.065)	(0.063)	(0.064)	(0.063)
Employment status (ref: Employed)					
Self-employed	0.069^{*}	0.069^{*}	0.065^{*}	0.063^{*}	0.067^{*}
	(0.037)	(0.037)	(0.036)	(0.036)	(0.036)
Retired	0.005	0.003	0.013	0.003	0.013
	(0.038)	(0.038)	(0.037)	(0.037)	(0.037)
Other	-0.063^{**}	-0.062^{*}	-0.059^{*}	-0.055^{*}	−0.044
	(0.032)	(0.032)	(0.031)	(0.032)	(0.031)
Annual HH income (ref: £20k - <£40k)					
Less than £20k	0.015	0.013	0.020	0.019	0.006
	(0.032)	(0.032)	(0.031)	(0.031)	(0.031)
£40k - <£70k	0.058^{**}	0.058^{**}	0.051^{*}	0.061^{**}	0.054^{**}
	(0.027)	(0.027)	(0.027)	(0.027)	(0.026)
£70k or more	0.060^{**}	0.060^{**}	0.057^{**}	0.064^{**}	0.059^{**}
	(0.029)	(0.029)	(0.028)	(0.028)	(0.028)
Info missing	0.025	0.022	0.020	0.019	0.002
	(0.029)	(0.030)	(0.029)	(0.029)	(0.029)
Risk averse	−0.022	−0.023	−0.028	−0.018	−0.026
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
High financial literacy	0.036^{*}	0.037^{*}	0.053^{***}	0.053^{***}	0.060^{***}
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Understands relative asset risk	0.003	0.004	0.003	0.004	0.010
	(0.022)	(0.022)	(0.022)	(0.022)	(0.021)
Confident working with numbers	0.027	0.027	0.028	0.033^{*}	0.023
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Survey not finished on same day	0.018	0.020	0.018	0.006	0.026
	(0.032)	(0.032)	(0.031)	(0.031)	(0.031)
Survey conducted face-to-face		0.080	0.044	0.040	0.031
		(0.055)	(0.054)	(0.054)	(0.054)
Straightlining, fully			0.437^{***}		
			(0.039)		
Straightlining, partially				0.238^{***}	
				(0.022)	
Bunching 50%/50%					0.400^{***}
					(0.049)
Bunching 100%					0.372^{***}
					(0.039)
Constant	0.238^{***}	0.240^{***}	0.224^{***}	0.193^{***}	0.210^{***}
	(0.072)	(0.072)	(0.071)	(0.071)	(0.070)
Mean DepVar	0.41	0.41	0.41	0.41	0.41
Adj. R-squared	0.01	0.01	0.05	0.05	0.06
N	2,878	2,878	2,878	2,878	2,878
Month Fixed Effects	✓	✓	✓	✓	✓

Notes: The table shows estimates from a linear probability model. The dependent variable is equal to 1 if an individual is “satisfying the return and risk constraint,” i.e., forming subjective expectations about three different investments in a way consistent with Equation 3.1 and Equation 3.4. ‘ref.’ indicates the omitted category. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

Table 3.7. Robustness: Survey duration, survey mode, and response behavior

		Percentage satisfying the		
	N	Return constraint	Risk constraint	Return and risk constraint
(1) Baseline	2,888	49.8	79.1	41.0
(2) Subsample excl. surveys not finished on same day	2,625	49.4	78.9	40.8
(3) Subsample excl. surveys conducted face-to-face	2,783	49.3	79.1	40.7
(4) Subsample excl. straightlining fully	2,721	47.6	78.0	38.5
(5) Subsample excl. straightlining partially	2,240	46.0	74.9	35.7
(6) Subsample excl. bunching 50%/50%	2,787	48.7	78.7	39.7
(7) Subsample excl. bunching 100%	2,724	48.3	77.8	39.1
(8) Subsample excl. bunching 50%/50% or 100%	2,623	47.1	77.3	37.6

Notes: The table shows the percentage of respondents satisfying the return or risk constraint for different subsamples. *Data Source*: Financial Lives 2020 survey.

consider bracket midpoints (instead of deriving bracket points using historical returns data). Following the literature (see e.g., Kuchler and Zafar 2019; Bailey, Cao, Kuchler, Stroebel, and Wong 2018; Giglio, Maggiori, Stroebel, and Utkus 2021), I make assumptions about the points assigned to the open-ended brackets. In a first robustness check, I define points in the open-ended brackets as -12.5 and 17.5, respectively, resulting in intervals of equal width between the points. More aligned with historical returns data, I conduct a second robustness check, where I define points in the open-ended brackets as -20 and +30, respectively. Importantly, for both robustness checks, the bracket points are held constant across all three investments. Finally, I employ a parametric estimation approach: Similar to Drerup, Enke, and von Gaudecker (2017) and Zimpelmann (2021), I fit a log-normal distribution to the cumulative distribution function of expectations to obtain the mean and standard deviation of the subjective return distribution.

For the different estimation approaches, the corresponding results (i.e., the percentage of respondents either satisfying the return constraint, the risk constraint, or both) are summarized in Table 3.8. Three key findings emerge: First, the number of observations for which the constraints are unsatisfiable (i.e., for which there exists no set of probabilities that can be assigned to the portfolio investment such that both constraints are satisfied) varies with the estimation approach (see column "*N* unsatisfiable"). Given the change in the bracket points for the different specifications, this is not surprising. For the specification using bracket midpoints, all observations are satisfiable, which is a direct consequence of having identical bracket points for all three investment types. In the parametric estimation approach, the 123 observations classified as unsatisfiable refer to respondents who put all the probability mass in the open-ended brackets for any of the three investments. Second, the percentage of respondents satisfying the risk constraint is robust across the different specifications, ranging between 78.2% and 80%. Third, and in contrast to the previous point, the percentage of respondents satisfying the return constraint shows high variability with respect to using different estimation approaches (percentages vary between 49.8% and 61.5%), also translating into high variability in the percentage of respondents satisfying both constraints.

Table 3.8. Robustness: Alternative estimation of the mean and standard deviation of the subjective return distributions

				Percentage satisfying the		
		N unsatisfiable	*N*	Return constraint	Risk constraint	Return and risk constraint
(1)	Baseline (approach by Hurd et al. 2011, historical returns from 5/2003–7/2019)	38	2,888	49.8	79.1	41.0
(2)	Approach by Hurd et al. 2011, historical returns from 3/1985–7/2019	24	2,902	50.9	78.4	41.6
(3)	Bracket midpoints with $(-12.5, 17.5)$ in open-ended brackets	0	2,926	61.5	78.2	50.8
(4)	Bracket midpoints with $(-20, 30)$ in open-ended brackets	0	2,926	60.8	79.0	51.3
(5)	Fitting log-normal distribution	123	2,803	58.1	80.0	49.9

Notes: The table shows the percentage of respondents satisfying the return or risk constraint considering alternative approaches in the estimation of the mean and standard deviation of the subjective return distributions. *N unsatisfiable* represents the number of observations for which an analysis of satisfying the constraints cannot be undertaken (for a discussion of satisfiability, see Sections 3.5 and 3.6.3). *Data Source*: Financial Lives 2020 survey.

In the specification using bracket midpoints, the high percentage of respondents satisfying the return constraint is driven by respondents applying "partial straightlining" in their response behavior, i.e., having a probability distribution for the portfolio investment that is identical to either the distribution for the housing or stock market investment, or both. Since in the "midpoint approach" bracket points are identical across all three investments, survey participants applying partial straightlining automatically satisfy both constraints. In Table C.8 in the Appendix, I show results excluding respondents applying partial straightlining: variability across the different specifications in satisfying the return constraint considerably decreases, with values varying between 45.8% and 49.9%.

Table C.6 in the Appendix shows that there is a strong association between individuals' financial literacy and their partial straightlining behavior: Individuals who correctly answer all four standard financial literacy questions (following Lusardi and Mitchell 2011b, 2014) are significantly less likely to exhibit partial straightlining behavior. Furthermore, individuals who answer all of the four financial literacy questions with "don't know" (which can be considered not only a lack of knowledge but also straightlining behavior in a different question type, indicating a general tendency towards straightlining) are significantly more likely to engage in partial straightlining. Hence, I believe that reporting results excluding respondents exhibiting partial straightlining behavior is a legitimate step to take. However, it is crucial to recognize that partial straightlining behavior does not necessarily indicate the adoption of mechanical decision rules; for some respondents, this behavior might well be the result of thorough analyses. Therefore, excluding the whole subsample of respondents adopting partial straightlining can result in underestimating the true percentage of respondents with a proper understanding of basic diversification properties. In Section 3.7, I discuss this issue in more detail, together with implications for future survey design.

To summarize, the results I obtain are to some extent sensitive to (i) the assumptions regarding the approach in the estimation of the mean and standard deviation of the subjective return distribution and (ii) the assumptions regarding the validity of certain response behavior in the probabilistic expectation questions. As pointed out by Armona, Fuster, and Zafar (2019), we do not know individuals' "mental model" when forming their return expectations. From the series of sensitivity checks conducted in Section 3.6, I can conclude that the true percentage of respondents satisfying the return constraint lies in the range between 45.8% and 61.5%, the percentage satisfying the risk constraint lies between 71.7% and 80%, and the percentage satisfying both constraints lies between 35.3% and 51.3%. Given that percentages at the upper end of the ranges are often driven by respondents showing response behavior such as straightlining or bunching, together with the empirical evidence that this kind of response behavior is particularly present among individuals with low education and/or low overall financial literacy (see Table C.6 in the Appendix), the percentage of respondents satisfying the constraints *and* having an actual understanding of the studied concepts of portfolio construction is likely closer to the lower end of the respective ranges.

3.7 Survey Design: Limitations and Implications

The survey module on asset return and risk included in the Financial Lives 2020 survey is unique in that it is the first to elicit subjective expectations about joint return distributions. However, it also exhibits limitations, which I want to discuss here in more detail together with implications for future survey design.

3.7.1 Non-Response and Self Selection

As outlined in Section 3.4.1, a large share of survey participants do not provide a response to the probabilistic expectation questions. Even though the Financial Lives survey includes a set of introductory examples to familiarize respondents with the probabilistic question type (see Figure C.1 in the Appendix), one out of four respondents reports having difficulties forming return expectations using subjective probabilities, and consequently drops out of the survey module altogether. Furthermore, the subsample of respondents that provides answers to the probabilistic expectation questions is a highly selected one, characterized by high socio-economic status.

One question for future research that arises from these results is whether there is a way to reduce item non-response and to mitigate the introduction of selection effects. To address this, a potential approach might be to employ an alternative question format for eliciting subjective expectations. For instance, Delavande and Rohwedder (2008) propose a question format that incorporates visualizations—specifically designed for the elicitation of subjective probability distributions in internet surveys. The main idea behind their visual question format is to provide respondents with a certain number of balls that they then assign to different "containers," which reflect different ranges of a certain variable of interest (in my context, different containers would reflect different ranges of asset returns). The more balls a respondent assigns to a specific container, the more likely they consider an outcome to materialize in the range represented by the container.

The main difference between the visual question format and the format in the Financial Lives survey is that instead of filling in the percent chances for each of the return brackets in a blank field, respondents are provided with a set of balls that they assign to different containers. Especially for individuals who have difficulties dealing with numerical expressions of probabilities, the visual question format might be easier to understand and less demanding. Another advantage of the visual question format is that the number of balls that are yet to be assigned to the different containers is explicitly visible, while in the question format used in the Financial Lives survey, respondents have to perform mental calculations, i.e., they have to keep track of how much of the 100 they have already assigned (and how much are still to be assigned) to the different brackets. Since in the Financial Lives survey, non-response is particularly high among individuals with low education, low financial literacy, and low confidence working with numbers, changing the question format by incorporating visual tools might lead to an increase in response rates among individuals with lower socio-economic status, possibly mitigating the introduction of selection effects.[36] A variation of the visual question format developed by Delavande and Rohwedder (2008) was already adopted successfully in household surveys eliciting individuals' subjective expectations about asset returns (Drerup, Enke, and von Gaudecker 2017; Drerup, Wibral, and Zimpelmann 2023; Zimpelmann 2021).[37]

Another visualization tool that has been applied in the elicitation of asset return expectations involves displaying percent chances assigned by respondents to different return brackets in the form of a histogram (see e.g., Giglio, Maggiori, Stroebel, and Utkus 2021). Thereby, when assigning probabilities to the different brackets, respondents immediately get a visual representation of their subjective probability distribution. However, whether or not this kind of visual aid is powerful enough to reduce non-response (particularly among individuals with lower socio-economic status) is unclear.

In the Financial Lives survey, respondents received a gift card worth £10 once they completed the whole survey. However, participation in the survey module on asset return and risk (including the probabilistic expectation questions) was not incentivized explicitly. Integrating such a financial incentive in the expectation-elicitation task could potentially also contribute to an increase in item response rates. This has been done, for instance, by Drerup, Enke, and von Gaudecker (2017), who provide payoffs to a randomized subset of respondents, with payoffs depending on the accuracy of respondents' expressed expectations about stock market returns.

To summarize, the incorporation of visual aids or financial incentives in the measurement of probabilistic return expectations might reduce non-response (and eventually also selection effects). However, it is important to note that individuals who do not provide answers to the probabilistic return expectation questions in the first place might just not have "well-defined probability distributions in their mind" (Binswanger and Salm 2017).

[36] Delavande and Rohwedder (2008) measure expectations about social security benefits using both the visual question format and the percent-chance format (eliciting cumulative probabilities), and find no significant difference in the non-response rates between the two formats. However, it is not clear to what extent these results can be generalized when eliciting expectations for a different outcome variable or when having a different question design in which percent chances refer to marginal rather than cumulative probabilities.

[37] In that particular household survey, responses to the expectation questions are financially incentivized, which is not the case in the Financial Lives survey data, upon which my analysis is based. Hence, comparing item non-response rates between the two surveys would be misleading.

3.7.2 Straightlining and Bunching

Some respondents apply response strategies such as straightlining or bunching when being asked about their subjective expected return distributions (Section 3.6). In total, 6% (23%) of the respondents are characterized by full (partial) straightlining, and 3% (7%) of the respondents exhibit bunching around 50%/50% (100%). My results show that respondents who exhibit such response strategies are generally more likely to satisfy the return and risk constraint (see Table 3.6). However, whether these respondents do indeed have a better understanding of the underlying asset-related concepts is unclear, as there can be various reasons for observing the above-described response strategies: It may be an expression of epistemic uncertainty, inability, or a lack of motivation (Krosnick 1991; Bruine de Bruin, Fischhoff, Millstein, and Halpern-Felsher 2000; Hurd 2009). For some respondents, however, the provided responses may be an accurate expression of their subjective expectations (Manski 2018). Unfortunately, with the information available in the Financial Lives survey, it is difficult to differentiate between these types of respondents.

For a more in-depth analysis of my results, I would require additional information. For instance, to assess the precision of the elicited subjective expectations data, Drerup, Enke, and von Gaudecker (2017) propose a set of measures which involve asking respondents about how difficult they perceived the expectation-elicitation task, or how certain or confident they were about their answers to the expectations questions. Further, they study the consistency of respondents' expectations. To do so, they elicit and compare (i) point estimates of the expected returns and (ii) the full return distribution using probabilistic assessments.[38] Similar measures are used by Giglio, Maggiori, Stroebel, and Utkus (2021), who study subjective expectations about stock market returns among retail investors. Besides analyzing additional survey data, collecting metadata on the time that respondents spend on answering a probabilistic expectation question could give valuable insights into how much effort of thought they put into it.

[38] Eliciting point estimates in addition to the full distribution of returns in the Financial Lives survey would have allowed me to also directly test Equation 3.1 (i.e., whether or not someone satisfied the return constraint), and to check internal consistency.

3.8 Conclusion

In this paper, I draw on data from a newly designed survey module to study the formation of expectations about univariate and joint return distributions using subjective probabilities. I find that a large share of individuals have difficulty dealing with the assessment of different outcomes in terms of probabilities. Furthermore, a non-negligible share only partially, or not at all, take into account basic diversification properties when forming expectations about the returns and risks of different investments. This is particularly the case for individuals with lower socio-economic status.

Individuals who do not form subjective expectations in line with basic diversification properties—especially those who underestimate the expected return of a portfolio investment and those who overestimate the risk of such—do not have much incentive to diversify their portfolios, which can result in sizeable return losses, particularly for individuals with low financial literacy (von Gaudecker 2015). My results suggest that understanding of probabilities and basic concepts of diversification are important topics to be covered in financial-education programs and information campaigns, which should specifically be targeted towards individuals with lower financial literacy and lower socio-economic status.

The survey data considered in my paper stems from an initial effort to measure expectations about joint return distributions. While the formation of return distributions for single assets is already considered challenging by many individuals, the formation of joint return distributions adds additional complexity as it requires individuals to also take into account the correlation of assets. In my analyses, I find that many survey participants do not provide responses to the expectation questions, or that they resort to simple response strategies, introducing measurement noise. Future efforts to refine and improve the measurement of joint return distributions will therefore be crucial to deepen our understanding and derive more definite conclusions. Providing an extensive training on probabilities before eliciting subjective expectations about joint return distributions could be an interesting next step. Finally, extending this analysis by investigating the role of joint return distributions in determining stock market participation, portfolio choice, and portfolio diversification is another important avenue for future work.

Appendix C1 Survey Questions and Survey Data

Appendix C1 contains information on the exact wording of the survey questions on subjective asset return expectations (Figure C.1) and financial literacy (Table C.1), a description and summary statistics of all the variables used in the analyses (Tables C.2 and C.3), and detailed summary statistics of the survey responses on subjective asset return expectations (Table C.4).

Figure C.1. Measurement of subjective expectations in the Financial Lives survey

RISK_INTRO2_2 **[STATE TO ALL]**

This example shows what we think the chances are of different numbers of days of **rain in Edinburgh** next July.

We don't know for sure how much it will rain, but we can give our best estimate.

0%
Still to use

0 to 5 days	5	% ←	5% chance of 0 to 5 days of rain
6 to 10 days	10	% ←	10% chance of 6 to 10 days of rain
11 to 15 days	30	% ←	30% chance of 11 to 15 days of rain
16 to 20 days	35	%	and so on…
21 to 25 days	15	%	
26 or more days	5	%	
TOTAL	**100**	**%**	

Note all answers must total to 100%.

RISK_INTRO2B **[STATE TO ALL]**

Here is another weather example. Again, you just need to read this.

This shows what we think the chances are of different numbers of days of **rain in Barcelona** next July (usually less rainy than Edinburgh!)

5%
Still to use

0 to 5 days	80	% ←	80% chance of 0 to 5 days of rain
6 to 10 days	10	% ←	10% chance of 6 to 10 days of rain
11 to 15 days	5	% ←	5% chance of 11 to 15 days of rain
16 to 20 days		%	and so on…
21 to 25 days		%	
26 or more days		%	
TOTAL	**95**	**%**	

Note all answers need to total to 100%. In this example, 5% points are still to be used.

Figure C.1. Measurement of subjective expectations in the Financial Lives survey (cont.)

RISK4_INTRO [STATE TO ALL]

Now you've seen the examples of the chances of rain next July in Edinburgh and Barcelona, we want to ask you about the chances of different investments making money.

For the next few questions, imagine you receive an unexpected inheritance of £100,000.

RISK5

Imagine you put the £100,000 towards buying **a house in your local area**.

What do you think are the percentage chances that the house will have gone up or down in value by the amounts given below in 12 months' time?

Write a percentage chance in each box-to reflect how likely you think different outcomes are. Make sure your percentages add up to 100%.

	0% *Still to use*	
Rise in house value		% chance of 15.1% or more **rise** in house value
		% chance of 10.1% to 15% **rise** in house value
		% chance of 5.1% to 10% **rise** in house value
		% chance of 0.1% to 5% **rise** in house value
Fall in house value		% chance of 0% to 4.9% **fall** in house value
		% chance of 5% to 9.9% **fall** in house value
		% chance of 10% or more **fall** in house value
Total		**% Note all answers must total to 100%.**

RISK6

Imagine instead you invest the £100,000 in **the FTSE 100, which is the main UK stock market index**.

What do you think are the percentage chances that your stock market investment will have gone up or down in value by the amounts given below in 12 months' time?

Write a percentage chance in each box to reflect how likely you think different outcomes are. Make sure your percentages add up to 100%.

	0% *Still to use*	
Rise in stock market value		% chance of 15.1% or more **rise** in stock market value
		% chance of 10.1% to 15% **rise** in stock market value
		% chance of 5.1% to 10% **rise** in stock market value
		% chance of 0.1% to 5% **rise** in stock market value
Fall in stock market value		% chance of 0% to 4.9% **fall** in stock market value
		% chance of 5% to 9.9% **fall** in stock market value
		% chance of 10% or more **fall** in stock market value
Total		**% Note all answers must total to 100%.**

Figure C.1. Measurement of subjective expectations in the Financial Lives survey (cont.)

RISK7

Now imagine you decide to split your investment and put:

- £50,000 towards buying a house in your local area and
- £50,000 in the main UK stock market index, called the FTSE 100 index

What do you think are the percentage chances that this total investment will have gone up or down in value by the amounts given below in 12 months' time?

Write a percentage chance in each box to reflect how likely you think different outcomes are. Make sure your percentages add up to 100%.

0%
Still to use

Rise in investment value		% chance of 15.1% or more **rise** in investment value
		% chance of 10.1% to 15% **rise** in investment value
		% chance of 5.1% to 10% **rise** in investment value
		% chance of 0.1% to 5% **rise** in investment value
Fall in investment value		% chance of 0% to 4.9% **fall** in investment value
		% chance of 5% to 9.9% **fall** in investment value
		% chance of 10% or more **fall** in investment value
TOTAL		% **Note all answers must total to 100%.**

Data Source: Financial Lives 2020 survey.

Table C.1. Measurement of financial literacy in the Financial Lives Survey

Concept	Survey question
Interest rate	Suppose you put £100 into a savings account with a guaranteed interest rate of 2% per year. There are no fees or tax to pay. You don't make any further payments into this account and you don't withdraw any money. How much would be in the account at the end of the first year, once the interest payment is made? Please type in your answer to the nearest pound.
Interest compound	And how much would be in the account at the end of five years (remembering that there are no fees or tax deductions)? 1. More than £110 2. Exactly £110 3. Less than £110 4. It is impossible to tell from the information given 5. Do not know
Inflation	If the inflation rate is 5% and the interest rate you get on your savings is 3%, will your savings have more, less or the same amount of buying power in a year's time? 1. More 2. The same 3. Less 4. Do not know
Risk diversification	Is the following statement true or false? Buying shares in a single company usually provides a safer return than buying shares in a range of companies. 1. True 2. False 3. Do not know

Notes: The table shows the financial literacy questions on interest rates, interest compound, inflation, and risk diversification included in the Financial Lives 2020 survey.

Table C.2. Description of variables

Label	Description
(a) Socio-demographic characteristics	
	Dummy equal to 1 if …
Gender: Female	Female.
Gender: Male	Male.
Gender: Other	Non-binary/gender-fluid, or gender not disclosed.
Age 18 to 34	Aged between 18 and 34.
Age 35 to 50	Aged between 35 and 50.
Age 51 to 64	Aged between 51 and 64.
Age 65 or older	Aged 65 or older.
Education: Higher	One of the following qualifications: (1) Higher degree, or (2) Degree or degree equivalent, or (3) Other Higher Education below degree level.

Continued on next page

Table C.2 (*Continued*)

Label	Description
Education: Lower or medium	One of the following qualifications: (1) A level, vocational level 3 and equivalents, or (2) Trade Apprenticeships, or (3) O level/ GCSE Grades 4-9/A*-C, vocational level 2 and equivalents, or (4) Qualifications at level 1 and below, or (5) Other qualifications including overseas.
Education: None	No qualifications, or question about qualifications answered with "don't know."
Education: Info missing	No information on respondent's education.
Employment status: Employed	Working for an employer (full-time or part-time) or agency work.
Employment status: Self-employed	Being self-employed (full-time or part-time).
Employment status: Retired	Being retired or semi-retired (drawing a pension or other income but still working).
Employment status: Other	Being unemployed, sick/disabled, student, carer, looking after the home, doing voluntary work, or answering "don't know."
	Total annual household income from all sources (including benefits) before taxes and other deductions …
Annual HH income: less than £20k	…less than £20,000 a year.
Annual HH income: £20k–<£40k	…£20,000 or more but less than £40,000 a year.
Annual HH income: £40k–<£70k	…£40,000 or more but less than £70,000 a year.
Annual HH income: £70k or more	…£70,000 or more a year.
Annual HH income: info missing	No information on respondents' total annual household income.
Risk averse	Willingness to take risks ≤ 4, when 0 is "not at all willing to take risks" and 10 is "very willing to take risks."
High financial literacy	All four standard financial-literacy questions (about interest rates, interest compounding, inflation, and risk diversification) are answered correctly.
All financial-literacy questions don't know	All four standard financial-literacy questions (about interest rates, interest compounding, inflation, and risk diversification) are answered with "don't know."
Understands relative asset risk	"Bank/savings account" ranked as least risky investment in a list together with two other investments such as "housing in local area" and "stock market."
Confident working with numbers	Confidence working with numbers equal to 10, when 0 is "not at all confidenct" and 10 is "completely confident."
Holding shares/equities	Owning shares or equities.
Homeowner	Owning the property currently living in (1) outright, or (2) with a mortgage (or a different kind of loan).
(*b*) *Survey duration and mode*	
Survey not finished on same day	Survey questionnaire not finished on the same day.
Survey conducted face-to-face	Survey conducted face-to-face, and 0 if survey conducted online.
(*c*) *Response behavior in probabilistic expectations questions*	
Straightlining, fully	Reported identical probability distributions to the probabilistic expectations questions on returns for all three investments (in the local housing market, the FTSE-100 stock-market index, and a portfolio).
Straightlining, partially	Reported probability distribution associated with the portfolio investment is identical to the probability distribution of the investment in the FTSE-100 stock-market index or to the probability distribution of the investment in the local housing market.
Bunching 50%/50%	Reported probabilities of 50/50 in the probabilistic expectations questions to any two return brackets, and 0 to the remaining ones for all three investments (in the local housing market, the FTSE-100 stock-market index, and a portfolio).
Bunching 100%	Reported a probability of 100 in the probabilistic expectations questions to any return bracket, and 0 to the remaining ones for all three investments (in the local housing market, the FTSE-100 stock-market index, and a portfolio).

Notes: The table shows a detailed description of the variables used in the analyses.

Table C.3. Summary Statistics

	Observations total			Observations with answers to expectation questions			Observations with answers to expectation questions and constraints being satisfiable		
	N	Mean	Std Dev	N	Mean	Std Dev	N	Mean	Std Dev
(a) Socio-demographic characteristics									
Gender: Male	3,843	0.48	0.50	2,926	0.52	0.50	2,888	0.52	0.50
Gender: Female	3,843	0.51	0.50	2,926	0.47	0.50	2,888	0.47	0.50
Gender: Other	3,843	0.01	0.11	2,926	0.01	0.09	2,888	0.01	0.09
Age 18 to 34	3,843	0.27	0.44	2,926	0.23	0.42	2,888	0.23	0.42
Age 35 to 50	3,843	0.29	0.45	2,926	0.30	0.46	2,888	0.30	0.46
Age 51 to 64	3,843	0.22	0.42	2,926	0.24	0.43	2,888	0.24	0.43
Age 65 or older	3,843	0.22	0.41	2,926	0.23	0.42	2,888	0.23	0.42
Higher education	3,843	0.56	0.50	2,926	0.63	0.48	2,888	0.63	0.48
Lower or medium education	3,843	0.31	0.46	2,926	0.30	0.46	2,888	0.30	0.46
No education	3,843	0.08	0.27	2,926	0.05	0.22	2,888	0.05	0.22
Education info missing	3,843	0.05	0.22	2,926	0.03	0.16	2,888	0.02	0.15
Employment status: Employed	3,843	0.54	0.50	2,926	0.55	0.50	2,888	0.55	0.50
Employment status: Self-employed	3,843	0.07	0.25	2,926	0.07	0.26	2,888	0.07	0.26
Employment status: Retired	3,843	0.25	0.43	2,926	0.27	0.44	2,888	0.26	0.44
Employment status: Other	3,843	0.14	0.35	2,926	0.11	0.32	2,888	0.11	0.31
Annual HH income: less than £20k	3,843	0.17	0.38	2,926	0.15	0.36	2,888	0.15	0.36
Annual HH income: £20k - <£40k	3,843	0.22	0.41	2,926	0.23	0.42	2,888	0.23	0.42
Annual HH income: £40k - <£70k	3,843	0.20	0.40	2,926	0.23	0.42	2,888	0.23	0.42
Annual HH income: £70k+	3,843	0.17	0.37	2,926	0.20	0.40	2,888	0.20	0.40
Annual HH income: info missing	3,843	0.24	0.43	2,926	0.18	0.39	2,888	0.18	0.39
Risk averse	3,806	0.44	0.50	2,916	0.42	0.49	2,878	0.42	0.49
High financial literacy	3,843	0.38	0.49	2,926	0.47	0.50	2,888	0.47	0.50
All financial-literacy questions don't know	3,843	0.12	0.32	2,926	0.05	0.22	2,888	0.05	0.22
Understands relative asset risk	3,843	0.71	0.45	2,926	0.75	0.43	2,888	0.76	0.43
Confident working with numbers	3,843	0.38	0.49	2,926	0.42	0.49	2,888	0.42	0.49
Holding shares/equities	3,843	0.22	0.42	2,926	0.27	0.44	2,888	0.27	0.44
Homeowner	3,843	0.66	0.48	2,926	0.72	0.45	2,888	0.72	0.45
(b) Survey duration and mode									
Survey not finished on same day	3,843	0.08	0.27	2,926	0.09	0.29	2,888	0.09	0.29
Survey conducted face-to-face	3,843	0.06	0.23	2,926	0.04	0.19	2,888	0.04	0.19
(c) Response behavior in probabilistic expectations questions									
Straightlining, fully	2,926	0.06	0.24	2,926	0.06	0.24	2,888	0.06	0.23
Straightlining, partially	2,926	0.23	0.42	2,926	0.23	0.42	2,888	0.22	0.42
Bunching 50%/50%	2,926	0.03	0.18	2,926	0.03	0.18	2,888	0.03	0.18
Bunching 100%	2,926	0.07	0.25	2,926	0.07	0.25	2,888	0.06	0.23

Notes: The table shows summary statistics for the total sample of observations, the subsample of observations that provides answers to the probabilistic expectation questions, and the subsample of observations that provides answers to the probabilistic expectation questions and for which the return and risk constraints are satisfiable (for a discussion of satisfiability, see Section 3.5). *Data Source*: Financial Lives 2020 survey.

Table C.4. Summary statistics of subjective expected one-year asset returns

	Mean	SD	P10	P25	P50	P75	P90
(*a*) *Probability 1yr local HP change in bracket* (%)							
B1: -10% or less	2.67	8.83	0	0	0	1	10
B2: -9.9% to -5%	4.09	8.45	0	0	0	5	10
B3: -4.9% to 0%	15.96	19.56	0	0	10	25	50
B4: 0.1% to 5%	45.46	31.76	0	20	48	70	95
B5: 5.1% to 10%	18.59	22.58	0	0	10	25	50
B6: 10.1% to 15%	7.54	14.50	0	0	0	10	20
B7: 15.1% or more	5.69	15.87	0	0	0	5	15
(*b*) *Probability 1yr stock market return in bracket* (%)							
B1: -10% or less	5.24	12.36	0	0	0	5	15
B2: -9.9% to -5%	7.81	11.79	0	0	5	10	20
B3: -4.9% to 0%	22.24	20.24	0	5	20	30	50
B4: 0.1% to 5%	34.61	26.95	0	15	30	50	75
B5: 5.1% to 10%	17.32	20.74	0	0	10	20	45
B6: 10.1% to 15%	7.03	12.18	0	0	0	10	20
B7: 15.1% or more	5.75	15.07	0	0	0	5	15
(*c*) *Probability 1yr portfolio return in bracket* (%)							
B1: -10% or less	3.50	10.07	0	0	0	3	10
B2: -9.9% to -5%	5.77	9.94	0	0	0	10	15
B3: -4.9% to 0%	19.24	18.43	0	3	15	30	50
B4: 0.1% to 5%	39.90	29.03	0	20	40	50	85
B5: 5.1% to 10%	18.56	22.50	0	0	10	25	50
B6: 10.1% to 15%	7.37	14.18	0	0	0	10	20
B7: 15.1% or more	5.67	15.43	0	0	0	5	15

Notes: The table shows summary statistics of the probabilities assigned to the different return brackets in the expectation elicitation for three different investments: (a) Housing, (b) stock, and (c) a two-asset portfolio including both. *Data Source*: Financial Lives 2020 survey, N=2,926.

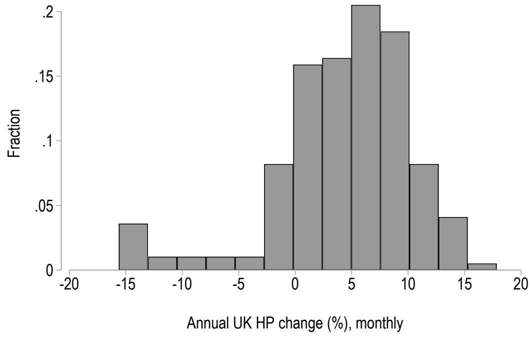
Appendix C2 Non-Parametric Estimation: Supplementary Material

Appendix C2 offers detailed information on the historical returns data utilized in the non-parametric estimation approach by Hurd, van Rooij, and Winter (2011), as well as detailed information on the obtained estimation results. Figure C.2 illustrates the historical trends and distribution of monthly-computed annual returns for the UK House Price Index, the FTSE-100 UK stock market index, and a balanced portfolio comprising equal parts of both. Figure C.3 shows the distribution of the estimated mean and standard deviation of the subjective return distributions of the housing, stock, and portfolio investment.

Figure C.2. Historical asset returns

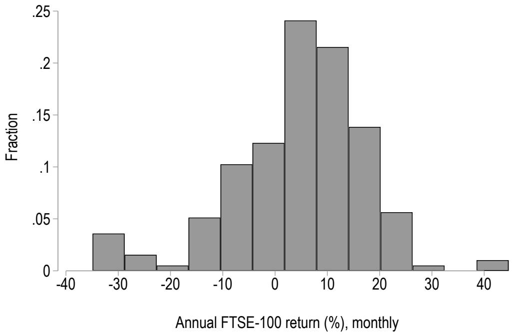
(a) UK HP change over time

(b) Distribution of UK HP change



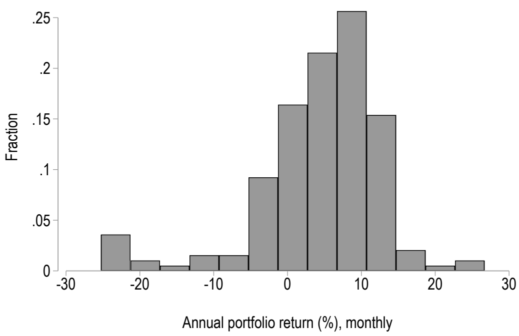
(c) FTSE-100 return over time

(d) Distribution of FTSE-100 return



(e) Portfolio return over time

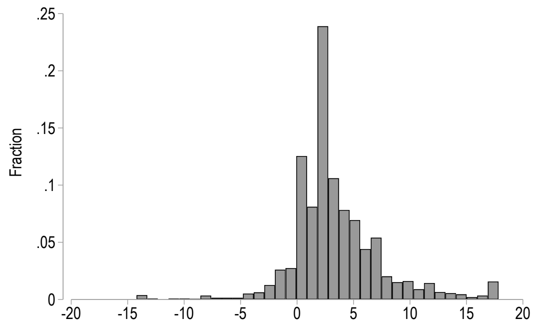
(f) Distribution of portfolio return



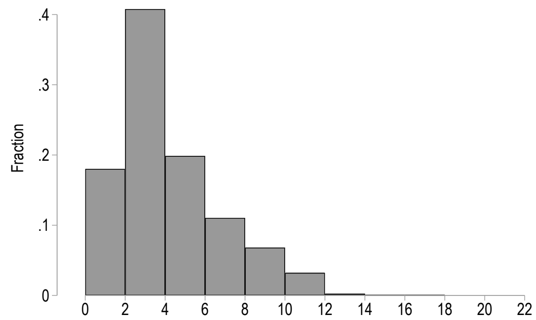
Notes: The figure shows the historical trends and the distribution of monthly-computed annual asset returns for the period from May 2003 to July 2019, used in the computation of $E(r|r \in B_j)$ in Section 3.3.3. In Panels (e) and (f), the "portfolio return" is computed for each month as the weighted average of the annual UK HP change (see Panels a and b) and the annual FTSE-100 return (see Panels c and d) with weights equal to 0.5. The horizontal lines in Panels (a), (c), and (e) correspond to the survey return brackets. *Data Source*: Historical values from the UK HPI and the FTSE-100 UK stock-market index.

Figure C.3. Estimates of the moments of the subjective return distributions

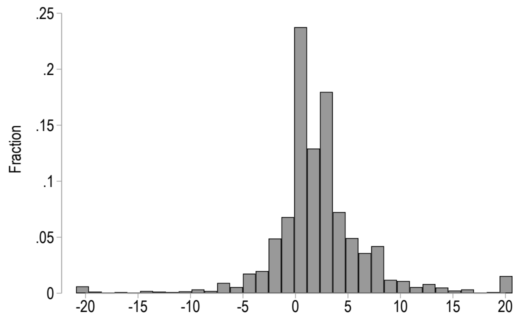
(a) Mean of the subjective return distribution for the housing investment



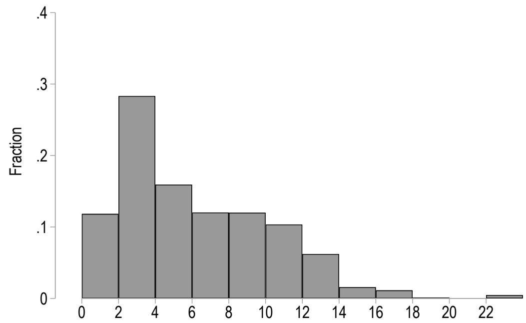
(b) Std Dev of the subjective return distribution for the housing investment



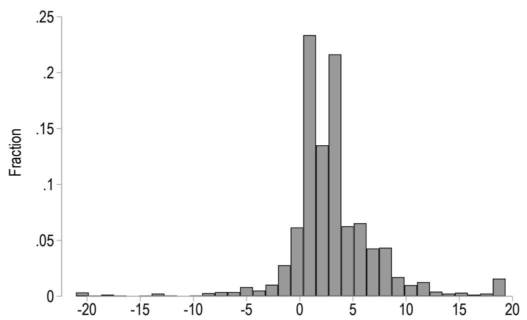
(c) Mean of the subjective return distribution for the stock investment



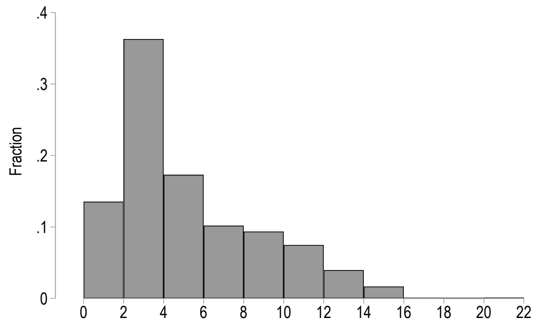
(d) Std Dev of the subjective return distribution for the stock investment



(e) Mean of the subjective return distribution for the portfolio investment



(f) Std Dev of the subjective return distribution for the portfolio investment



Notes: The figures show histograms of the distribution of the mean, $E(r)$, and standard deviation, $SD(r)$, of the subjective return distributions (obtained using the non-parametric estimation approach by Hurd, van Rooij, and Winter, 2011) for three different investments: Housing, stock, and a two-asset portfolio including both. *Data Source*: UK HPI, FTSE-100 UK stock-market index, and Financial Lives 2020 survey, N=2,926.

Appendix C3 Robustness: Supplementary Material

Appendix C3 provides supplementary material that complements the discussions on robustness in Section 3.6 of the main paper. Table C.5 presents regression results, including additional controls for survey duration, survey mode, and response behavior, with “satisfying the return constraint” (columns 1–4) and “satisfying the risk constraint” (columns 5–8) as dependent variables. Table C.6 presents correlates of straightlining and bunching response behavior in the survey questions on subjective asset return expectations. Table C.7 reports conditional averages of historical returns data (“bracket points”) for the period from March 1985 to July 2019. Table C.8 shows the percentage of respondents who satisfy one or both of the constraints, excluding those who exhibit partial straightlining behavior. Tables C.9 and C.10 report correlation coefficients of the means and standard deviations of the subjective return distributions obtained using alternative estimation approaches.

Table C.5. Satisfying the return/risk constraint: Role of survey duration, survey mode, and response behavior

Dependent variable	Satisfying the return constraint				Satisfying the risk constraint			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Gender (ref: Male)								
Female	0.015	0.017	0.020	0.019	0.009	0.009	0.011	0.011
	(0.020)	(0.020)	(0.020)	(0.020)	(0.016)	(0.016)	(0.016)	(0.016)
Other	−0.127	−0.123	−0.097	−0.104	0.088	0.088	0.102	0.109
	(0.113)	(0.113)	(0.111)	(0.111)	(0.092)	(0.092)	(0.092)	(0.091)
Age (ref: 35–50)								
18 to 34	0.028	0.028	0.030	0.030	−0.021	−0.021	−0.020	−0.020
	(0.026)	(0.026)	(0.026)	(0.026)	(0.021)	(0.021)	(0.021)	(0.021)
51 to 64	0.041	0.041	0.029	0.035	0.030	0.030	0.023	0.026
	(0.027)	(0.027)	(0.027)	(0.027)	(0.022)	(0.022)	(0.022)	(0.022)
65 or older	0.009	−0.001	−0.014	−0.005	0.001	−0.001	−0.008	−0.004
	(0.043)	(0.043)	(0.043)	(0.043)	(0.035)	(0.035)	(0.035)	(0.035)
Education (ref: Higher)								
Lower or medium	0.013	0.011	0.013	0.013	0.001	0.001	0.002	0.002
	(0.021)	(0.021)	(0.021)	(0.021)	(0.017)	(0.017)	(0.017)	(0.017)
None	0.041	0.028	0.026	0.017	0.047	0.044	0.044	0.033
	(0.045)	(0.046)	(0.045)	(0.045)	(0.037)	(0.037)	(0.037)	(0.037)
Info missing	−0.016	−0.012	−0.031	−0.028	−0.024	−0.023	−0.033	−0.044
	(0.066)	(0.066)	(0.065)	(0.065)	(0.054)	(0.054)	(0.053)	(0.053)
Employment status (ref: Employed)								
Self-employed	0.053	0.054	0.051	0.052	0.067^{**}	0.067^{**}	0.066^{**}	0.066^{**}
	(0.037)	(0.037)	(0.037)	(0.037)	(0.030)	(0.030)	(0.030)	(0.030)
Retired	0.033	0.030	0.038	0.038	0.002	0.002	0.006	0.009
	(0.038)	(0.038)	(0.038)	(0.038)	(0.031)	(0.031)	(0.031)	(0.031)
Other	-0.078^{**}	-0.076^{**}	-0.074^{**}	-0.062^{*}	−0.003	−0.003	−0.002	0.006
	(0.033)	(0.033)	(0.032)	(0.032)	(0.027)	(0.027)	(0.027)	(0.026)
Annual HH income (ref: £20k - <£40k)								
Less than £20k	0.021	0.017	0.024	0.012	−0.022	−0.023	−0.020	−0.027
	(0.032)	(0.032)	(0.032)	(0.032)	(0.026)	(0.026)	(0.026)	(0.026)
£40k - <£70k	0.053^{*}	0.052^{*}	0.046^{*}	0.049^{*}	0.014	0.014	0.011	0.012
	(0.028)	(0.028)	(0.027)	(0.027)	(0.022)	(0.022)	(0.022)	(0.022)
£70k or more	0.057^{**}	0.057^{*}	0.055^{*}	0.056^{*}	0.025	0.025	0.024	0.025
	(0.029)	(0.029)	(0.029)	(0.029)	(0.024)	(0.024)	(0.024)	(0.023)
Info missing	0.057^{*}	0.052^{*}	0.051^{*}	0.038	−0.038	−0.038	−0.039	-0.049^{**}
	(0.030)	(0.030)	(0.030)	(0.030)	(0.024)	(0.025)	(0.024)	(0.024)
Risk averse	−0.023	−0.023	−0.028	−0.026	−0.012	−0.013	−0.015	−0.013
	(0.019)	(0.019)	(0.019)	(0.019)	(0.016)	(0.016)	(0.016)	(0.016)
High financial literacy	0.044^{**}	0.046^{**}	0.059^{***}	0.063^{***}	0.044^{***}	0.045^{***}	0.051^{***}	0.057^{***}
	(0.021)	(0.021)	(0.021)	(0.021)	(0.017)	(0.017)	(0.017)	(0.017)
Understands relative asset risk	0.004	0.005	0.005	0.009	0.030^{*}	0.030^{*}	0.030^{*}	0.035^{*}
	(0.022)	(0.022)	(0.022)	(0.022)	(0.018)	(0.018)	(0.018)	(0.018)
Confident working with numbers	0.045^{**}	0.044^{**}	0.045^{**}	0.042^{**}	0.018	0.018	0.018	0.016
	(0.020)	(0.020)	(0.019)	(0.019)	(0.016)	(0.016)	(0.016)	(0.016)
Interview not finished on same day	0.035	0.037	0.036	0.042	0.013	0.014	0.013	0.017
	(0.032)	(0.032)	(0.032)	(0.032)	(0.026)	(0.026)	(0.026)	(0.026)
Interview conducted face-to-face		0.112^{**}	0.083	0.078		0.019	0.003	−0.018
		(0.055)	(0.055)	(0.055)		(0.045)	(0.045)	(0.045)
Straightlining			0.360^{***}				0.193^{***}	
			(0.040)				(0.033)	
Bunching 50%/50%				0.331^{***}				0.138^{***}
				(0.050)				(0.041)
Bunching 100%				0.266^{***}				0.258^{***}
				(0.041)				(0.033)
Constant	0.308^{***}	0.311^{***}	0.298^{***}	0.289^{***}	0.706^{***}	0.707^{***}	0.700^{***}	0.691^{***}
	(0.073)	(0.073)	(0.072)	(0.072)	(0.060)	(0.060)	(0.059)	(0.059)
Mean DepVar	0.50	0.50	0.50	0.50	0.79	0.79	0.79	0.79
Adj. R-squared	0.01	0.01	0.04	0.04	0.01	0.01	0.02	0.03
N	2,878	2,878	2,878	2,878	2,878	2,878	2,878	2,878
Month Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The table shows estimates from a linear probability model. In columns 1–4, the dependent variable is equal to 1 if an individual is "satisfying the return constraint," i.e., forming subjective expectations about three different investments in a way consistent with Equation 3.1. In columns 5–8, the dependent variable is equal to 1 if an individual is "satisfying the risk constraint," i.e., forming subjective expectations about three different investments in a way consistent with Equation 3.4. 'ref.' indicates the omitted category. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

Table C.6. Correlates of straightlining and bunching response behavior

Dependent variable	Straightlining fully	Straightlining partially	Bunching 50%/50%	Bunching 100%
	(1)	(2)	(3)	(4)
Gender (ref: Male)				
Female	−0.011	0.011	−0.004	−0.005
	(0.009)	(0.017)	(0.007)	(0.009)
Other	−0.073	−0.040	0.010	-0.086^{*}
	(0.053)	(0.094)	(0.042)	(0.052)
Age (ref: 35–50)				
18 to 34	−0.007	−0.029	−0.000	−0.009
	(0.012)	(0.022)	(0.010)	(0.012)
51 to 64	0.035^{***}	0.050^{**}	0.005	0.015
	(0.013)	(0.023)	(0.010)	(0.012)
65 or older	0.045^{**}	0.063^{*}	−0.001	0.029
	(0.020)	(0.036)	(0.016)	(0.020)
Education (ref: Higher)				
Lower or medium	−0.005	0.012	−0.003	0.001
	(0.010)	(0.018)	(0.008)	(0.010)
None	0.011	−0.016	−0.009	0.061^{***}
	(0.021)	(0.038)	(0.017)	(0.021)
Info missing	0.029	0.041	−0.028	0.054^{*}
	(0.032)	(0.057)	(0.025)	(0.031)
Employment status (ref: Employed)				
Self-employed	0.008	0.023	0.000	0.002
	(0.017)	(0.031)	(0.014)	(0.017)
Retired	−0.020	0.003	−0.003	−0.021
	(0.018)	(0.032)	(0.014)	(0.017)
Other	−0.010	−0.039	-0.026^{**}	-0.027^{*}
	(0.015)	(0.027)	(0.012)	(0.015)
Annual HH income (ref: £20k - <£40k)				
Less than £20k	−0.016	−0.021	0.004	0.016
	(0.015)	(0.027)	(0.012)	(0.015)
£40k - <£70k	0.015	−0.011	0.002	0.008
	(0.013)	(0.023)	(0.010)	(0.013)
£70k or more	0.005	−0.019	0.004	−0.003
	(0.014)	(0.024)	(0.011)	(0.013)
Info missing	0.004	0.010	0.016	0.036^{***}
	(0.014)	(0.025)	(0.011)	(0.014)
Risk averse	0.013	−0.018	0.016^{**}	−0.007
	(0.009)	(0.016)	(0.007)	(0.009)
High financial literacy	-0.034^{***}	-0.060^{***}	-0.025^{***}	-0.031^{***}
	(0.010)	(0.017)	(0.008)	(0.010)
All financial-literacy questions don't know	0.055^{**}	0.108^{***}	−0.009	0.105^{***}
	(0.021)	(0.038)	(0.017)	(0.021)
Understands relative asset risk	0.003	0.003	0.007	-0.021^{**}
	(0.010)	(0.019)	(0.008)	(0.010)
Confident working with numbers	−0.002	-0.027^{*}	0.004	0.005
	(0.009)	(0.016)	(0.007)	(0.009)
Constant	0.030	0.184^{***}	0.032	0.032
	(0.034)	(0.061)	(0.027)	(0.034)
Mean DepVar	0.06	0.22	0.03	0.06
Adj. R-squared	0.01	0.01	0.00	0.04
N	2,878	2,878	2,878	2,878
Month Fixed Effects	✓	✓	✓	✓

Notes: The table shows estimates from linear probability models. The dependent variable is equal to 1 if an individual exhibits the following response behavior: "straightlining, fully" (column 1), "straightlining, partially" (column 2), "bunching 50%/50%" (column 3), and "bunching 100%' (column 4). 'ref.' indicates the omitted category. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data Source*: Financial Lives 2020 survey.

Table C.7. Conditional averages of historical asset returns: Alternative time period

Bracket	Housing	Stock	Portfolio
B1: -10% or less	−14.21	−19.10	−21.05
B2: -9.9% to -5%	−7.15	−7.24	−7.12
B3: -4.9% to 0%	−1.78	−2.39	−2.05
B4: 0.1% to 5%	2.51	2.87	2.70
B5: 5.1% to 10%	7.70	7.37	7.66
B6: 10.1% to 15%	12.03	12.65	11.92
B7: 15.1% or more	21.41	23.12	20.34

Notes: The table shows historical averages of one-year rates of return conditional on the returns being in the respective brackets, $E(r|r \in B_j)$, for three different investments (following the non-parametric estimation approach by Hurd, van Rooij, and Winter 2011). Historical returns data from 3/1985–7/2019. *Data Source*: Historical values from the UK HPI and the FTSE-100 UK stock-market index.

Table C.8. Percentage satisfying the return and risk constraint: Excluding partial straight-liners

				Percentage satisfying the		
		N unsatisfiable	*N*	Return constraint	Risk constraint	Return and risk constraint
(1)	Baseline (approach by Hurd et al. 2011, historical returns from 5/2003–7/2019)	11	2,240	46.0	74.9	35.7
(2)	Approach by Hurd et al. 2011, historical returns from 3/1985–7/2019	6	2,245	46.9	74.5	36.2
(3)	Bracket midpoints with $(-12.5, 17.5)$ in open-ended brackets	0	2,251	49.9	71.7	36.1
(4)	Bracket midpoints with $(-20, 30)$ in open-ended brackets	0	2,251	49.0	72.7	36.7
(5)	Fitting log-normal distribution	84	2,167	45.8	74.1	35.3

Notes: The table shows the percentage of respondents satisfying the return or risk constraint considering alternative approaches in the estimation of the mean and standard deviation of the subjective return distributions and excluding respondents exhibiting partial straightlining behavior. *N unsatisfiable* represents the number of observations for which an analysis of satisfying the constraints cannot be undertaken (for a discussion of satisfiability, see Sections 3.5 and 3.6.3). *Data Source*: Financial Lives 2020 survey.

Table C.9. Correlation coefficients of the estimated means of the subjective return distributions

		(1)	(2)	(3)	(4)	(5)
Housing						
(1)	Baseline (approach by Hurd et al. 2011, historical returns from 5/2003–7/2019)	1	.	.	.	.
(2)	Approach by Hurd et al. 2011, historical returns from 3/1985–7/2019	0.99	1	.	.	.
(3)	Bracket midpoints with (-12.5,17.5) in open-ended brackets	0.99	0.99	1	.	.
(4)	Bracket midpoints with (-20,30) in open-ended brackets	0.98	0.99	0.98	1	.
(5)	Fitting log-normal distribution	0.94	0.94	0.93	0.93	1
Stock						
(1)	Baseline (approach by Hurd et al. 2011, historical returns from 5/2003–7/2019)	1	.	.	.	.
(2)	Approach by Hurd et al. 2011, historical returns from 3/1985–7/2019	0.99	1	.	.	.
(3)	Bracket midpoints with (-12.5,17.5) in open-ended brackets	0.98	0.99	1	.	.
(4)	Bracket midpoints with (-20,30) in open-ended brackets	0.98	0.99	0.98	1	.
(5)	Fitting log-normal distribution	0.91	0.91	0.89	0.90	1
Portfolio						
(1)	Baseline (approach by Hurd et al. 2011, historical returns from 5/2003–7/2019)	1	.	.	.	.
(2)	Approach by Hurd et al. 2011, historical returns from 3/1985–7/2019	0.99	1	.	.	.
(3)	Bracket midpoints with (-12.5,17.5) in open-ended brackets	0.98	0.98	1	.	.
(4)	Bracket midpoints with (-20,30) in open-ended brackets	0.97	0.98	0.97	1	.
(5)	Fitting log-normal distribution	0.90	0.90	0.87	0.88	1

Notes: The table shows the correlation coefficients of the means of the subjective return distributions obtained using alternative estimation approaches. *Data Source*: Financial Lives 2020 survey, N=2,767.

Table C.10. Correlation coefficients of the estimated standard deviations of the subjective return distributions

		(1)	(2)	(3)	(4)	(5)
Housing						
(1)	Baseline (approach by Hurd et al. 2011, historical returns from 5/2003–7/2019)	1	.	.	.	.
(2)	Approach by Hurd et al. 2011, historical returns from 3/1985–7/2019	0.98	1	.	.	.
(3)	Bracket midpoints with (-12.5,17.5) in open-ended brackets	0.99	0.98	1	.	.
(4)	Bracket midpoints with (-20,30) in open-ended brackets	0.96	0.99	0.96	1	.
(5)	Fitting log-normal distribution	0.87	0.87	0.87	0.87	1
Stock						
(1)	Baseline (approach by Hurd et al. 2011, historical returns from 5/2003–7/2019)	1	.	.	.	.
(2)	Approach by Hurd et al. 2011, historical returns from 3/1985–7/2019	0.99	1	.	.	.
(3)	Bracket midpoints with (-12.5,17.5) in open-ended brackets	0.97	0.97	1	.	.
(4)	Bracket midpoints with (-20,30) in open-ended brackets	0.98	0.99	0.97	1	.
(5)	Fitting log-normal distribution	0.89	0.88	0.87	0.87	1
Portfolio						
(1)	Baseline (approach by Hurd et al. 2011, historical returns from 5/2003–7/2019)	1	.	.	.	.
(2)	Approach by Hurd et al. 2011, historical returns from 3/1985–7/2019	0.99	1	.	.	.
(3)	Bracket midpoints with (-12.5,17.5) in open-ended brackets	0.98	0.98	1	.	.
(4)	Bracket midpoints with (-20,30) in open-ended brackets	0.98	0.98	0.97	1	.
(5)	Fitting log-normal distribution	0.71	0.71	0.70	0.71	1

Notes: The table shows the correlation coefficients of the standard deviations of the subjective return distributions obtained using alternative estimation approaches. *Data Source*: Financial Lives 2020 survey, N=2,767.

Bibliography

Adam, Klaus, Oliver Pfäuti, and Timo Reinelt (2022). *Subjective Housing Price Expectations, Falling Natural Rates and the Optimal Inflation Target.* Discussion Paper DP17187. CEPR.

Adelino, Manuel, Antoinette Schoar, and Felipe Severino (2018). *Perception of House Price Risk and Homeownership.* Working Paper 25090. National Bureau of Economic Research.

Agarwal, Sumit and Bhashkar Mazumder (2013). "Cognitive Abilities and Household Financial Decision Making." *American Economic Journal: Applied Economics*, 5 (1), 193–207.

Agnew, Julie R., Hazel Bateman, and Susan Thorp (2013). "Financial Literacy and Retirement Planning in Australia." *Numeracy*, 6 (2), Article 7.

Agrawal, Sarthak and David Phillips (2020). *Catching up or falling behind? Geographical inequalities in the UK and how they have changed in recent years.* The Institute for Fiscal Studies.

Ahn, Hie Joo and Choongryul Yang (2022). *Effects of Monetary Policy on Household Expectations: The Role of Homeownership.* Finance and Economics Discussion Series 2022-065. Board of Governors of the Federal Reserve System (U.S.)

Aigner, Dennis J. (1973). "Regression with a binary independent variable subject to errors of observation." *Journal of Econometrics*, 1 (1), 49–59.

Almenberg, Johan, Annamaria Lusardi, Jenny Säve-Söderbergh, and Roine Vestman (2020). "Attitudes toward Debt and Debt Behavior." *The Scandinavian Journal of Economics.*

Ameriks, John, Andrew Caplin, and John Leahy (2003). “Wealth Accumulation and the Propensity to Plan.” *The Quarterly Journal of Economics*, 118 (3), 1007–1047.

Ameriks, John, Gábor Kézdi, Minjoon Lee, and Matthew D. Shapiro (2020). “Heterogeneity in Expectations, Risk Tolerance, and Household Stock Shares: The Attenuation Puzzle.” *Journal of Business & Economic Statistics*, 38 (3), 633–646.

Amromin, Gene and Steven A. Sharpe (2014). “From the Horse’s Mouth: Economic Conditions and Investor Expectations of Risk and Return.” *Management Science*, 60 (4), 845–866.

Anderson, Anders, Forest Baker, and David T. Robinson (2017). “Precautionary savings, retirement planning and misperceptions of financial literacy.” *Journal of Financial Economics*, 126 (2), 383–398.

Andre, Peter, Carlo Pizzinelli, Christopher Roth, and Johannes Wohlfart (2022). “Subjective Models of the Macroeconomy: Evidence From Experts and Representative Samples.” *The Review of Economic Studies*.

Armantier, Olivier, Wändi Bruine de Bruin, Simon Potter, Giorgio Topa, Wilbert van der Klaauw, and Basit Zafar (2013). “Measuring Inflation Expectations.” *Annual Review of Economics*, 5 (1), 273–301.

Armona, Luis, Andreas Fuster, and Basit Zafar (2019). “Home Price Expectations and Behaviour: Evidence from a Randomized Information Experiment.” *The Review of Economic Studies*, 86 (4), 1371–1410.

Bachas, Natalie, Olivia S. Kim, and Constantine Yannelis (2021). “Loan guarantees and credit supply.” *Journal of Financial Economics*, 139 (3), 872–894.

Bachmann, Rüdiger, Giorgio Topa, and Wilbert van der Klaauw (2023). *Handbook of Economic Expectations.* Elsevier.

Badarinza, Cristian, John Y. Campbell, and Tarun Ramadorai (2016). “International Comparative Household Finance.” *Annual Review of Economics*, 8 (1), 111–144.

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong (2018). “Social Connectedness: Measurement, Determinants, and Effects.” *Journal of Economic Perspectives*, 32 (3), 259–80.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel (2018). "The Economic Effects of Social Networks: Evidence from the Housing Market." *Journal of Political Economy*, 126 (6), 2224–2276.

Bailey, Michael, Eduardo Dávila, Theresa Kuchler, and Johannes Stroebel (2019). "House Price Beliefs And Mortgage Leverage Choice." *The Review of Economic Studies*, 86 (6), 2403–2452.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer (2015). "X-CAPM: An extrapolative capital asset pricing model." *Journal of Financial Economics*, 115 (1), 1–24.

Beatty, Christina and Stephen Fothergill (2016). *The uneven impact of welfare reform. The financial losses to places and people.* Project Report. Sheffield Hallam University.

Beckmann, Elisabeth, Christa Hainz, and Sarah Kiesl-Reiter (2022). *Third-Party Loan Guarantees: Measuring Literacy and its Effect on Financial Decisions.* Working Paper Series 237. Oesterreichische Nationalbank.

Beckmann, Elisabeth and Sarah Kiesl-Reiter (2020). "How financially literate is CESEE? Insights from the OeNB Euro Survey." *Focus on European Economic Integration*, Q3/20, 36–49.

— (2023). "Financial literacy and financial wellbeing: Evidence from Eastern Europe in a high inflation environment." *Journal of Financial Literacy and Wellbeing*, 1 (2), 263–367.

Beckmann, Elisabeth, Sarah Kiesl-Reiter, and Helmut Stix (2018). "A geographic perspective on banking in Central, Eastern and Southeastern Europe." *Focus on European Economic Integration*, Q1/18, 26–47.

Beckmann, Elisabeth and Helmut Stix (2015). "Foreign currency borrowing and knowledge about exchange rate risk." *Journal of Economic Behavior & Organization*, 112, 1–16.

Bianchi, Milo (2018). "Financial Literacy and Portfolio Dynamics." *The Journal of Finance*, 73 (2), 831–859.

BIK (2018). *Poręczyciele w odwrocie (*translated*: Guarantors in the background).* URL: https://media.bik.pl/informacje-prasowe/388223/bik-poreczyciele-w-odwrocie (last accessed on 04/12/2020).

Binswanger, Johannes and Martin Salm (2017). "Does everyone use probabilities? The role of cognitive skills." *European Economic Review*, 98, 73–85.

Blume, Marshall E. and Irwin Friend (1975). "The Asset Structure of Individual Portfolios and Some Implications for Utility Functions." *The Journal of Finance*, 30 (2), 585–603.

Bonin, Holger, Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde (2007). "Cross-sectional earnings risk and occupational sorting: The role of risk attitudes." *Labour Economics*, 14 (6), 926–937.

Bottan, Nicolas L. and Ricardo Perez-Truglia (2020). *Betting on the House: Subjective Expectations and Market Choices.* NBER Working Papers 27412. National Bureau of Economic Research.

Brown, Jeffrey R., Zoran Ivković, Paul A. Smith, and Scott Weisbenner (2008). "Neighbors Matter: Causal Community Effects and Stock Market Participation." *The Journal of Finance*, 63 (3), 1509–1531.

Brown, Martin and Roman Graf (2013). "Financial Literacy and Retirement Planning in Switzerland." *Numeracy*, 6 (2), Article 6.

Bruine de Bruin, Wändi and Katherine G. Carman (2012). "Measuring Risk Perceptions: What Does the Excessive Use of 50% Mean?" *Medical Decision Making*, 32 (2), 232–236.

Bruine de Bruin, Wändi, Baruch Fischhoff, Susan G. Millstein, and Bonnie L. Halpern-Felsher (2000). "Verbal and Numerical Expressions of Probability: "It's a Fifty–Fifty Chance"." *Organizational Behavior and Human Decision Processes*, 81 (1), 115–131.

Bucher-Koenen, Tabea and Annamaria Lusardi (2011). "Financial literacy and retirement planning in Germany." *Journal of Pension Economics and Finance*, 10 (4), 565–584.

Bucher-Koenen, Tabea, Annamaria Lusardi, Rob Alessie, and Maarten van Rooij (2017). “How Financially Literate Are Women? An Overview and New Insights.” *Journal of Consumer Affairs*, 51 (2), 255–283.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo (2016). “Understanding Booms and Busts in Housing Markets.” *Journal of Political Economy*, 124 (4), 1088–1147.

Campbell, John Y. (2016). “Restoring Rational Choice: The Challenge of Consumer Financial Regulation.” *American Economic Review*, 106 (5), 1–30.

Case, Karl E. and Robert J. Shiller (1989). “The Efficiency of the Market for Single-Family Homes.” *The American Economic Review*, 79 (1), 125–137.

Case, Karl E., Robert J. Shiller, and Anne K. Thompson (2012). “What Have They Been Thinking? Homebuyer Behavior in Hot and Cold Markets.” *Brookings Papers on Economic Activity*, 43 (2), 265–315.

Cavallo, Alberto, Guillermo Cruces, and Ricardo Perez-Truglia (2017). “Inflation Expectations, Learning, and Supermarket Prices: Evidence from Survey Experiments.” *American Economic Journal: Macroeconomics*, 9 (3), 1–35.

Chiburis, Richard C., Jishnu Das, and Michael Lokshin (2012). “A practical comparison of the bivariate probit and linear IV estimators.” *Economics Letters*, 117 (3), 762–766.

Chopra, Felix, Christopher Roth, and Johannes Wohlfart (2023). *Home Price Expectations and Spending: Evidence from a Field Experiment.* CESifo Working Paper 10450.

Clark, Robert, Annamaria Lusardi, and Olivia S. Mitchell (2017). “Financial knowledge and 401(k) investment performance: a case study.” *Journal of Pension Economics and Finance*, 16 (3), 324–347.

Coibion, Olivier, Dimitris Georgarakos, Yuriy Gorodnichenko, and Maarten van Rooij (2023). “How Does Consumption Respond to News about Inflation? Field Evidence from a Randomized Control Trial.” *American Economic Journal: Macroeconomics*, 15 (3), 109–52.

Coibion, Olivier, Yuriy Gorodnichenko, and Rupal Kamdar (2018). "The Formation of Expectations, Inflation, and the Phillips Curve." *Journal of Economic Literature*, 56 (4), 1447–91.

Cole, Shawn, Xavier Giné, Jeremy Tobacman, Petia Topalova, Robert Townsend, and James Vickery (2013). "Barriers to Household Risk Management: Evidence from India." *American Economic Journal: Applied Economics*, 5 (1), 104–35.

Creditreform Wirtschaftsforschung (2020). *SchuldnerAtlas Deutschland 2020. Überschuldung von Verbrauchern.* url: https://www.boniversum.de/aktuelles-studien/schuldner-atlas/schuldneratlas-downloads (last accessed on 02/02/2024).

Crossley, Thomas F., Tobias Schmidt, Panagiota Tzamourani, and Joachim K. Winter (2020). "Interviewer effects and the measurement of financial literacy." *Journal of the Royal Statistical Society: Series A (Statistics in Society).*

Cupák, Andrej, Pirmin Fessler, Alyssa Schneebaum, and Maria Silgoner (2018). "Decomposing gender gaps in financial literacy: New international evidence." *Economics Letters*, 168 (C), 102–106.

D'Acunto, Francesco, Daniel Hoang, Maritta Paloviita, and Michael Weber (2019). "Cognitive Abilities and Inflation Expectations." *AEA Papers and Proceedings*, 109, 562–66.

D'Acunto, Francesco, Ulrike Malmendier, and Michael Weber (2021). "Gender roles produce divergent economic expectations." *Proceedings of the National Academy of Sciences*, 118 (21).

De Blasio, Guido, Stefania De Mitri, Alessio D'Ignazio, Paolo Finaldi Russo, and Lavinia Stoppani (2018). "Public guarantees to SME borrowing. A RDD evaluation." *Journal of Banking & Finance*, 96, 73–86.

De Haas, Ralph and Matteo Millone (2020). "The Impact of Information Sharing on the Use of Collateral versus Guarantees." *World Bank Economic Review*, 34 (Supplement), 14–19.

De Moura, Leonardo and Nikolaj Bjørner (2008). "Z3: An Efficient SMT Solver." *Tools and Algorithms for the Construction and Analysis of Systems.* Ed. by C. R. Ramakrishnan and Jakob Rehof. Berlin, Heidelberg: Springer Berlin Heidelberg, 337–340.

De Stefani, Alessia (2021). "House price history, biased expectations, and credit cycles: The role of housing investors." *Real Estate Economics*, 49 (4), 1238–1266.

Delavande, Adeline and Susann Rohwedder (2008). "Eliciting Subjective Probabilities in Internet Surveys." *Public Opinion Quarterly*, 72 (5), 866–891.

Disney, Richard and John Gathergood (2013). "Financial literacy and consumer credit portfolios." *Journal of Banking & Finance*, 37 (7), 2246–2254.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner (2011). "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences." *Journal of the European Economic Association*, 9 (3), 522–550.

Dominitz, Jeff (1998). "Earnings Expectations, Revisions, and Realizations." *The Review of Economics and Statistics*, 80 (3), 374–388.

Dominitz, Jeff and Charles F. Manski (1997a). "Perceptions of Economic Insecurity: Evidence from the Survey of Economic Expectations." *Public Opinion Quarterly*, 61 (2), 261–287.

— (1997b). "Using Expectations Data To Study Subjective Income Expectations." *Journal of the American Statistical Association*, 92 (439), 855–867.

— (2004). "How Should We Measure Consumer Confidence?" *Journal of Economic Perspectives*, 18 (2), 51–66.

— (2007). "Expected Equity Returns and Portfolio Choice: Evidence from the Health and Retirement Study." *Journal of the European Economic Association*, 5 (2/3), 369–379.

— (2011). "Measuring and Interpreting Expectations of Equity Returns." *Journal of Applied Econometrics*, 26 (3), 352–370.

Drerup, Tilman H. (2019). "Eliciting subjective expectations for bivariate outcomes." *Journal of Behavioral and Experimental Finance*, 23, 29–45.

Drerup, Tilman H., Matthias Wibral, and Christian Zimpelmann (2023). "Skewness expectations and portfolio choice." *Experimental Economics*, 26, 107–144.

Drerup, Tilman, Benjamin Enke, and Hans-Martin von Gaudecker (2017). "The precision of subjective data and the explanatory power of economic models." *Journal of Econometrics*, 200 (2), 378–389.

FCA (2017). *High-Cost Credit Review Technical Annex 1: Credit Reference Agency (CRA) Data Analysis of UK Personal Debt.* URL: https://www.fca.org.uk/publications/feedback-statements/fs17-2-high-cost-credit (last accessed on 04/12/2020).

Financial Conduct Authority (2021). *The Financial Lives 2020 survey (Wave 2): Technical report.* URL: https://www.fca.org.uk/publication/research/financial-lives-survey-2020-technical-report.pdf (last accessed on 01/03/2024).

Fuster, Andreas, David Laibson, and Brock Mendel (2010). "Natural Expectations and Macroeconomic Fluctuations." *Journal of Economic Perspectives*, 24 (4), 67–84.

Fuster, Andreas, Ricardo Perez-Truglia, Mirko Wiederholt, and Basit Zafar (2022). "Expectations with Endogenous Information Acquisition: An Experimental Investigation." *The Review of Economics and Statistics*, 104 (5), 1059–1078.

Gathergood, John and Jörg Weber (2017). "Financial literacy, present bias and alternative mortgage products." *Journal of Banking & Finance*, 78, 58–83.

Gelain, Paolo and Kevin J. Lansing (2014). "House prices, expectations, and time-varying fundamentals." *Journal of Empirical Finance*, 29, 3–25.

Gerardi, Kristopher, Lorenz Goette, and Stephan Meier (2013). "Numerical ability predicts mortgage default." *Proceedings of the National Academy of Sciences*, 110 (28), 11267–11271.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus (2021). "Five Facts about Beliefs and Portfolios." *American Economic Review*, 111 (5), 1481–1522.

Glaeser, Edward L. and Charles G. Nathanson (2017). "An extrapolative model of house price dynamics." *Journal of Financial Economics*, 126 (1), 147–170.

Goetzmann, William N. and Alok Kumar (2008). "Equity Portfolio Diversification." *Review of Finance*, 12 (3), 433–463.

Gohl, Niklas, Peter Haan, Claus Michelsen, and Felix Weinhardt (2024). "House price expectations." *Journal of Economic Behavior & Organization*, 218, 379–398.

Gomes, Francisco, Michael Haliassos, and Tarun Ramadorai (2021). "Household Finance." *Journal of Economic Literature*, 59 (3), 919–1000.

Guiso, Luigi and Eliana Viviano (2015). "How Much Can Financial Literacy Help?" *Review of Finance*, 19 (4), 1347–1382.

Guren, Adam M. (2018). "House Price Momentum and Strategic Complementarity." *Journal of Political Economy*, 126 (3), 1172–1218.

Haliassos, Michael, Thomas Jansson, and Yigitcan Karabulut (2020). "Financial Literacy Externalities." *The Review of Financial Studies*, 33 (2), 950–989.

Hastings, Justine and Olivia S. Mitchell (2020). "How financial literacy and impatience shape retirement wealth and investment behaviors." *Journal of Pension Economics and Finance*, 19 (1), 1–20.

Heiss, Florian, Michael Hurd, Maarten van Rooij, Tobias Rossmann, and Joachim Winter (2022). "Dynamics and heterogeneity of subjective stock market expectations." *Journal of Econometrics*, 231 (1), 213–231.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil (2012). "Measuring Economic Growth from Outer Space." *American Economic Review*, 102 (2), 994–1028.

HM Land Registry (2021). *Guidance: Quality and methodology.* url: https://www.gov.uk/government/publications/about-the-uk-house-price-index/quality-and-methodology (last accessed on 05/06/2022).

Hu, Zhongchen (2022). "Social interactions and households' flood insurance decisions." *Journal of Financial Economics*, 144 (2), 414–432.

Hudomiet, Péter, Gábor Kézdi, and Robert J. Willis (2011). "Stock market crash and expectations of American households." *Journal of Applied Econometrics*, 26 (3), 393–415.

Hudomiet, Péter and Robert J. Willis (2013). "Estimating Second Order Probability Beliefs from Subjective Survival Data." *Decision Analysis*, 10 (2), 152–170.

Hurd, M. D. and Kathleen McGarry (2002). "The Predictive Validity of Subjective Probabilities of Survival." *The Economic Journal*, 112 (482), 966–985.

Hurd, Michael D. (2009). "Subjective Probabilities in Household Surveys." *Annual Review of Economics*, 1 (1), 543–562.

Hurd, Michael, Maarten van Rooij, and Joachim Winter (2011). "Stock Market Expectations of Dutch Households." *Journal of Applied Econometrics*, 26 (3), 416–436.

International Monetary Fund (2017). "Household Debt and Financial Stability." *Global Financial Stability Report October 2017: Is Growth at Risk?* IMF, Chapter 2.

Jaeger, David A, Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, and Holger Bonin (2010). "Direct Evidence on Risk Attitudes and Migration." *The Review of Economics and Statistics*, 92 (3), 684–689.

Kaiser, Tim, Annamaria Lusardi, Lukas Menkhoff, and Carly Urban (2022). "Financial education affects financial knowledge and downstream behaviors." *Journal of Financial Economics*, 145 (2, Part A), 255–272.

Kaplan, Greg, Kurt Mitman, and Giovanni L. Violante (2020). "The Housing Boom and Bust: Model Meets Evidence." *Journal of Political Economy*, 128 (9), 3285–3345.

Kaustia, Markku and Samuli Knüpfer (2012). "Peer performance and stock market entry." *Journal of Financial Economics*, 104 (2), 321–338.

Kézdi, Gábor and Robert J. Willis (2011). *Household Stock Market Beliefs and Learning*. Working Paper 17614. National Bureau of Economic Research.

Kiesl-Reiter, Sarah, Melanie Lührmann, Jonathan Shaw, and Joachim Winter (2024). *The Formation of Subjective House Price Expectations*. Rationality and Competition Discussion Paper Series 491. CRC TRR 190 Rationality and Competition.

Kim, Yujin, Jennifer Dykema, John Stevenson, Penny Black, and D. Paul Moberg (2019). "Straightlining: Overview of Measurement, Comparison of Indicators, and Effects in Mail-Web Mixed-Mode Surveys." *Social Science Computer Review*, 37 (2), 214–233.

Kindermann, Fabian, Julia Le Blanc, Monika Piazzesi, and Martin Schneider (2021). *Learning about Housing Cost: Survey Evidence from the German House Price Boom*. Working Paper 28895. National Bureau of Economic Research.

Klapper, Leora F., Annamaria Lusardi, and Georgios A. Panos (2012). *Financial Literacy and the Financial Crisis*. Working Paper 17930. National Bureau of Economic Research.

Klapper, Leora and Annamaria Lusardi (2020). "Financial literacy and financial resilience: Evidence from around the world." *Financial Management*, 49 (3), 589–614.

Klapper, Leora, Annamaria Lusardi, and Georgios A. Panos (2013). "Financial literacy and its consequences: Evidence from Russia during the financial crisis." *Journal of Banking & Finance*, 37 (10), 3904–3923.

Kleinjans, Kristin J. and Arthur van Soest (2014). "Rounding, Focal Point Answers and Nonresponse to Subjective Probability Questions." *Journal of Applied Econometrics*, 29 (4), 567–585.

Koşar, Gizem and Cormac O'Dea (2023). "Expectations data in structural microeconomic models." *Handbook of Economic Expectations*. Elsevier, 647–675.

Krosnick, Jon A. (1991). "Response strategies for coping with the cognitive demands of attitude measures in surveys." *Applied Cognitive Psychology*, 5 (3), 213–236.

Kuchler, Theresa, Monika Piazzesi, and Johannes Stroebel (2023). "Housing market expectations." *Handbook of Economic Expectations*. Elsevier, 163–191.

Kuchler, Theresa and Basit Zafar (2019). "Personal Experiences and Expectations about Aggregate Outcomes." *The Journal of Finance*, 74 (5), 2491–2542.

Kuhnen, Camelia M. and Andrei C. Miu (2017). "Socioeconomic status and learning from financial information." *Journal of Financial Economics*, 124 (2), 349–372.

Landvoigt, Tim (2017). "Housing demand during the boom: The role of expectations and credit constraints." *The Review of Financial Studies*, 30 (6), 1865–1902.

Laudenbach, Christine, Benjamin Loos, Jenny Pirschel, and Johannes Wohlfart (2021). "The trading response of individual investors to local bankruptcies." *Journal of Financial Economics*, 142 (2), 928–953.

Lee, David S., Justin McCrary, Marcelo J. Moreira, and Jack Porter (2022). "Valid t-Ratio Inference for IV." *American Economic Review*, 112 (10), 3260–90.

Lusardi, Annamaria and Carlo de Bassa Scheresberg (2013). *Financial Literacy and High-Cost Borrowing in the United States*. Working Paper 18969. National Bureau of Economic Research.

Lusardi, Annamaria, Pierre-Carl Michaud, and Olivia S. Mitchell (2017). "Optimal Financial Knowledge and Wealth Inequality." *Journal of Political Economy*, 125 (2), 431–477.

Lusardi, Annamaria and Olivia S. Mitchell (2008). “Planning and Financial Literacy: How Do Women Fare?” *The American Economic Review*, 98 (2), 413–417.

— (2011a). “Financial literacy and retirement planning in the United States.” *Journal of Pension Economics and Finance*, 10 (4), 509–525.

— (2011b). “Financial literacy around the world: an overview.” *Journal of Pension Economics and Finance*, 10 (4), 497–508.

— (2014). “The Economic Importance of Financial Literacy: Theory and Evidence.” *Journal of Economic Literature*, 52 (1), 5–44.

— (2017). “How Ordinary Consumers Make Complex Economic Decisions: Financial Literacy and Retirement Readiness.” *Quarterly Journal of Finance*, 7 (3), 1750008.

— (2023). “The Importance of Financial Literacy: Opening a New Field.” *Journal of Economic Perspectives*, 37 (4), 137–54.

Lusardi, Annamaria, Olivia S. Mitchell, and Noemi Oggero (2020). “Debt and Financial Vulnerability on the Verge of Retirement.” *Journal of Money, Credit and Banking*, 52 (5), 1005–1034.

Lusardi, Annamaria and Peter Tufano (2015). “Debt literacy, financial experiences, and overindebtedness.” *Journal of Pension Economics and Finance*, 14 (4), 332–368.

Malmendier, Ulrike (2021a). “Experience Effects in Finance: Foundations, Applications, and Future Directions.” *Review of Finance*, 25 (5), 1339–1363.

— (2021b). “FBBVA Lecture 2020 Exposure, Experience, and Expertise: Why Personal Histories Matter in Economics.” *Journal of the European Economic Association*, 19 (6), 2857–2894.

Malmendier, Ulrike and Stefan Nagel (2011). “Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?” *The Quarterly Journal of Economics*, 126 (1), 373–416.

Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers (2003). “Disagreement about Inflation Expectations.” *NBER Macroeconomics Annual*, 18, 209–248.

Manski, Charles F. (2004). “Measuring Expectations.” *Econometrica*, 72 (5), 1329–1376.

— (2018). "Survey Measurement of Probabilistic Macroeconomic Expectations: Progress and Promise." *NBER Macroeconomics Annual*, 32, 411–471.

Merkle, Christoph and Martin Weber (2014). "Do investors put their money where their mouth is? Stock market expectations and investing behavior." *Journal of Banking & Finance*, 46 (1), 372–386.

Mian, Atif and Amir Sufi (2010). "The Great Recession: Lessons from Microeconomic Data." *The American Economic Review*, 100 (2), 51–56.

— (2018). "Finance and Business Cycles: The Credit-Driven Household Demand Channel." *Journal of Economic Perspectives*, 32 (3), 31–58.

OECD (2015). *National Strategies for Financial Education. OECD/INFE Policy Handbook.* OECD.

Office for National Statistics (2022). *M01 Regional labour market: Modelled unemployment for local and unitary authorities.* url: https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/datasets/modelledunemploymentforlocalandunitaryauthoritiesm01 (last accessed on 05/06/2022).

— (2023). *Residential property sales for administrative geographies: HPSSA dataset 6.* url: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/datasets/numberofresidentialpropertysalesfornationalandsubnationalgeographiesquarterlyrollingyearhpssadataset06 (last accessed on 11/11/2023).

Overman, Henry and Xiaowei Xu (2022). *Spatial disparities across labour markets.* The Institute for Fiscal Studies.

Piazzesi, Monika and Martin Schneider (2009). "Momentum Traders in the Housing Market: Survey Evidence and a Search Model." *The American Economic Review*, 99 (2), 406–411.

Piketty, Thomas (2014). *Capital in the Twenty-First Century.* Harvard University Press.

Stango, Victor and Jonathan Zinman (2009). "Exponential Growth Bias and Household Finance." *The Journal of Finance*, 64 (6), 2807–2849.

Townsend, Robert M. (1994). "Risk and Insurance in Village India." *Econometrica*, 62 (3), 539–591.

van Ooijen, Raun and Maarten C.J. van Rooij (2016). "Mortgage Risks, Debt Literacy and Financial Advice." *Journal of Banking & Finance*, 72, 201–217.

van Rooij, Maarten, Annamaria Lusardi, and Rob Alessie (2011). "Financial literacy and stock market participation." *Journal of Financial Economics*, 101 (2), 449–472.

— (2012). "Financial Literacy, Retirement Planning and Household Wealth." *The Economic Journal*, 122 (560), 449–478.

Vellekoop, Nathanael and Mirko Wiederholt (2019). *Inflation Expectations and Choices of Households.* SAFE Working Paper Series 250. Leibniz Institute for Financial Research SAFE.

Vissing-Jorgensen, Annette (2003). "Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions." *NBER Macroeconomics Annual*, 18, 139–194.

von Gaudecker, Hans-Martin (2015). "How Does Household Portfolio Diversification Vary with Financial Literacy and Financial Advice?" *The Journal of Finance*, 70 (2), 489–507.

World Bank (2014). *Financial Education Programs and Strategies.* url: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/901211472719528753/financial-education-programs-and-strategies-approaches-and-available-resources (last accessed on 04/29/2020).

YouGov (2021). *Informal financial networks.* Survey commissioned by ifo Institute, unpublished.

Zimpelmann, Christian (2021). *Stock Market Beliefs and Portfolio Choice in the General Population.* CRC TR 224 Discussion Paper Series 258. University of Bonn and University of Mannheim, Germany.

Zivot, Eric (2021). *Introduction to Computational Finance and Financial Econometrics with R.* url: https://bookdown.org/compfinezbook/introcompfinr/ (last accessed on 03/02/2023).

Eidesstattliche Versicherung

Ich versichere hiermit eidesstattlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie mir gegebene Anregungen sind als solche kenntlich gemacht. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Sofern ein Teil der Arbeit aus bereits veröffentlichten Papers besteht, habe ich dies ausdrücklich angegeben.

München, 11. März 2024

Sarah Maria Kiesl-Reiter