Journal of Banking and Finance 148 (2023) 106743

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

What drives stock market participation? The role of institutional, traditional, and behavioral factors☆

Markku Kaustia[a,*], Andrew Conlin[b], Niilo Luotonen[a]

[a] Aalto University School of Business, Finland
[b] Oulu Business School, Finland

ARTICLE INFO

Article history:
Received 7 February 2021
Accepted 2 December 2022
Available online 5 December 2022

JEL:
G5
G4
G15

Keywords:
Household finance
Stock market participation
Investor behavior
Financial decision-making

ABSTRACT

We analyze stock market participation in 19 European countries, providing a composite view of the interplay and relative importance of established participation drivers. We jointly control for nearly all relevant drivers found in prior studies, which tend to introduce one novel variable at a time and often omit risk-aversion. Excellent full model predictive power decomposes into institutional (country) fixed effects (about 30% of total), traditional individual-level variables (50%), and more recently identified behavioral variables (20%). We sketch a hierarchical framework where factors' effects vary by agents' proneness to participate. We also challenge and complement existing interpretations given to sociability, IQ, trust, and life experiences.

1. Introduction

Limited stock market participation has been the quintessential topic in the emerging field of household finance.[1] Differences in stockholding propensity across countries, and between people of comparable wealth, show that country factors related to economic or cultural institutions are important. On the individual level, wealth, income, and education are highly influential. Besides these traditional factors, more recent studies have uncovered an interesting set of behavioral factors that explain stock market participation. What is missing, however, is a composite view of the interplay and relative importance of the various drivers of participation.

In this paper, we take a novel, big-picture approach. Whereas the literature has tended to focus on testing the effect of a single new factor after controlling for traditional ones, we combine an extensive set of these variables in one model and explore the relative contribution of three types of factors – institutional, traditional, and new behavioral ones. We jointly estimate the effects of all variables while including a directly queried financial risk aversion measure known to predict actual financial risk-taking well (Dohmen et al., 2011; Halko et al., 2012; Guiso et al., 2018; Laudenbach et al., 2020). Then, based on the contribution of individual explanatory variables in various regulatory environments and subsamples of varying sophistication, we sketch a hierarchical model that describes the likely importance of each factor in different contexts.

The data are from the first four main waves of the Survey of Health, Aging and Retirement in Europe (SHARE). It is a multidisciplinary survey targeted at individuals aged 50 and above as well as their spouses. Our sample consists of more than 100,000 individuals from 19 European countries, interviewed between years 2004 and 2013. The SHARE data are extremely suit-

☆ We thank Petri Jylhä, Mikko Niemenmaa, Joni Kokkonen, Elias Rantapuska, as well as seminar participants at the 2nd Research in Behavioral Finance Conference in Amsterdam, Goethe University Frankfurt House of Finance, LMU Institute for Finance & Banking and Institute for Capital Markets and Corporate Finance Joint Ski Conference 2018, Aalto Finance lunch seminar, and Bank of Finland Economics Seminar for comments. We also want to extend special thanks to Elina Oinonen for comments and assistance with the data. Luotonen gratefully acknowledges funding from the OP Group Research Foundation and the Finnish Foundation for the Advancement of Securities Markets. Conlin thanks the OP Group Research Foundation for generous funding. This paper uses data from SHARE Waves 1, 2, 4, and 5, see Börsch-Supan et al. (2013) for methodological details. The SHARE data collection has been funded by the European Commission. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging and from various national funding sources is gratefully acknowledged (see www.share-project.org).

* Corresponding author.
E-mail address: markku.kaustia@aalto.fi (M. Kaustia).

[1] See Campbell (2006) and Guiso and Sodini (2013) for reviews. An individual's stock market participation status has major implications on her wealth accumulation and consumption (Mehra and Prescott, 1985; Dynan and Maki, 2001; Cocco et al., 2005; Guvenen, 2006).

https://doi.org/10.1016/j.jbankfin.2022.106743

able for our research questions, with a comprehensive nature, a broad geographical scope, a longitudinal dimension, and a measure of risk attitude. While respondent age is relatively high, the effects we find should plausibly generalize to younger cohorts. In particular, the related studies of Choi and Robertson (2020) and Bender et al. (2022) show that demographics like age and wealth are largely orthogonal to what drives investment decisions. Furthermore, our results generally line up well with studies whose sample age distributions differ from ours, including recent work by Changwony et al. (2021) and Fey et al. (2021).

We limit our choice of explanatory variables to those used in published studies, which typically examine one new variable at a time, with some controls.[2] What we refer to as the six traditional variables in this study comprise gender, age, education, income, wealth, and risk aversion. The new variables we utilize are sociability (and its subcomponents), trust, cognitive skills (and its subcomponents), health, religiosity, political orientation, hard life experiences, optimism, bequest motive, body mass index, and height (see Appendix 1 for variable definitions, and Table 1 for prior studies analyzing one of these factors at a time).

Our baseline logit model explains stock market participation very well: it produces a value of 0.85 for the area under the Receiver Operating Characteristic (ROC) curve, where values above 0.8 are considered excellent (Hosmer and Lemeshow, 2000). Ranging from 0.5 (coin toss) to 1.0 (perfect classification), the area under the ROC gives the average probability of the model correctly identifying participants on a randomly drawn pair where one participates and the other one does not. To address the issue of possible overfitting with the use of 17 covariates, we investigate the robustness of the results by experimenting with lasso regressions (see the Internet Appendix for details and results).

We explore the relative contribution of the institutional, traditional, and new behavioral factors by using the additive Shapley decomposition of the model's R-squared. We find that institutional/other country-level factors (captured by country fixed effects) jointly account for 32% of the variation explained. The six traditional individual-level variables contribute 49%, or about half of the total. The set of new variables, although more numerous at eight, jointly contribute only 19%. Our results are in line with Choi and Robertson (2020), who find that investors are strongly influenced by classical factors, with behavioral factors having less, but still significant influence. We also conduct a latent common factor analysis. Most individual variables, traditional as well as some new, load on a single common factor. As the sole explanatory variable along with country fixed effects, this single factor achieves an R-squared of 27.3%, or 93% of that of the full model.

The hierarchical taxonomy in which we consolidate our findings is inspired by, and loosely modeled after, Maslow's (1943) theory for the hierarchy of needs. To build the hierarchy, we examine whether each variable becomes stronger, or weaker, when moving toward more sophisticated subsamples where participation is more common. We analyze this at the country level, grouping countries on institutional quality, and at the individual level, grouping individuals on wealth and education. Two variables always become stronger where baseline participation is higher: religiosity, which has a negative effect on participation, and political orientation, where more right-leaning preferences have a positive effect. These same variables, religiosity and political orientation, do not load on the single latent factor while most of the other variables do.

The conceptual framework for the hierarchy of different effects is in Fig. 1. The idea is that low-level factors (basic requirements for participation, such as wealth) need to be favorable before higher-level factors (related to information processing – such as cognitive skills and willpower[3]) have more bite. And, moving to settings where both such sets of factors already are favorable, the highest level of self-actualizing considerations (values and attitudes) may then help separate stock market participants from non-participants. We conduct analysis that supports the hierarchical model, but this is just a beginning step. Our hope is that this framework will lead to specific testable ideas, and that it ultimately helps target suitable policy tools at different household segments.

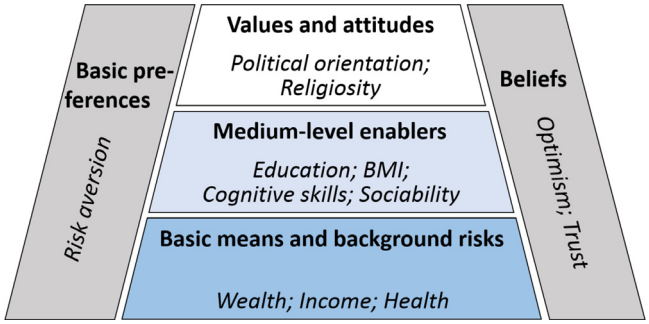


Fig. 1. Conceptual hierarchical model of stock market participation drivers. Boldface group headings (e.g., 'medium-level enablers') describe the type of influence, while italicized concepts represent empirical variables proxying for effects. Better empirical proxies, or entirely new contributing variables, could of course be envisioned, and used in the future. However, in this graph we limit to those empirical proxies employed in this study, serving as examples of how each type of influence can be measured.

We also take a fresh look at the new behavioral variables, reporting several novel stylized facts while challenging and complementing earlier interpretations. Splitting our measure of sociability (the strongest predictor of participation among the behavioral variables) into its subcomponents, we show that activities related to religion, as studied by Hong et al. (2004), likely provide a lower bound estimate for the impact that different types of social activities have on participation. Among cognitive skills, our results highlight the importance of recall ability, a skill shown to associate with positive life outcomes in the psychology literature (see, e.g., Packiam Alloway and Alloway, 2010). We demonstrate that omitting risk aversion is likely the reason why numeracy has been found more influential in prior studies (Christelis et al., 2010; Grinblatt et al., 2011). When we investigate the channel through which trust associates with participation, we find support for trust toward financial advisors and intermediaries being relevant (see Guiso et al., 2008), while trust toward listed companies (Giannetti and Wang, 2016) does not drive the effect. Finally, when it comes to past experiences (Malmendier and Nagel, 2011; Knüpfer et al., 2017), financial hardship matters, whereas other negative life events do not show any effect in our data. These findings, along with ones related to other individual predictors of participation, are discussed further in Section 6.

The stock market participation literature has been building over the last 30 years, brick by brick. The literature lacks, however, a consolidated view. This paper fills that gap. Using an extensive data set, we show how well the known factors explain participation overall and how much each set of factors contributes to the explanation. Our hierarchy of factors provides a framework for future work in the field.

[2] The richness of the SHARE data has its limits, and we cannot include every factor that has been shown to predict participation. Omitted factors include, e.g., savings goals (Changwony et al., 2021) and financial advice (Changwony et al., 2021; Gennaioli et al., 2015).

[3] Willpower can be proxied by the Body Mass Index (BMI), which is known to be associated with impatience and procrastination (Ikeda et al., 2010), and difficulty in managing expenditures (Borghans and Golsteyn, 2006).

Table 1
Summary of risk aversion measures used in published studies on the new variables.
Journal name abbreviations are: *EER* for European Economic Review; *JBF* for Journal of Banking and Finance; *JEEA* for Journal of the European Economic Association; *JF* for Journal of Finance; *JFE* for Journal of Financial Economics; *NBER* for National Bureau of Economic Research; *MS* for Management Science; *QJE for Quarterly Journal of Economics; RoF* for Review of Finance. Data source abbreviations are: *BHPS* for British Household Panel Survey; *DHS* for Dutch National Bank household survey; *FCSD* for Finnish Central Securities Depository; *HRS* for Health and Retirement Study; *IRS* for Internal Revenue Service; *NLSY* for National Longitudinal Study of Youth; *PSID* for Panel Study of Income Dynamics; *SCF* for Survey of Consumer Finances; *SHARE* for Survey of Health and Retirement in Europe; *WVS* for World Values Survey.

Authors	Title	Journal, year	Data	Risk aversion proxy
Laudenbach, Malmendier, Niessen-Ruenzi	The long-lasting effecs of living under communism on attitudes towards finance markets	NBER, 2020	German brokerage firm, 2004–2012	Brokerage firm's question, response 1(conservative) – 3 (speculative);
Addoum, Korniotis, Kumar	Stature, obesity, and portfolio choice	MS, 2017	US: HRS 1992–2008 & NLSY 1992–2008; SHARE W1 (2004), W2 (2006); Netherlands: DHS 1993–2009	HRS: Measure based on income gambles like the one used by Hong et al. (2004), created by Kimball, Sahm, and Shapiro (2008). NLSY: Proxy based on three gambles SHARE & DHS: None; control for gender, education, wealth.
Knüpfer, Rantapuska, Sarvimäki	Formative Experiences and Portfolio Choice: Evidence from the Finnish Great Depression	JF, 2017	Statistics Finland 1990–2005; Finnish Tax Authority 1987–2005	None; control for labor market conditions, age, gender, education, income…
Giannetti, Wang	Corporate scandals and household stock market participation	JF, 2016	US: PSID 1984–2009, combined with brokerage and fraud data	Dummy for below median risk tolerance based on income gamble questions in 1996 wave; included 1996 and onwards. Also control for wealth, age, education, … Gender control unclear.
Changwony, Campbell, Tabner	Social engagement and stock market participation	RoF, 2015	UK: BHPS 1995, 2000, 2005	None; control for gender, education, wealth, health, having debt, self-employment, …
Atella, Brunetti, Maestas	Household portfolio choices, health status, and health care systems: A cross-country analysis based on SHARE	JBF, 2012	SHARE W1 (2004)	None; control for gender, education, wealth, sociability, religiosity, cognitive skills, …
Georgarakos, Pasini	Trust, sociability, and stock market participation	RoF, 2011	SHARE W1 (2004) added with trust data from WVS (W2 data used in online appendix)	None; control for gender, education, wealth, …
Grinblatt, Keloharju, Linnainmaa	IQ and stock market participation	JF, 2011	Finland: Six data sets merged, holdings data from end-2000	None; control for gender, education, wealth, …
Kaustia, Torstila	Stock market aversion? Political preferences and stock market participation	JFE, 2011	Finland: FCSD ownership data 1995–2002; Statistics Finland 1996, 2002; Survey data 2003, 2006, 2009	None; control for gender, education, wealth.
Malmendier, Nagel	Depression babies: do macroeconomic experiences affect risk taking?	QJE, 2011	US: SCF 1960–2007	Same question that we use, responses in reverse order (i.e., risk tolerance)
Christelis, Jappelli, Padula	Cognitive abilities and portfolio choice	EER, 2010	SHARE W1 (2004)	None; control for education, income, wealth, religiosity, … Gender control unclear.
Brown, Ivković, Smith, Weisbenner	Neighbors matter: Causal community effects and stock market participation	JF, 2008	US: IRS panel data on tax returns 1987–1996	None; control for individual fixed effects, income, age.
Guiso, Sapienza, Zingales	Trusting the stock market	JF, 2008	Netherlands: DHS 2003; Italian bank customers	DHS: Risky gamble question. Also control for ambiguity aversion, optimism. Italy: Dummy based on claim: Risk is (1) an uncertain event from which one can profit, or (2) an uncertain event one should protect himself from.
Dominitz, Manski	Expected equity returns and portfolio choice: Evidence from the Health and Retirement Study	JEEA, 2007	US: HRS 2004	None; control for gender and marital status.
Puri, Robinson	Optimism and economic choice	JFE, 2007	US: SCF (all suitable waves since 1989, unclear which is the latest they use)	Question on willingness to take financial risk: "Take substantial risk for substantial reward", …, "not willing to take any risk."
Hong, Kubik, Stein	Social interaction and stock-market participation	JF, 2004	US: HRS 1992	Dummy indicating if at least one household member would take a job that would double their income with a 50% chance and cut it by a third with a 50% chance.
Rosen, Wu	Portfolio choice and health status	JFE, 2004	US: HRS 1992, 1994, 1996, 1998	Dummy indicating if respondent would take a job that would double their income with a 50% chance and cut it in half with a 50% chance.

2. Determinants of risky asset holdings

The discussion in this section is mostly driven by empirical findings. For comprehensive treatments of the theory, see Guiso et al. (2003a), Campbell (2006), and Curcuru et al. (2010). Our own empirical focus is mainly on the dichotomous participation decision, and then stock market allocation separately; some earlier work analyzes them jointly. We discuss these papers as well, even if evidence is not separately available for the participation decision.

2.1. "Country factors" – the institutional and cultural environment

Macroeconomic factors should have an impact on the probability of participation. There could also be a cultural channel whereby a culture of stock ownership in a country takes time to develop, and is likely promoted by a host of intertwined factors (Guiso et al., 2003b). Also, investor protection and regulation are key in promoting household equity ownership (Giannetti and Koskinen, 2010; Pagano and Volpin, 2006). The various economic factors are, in turn, shaped by non-economic country characteristics, such as religion (Campante and Yanagizawa-Drott, 2015; Barro and McCleary, 2003; Stulz and Williamson, 2003).

Christelis et al. (2013) provide empirical evidence on the importance of country factors. Compared to Europe, stock market participation is more likely, and mortgages both larger and more common, among U.S. households. As pointed out by Georgarakos and Pasini (2011) and others, differences in financial market participation are also substantial across countries within Europe. We use country fixed effects to capture all these effects.

2.2. "Traditional variables" – individual-level determinants of the early literature

In the absence of participation constraints, everyone should invest a strictly positive amount in stocks (see, e.g., Merton, 1969; 1971). With a fixed participation cost, however, the combination of risk aversion and low wealth can lead to rational non-participation (Vissing-Jørgensen, 2002, 2004). Alternatively, non-participation can be generated in theoretical models where preferences exhibit loss aversion or disappointment aversion (see Ang et al., 2005). Empirically, wealth explains participation, as do measures of risk aversion (Haliassos and Bertaut, 1995; Bertaut, 1998).[4] The average individual investor has decreasing relative risk aversion (Cohn et al., 1975; Calvet and Sodini, 2014), and the effect of wealth becomes stronger with high levels of wealth (Riley and Chow, 1992). However, investing in stocks is surprisingly uncommon even among wealthy households, suggesting that non-economic costs may also be important (Mankiw and Zeldes, 1991; Campbell, 2006; Curcuru et al., 2010).

Consistent with information costs, education has a strong effect on participation, even controlling for wealth and income (Haliassos and Bertaut, 1995; Cole et al., 2014). Education can change decision making in several ways; through increasing financial literacy and cognitive skills, or affecting social networks, job opportunities, and beliefs and attitudes. Each of these channels can have a direct effect as well.

Background risk—uncertain labor market income, entrepreneurial income, or fixed assets like real estate—can cause a need to reduce total risk by avoiding stocks (Heaton and Lucas, 2000a). Guiso et al. (1996) find that the variability of wage income reduces stock holdings in Italy. In the U.S., the effect is statistically significant only for proprietary business income (Heaton and Lucas, 2000b). Frantonini (1998) finds that mortgage payments reduce risky asset holdings, while Vestman (2019) and Kullmann and Siegel (2005) show that homeowners are more likely to invest in stocks than renters. Basic demographics are relevant if they proxy for risk aversion or background risk. Pure age effect, controlling for both cohort effects and time effects, is usually impossible to estimate (Poterba and Samwick, 2001). After addressing this issue to the best extent that their data allows, Ameriks and Zeldes (2004) find that participation does decline with age to some degree. Men are more likely to invest in stocks and the gender difference is larger in single households (Jianakoplos and Bernasek, 1998; Sundén and Surette, 1998; Barber and Odean, 2001). However, Halko et al. (2012) show that this effect vanishes after controlling for other important determinants in Finland, a country often considered the most gender equal in the world.

2.3. "New variables" – behavioral and other recently identified determinants

Social interaction affects households' investment decisions. Participation in retirement plans is influenced by the choices of co-workers (Madrian and Shea, 2001; Duflo and Saez, 2002). Hong et al. (2004) find that households interacting with their neighbors or attending church are more likely to invest in stocks. Georgarakos and Pasini (2011) and Changwony et al. (2015) present similar findings for other social activities. The evidence reported by Brown et al. (2008) and Kaustia and Knüpfer (2012) suggests that the effect of social interaction is causal.

The concept of social capital has been linked to stock market participation by Guiso et al. (2004). The idea is that high social capital enhances the level of trust in a society, which in turn boosts financial development and increases household stock market participation. To measure social capital, researchers have often turned to survey evidence on generalized trust (see Sapienza et al., 2013). Guiso et al. (2008) find that trusting individuals are significantly more likely to hold stocks. The effect of trust remains equally strong for wealthier households. In addition to generalized trust, Guiso et al. (2008) show that trust particularly towards banks also predicts participation. Giannetti and Wang (2016) show that the incidence of corporate fraud in a U.S. state makes the inhabitants of that state invest less in stocks, also in the non-fraudulent firms. They explain this result by loss of trust toward the stock market among households.

Kaustia and Torstila (2011) propose that some people perceive the stock market as incongruent with their personal values. For them, investing in stocks would create a discrepancy between actions and values—cognitive dissonance—and so cause an additional mental participation cost. Accordingly, Kaustia and Torstila (2011) find a positive association between voting right of center and investing in stocks. Changwony et al. (2015) find a similar result.

Frederick (2005) shows that individuals with high cognitive reflection test (CRT) scores are less loss averse compared to their peers with lower CRT scores. Benjamin et al. (2013) find that more cognitively able individuals are more risk neutral over small stakes and more patient over short time horizons. Dohmen et al. (2010) study a random sample of 1000 German adults and report that lower cognitive abilities are associated with greater risk aversion and impatience even controlling for education, income, and credit constraints. Higher cognitive abilities can thus affect stock market participation through risk preferences, in addition to lowering information costs. Christelis et al. (2010) find that cognitive abilities are strongly associated with equity ownership using the first wave of SHARE data. Grinblatt et al. (2011) find

[4] In theory, the likelihood of participation can also increase with risk aversion. Gomes and Michaelides (2005) show that this happens for plausible parameter values with uninsurable labor income risk and fixed entry costs.

that IQ scores measured for young adults predict stock market participation later in life, and mathematical skill is the most influential subcomponent of IQ. Van Rooij et al. (2007) show the same for a specific skill, namely scores on a financial literacy test. However, Cole et al. (2014) do not find evidence of financial literacy education affecting participation in a natural experiment.

Several studies have investigated the effect of health on portfolio choice, specifically whether being in poor health reduces investment in risky assets (Rosen and Wu, 2004; Edwards, 2008). Elliott and Zhao (2009) and Love and Smith (2010) find evidence consistent with a causal effect, although the latter paper only for married households. Atella et al. (2012) show that the negative effect of bad self-assessed health on the probability to hold stocks is only present in countries with no national healthcare system, while Bressan et al. (2014) show that health measures other than self-assessments are irrelevant for portfolio choice.

The relationship between religiosity and stockholding has received a fair amount of attention, but the results are mixed. While going to church is positively associated with stock market participation (Hong et al., 2004), Georgarakos and Pasini (2011) finds religiosity has a negative effect, and Changwony et al. (2015) find little effect when controlling for other types of social engagement. Differences between religions may also be relevant. Halek and Eisenhauer (2001) show Catholics and Jews to be more tolerant toward speculative risk-taking, and accordingly, Kumar et al. (2011) find more risk-seeking investment behavior to be more prevalent in regions where the Catholic-Protestant ratio is high.

Formative childhood or youth experiences, or dire conditions in the (local and/or global) economy, can leave a trace in an individual's memory for a long time and reduce participation. Malmendier and Nagel (2011) and Knüpfer et al. (2017) find empirical evidence consistent with that idea. Finally, an individual's physique has been connected to financial decisions by Addoum et al. (2017). They show taller and less obese people are more likely to hold stocks.

3. Data

We employ data from the first four main waves of the cross-national Survey of Health, Aging and Retirement in Europe (SHARE).[5] Waves 1, 2, 4, and 5 were conducted primarily in 2004, 2006, 2011, and 2013, respectively, interviewing more than 100,000 individuals from 20 countries. Dropping Ireland due to a data conformity issue, our sample includes 19 countries: Austria, Belgium, Czech Republic, Denmark, Estonia, France, Germany, Greece, Hungary, Israel, Italy, Luxembourg, Netherlands, Poland, Portugal, Slovenia, Spain, Sweden, Switzerland.

The survey collects information on health, socio-economic status and social networks of individuals aged 50 and above and their spouses who can be of any age. SHARE achieves wave-to-wave retention rates of 70% or more for nearly all countries (Bergmann et al., 2019) The main questionnaire is partly based on the Health and Retirement Study (HRS) and the English Longitudinal Study of Aging (ELSA). All questions are standardized across countries. The survey has largely held its form through the waves, allowing us to pool together information over time. Most of the questions (e.g., those related to cognitive abilities, health, social activities, employment status and future expectations) are posed separately to household members. However, it is possible for a household to appoint one household member as a "finance responsible" to answer questions related to assets and financial transfers on behalf of the entire household. We use the household-level figures. We do not use survey weights in our analyses.[6]

To diminish data loss due to non-response, SHARE data have been completed using a method called multiple imputation. If a respondent has not reported her annual income, for instance, the imputation method generates five estimates based on her other characteristics, and the income and characteristics of other respondents.[7] For each wave of SHARE, the imputation module contains five sets of data, each representing a different draw from the distribution of missing values.

3.1. Variables

The zero/one stock market participation variables we aim to explain are *Direct participation*, indicating whether a respondent owns stocks directly, and *Total participation*, including stockholdings also through mutual funds or IRAs on top of direct ones. We also analyze *Equity share*, which expresses the fraction of a respondent's financial assets accounted for by direct and indirect stockholdings. Our main focus is on *Total participation*. The explanatory variables are age, gender, education, total assets, total liabilities, income, risk aversion, sociability, trust, political preference, cognitive skills, health, body mass index (BMI), and height.

Perhaps most important among the explanatory variables, separating our study from several earlier ones, is risk aversion. It is measured with the following question, similar to the US Survey of Consumer Finances (SCF; see, e.g., Haliassos and Bertaut, 1995, p. 1121). "Thinking about financial risk that you are willing to take, do you (a) Take substantial financial risks expecting to earn substantial returns, (b) Take above average financial risks expecting to earn above average returns, (c) Take average financial risks expecting to earn average returns, and (d) Not willing to take any financial risks." We convert these answers to a four-point scale where 4 indicates maximum risk aversion (answer d) and 1 is least risk averse (answer a). As shown by Dohmen et al. (2011), this type of a directly queried measure is a good proxy for actual risk taking.

In the interest of space, we omit here the detailed definitions of the other explanatory variables. Instead, they are provided in Appendix 1.

3.2. Descriptive analysis

Fig. 2 shows the wave-to-wave development of total and direct stock market participation rates in the ten countries that took part in all four waves. The rates are clearly highest in the Nordic countries of Sweden and Denmark, where total participation is around 60%. In comparison, in the Mediterranean countries of Spain and Italy, about 10% own any stocks. Typically, about half of total participation comes from indirect stock investments.

Fig. 2 also illustrates a general increasing preference for indirect stockholdings and a decline of direct holdings. To get a rough idea of a potential relation between equity prices and participation over time, Fig. 3 plots the cumulative returns of the Euro Stoxx 50 index during the sample period together with the participation rates aggregated across countries covered in all waves. Changes in par-

[5] As of March 2022, eight waves have been collected in total We use primarily the first, second, fourth, and fifth. We make limited use of the third wave, collected in 2009, because of its focus on early-life experiences.

[6] We base our choice on Solon et al. (2015): we do not seek population descriptive statistics; we use probit with clustered standard errors to account for heteroskedasticity; the sampling procedure is unlikely to have led to excess inclusion or omission of stock market participants.

[7] A detailed description of the imputation procedure is given by Christelis (2011) and in the SHARE Wave 4 Release Guide 1.1.1, available at http://www.share-project.org/.

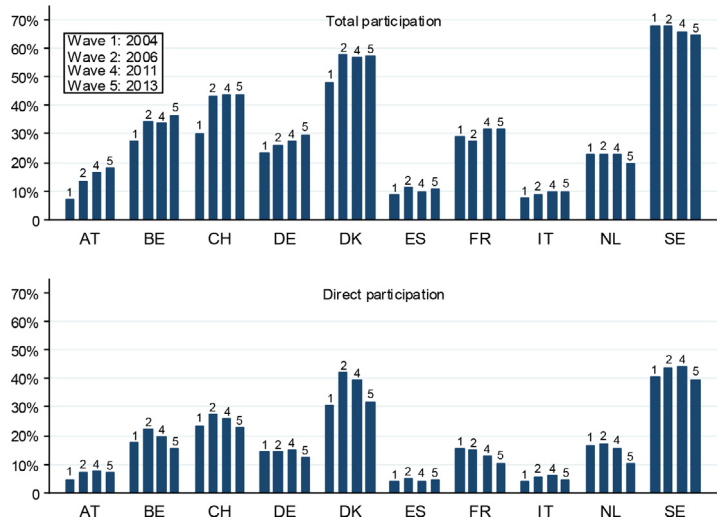


Fig. 2. Total and direct stock market participation by country in the four waves.
The figure shows the fraction of respondents who have invested in stocks either directly or indirectly through mutual funds or IRAs (total participation combines direct and indirect participation). Only countries with data available for all four main waves are included. *AT* is short for Austria, *BE* for Belgium, *CH* for Switzerland, *DE* for Germany, *DK* for Denmark, *ES* for Spain, *FR* for France, *IT* for Italy, *NL* for the Netherlands, and *SE* for Sweden.

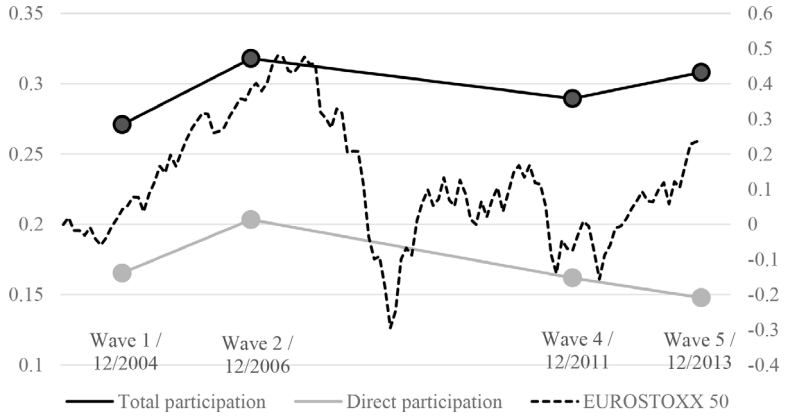


Fig. 3. Participation in the individual survey waves and equity prices in Europe.
The fraction of respondents who own stocks in each wave is shown on the left axis. Cumulative returns of the Euro Stoxx 50 index, starting January 2004, are shown on the right axis. The index covers 50 blue-chip stocks from 12 Eurozone countries.

ticipation rates are congruent with changes in stock prices, except for the decline in direct participation between the 2011 and 2013 waves.

Table 2 summarizes respondent demographics by stock market participation status. Participants are younger than non-participants, and those with indirect stockholdings form the youngest group. As the sampled individuals are aged 50 and above, this is in line with Choi and Robertson (2020) who find that US households are most likely to participate in their 50 s. The pattern can be due to pension savings indirectly allocated to stocks through mutual funds or IRAs that are gradually depleted or turned into safer assets after retirement. Participants are more often men, have more education, earn more, and have accumulated more wealth. All the differences are greater when comparing non-stockholders to those who directly own stocks.

Table 3 turns to the psychological and physical attributes of the respondents. Risk aversion is lower, trust in others higher, political preferences right of center, religiosity lower, and scores in cognitive skills higher, among stock investors. The median stockholder reports three social activities compared to one reported by the median non-stockholder, is taller both in absolute terms and proportionally to her weight, and considers herself to be in better health.

4. Regression results

4.1. Drivers of stock market participation

Table 4 presents our baseline analysis of drivers of stock market participation. It reports results from probit regressions where the dependent variable takes the value of one if an individual holds stocks either directly or through equity mutual funds (columns 1 through 6). To enable comparisons across regressors, the average marginal effect of a one-standard-deviation increase is reported for each non-dummy variable.[8] Squared terms are included for the effects of age and relative height, consistent with prior studies, though not separately tabulated.

In Column 1, we include only the traditional determinants for risky asset holdings: demographics and risk aversion, as well as country fixed effects. Column 2 deviates from the others in presentation format. Instead of a single regression, it summarizes the results of nine separate regressions. In each regression, only one of the new determinants at a time is present (i.e., added to the

[8] For the dummy regressors Male, Hard life experience, and Financial hardship, the coefficient gives the marginal effect of change from 0 to 1.

specification of col. 1), and the coefficient of that determinant is displayed in col. 2. For example, the coefficient of 0.023 for sociability is from a regression like col. 1, except that also sociability is included. Column 3-BL is the baseline specification, including all variables of interest and the full set of controls. In column 4 we drop risk aversion, to see the effects of omitting this key variable. Column 5 adds two life experience variables, hard life experience and financial hardship, that are only available from Wave 3 for a smaller set of countries. In column 6, we drop the country fixed effects. Finally, in column 7, the dependent variable is the *Direct participation* dummy, while the explanatory variables are the same as in the baseline column 3-BL.

We cluster standard errors at the household level. This is particularly important since some of the variables are not collected from a respondent in all waves. In these cases we update missing values using information from other waves. Characteristics for which information has been updated include education, risk aversion, cognitive abilities, trust, political preferences, and religiosity. These variables can also be quite static, even if collected in each wave. To investigate any remaining effect on standard errors, we replicate the main analysis using a single cross-section (Wave 5). The findings remain the same, although the smaller number of observations leads to slightly less precise estimates.

Controlling for country effects, the traditional variables all have anticipated effects on stock market participation, as seen in column 1. Adding each of the new variables to this specification on their own in column 2, we see that all of them behave as suggested by prior studies. The baseline specification, with all variables simultaneously included, is presented in column 3-BL. A number of coefficients do decrease in both size and significance in this horse race, but height is the only new variable that loses its predictive power altogether. However, as the body mass index describes a respondent's weight in proportion to her height, height still plays a role through that channel. In sum, the findings in column 3-BL show that even in the presence of a comprehensive set of controls, each new variable provides some additional information. Sociability (2.3 percentage point effect size), cognitive skills (1.6.), and religiosity (−1.3) are the most economically significant new variables, where effects are per one-standard-deviation increase, on a 27% base rate frequency of participation.

As shown in Table 1, measures of risk aversion have been either indirectly derived, or not available at all, in many prior studies explaining participation with new variables. Column 4 drops the risk aversion control to examine any bias this might cause in other coefficients. Of the basic demographics, particularly being male, relatively young, and well-paid appear to be positively associated with risk tolerance, as the marginal effects attributed to these variables are inflated vis-à-vis column 3-BL. Of the new variables, the same applies to trust, right of center political orientation, health, and, to a slightly lesser extent, cognitive skills and the body mass index.

Participation studies have traditionally been conducted with national data sets. Databases like SHARE allow us to see whether the drivers of participation vary across different cultural and regulatory environments. In column 6, we take a first look at this issue by omitting country dummies. This alters a number of the estimates quite substantially. We examine country differences further in Section 5.

Column 7 limits to directly held stocks only. Compared to other households in the sample, those with direct stockholdings are on average older, not younger. Most interestingly, individuals who are relatively trusting toward others are not more likely to make direct stock investments. If purchasing a company's stock directly is indicative of a person's trust toward that company, this finding would suggest that trust toward companies, as proposed by Guiso et al. (2008) and Giannetti and Wang (2016), plays little part in the overall effect trust has for stock market participation.

4.2. Overall model success

How well do these models identify stock market participants? In this subsection, we take a first look at the big-picture explanatory power (we return to the issue with more detail in Section 5). The baseline model, with 35 regressors,[9] has a pseudo-R-squared of 0.298, so its log-likelihood is 0.7 times that of a model predict-

[9] In addition to the 15 variables shown in Table 4, the set of 35 includes the squared terms of age and height, and 18 country dummies.

Table 2
Summary of demographics by participation status
For individuals who participated in more than one survey wave, only the latest observation is used. *ISCED-97* stands for the International Standard Classification of Education 1997, a 0–6 scale allowing international comparisons of education level. *Income* is annual, and both income and wealth are reported at the household level. *Financial wealth* includes bank accounts, bonds, stocks, mutual funds, and long-term savings, while *Total wealth* also includes real estate, business wealth, and cars. *Any stockholdings* means that respondent who own stocks in any form, either directly or indirectly, are included in the sample.

	Age	Male	Education (ISCED-97)	Income ('000 €)	Net wealth ('000 €)	
					Financial	Total
No stockholdings						
Mean	68.03	0.40	2.53	26.48	19.75	179.51
Median	67	0	3	15.80	2.43	95.34
Standard deviation	11.02	0.49	1.46	69.79	86.28	539.63
# observations	55,299	55,299	54,785	55,299	55,299	55,299
Any stockholdings						
Mean	64.46	0.53	3.46	55.12	120.33	420.18
Median	63	1	3	33.89	57.27	262.97
Standard deviation	9.72	0.50	1.41	102.19	224.00	633.69
# observations	16,182	16,182	16,023	16,182	16,182	16,182
Direct stockholdings						
Mean	65.66	0.58	3.57	61.86	150.98	506.26
Median	65	1	3	37.79	76.40	321.93
Standard deviation	9.93	0.49	1.41	118.02	259.45	721.90
# observations	7679	7679	7588	7679	7679	7679
Full sample						
Mean	66.47	0.44	2.74	34.15	45.57	249.59
Median	65	0	3	20.38	6.71	135.84
Standard deviation	10.74	0.50	1.49	80.81	143.29	593.36
# observations	107,894	107,894	106,277	107,894	107,894	107,894

Table 3

Summary of psychological and physical attributes by participation status

Only the latest observation is used from participants in more than one wave. *Risk aversion* (1–4 scale) indicates willingness to a) take substantial financial risks expecting to earn substantial returns, b) take above average financial risks expecting to earn above average returns, c) take average financial risks expecting to earn average returns, or d) not take any financial risks. *Sociability* counts the social activities a respondent reports during the past year (Waves 4, 5) or month (Waves 1, 2). *Trust* is response to, 0–10 scale, "Generally speaking, would you say that most people can be trusted or that you can't be too careful?". *Right of c.* is response to, scale 0–10, where zero means left and ten means right, where would you place yourself?". *Cognitive skills* is an index on scores of numeracy, fluency, and recall tests, where scores are first demeaned and divided by their sample s. dev., then averaged. *Health* is 1–5 scale on "Would you say your health is (1) excellent - (5) poor?". *Religiosity* is a 0–5 scale measuring the frequency of praying. *Height* is in centimeters, and *BMI* is Weight (kg) / [Height (m)]^2; both standardized relative to country and gender. For *Hard life experience* and *Financial hardship* are indicators, please see Appendix 1 for definitions. *Any stockholdings* means stocks held either directly or indirectly.

	Risk avers.(1–4)	Sociability (0–5)	Trust (0–10)	Right of center (0–10)	Cognitive skills	Health (1–5)	Religiosity (0–5)	Height (cm)	BMI	Hard life exper. (0–1)	Financial hardship (0–1)
No stockholdings											
Mean	3.82	0.51	5.52	4.99	−0.12	2.68	1.91	166.81	27.14	0.10	0.27
Median	4	0	5	5	−0.12	3	1	166	26.48	0	0
Standard deviation	0.47	0.80	2.46	2.28	0.60	1.08	1.92	9.03	4.77	0.30	0.45
# observations	49,170	54,215	49,503	43,344	54,792	55,299	49,085	55,091	54,958	13,863	13,890
Any stockholdings											
Mean	3.36	1.02	6.33	5.27	0.18	3.28	1.26	170.84	26.31	0.09	0.27
Median	3	1	7	5	0.18	3	0	170	25.71	0	0
Standard deviation	0.72	1.02	2.24	2.24	0.55	1.06	1.68	9.28	4.24	0.28	0.44
# observations	14,909	16,065	14,987	14,298	16,138	16,182	14,770	16,164	16,144	4519	4526
Direct stockholdings											
Mean	3.23	1.06	6.44	5.43	0.22	3.35	1.29	171.54	26.04	0.09	0.26
Median	3	1	7	5	0.22	3	0	172	25.51	0	0
Standard deviation	0.76	1.02	2.23	2.23	0.57	1.06	1.68	9.15	4.03	0.29	0.44
# observations	6919	7626	6969	6657	7651	7679	6864	7667	7655	2362	2366
Full sample											
Mean	3.71	0.61	5.72	5.07	−0.04	2.82	1.72	167.96	26.88	0.09	0.26
Median	4	0	6	5	−0.03	3	1	168	26.23	0	0
Standard deviation	0.57	0.87	2.42	2.28	0.60	1.10	1.87	9.15	4.62	0.29	0.44
# observations	90,729	105,229	96,157	86,026	106,681	107,809	95,313	107,376	107,071	27,177	27,239

ing participation with a constant only.[10] It is not obvious, however, what this implies in terms of absolute predictive power.

As a more intuitive measure of model performance, we use "*c*" – the area under the receiver operating characteristic (ROC) curve.[11] *c* indicates how well the model can discriminate between stock market participants and nonparticipants. Ranging from 0.5 to 1.0, it gives the probability of the model predicting a higher probability for participants than for nonparticipants. For example, 0.5 would indicate the model is equal to a coin toss, and 1.0 indicates perfect classification. In Table 4 column 3, we see that c is 0.85; looking across all participant-nonparticipant pairs in the sample, the model in column 3 predicts a higher probability of participation for the participant in 85.2% of the pairs. Values of $c > 0.8$ are considered excellent (Hosmer and Lemeshow, 2000). The high values of c across the columns in Table 4 provide further evidence of the model's strength.

Fig. 4, Panel A shows the ROC curve plotting the model's sensitivity against (1-specificity). Sensitivity, also known as the true positive rate, is the percentage of stock market participants for whom the model generates a predicted probability higher than a given cutoff point. Specificity, also known as the true negative rate, is the percentage of nonparticipants for whom the model generates a predicted probability lower than a given cutoff point. There are many points in Panel A, due to multiple imputations in the data. Panel B presents another view of sensitivity and specificity. It rounds each observation's predicted probability to the nearest 0.01 and then averages the sensitivity and specificity across observations at each predicted probability level.

5. The role of institutional, traditional, and new factors

It is natural that participation rates are higher in more sophisticated environments and subsamples of individuals. In this section we are interested in testing whether there are any systematic changes in the effects of individual variables as one moves across such subsamples. This also works as a check of robustness and stability of the effects. However, we do not think all variables should necessarily have the same effect across the subsamples, an issue we return to later with a discussion of a hierarchical model. Here we use country regulatory environment, and individual wealth and education, as indicators of sophistication.

5.1. Determinants of participation in different regulatory environments

We rank countries by their regulatory quality, as measured by the World Bank's Worldwide Governance Indicators. The measure aims to capture governmental ability to "permit and promote private sector development" through policy, including investment freedom and transparency of institutions. Country-level analyses find investor protection and regulation to be important in promoting household equity ownership (Pagano and Volpin, 2006; Giannetti and Koskinen, 2010).

Panel A of Table 5 shows the country groups of high, medium, and low regulatory quality, and Panel B presents regression results for each group. Participation increases strongly and monotonically with regulatory quality. As discussed above, the model's ability to correctly identify a stockholder improves in subsamples with higher participation rates, here corresponding to moving up

[10] The McFadden pseudo-R-squared used in our estimations is defined as follows: pseudo-R^2 = 1 – (log-pseudo-likelihood of full model / log-pseudo-likelihood of constant-only model). As likelihoods are between zero and one, their logarithms are negative, and a smaller log-likelihood means better fit.

[11] For a good introduction to the *c* statistic, see e.g., Hosmer and Lemeshow (2000).

Table 4
Determinants of stock market participation

Average marginal effects on the probability of stockholding of a one-standard-deviation change in an explanatory variable (zero-to-one for dummies), obtained through probit regressions. *Any holdings* means that a respondent holds stocks either directly or indirectly, *Direct* means that stocks are owned directly. Column 1 only includes traditional individual level variables. Column 2 summarizes results from ten separate regressions. Each of them is like the one presented in column 1, i.e., having all traditional variables, except that one of the ten new variables is added as an explanatory variable. This is done with each new variable, one at a time. In column 2, the coefficient of the new variable is shown, while the rest of the output is omitted. Column 3-BL is the baseline specification. Column 4 drops risk aversion, and Column 5 adds two life experience variables only available in survey wave 3 covering a different set of countries . Column 6 drops country fixed effects. Column 7 is like baseline (column 3-BL), but the dependent variable is direct holdings only. *Education* has six categories based on the International Standard Classification of Education 1997 (ISCED-97). *Income* (annual, net) and wealth measures are on the household level, in euro. *Assets* include gross financial wealth (bank accounts, bonds, stocks, mutual funds, long-term savings) and real wealth (real estate, business, cars). *Liabilities* include mortgages and any financial liabilities. *Risk avers.* is a 1–4 scale indicating whether a respondent is willing to (a) take substantial financial risks expecting to earn substantial returns, (b) take above average financial risks expecting to earn above average returns, (c) take average financial risks expecting to earn average returns, or (d) not take any financial risks. *Sociability* is a variable that counts the social activities a respondent reports to have engaged in during the previous year (Waves 4 and 5) or month (Waves 1 and 2). *Trust* equals the response, on a 0–10 scale, to "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?". *Right of cen.* equals the response to "In politics people sometimes talk of 'left' and 'right'. On a scale from zero to ten, where zero means left and ten means right, where would you place yourself?". *Cog. skills* is an index based on scores of numeracy, fluency, and recall tests, where the three scores are first demeaned and divided by their sample standard deviations, and then averaged. *Health* is a 1–5 scale based on "Would you say your health is (1) excellent, ..., (5) poor?". *Religiosity* is a 0–5 scale measuring the frequency of praying, where zero means never and five means more than once a day. *Height* is reported in centimeters, and *BMI* is calculated as Weight (kg) / [Height (m)]^2; both *Height* and *BMI* are standardized relative to country, gender, age, and survey wave. *Hard life experience* is an indicator variable for ever having an adverse event in one's living situation (see Appendix 1 for the list of events). *Financial hardship* is an indicator variable for having a distinct period of financial hardship starting before 1995. Standard errors are clustered by household, and t-statistics shown in parentheses below the marginal effects. *c* is the estimated area under the ROC curve. *, **, and *** stand for statistical significance at the 10%, 5%, and 1% levels, respectively.

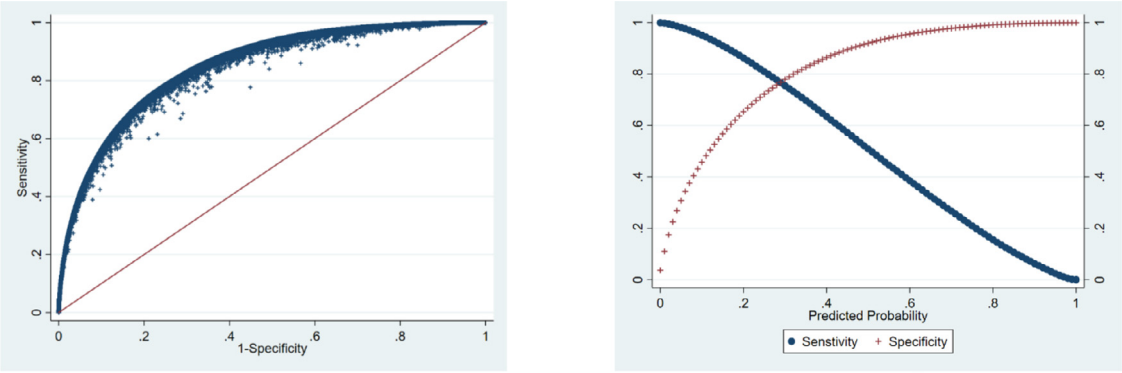
	Any stock holdings						Direct
	(1)	(2)	(3)-BL	(4)	(5)	(6)	(7)
Male	0.025***		0.021***	0.034***	0.027***	0.006**	0.012***
	(10.61)		(7.79)	(12.76)	(6.07)	(2.10)	(5.55)
Age	−0.032***		−0.021***	−0.032***	−0.021***	0.000	0.011***
	(−19.61)		(−14.76)	(−17.43)	(−5.89)	(0.91)	(4.98)
Education	0.042***		0.031***	0.034***	0.029***	0.027***	0.018***
	(31.86)		(20.35)	(22.26)	(11.19)	(16.57)	(13.25)
Ln(Income)	0.025***		0.021***	0.029***	0.019***	0.052***	0.015***
	(15.95)		(13.19)	(17.09)	(8.9)	(28.2)	(10.6)
Ln(Assets)	0.122***		0.120***	0.126***	0.154***	0.110***	0.094***
	(69.87)		(62.16)	(63.16)	(40.47)	(55.40)	(43.43)
Ln(Liab.)	−0.019***		−0.019***	−0.019***	−0.01***	0.005**	−0.019***
	(−13.97)		(−14.37)	(−14.05)	(−10.46)	(2.52)	(−14.15)
Risk avers.	−0.046***		−0.045***		−0.045***	−0.055***	−0.032***
	(−46.45)		(−42.40)		(−26.51)	(−50.29)	(−39.27)
Sociability		0.023***	0.023***	0.024***	0.025***	0.03***	0.011***
		(22.21)	(20.02)	(21.15)	(12.56)	(25.02)	(12.27)
Trust		0.010***	0.007***	0.01***	0.007***	0.019***	0.002
		(8.7)	(5.85)	(6.52)	(2.81)	(14.14)	(1.17)
Right of cen.		0.007***	0.007***	0.011***	0.009***	0.011***	0.007***
		(4.48)	(5.8)	(7.77)	(3.91)	(8.64)	(6.74)
Cog. skills		0.018***	0.016***	0.02***	0.021***	0.028***	0.015***
		(14.39)	(11.37)	(13.94)	(9.12)	(19.14)	(12.24)
Health		0.013***	0.008***	0.012***	0.005**	0.019***	0.003***
		(10.42)	(5.66)	(8.48)	(2.40)	(12.78)	(2.67)
Religiosity		−0.007***	−0.013***	−0.013***	−0.015***	−0.037***	−0.007***
		(−6.05)	(−8.93)	(−8.79)	(−5.95)	(−25.09)	(−5.75)
Height		0.003**	0.000	0.002	−0.001	−0.003**	0.000
		(2.52)	(0.36)	(1.39)	(−0.36)	(−2.28)	(−0.07)
BMI		−0.007***	−0.005***	−0.006***	−0.004*	−0.003*	−0.005***
		(−5.72)	(−3.82)	(−3.94)	(−1.82)	(−1.75)	(−3.87)
Hard life exp		−0.001			0.002		
		(−0.13)			(0.30)		
Fin. hardship		−0.012**			−0.016***		
		(−2.56)			(−3.01)		
Country eff.	Yes	Yes	Yes	Yes	Yes	No	Yes
Pseudo R^2	0.297		0.298	0.281	0.320	0.232	0.280
c	0.853		0.852	0.844	0.861	0.817	0.86
N	118,106		104,610	106,113	42,709	104,610	103,836

in regulatory quality. While most of the determinants remain significant in all three groups, they are often stronger in countries with higher regulatory quality.

Interpersonal trust, political orientation, and self-assessed health do not predict participation in countries of low regulatory quality. The results for trust contrast the findings of Guiso et al. (2008), who find trust to predict stock ownership in both Italy (low regulatory quality) and the Netherlands (high regulatory quality). Kaustia and Torstila's (2011) and Changwony's et al.'s (2015) findings of right of center political preferences predicting participation come from countries of high regulatory quality and appear not to generalize to countries of low regulatory quality.[12]

Regarding health, Atella et al. (2012) use SHARE data to show that health only affects participation in countries with no national healthcare system. Their group of countries with a national

[12] Kaustia and Torstila's (2011) data come from Finland, whose regulatory quality score is highest in Europe at 1.90. Changwony et al.'s (2015) data are from the UK, whose score of 1.83 is still a tad higher than Switzerland's 1.82, the top-ranking country in the SHARE sample.



Panel A. ROC curve

Panel B. Sensitivity vs. Specificity

Fig. 4. ROC curve, sensitivity and specificity for the baseline model.
The figure displays two ways of viewing the baseline model's ability to correctly classify participants and nonparticipants. Panel A shows the Receiver Operating Characteristics curve. The curve plots sensitivity (the true positive rate) against the false positive rate (calculated as 1-specificity) across cutoff points for the probability of stock market participation ranging from 0 to 1. The curve is not a smooth line due to the multiple imputations in the data. Panel B plots the sensitivity (the true positive rate) and specificity (the true negative rate) for predicted probabilities of stock market participation ranging from 0 to 1. To generate the smooth lines, we first round each observation's predicted probability to the nearest 0.01, and then average the sensitivity and specificity values across observations at each predicted probability level from 0 to 1 at increments of 0.01.

healthcare system, where no effect is found, consists of Denmark, Italy, Sweden, and Spain. Atella et al. (2012) show no analysis of the countries individually. In our analysis, the effect of health is strong in countries with high regulatory quality (including both Denmark and Sweden individually), but not present in countries of low regulatory quality (including both Spain and Italy individually). The effect of health may have more to do with general institutions, rather than, or in addition to, the healthcare system per se.

5.2. Determinants of participation across wealth, education, and gender

Wealth and education are highly significant predictors of stock market participation. Next, we use these variables to split the sample into three groups of households and study the determinants of participation therein. To account for the varying levels of average wealth and education across countries, we sort by wealth and education within each country.

Table 6 shows that two of the key traditional variables, gender and education, lose their predictive power in some of the subsamples. Among those in their country's top terciles of both wealth and education, an additional year of education makes no difference for owning stocks. Of the new variables, interpersonal trust, self-assessed health, and the body mass index lose their significance in some of the subsamples. Trust does not increase participation among those with high education combined with high wealth. Health is insignificant among the highly educated, and BMI is insignificant among those with lower wealth.

Perhaps the most interesting observation from Table 6 is that even among those who are both wealthy and highly educated, personal characteristics really matter. This issue is at the core of the stock market participation puzzle: why are many of those who are likely to know about the stock market and have enough wealth not participating? The gender gap remains large, men being four percentage points more likely to own stocks. Willingness to avoid financial risk is another significant reason to stay out, implying that even in this group, many are not convinced with the historically favorable long-term risk-return relationship of stocks. For many, political values appear to provide an argument against the stock market. Interestingly, even in this group that at the outset would appear financially sophisticated, cognitive skills have a highly significant effect.

We consider variation in the effects across gender in unreported tests. Most variables get coefficients within 20% of their baseline values. Two noteworthy patterns nevertheless emerge. The positive effect of right of center political views is 63% greater for women compared to men. On the other hand, the negative effect of religiosity is 46% greater for men. If this pair of variables measures attitude toward stocks, then both genders are approximately equally exposed to such an effect overall, but with a different split between the specific variables. The other variable pair where we see a similar shift is Health and BMI. Specifically, the positive effect of health is 58% greater for women, and the negative effect of BMI is 68% greater for men. We have argued that BMI could proxy for self-control, but obviously it also has a health component to it. The two variables are not very correlated though, and especially for men the BMI picks up variation unrelated to subjective health.

5.3. The explanatory power of each factor

While the area under Receiver Operating Characteristic (ROC) – curve, denoted by *c*, is our preferred metric for overall predictive success, the analysis of R-squared is well-suited for analyzing the contribution of the institutional, traditional, and behavioral factors. To do this, we use the Shapley decomposition of the McFadden pseudo-R-squared from the probit regression.[13] This avoids deciding in which order to add regressors, something that would significantly affect the results. The Shapley method produces an additive decomposition.

Fig. 5 presents the results. Traditional individual variables typically have more explanatory power as a group compared to country effects and new variables. They account for about half of the increase in log-likelihood compared to a constant-only model. Country effects account for roughly a third, and new variables about a fifth. Considering that the traditional variables are fewest in number, at seven, compared to eight new variables and 19 country dummies, their relative influence is strong. Only in two exceptional subsamples are country effects jointly more important:

[13] The total is 29.8%, which is the McFadden pseudo-R-squared from a probit regression. OLS R-squared is almost identical at 30.0%, and the contributions of the factors are very similar. We feel it is cleaner to use R-squared unadjusted for the number of variables in the model, as our goal is not to see if the model fit improves with additional variables. Using adjusted R-squared leads to the similar conclusions. For details on the Shapley method, see Stanislav Kolenikov's STATA module documentation, https://ideas.repec.org/c/boc/bocode/s411401.html.

Table 5
Determinants of stock market participation by country regulatory quality
Panel A shows how the sample countries score on regulatory quality according to a measure provided by the World Bank's Worldwide Governance Indicators. The scores are used to divide the countries into terciles. Panel B shows average marginal effects on the probability of direct or indirect stockholding of a one-standard-deviation change in an explanatory variable (zero-to-one for dummies), obtained through probit regressions in each tercile of countries. *Education* has six categories based on the International Standard Classification of Education 1997 (ISCED-97). *Income* (annual, net) and wealth measures are on the household level, in euro. *Total assets* include gross financial wealth (bank accounts, bonds, stocks, mutual funds, long-term savings) and real wealth (real estate, business, cars). *Total liabilities* include mortgages and any financial liabilities. *Risk aversion* is a 1–4 scale indicating whether a respondent is willing to a) take substantial financial risks expecting to earn substantial returns, b) take above average financial risks expecting to earn above average returns, c) take average financial risks expecting to earn average returns, or d) not take any financial risks. *Sociability* is a variable that counts the social activities a respondent reports to have engaged in during the previous year (Waves 4 and 5) or month (Waves 1 and 2). *Trust* equals the response, on a 0–10 scale, to "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?". *Right of center* equals the response to "In politics people sometimes talk of 'left' and 'right'. On a scale from zero to ten, where zero means left and ten means right, where would you place yourself?". *Cognitive skills* is an index based on scores of numeracy, fluency, and recall tests, where the three scores are first demeaned and divided by their sample standard deviations, and then averaged. *Health* is a 1–5 scale based on "Would you say your health is (1) excellent, ..., (5) poor?". *Religiosity* is a 0–5 scale measuring the frequency of praying, where zero means never and five means more than once a day. *Height* is reported in centimeters, and *BMI* is calculated as Weight (kg) / [Height (m)]^2; both are standardized relative to country, gender, age, and survey wave. Standard errors are clustered by household, and t-statistics shown in parentheses below the marginal effects. *c* is the estimated area under the ROC curve. *, **, and *** stand for statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Regulatory quality scores from World Bank's Worldwide Governance Indicators

Top tercile	Middle tercile	Bottom tercile
Country (score)	Country (score)	Country (score)
Switzerland (1.82)	Luxembourg (1.65)	Spain (0.78)
Sweden (1.80)	Austria (1.49)	Hungary (0.77)
Netherlands (1.78)	Israel (1.21)	Portugal (0.77)
Denmark (1.72)	Belgium (1.17)	Italy (0.66)
Germany (1.70)	France (1.09)	Slovenia (0.66)
Estonia (1.67)	Poland (1.06)	Greece (0.34)
	Czech Republic (1.02)	

Panel B: Determinants of participation in regulatory quality terciles

	Top tercile	Middle tercile	Bottom tercile
Participation rate	37.6%	25.7%	10.0%
Male (0/1)	0.020***	0.021***	0.021***
	(4.41)	(4.84)	(5.02)
Age (years)	−0.032***	−0.021***	−0.021***
	(−10.94)	(−8.67)	(−6.29)
Education (ISCED)	0.028***	0.034***	0.028***
	(9.89)	(13.98)	(12.18)
Ln(Income (€))	0.007***	0.032***	0.012***
	(2.71)	(7.21)	(4.75)
Ln(Total assets (€))	0.14***	0.112***	0.075***
	(45.00)	(33.93)	(16.20)
Ln(Total liab. (€))	−0.023***	−0.023***	−0.009***
	(−10.99)	(−8.04)	(−4.66)
Risk aversion (1–4)	−0.055***	−0.046***	−0.028***
	(−30.08)	(−25.78)	(−16.75)
Sociability (0–5)	0.023***	0.027***	0.017***
	(11.55)	(14.33)	(8.39)
Trust (0–10)	0.010***	0.010***	0.000
	(4.00)	(4.81)	(0.10)
Right of center (0–10)	0.014***	0.009***	−0.002
	(5.70)	(3.70)	(−1.04)
Cognitive skills	0.018***	0.016***	0.011***
	(7.67)	(6.83)	(5.07)
Health (1–5)	0.010***	0.009***	0.003
	(4.07)	(3.75)	(1.58)
Religiosity (0–5)	−0.015***	−0.015***	−0.007***
	(−5.72)	(−5.61)	(−3.12)
Height (relative)	−0.002	0.002	0.003
	(−1.07)	(0.95)	(1.33)
BMI (relative)	−0.004*	−0.006**	−0.006***
	(−1.82)	(−2.53)	(−2.77)
Country effects	Yes	Yes	Yes
Pseudo R2	0.325	0.208	0.204
c	0.840	0.838	0.845
N	41,036	40,904	22,670

among those with high wealth, and among those with both low wealth and low education. We also perform the Shapley decompositions on the R-squares from OLS regressions, and by and large find results very similar to Fig. 5 (results not reported).

One might expect that when we look at people from countries with similar regulatory environments, individual-level differences would mainly determine participation. This is indeed what we find for countries with intermediate or low regulatory quality. Surprisingly, however, among countries where regulatory quality is high, the effect of macro-level differences is substantial. Estonia is an outlier in this group – its stock market opened in 1996 and its participation rate in the sample is about 8%. If we drop Estonia from the analysis, the contribution of country factors for high regulatory quality countries drops to 8 percentage points.

Table 7 tabulates the relative explanatory power of the three factors. We use the values shown in Fig. 5, and divide the new variables' pseudo-R-squared contribution with that of traditional variables. In the full sample baseline specification, this "explanatory power ratio" is slightly below 40 percent. Level of education affects the explanatory power ratio between new and traditional

Table 6
Determinants of stock market participation by wealth and education

Average marginal effects on the probability of direct or indirect stockholding of a one-standard-deviation change in an explanatory variable (zero-to-one for dummies), obtained through probit regressions. In Panel A, we use our full sample of responses from SHARE Waves 1, 2, 4, and 5. In Panel B, we include two life experience variables from Wave 3, which significantly reduces the sample size. The sample is divided into terciles based on *Net wealth* (financial + real wealth – mortgages and financial liabilities) and *Education* (six categories based on the International Standard Classification of Education 1997 (ISCED-97)). The divisions are made on the country level: *Low* means belonging to the bottom tercile of one's country, *High* means the top tercile. *Income* (annual, net) and wealth measures are on the household level, in euro. *Total assets* include gross financial wealth (bank accounts, bonds, stocks, mutual funds, long-term savings) and real wealth (real estate, business, cars). *Total liabilities* include mortgages and any financial liabilities. *Risk aversion* is a 1–4 scale indicating whether a respondent is willing to (a) take substantial financial risks expecting to earn substantial returns, (b) take above average financial risks expecting to earn above average returns, (c) take average financial risks expecting to earn average returns, or d() not take any financial risks. *Sociability* is a variable that counts the social activities a respondent reports to have engaged in during the previous year (Waves 4 and 5) or month (Waves 1 and 2). *Trust* equals the response, on a 0–10 scale, to "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?". *Right of center* equals the response to "In politics people sometimes talk of 'left' and 'right'. On a scale from zero to ten, where zero means left and ten means right, where would you place yourself?". *Cognitive skills* is an index based on scores of numeracy, fluency, and recall tests, where the three scores are first demeaned and divided by their sample standard deviations, and then averaged. *Health* is a 1–5 scale based on "Would you say your health is (1) excellent, ..., (5) poor?". *Religiosity* is a 0–5 scale measuring the frequency of praying, where zero means never and five means more than once a day. *Height* is reported in centimeters, and *BMI* is calculated as Weight (kg) / [Height (m)]^2; both *Height* and *BMI* are standardized relative to country, gender, age, and survey wave. *Hard life experience* is an indicator variable for ever having an adverse event in one's living situation (see Appendix 1 for the list of events). *Financial hardship* is an indicator variable for having a distinct period of financial hardship starting before 1995. Standard errors are clustered by household, and t-statistics shown in parentheses below the marginal effects. *c* is the estimated area under the ROC curve. *, **, and *** stand for statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A.	Net wealth		Education		Both low	Both high
	Low	High	Low	High		
Participation rate	13.0%	42.4%	19.3%	38.8%	10.4%	50.9%
Male (0/1)	0.002	0.035***	0.010***	0.038***	0.004	0.042***
	(0.61)	(6.21)	(2.87)	(6.28)	(0.83)	(4.44)
Age (years)	−0.021***	−0.032***	−0.021***	−0.032***	−0.021***	−0.032***
	(−9.23)	(−6.95)	(−11.26)	(−7.75)	(−7.18)	(−4.20)
Education (ISCED)	0.018***	0.039***	0.043***	0.022**	0.021***	0.018
	(8.35)	(13.34)	(9.55)	(2.38)	(3.53)	(1.34)
Ln(Income (€))	0.007***	0.032***	0.015***	0.021***	0.007***	0.033***
	(3.90)	(7.18)	(6.71)	(4.80)	(3.32)	(4.16)
Ln(Total assets (€))	0.073***	0.126***	0.092***	0.157***	0.063***	0.175***
	(32.29)	(13.95)	(39.93)	(32.27)	(23.74)	(11.14)
Ln(Total liab. (€))	−0.019***	−0.023***	−0.019***	−0.028***	−0.019***	−0.028***
	(−8.63)	(−8.16)	(−8.94)	(−8.54)	(−6.57)	(−6.22)
Risk aversion (1–4)	−0.021***	−0.063***	−0.034***	−0.065***	−0.016***	−0.072***
	(−14.12)	(−28.87)	(−23.83)	(−27.98)	(−8.73)	(−19.98)
Sociability (0–5)	0.017***	0.027***	0.023***	0.028***	0.017***	0.030***
	(9.57)	(12.17)	(14.10)	(11.96)	(7.37)	(8.20)
Trust (0–10)	0.005***	0.010***	0.010***	0.007**	0.005**	0.005
	(2.90)	(3.09)	(5.50)	(2.19)	(2.29)	(1.02)
Right of center (0–10)	0.005***	0.009***	0.005***	0.014***	0.005**	0.018***
	(3.22)	(3.51)	(3.31)	(4.52)	(2.31)	(3.76)
Cognitive skills	0.008***	0.022***	0.017***	0.016***	0.011***	0.020***
	(4.19)	(7.11)	(9.48)	(4.93)	(4.67)	(3.94)
Health (1–5)	0.009***	0.006*	0.009***	0.002	0.008***	−0.002
	(4.45)	(1.90)	(5.06)	(0.71)	(3.61)	(−0.54)
Religiosity (0–5)	−0.009***	−0.015***	−0.013***	−0.019***	−0.007***	−0.013***
	(−4.26)	(−4.59)	(−7.24)	(−5.40)	(−3.06)	(−2.59)
Height (relative)	−0.003	0.004	0.000	0.001	−0.005**	0.000
	(−1.43)	(1.39)	(−0.17)	(0.47)	(−2.41)	(−0.07)
BMI (relative)	−0.002	−0.012***	−0.004**	−0.014***	−0.002	−0.017***
	(−0.96)	(−4.31)	(−2.23)	(−4.24)	(−0.93)	(−3.36)
Country eff	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.278	0.255	0.308	0.242	0.283	0.197
c	0.856	0.824	0.863	0.816	0.866	0.789
N	35,624	34,822	54,173	25,671	22,342	12,608
Panel B.						
Participation rate	0.14	0.44	0.21	0.41	0.11	0.53
Male (0/1)	0.007	0.043***	0.017***	0.050***	0.004	0.059***
	(1.07)	(5.46)	(2.91)	(4.96)	(0.53)	(4.04)
Age (years)	−0.010***	−0.021***	−0.021***	−0.010*	−0.010***	−0.010
	(−3.62)	(−3.12)	(−5.27)	(−1.75)	(−3.21)	(−1.27)
Education (ISCED)	0.015***	0.038***	0.042***	0.024	0.020**	0.009
	(3.82)	(7.60)	(5.29)	(1.48)	(1.99)	(0.42)
Ln(Income (€))	0.011***	0.030***	0.012***	0.030***	0.010***	0.039***
	(4.16)	(6.41)	(4.78)	(5.6)	(2.78)	(4.48)
Ln(Total assets (€))	0.097***	0.132***	0.121***	0.203***	0.09***	0.192***
	(21.73)	(8.29)	(26.49)	(19.73)	(15.62)	(6.42)
Ln(Total liab. (€))	−0.010***	−0.012***	−0.01***	−0.012***	−0.01***	−0.013***
	(−6.57)	(−5.98)	(−6.92)	(−5.45)	(−4.84)	(−4.44)
Risk aversion (1–4)	−0.021***	−0.061***	−0.033***	−0.065***	−0.017***	−0.067***
	(−8.72)	(−17.88)	(−14.83)	(−17.12)	(−5.52)	(−11.7)
Sociability (0–5)	0.016***	0.027***	0.025***	0.026***	0.02***	0.028***
	(5.40)	(7.32)	(9.13)	(6.87)	(5.14)	(4.93)
Trust (0–10)	0.005	0.007*	0.01***	0.005	0.002	0.007
	(1.38)	(1.86)	(3.04)	(0.94)	(0.9)	(0.88)

(*continued on next page*)

Table 6 (*continued*)

Right of center (0–10)	0.009***	0.011**	0.009***	0.018***	0.009***	0.023***
	(2.75)	(2.45)	(2.76)	(3.24)	(2.65)	(2.83)
Cognitive skills	0.016***	0.027***	0.024***	0.026***	0.017***	0.029***
	(4.79)	(5.80)	(8.15)	(4.84)	(4.37)	(3.53)
Health (1–5)	0.004	−0.001	0.008***	−0.011**	0.005	−0.014*
	(1.38)	(−0.36)	(2.72)	(−1.96)	(1.31)	(−1.72)
Religiosity (0–5)	−0.011***	−0.015***	−0.013***	−0.031***	−0.01**	−0.021**
	(−3.13)	(−3.24)	(−4.03)	(−5.35)	(−2.24)	(−2.46)
Height (relative)	−0.007**	0.005	−0.003	−0.006	−0.010***	−0.009
	(−2.19)	(1.24)	(−0.89)	(−1.23)	(−2.75)	(−1.26)
BMI (relative)	−0.001	−0.011**	−0.003	−0.014**	0.000	−0.014*
	(−0.39)	(−2.51)	(−1.02)	(−2.53)	(−0.03)	(−1.77)
Hard life exp.	0.001	0.006	0.003	0.021	0.005	0.016
	(0.08)	(0.38)	(0.32)	(1.25)	(0.49)	(0.66)
Fin. hardship	−0.009	−0.016	−0.009	−0.031**	−0.005	−0.030*
	(−1.37)	(−1.61)	(−1.34)	(−2.50)	(−0.58)	(−1.67)
Country eff.	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.292	0.282	0.326	0.267	0.293	0.222
c	0.862	0.838	0.869	0.829	0.866	0.804
N	14,043	14,745	22,447	9820	8964	5022

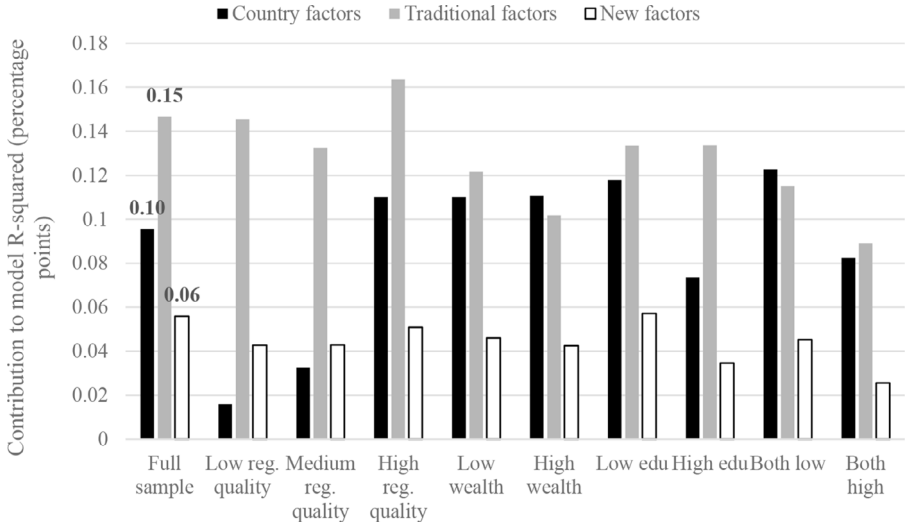


Fig. 5. Relative explanatory power of different factors.
In the graph, the pseudo-R-squared of our main specification in all subsamples is divided into three parts via a Shapley decomposition. Each part represents the contribution to explanatory power of a group of variables. Absolute values are shown, not fractions: the sum of the three components equals the pseudo-R-squared of the regression. The first variable group is *Country factors*, including a dummy for each of the 19 countries (one omitted). The second group is *Traditional factors*, including gender, age, education, income, wealth, and risk aversion. The third group is *New factors*, including sociability, trust, political orientation, cognitive skills, health, religiosity, height, and the body mass index.

variables. Among the less educated, the new variables contribute 43% of what the traditional variables do. Among the highly educated, the fraction is 26%. This is one demonstration of the challenge in explaining participation in the high end of households.

5.4. Which new variables appear the most important overall?

Table 8 shows that the set of new variables for which we estimate the predicted effect with statistical significance in all analyses comprises sociability, cognitive skills, and religiosity. The consistently robust variables are also those with the largest effect sizes in the baseline specification. Looking at the estimated changes in stockholding probability resulting from a one-standard-deviation increase, sociability ranks highest with 2.4 percentage points. This corresponds to almost 10% of the sample propensity to hold stocks. Cognitive skills rank second with an effect size of 1.6 pp. and religiosity is third with −1.3 pp. The effect sizes for trust and body mass index (BMI) are rather small, but they are consistent throughout all the specifications.

The effect size of financial hardship is also large at 1.6 pp. Besides coming from a smaller set of countries (40% of observations in the baseline model), one caveat is that it is probably the most endogenous among the new variables in our specifications, even if it is measured controlling for wealth and income and at 10 years prior to recording stock ownership for the first time. We are not going to argue that other variables, such as sociability and cognitive skill, are totally exogenous either, but probably more so.

5.5. Unique individual determinants or broad latent factors?

So far, we have shown how the new variables are related to stock market participation, both as individual drivers (Table 4) and grouped into factors (Fig. 5). Is some of this variation within the individual variables common to one or more latent factors, and is there overlap in loadings across the traditional and new variables? To answer these questions, in unreported analysis we run exploratory factor analysis on all fifteen variables (seven traditional, eight new). We first find the Pearson correlations for the continuous variables (income, etc.) and polychoric correlations for the categorical variables (education, etc.) and then use iterated principal factor extraction (see Table A2.1 for correlations).

In almost all specifications (including sample splits by education and wealth, individual countries, varying number of factors) we find consistent evidence for a single latent factor, and some

evidence for a second factor.[14] Using the cutoff of |0.32| as the minimum loading (Tabachnick et al., 2007), all traditional variables other than gender load on this factor. Age and risk aversion have negative loadings while other traditional variables have positive loadings. For the new variables, sociability, cognitive skills, and health also load on this factor. However, none of these factor loadings is very large, as the maximum loading is only |0.61|. Trust, right of center political orientation, religiosity, height, and BMI do not have loadings above the minimum. When keeping a second factor in the analysis, all variables retain similar loadings on the first factor. Religiosity now shows a loading of 0.91 on the second factor; no other variables load on the second factor. These five new variables thus do not share an underlying source of variation with the traditional variables, implying they are unique.

While 9 of the 15 variables have sizable loadings on the first factor, this factor explains only 15% of the total variation of the variables. To see how this single factor explains participation, we estimate factor scores for each individual. We use the factor scores as the main explanatory variable, include country dummies, and run OLS. The coefficient of the factor score has a *t*-statistic of 103, and the model has an R-squared of 27.3%, which is 93% of the R-squared of the full baseline model. There is thus a core common latent factor that, despite capturing only a small part of the heterogeneity across individuals, still captures nearly all of the so-far explained variation in stock market participation.

5.6. Hierarchical structure?

Our intention here is to present a general hypothesis, not to propose a specific econometric test, let alone carry out such a test. The idea is to see which drivers become more important as one moves toward more sophisticated subsamples. Table 5 broke the sample by country regulatory quality, and Table 6 did that for wealth and education. Of the new variables, the ones that always become stronger as regulatory quality improves are right of center, cognitive skills, health, and religiosity. When moving to higher wealth and education, and both, the variables that become stronger are sociability, right of center, religiosity, and BMI. The overlap of these two lists consists of variables that always become stronger in more sophisticated environments: right of center and religiosity. The effect of cognitive skills is similar for the wealth subsamples but differs for the education subsamples.

In response to these patterns, as well as our read of prior literature on other individual factors, we sketch a hierarchical structure of the factors affecting stock market participation (Fig. 1). The idea is that once the basic "needs" of stock market participation are met, higher order factors have a greater impact on the participation decision. Some individuals may participate as soon as they cross the direct fixed monetary costs hurdle (income, wealth). The threshold itself may vary by background risks (health). Others may require more knowledge (cognitive skills) or willpower (BMI) to do it. A few may still not participate, if they feel their personal values and attitudes are not in line with the idea of investing in stocks (political orientation, religiosity).

The results of the previous section's latent factor analysis concur with those from Tables 5 and 6. The variables income, wealth, and health all load on the first common factor, and this factor explains a large part of the so-far explained variation in stock market participation. BMI, right of center political orientation, and religiosity do not load on this factor yet their effect on participation increases in more sophisticated samples. Sociability and cognitive skills also load on the first factor, which does not strictly follow the model.

We also use cross-country variation to find evidence of the hierarchical structure of the participation drivers. We run the baseline regression separately for each country and then plot the marginal effects against the participation rates for each variable (see Fig. A2.1). Participation drivers at the top of the hierarchy should show a stronger effect in countries with higher participation rates. For income, a base-level driver, the relation is flat. For right of center political orientation, the relation is steep. While not conclusive, the overall evidence is suggestive of a hierarchical structure.

The purpose of this discussion is to raise the idea that the marginal non-participation driver, i.e., the most important factor keeping an individual out of stocks, can be different among different groups of people, and that there may be a hierarchical structure to such factors. Acknowledging the heterogeneity in this way can open a way forward in learning more about nonparticipation, as well as help develop and match the right policy tools for different groups.

6. Further findings complementing and challenging the participation literature

6.1. Factor subcomponents and variables with incomplete data

Our data allow a more granular look into two of the behavioral variables most strongly associated with participation, namely sociability and cognitive skills. The variable *Sociability* counts the number of different activities a respondent has engaged in. To see how various types of social activity associate with participation, in Table 9 we report results where the *Sociability*-variable is decomposed into different activities.

Four out of the five activities have significantly positive effect on participation. However, unlike Hong et al. (2004), we find that taking part in the activities of a religious organization, such as a church or synagogue, does not predict participation. In fact, it is the only type of social activity not associated with participation. The reason for this discrepancy in results is mainly not that we include a separate control for religiosity, but that we control for other types of sociability. In a regression where the indicator of religion-related activities is entered as the only measure of sociability, its coefficient is positive and significant. The effects of social activities related to education or training, and sports and social clubs, are remarkably strong, even after controlling for education, health, and body mass index. These findings seem consistent with prior literature emphasizing information about stocks as the mechanism for sociability's effect. As religious individuals are less likely to participate, religion-related activities should provide less exposure to information than other social activities.

Christelis et al. (2010) and Grinblatt et al. (2011) both find that numerical skills are the most important component of cognitive ability in predicting participation. Columns 2–4 in Table 9 break cognitive ability into its subcomponents, that is, numeracy, fluency, and recall. As seen in the final column, we find both numeracy and fluency to have zero effects in the baseline specification. Recall ability is the only positive and significant component of the *Cognitive skills*-variable.[15] Our more comprehensive

[14] When including a second factor, we use promax rotation.

[15] Recall ability may proxy for crystallized intelligence while numeracy and fluency are measures of fluid intelligence in our sample. Higher crystallized intelligence in older adults may compensate for lower fluid intelligence in economic decision making (Li et al., 2013).

Table 7
"Explanatory power ratio": new variables over traditional variables.
The table shows the fraction of the explanatory power of *Traditional variables* accounted for by the explanatory power of *New variables* in our main specification in all subsamples. The explanatory power of each group is retrieved via a Shapley decomposition of the pseudo-R^2, and the values are shown in Fig. 5. *Traditional variables* include gender, age, education, income, wealth, and risk aversion, while *New variables* include sociability, trust, political orientation, cognitive skills, health, religiosity, height, and the body mass index.

Full sample	Regulatory quality			Wealth		Education		Both low	Both high
	Low	Medium	High	Low	High	Low	High		
0.382	0.294	0.324	0.311	0.378	0.418	0.428	0.259	0.393	0.287

Table 8
Summary of explanatory power through specifications
Specifications where a variable was estimated to have a significant effect as predicted by prior literature are marked with an "X". In the next-to-last column, a variable whose estimate had the predicted sign in all specifications, even if not always statistically significant, is marked. In the final column, the percentage-point change in stockholding probability resulting from a one-standard-deviation increase in an explanatory variable, as estimated in our baseline specification in column (3) of Table 4, is shown. *Sociability* is a variable that counts the social activities a respondent reports to have engaged in during the previous year (Waves 4 and 5) or month (Waves 1 and 2). *Trust* equals the response, on a 0–10 scale, to "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?". *Right of center* equals the response to "In politics people sometimes talk of 'left' and 'right'. On a scale from zero to ten, where zero means left and ten means right, where would you place yourself?". *Cognitive skills* is an index based on scores of numeracy, fluency, and recall tests, where the three scores are first demeaned and divided by their sample standard deviations, and then averaged. *Health* is a 1–5 scale based on "Would you say your health is (1) excellent, ..., (5) poor?". *Religiosity* is a 0–5 scale measuring the frequency of praying, where zero means never and five means more than once a day. *Height* is reported in centimeters, and *BMI* is calculated as Weight (kg)/[Height (m)]^2; both are standardized relative to country, gender, age, and survey wave.

	Predicted, significant effect					Predicted sign throughout	Effect size (p.p.)
	Baseline	Direct partic.	All reg. en-vironments	All wealth levels	All educ. levels		
Sociability	X	X	X	X	X	X	2.35
Trust	X			X	X	X	0.73
Right of center	X	X		X	X		0.69
Cognitive skills	X	X	X	X	X	X	1.57
Health	X	X		X			0.77
Religiosity	X	X	X	X	X	X	−1.31
Height							0.00
Body mass index	X	X	X		X	X	−0.5

set of controls again appears to explain this discrepancy: Trying to match the specifications of Grinblatt et al. (2011) (only controlling for gender, age, education, income, and wealth) and Christelis et al. (2010) (omitting risk aversion from our baseline specification), we, too, find numeracy to be the strongest predictor of participation. The overall cognitive skills variable thus seems, to a large extent, to be picking up the effect of risk aversion, a result in line with Dohmen et al. (2010) who find that subjects with lower IQ test scores have higher experimentally elicited risk aversion.

In Table 10, we analyze two variables previously found to be important, but for which data is only available in some of the waves of SHARE. Puri and Robinson (2007) show optimistic individuals participate more.[16] Ameriks et al. (2011) demonstrate the importance of bequest considerations for financial decisions later in life, and correspondingly, Christelis et al. (2010) and Georgarakos and Pasini (2011) show that those who plan to leave a bequest are more likely to own stocks. In Table 10, the *Optimist*-dummy indicates if a respondent says she "sometimes" or "often" feels that the future looks good for her, as opposed to saying "rarely" or "never". The *Plans bequest*-dummy indicates if a respondent considers her probability for leaving a bequest to be positive. When individually adding the dummies to our baseline specification in all available waves, the results are in line with prior studies. However, when both are simultaneously included, which limits the observations to Wave 2 only, the *Optimist*-dummy loses its significance. The estimated effects of other variables are not notably changed from the baseline in any of the specifications and are not reported.

6.2. Long-term effects of macroeconomic experiences

We report the effects of earlier life experience in this separate subsection. That channel is among the most celebrated recent findings, but our approach here is slightly different. We run a regression similar to the baseline adding experience variables, using a subsample of the data. We utilize SHARE Wave 3, which has a different set of questions. This allows us to add two variables: dummies for *Hard life experience* and *Financial hardship* (see Appendix 1 for detailed definitions). However, this wave only partially covers the same set of countries, and we end up with a sample size of about 40% of the baseline.

The results are reported in Columns 2 and 5 of Table 4. We see that hard life experiences show no relationship with stock market participation. To the extent hard life experiences influence the perception of rare disaster risk, our results contrast with Choi and Robertson (2020). We find financial hardship to be negative and significant, with a large effect of 1.6 percentage points. When we omit risk aversion, we find, surprisingly, that the coefficient remains identical (results not reported). In contrast, when we drop country fixed effects, also the effect of financial hardship becomes essentially zero (results not reported). The fact that financial hardship only picks up relevant variation within countries (i.e., when country fixed effects are included), shows that countries do not line up as having more hardship and less participation, conditional on individual-level covariates. This makes sense, because over the long-term, the average financial hardship in a country should show

[16] Puri and Robinson's (2007) measure of optimism is subjective life-expectancy minus actuarial life-expectancy. Dominitz and Manski (2007) show that individuals with more optimistic expectations of equity returns are more likely to participate.

Table 9
Breakdown of *Sociability* and *Cognitive skills* into subcomponents
Average marginal effects on the probability of stockholding of a one-standard-deviation change in an explanatory variable (zero-to-one for dummies), obtained through probit regressions. *Any holdings* means that a respondent holds stocks either directly or indirectly. The indicators for different types of social activity show if a respondent engaged in an activity in the year (Waves 4 and 5) or month (Waves 1 and 2) preceding the interview. In the survey, the exact wordings describing the activities were (1) Taken part in a political or community-related organization; (2) Taken part in activities of a religious organization (church, synagogue, mosque etc.); (3) Done voluntary or charity work; (4) Attended an educational or training course; and (5) Gone to a sport, social, or other kind of club. *Fluency (0–10)*, *Numeracy (0–9)*, and *Recall (0–10)* are standardized scores from tests measuring verbal ability, numerical ability, and memory, respectively. GKL 2011 is short for Grinblatt et al. (2011), and CJP 2010 is short for Christelis et al. (2010). Standard errors are clustered by household, and t-statistics shown in parentheses below the marginal effects. *, **, and *** stand for statistical significance at the 10%, 5%, and 1% levels, respectively.

	Dependent var. (0/1): Any stock holdings			
Type of social activity (0/1):				
- Political organization	0.011**			
	(2.32)			
- Religious organization	0.004			
	(0.95)			
- Voluntary or charity work	0.020***			
	(6.13)			
- Education or training	0.040***			
	(11.40)			
- Sports or social club	0.038***			
	(14.28)			
Numeracy (0–9)		0.013***	0.010***	0.001
		(8.52)	(6.91)	(0.95)
Fluency (0–10)		0.009***	0.007***	0.001
		(6.17)	(5.00)	(0.30)
Recall (0–10)		−0.001	0.001	0.003**
		(−0.75)	(1.50)	(2.55)
Other controls as in...	Baseline	GKL 2011 (Demographicsonly)	CJP 2010 (Risk aversion omitted)	Baseline
Country effects	Yes	Yes	Yes	Yes
Pseudo R^2	0.299	0.272	0.280	0.297
N	104,787	104,610	104,610	104,610

Table 10
Variables with incomplete data: Optimism and bequest motive
Average marginal effects on the probability of stockholding of zero-to-one changes in the two explanatory variables of interest, obtained through probit regressions. *Any holdings* means that a respondent holds stocks either directly or indirectly. *Optimist (0/1)* indicates if a respondent says she "sometimes" or "often", as opposed to "rarely" or "never", feels that the future looks good for her. *Plans bequest (0/1)* indicates if a respondent considers her probability for leaving a bequest to be positive. In all three specifications, control variables include the full baseline specification, shown in column 3 of Table 4. Standard errors are clustered by household, and t-statistics shown in parentheses below the marginal effects. *, **, and *** stand for statistical significance at the 10%, 5%, and 1% levels, respectively.

	Dependent var. (0/1): Any stock holdings		
Optimist (0/1)	0.015***		0.008
	(4.42)		(1.47)
Plans bequest (0/1)		0.044***	0.032***
		(5.40)	(3.53)
Data from waves...	2, 4, 5	1, 2	2
Country effects	Yes	Yes	Yes
Pseudo R^2	0.300	0.324	0.333
N	92,295	54,224	41,952

up in wealth, which we control for. So, any experience effect remaining after controlling for wealth should operate within countries.

The within-country variation can come from individual heterogeneity, or differences in generational experiences, which is the main story in Malmendier and Nagel (2011). We next investigate the generational channel with two approaches. First, we use the three-year-long stock market downturn in 2000 to 2003, common in all of Europe. The data for the first Wave 1 of SHARE was collected right after, in 2004. As a result, participants in that wave have a more vivid memory of the downturn compared to those in later waves. We expect the effect to be stronger for the relatively younger due to its larger weight in their inventory of experience. We take the relatively young to be individuals between ages 45 and 55 at the time of the survey, so in their 40′s at the time of the run up and crash. The results reported in Table 11 are consistent with a generational experience effect: stock market participation rate is lower in Wave 1, and especially among the young relative to the old. The drawback of this test is that the shock is common to all countries.

The second test relates to a 1981 market crash that was confined to Italy. In late 1980 and early 1981, stocks in the Milan exchange rose rapidly, then prices started to plummet in June. Some of the largest companies lost over 20% of their value overnight, and the government suspended all trading. The crash closely followed the arrest of high-profile Milan bankers, later found guilty of illegal capital export. The case spread to include politicians, and "seems to have undermined the confidence of many small stockholders, not only in [concerned] banks but in the entire Milan financial community" (Tanner, 1981). The "young" Italians in our sample, aged 45–55 in Wave 1, were 22–32 years old at the time of the 1981 crash. That is, they were probably old enough to notice the extraordinary market events, and vulnerable to negative economic shocks due to their relatively short employment history. Thus, the crash may well have left a long-lasting trace in their memory. We therefore hypothesize that in Italy, the old-to-young ratio of stock market participation will be high relative to comparable countries. We test this by comparing Italy to other large, South-Western European countries, namely Spain and France, where the same cohorts did not experience a stock market crash in their early adulthood. While we find a large negative effect for Italian respondents overall, we are unable to identify a significant effect for young Italians in the sample.

Overall, these results are consistent with the lifetime experience model of Malmendier and Nagel (2011), but three issues are worth

Table 11
The effect of negative macroeconomic past experiences on stock market participation
Average marginal effects on the probability of direct or indirect stockholding of a zero-to-one change for the set of dummy variables of interest, obtained through probit regressions. In Columns 1–2, we use dummy variables for each wave of the survey. Wave 1 is the omitted group, as Wave 1 of the survey was conducted closest to the bear market years of 2000–2003. We do not use data from Wave 3 (see Section 3 of the paper). Columns 1–2 include all 19 countries in the analysis. In Columns 3–5, we use respondents from only Italy, Spain, and France. *Young (0/1)* indicates if the respondent is aged 55 or younger. In all columns, control variables include the full baseline specification, shown in column 3 of Table 4. Standard errors are clustered by household, and t-statistics shown in parentheses below the marginal effects. *, **, and *** stand for statistical significance at the 10%, 5%, and 1% levels, respectively.

Dependent variable:	All countries		Italy, Spain, and France only		
Wave 2 (dummy)	0.079***	0.079***	0.047***	0.048***	
	(21.1)	(18.75)	(6.32)	(5.76)	
Wave 4 (dummy)	0.090***	0.087***	0.059***	0.061***	
	(24.18)	(20.91)	(8.54)	(7.79)	
Wave 5 (dummy)	0.086***	0.084***	0.065***	0.064***	
	(23.12)	(20.47)	(9.33)	(8.21)	
Young (dummy)		−0.022***		0.004	0.018
		(−2.65)		(0.24)	(0.59)
Young*Wave 2		0.001		−0.004	
		(0.13)		(−0.23)	
Young*Wave 4		0.015		−0.008	
		(1.64)		(−0.47)	
Young*Wave 5		0.007		0.009	
		(0.77)		(0.50)	
Italy (dummy)			−0.041***	−0.041***	−0.040**
			(−5.49)	(−5.49)	(−2.17)
France (dummy)			0.113***	0.113***	0.121***
			(16.79)	(16.83)	(6.48)
Young*Italy					−0.017
					(−0.50)
Young*France					−0.010
					(−0.30)
Country dummies	Yes	Yes			
Wave 1 only					Yes
Pseudo R2	0.303	0.303	0.257	0.257	0.257
N	104,610	104,610	22,064	22,064	22,064

noting. First, the effect does not seem to operate through individual general experiences, such as wartime atrocities included in the definition of the hard life experience variable. Rather, the effect comes from the specific experience of financial hardship. Second, the effect does not seem to operate through modifying risk aversion. Third, between-country differences in the population-wide stock of experiences are picked up by other covariates such as wealth.

6.3. Which factors explain the share of financial wealth held in equities?

As our final set of tests, we look the fraction of a stockholders' financial wealth held in equities, or the conditional *equity share*. We define the value of equity holdings as the value of directly held stocks, plus the value of equity's fraction in mutual fund and IRA holdings, estimated as follows. When a respondent says she owns mutual funds or IRAs, she is asked whether they include (a) mostly stocks, (b) half stocks, half bonds, or (c) mostly bonds. In case (a), we multiply the fund or IRA holding value by 4/5 and include this in equity holdings. In cases (b) and (c), the respective coefficients are 1/2 and 1/5. The equity share is then found by dividing the equity holdings by the combined value of equities, bonds, mutual funds, IRAs, and bank accounts.

The results presented in Table 12 show patterns quite different from the 0/1 participation results. In the full model (column 3), only political preference is significant among the new vari-

Table 12
Determinants of equity's share in financial assets
The output of OLS regressions explaining the equity share, defined as the value of equity holdings divided by total financial assets, conditional on owning stocks. Equity holdings include directly held stocks and the equity fraction of mutual fund and IRA holdings. Total financial assets include equities, bonds, mutual funds, IRAs, and bank accounts. Values in euro as reported by respondents. *Education* has six categories based on the International Standard Classification of Education 1997 (ISCED-97). *Income* (annual, net) and wealth measures are on the household level, in euro. *Total assets* include gross financial wealth (bank accounts, bonds, stocks, mutual funds, long-term savings) and real wealth (real estate, business, cars). *Total liabilities* include mortgages and any financial liabilities. *Risk aversion* is a 1–4 scale indicating whether a respondent is willing to a) take substantial financial risks expecting to earn substantial returns, b) take above average financial risks expecting to earn above average returns, c) take average financial risks expecting to earn average returns, or d) not take any financial risks. *Sociability* is a variable that counts the social activities a respondent reports to have engaged in during the previous year (Waves 4 and 5) or month (Waves 1 and 2). *Trust* equals the response, on a 0–10 scale, to "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?". *Right of center* equals the response to "In politics people sometimes talk of 'left' and 'right'. On a scale from zero to ten, where zero means left and ten means right, where would you place yourself?". *Cognitive skills* is an index based on scores of numeracy, fluency, and recall tests, where the three scores are first demeaned and divided by their sample standard deviations, and then averaged. *Health* is a 1–5 scale based on "Would you say your health is (1) excellent, ..., (5) poor?". *Religiosity* is a 0–5 scale measuring the frequency of praying, where zero means never and five means more than once a day. *Height* is reported in centimeters, and *BMI* is calculated as Weight (kg) / [Height (m)]^2; both are standardized relative to country, gender, age, and survey wave. Standard errors are clustered by household, and t-statistics shown in parentheses below the marginal effects. *, **, and *** stand for statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)
Male (0/1)	−0.001	0.001	0.001	0.013***
	(−0.19)	(0.32)	(0.18)	(2.90)
Age (years)	−0.010***	−0.011***	−0.013***	−0.014***
	(−3.60)	(−3.94)	(−4.34)	(−4.56)
Age^2 (/ 1000)	0.083***	0.090***	0.102***	0.105***
	(3.89)	(4.25)	(4.58)	(4.67)
Education (ISCED)	−0.003*	−0.003**	−0.003	−0.001
	(−1.80)	(−1.98)	(−1.59)	(−0.60)
Ln(Income (€))	−0.011***	−0.009***	−0.01***	−0.003
	(−5.50)	(−4.46)	(−4.70)	(−1.36)
Ln(Total assets (€))	0.004**	0.007***	0.008***	0.011***
	(2.18)	(3.56)	(3.50)	(4.81)
Ln(Total liab. (€))	0.002***	0.003***	0.003***	0.003***
	(5.71)	(5.89)	(5.96)	(5.92)
Risk aversion (1–4)	−0.061***	−0.059***	−0.058***	
	(−21.83)	(−20.68)	(−20.06)	
Sociability (0–8)			−0.002	0.003***
			(−0.79)	(5.92)
Trust (0–10)			−0.002*	−0.001
			(−1.74)	(−0.39)
Right of center (0–10)			0.005***	−0.002
			(4.59)	(−1.54)
Cognitive skills			0.006	0.006***
			(1.45)	(5.86)
Health (1–5)			0.001	0.012***
			(0.63)	(3.11)
Religiosity (0–5)			0.003*	0.004*
			(1.75)	(1.75)
Height (relative)			0.001	0.002
			(0.42)	(1.48)
Height^2			0.001	0.002
			(0.50)	(0.79)
BMI (relative)			0.001	0.001
			(0.32)	(0.63)
Constant	0.967***	0.981***	1.018***	0.722***
	(10.08)	(10.19)	(10.07)	(7.12)
Country effects	No	Yes	Yes	Yes
R^2	0.035	0.050	0.054	0.029
N	19,799	19,799	19,040	19,244

ables. A move from the center to the very right on the 0–10 scale leads to an increase in equity share of 2.5 percentage points. When dropping risk aversion (column 4), sociability, cognitive skills, and health also become significant. This adds to the earlier findings suggesting that when risk aversion is not directly controlled for, these variables may partially act as risk aversion proxies. A significant gender effect also appears in the absence of risk aversion control. Dropping country effects (column 1) has little effect on traditional variables here.

7. Conclusion

The stock market participation literature has progressed steadily over the past few decades, but in step-by-step fashion, usually by adding one novel explanatory variable at a time. We take a consolidating approach by examining the roles of institutional, traditional, and behavioral factors, independently and jointly. We use data from SHARE, which comprehensively cover traditional drivers of participation, including risk aversion, and nearly all of the behavioral drivers from the recent stock market participation literature. We take full advantage of the data, exploiting cross-country variation in regulatory quality and varying levels of wealth and education across individuals.

Institutions clearly matter. Of the explanatory power, institutional factors captured by country fixed effects account for about a third. The traditional individual-level variables capture about 50%, and less than a fifth of the variation is explained by the new behavioral variables. We show that the effects of traditional variables and new behavioral variables change across levels of regulatory quality and across levels of wealth and education. In environments where participation rates are higher, such as high regulatory quality countries and among the wealthy and highly educated, variables reflecting values and attitudes have larger effects. In environments with lower levels of participation these variables have smaller or even non-significant effects. To complete our comprehensive approach, we also take a look at individual behavioral drivers and document complementing and contrasting results with previous literature.

Based partly on our results, and partly on prior research, we suggest a hierarchical model of participation drivers. In this model, similar in spirit to Maslow's (1943) hierarchy of needs, the low-level factors, such as wealth, have to be on a sufficient level before high-level, e.g., psychological, factors can come into play. Further progress in explaining non-participation, particularly among the well-off, would likely benefit from a focus in the high end of the model.

Declaration of Competing Interest

None.

CRediT authorship contribution statement

Markku Kaustia: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Andrew Conlin:** Software, Formal analysis, Data curation, Validation, Methodology, Writing – original draft, Writing – review & editing, Visualization, Project administration, Funding acquisition. **Niilo Luotonen:** Software, Formal analysis, Data curation, Validation, Methodology, Writing – original draft, Writing – review & editing, Visualization, Project administration, Funding acquisition.

Data availability

Data will be made available on request.

Appendix 1. Variable definitions

A. SHARE questions on financial assets

The survey asks the yes/no question "Do you own…" separately for each of stocks, mutual funds and retirement accounts and the follow-up "About how much…" for each of the categories. For mutual funds and retirement accounts, the survey also asks if the accounts are mostly stocks, half stocks and half bonds, or mostly bonds. We define respondents with mutual funds and/or retirement accounts with at least half the amount invested in stocks to be indirect stockholders.

B. Baseline variables

Education has six categories based on the International Standard Classification of Education 1997 (ISCED-97), facilitating comparison across countries.

Total assets, Total liabilities, and *Income*, are reported at the household level. *Total assets* include gross financial wealth (bank accounts, bonds, stocks, mutual funds, long-term savings) and real wealth (real estate, business, cars). *Total liabilities* include mortgages and any financial liabilities. *Income* is reported net of taxes, and includes revenues from employment, pension, and assets owned. The income figures from Wave 1 are reported in gross terms.

Risk aversion is measured with the following question, similar to the US Survey of Consumer Finances (SCF; see, e.g., Haliassos and Bertaut, 1995, p. 1121). "Thinking about financial risk that you are willing to take, do you a) Take substantial financial risks expecting to earn substantial returns, b) Take above average financial risks expecting to earn above average returns, c) Take average financial risks expecting to earn average returns, and d) Not willing to take any financial risks." We convert these answers to a four-point scale where 4 indicates maximum risk aversion (answer d) and 1 is least risk averse (answer a).

Sociability counts the social activities a respondent reports to have engaged in during the previous year (Waves 4 and 5) or month (Waves 1 and 2), with the following options: (a) Voluntary or charity work, (b) Attendance of an educational or training course, (c) Participation in a sports, social or other kind of club, (d) Taking part in the activities of a religious organization (church, synagogue, mosque, etc.), and (e) Taking part in a political or community-related organization.

Trust is the response to the question "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?" Responses are given on a 0–10 scale, where zero means "one can't be too careful" and ten means "most people can be trusted".

Right of center is the response to the question: "In politics people sometimes talk of 'left' and 'right'. On a scale from zero to ten, where zero means left and ten means right, where would you place yourself?"

For *Cognitive skills*, we construct an index based on three different dimensions of cognitive ability measured by SHARE. The first one, numeracy, is based on four calculations: (1) 10% of 1000, (2) half of 300, (3) 1.5 times 6000, (4) the value of 2000 euro after two years of compound interest of 10%. If (1) is answered correctly, the interviewer jumps to (3). If (3) is again correctly responded to, (4) will be asked as the final question. (2) is only asked if (1) is answered incorrectly, and after that no further question is asked. For details on scoring, see Dewey and Prince (2005). The second one, fluency, is measured by the number of different animals a respondent can name in one minute, with a cap at 70.[17] The third component, recall, equals the number of words a respondent can remember from a list of ten after a short time. We create the in-

[17] Energetic finance researchers should try this test: can you make the 70 mark?

dex by averaging the three components after standardizing them into z-scores.

Health is the response to the question: "Would you say your health is (1) excellent, ..., (5) poor?"

Religiosity is measured by how often a respondent prays. Responses are coded onto a 0–5 scale, where zero means never and five means more than once a day. Compared to church attendance as used by Hong et al. (2004), this measure should be more descriptive of belief intensity and less so of social activity.

The *Body mass index* (*BMI*) is calculated as follows: BMI = Weight (kg)/[Height (m)]^2. Like Addoum et al. (2017), we use relative measures standardized by country, gender, age, and survey wave for these variables.

C. Additional life experience variables from Wave 3

We add two variables from Wave 3, which covers a different set of countries and questions, with partial overlap to the main survey waves we utilize.

Hard life experience is dummy taking the value of 1 if the respondent has lived in a children's home; been fostered with another family; evacuated or relocated during a war; lived in a prisoner of war camp; lived in prison; lived in a labor camp; lived in a concentration camp; been an inpatient in a TB institution; stayed in a psychiatric hospital; been homeless for 1 month or more.

Financial hardship is an indicator variable for having experienced a distinct period of financial hardship starting before 1995. We choose 1995 as the cut-off to capture a long-term experience effect and not a contemporaneous wealth or income effect (Wave 1 was conducted in 2004).

Appendix 2. Additional Figures and Tables

Table A2.1
. Correlations. We calculate Pearson correlations for pairs with at least one continuous variable (*Ln(Income), Ln(Assets), Ln(Liabilities), Height, BMI*). We calculate Polychoric correlations for pairs of categorical or ordinal variables (*Male, Age, Education, Risk aversion, Sociability, Trust, Right of center, Cognitive skills, Health, Religiosity*). Bolded correlations are those NOT significant at 1% level (e.g., *Male* and *Trust*). All other correlations are significant at the 1% level.

	Male	Age	Educ.	Ln(Inc.)	Ln(Assets)	Ln(Liab.)	Risk avers.	Sociab.	Trust	Right of ctr	Cog. Skills	Health	Relig.	Height
Age	0.05	1.00												
Educ.	0.11	−0.26	1.00											
Ln(Income)	0.07	−0.14	0.25	1.00										
Ln(Assets)	0.07	−0.14	0.20	0.31	1.00									
Ln(Liab.)	0.04	−0.32	0.19	0.25	0.22	1.00								
Risk avers.	−0.21	0.17	−0.29	−0.17	−0.16	−0.15	1.00							
Sociability	0.02	−0.15	0.31	0.19	0.18	0.17	−0.24	1.00						
Trust	**−0.01**	−0.03	0.17	0.12	0.09	0.11	−0.14	0.18	1.00					
Right of center	**0.01**	0.06	**0.00**	**0.02**	0.06	**0.00**	−0.06	**0.02**	**0.01**	1.00				
Cog. Skills	0.07	−0.28	0.41	0.23	0.18	0.19	−0.18	0.23	0.16	**−0.01**	1.00			
Health	0.05	−0.28	0.27	0.22	0.24	0.18	−0.23	0.30	0.17	0.04	0.21	1.00		
Religiosity	−0.28	0.16	−0.20	−0.14	−0.06	−0.12	0.14	0.07	−0.04	0.13	−0.17	−0.10	1.00	
Height	**0.00**	**0.00**	0.11	0.05	0.08	**0.01**	−0.05	0.04	0.03	**0.01**	0.08	0.06	−0.03	1.00
BMI	**0.00**	**0.00**	−0.11	−0.03	−0.07	**0.01**	0.03	−0.04	−0.04	**0.00**	−0.05	−0.14	**0.01**	−0.13

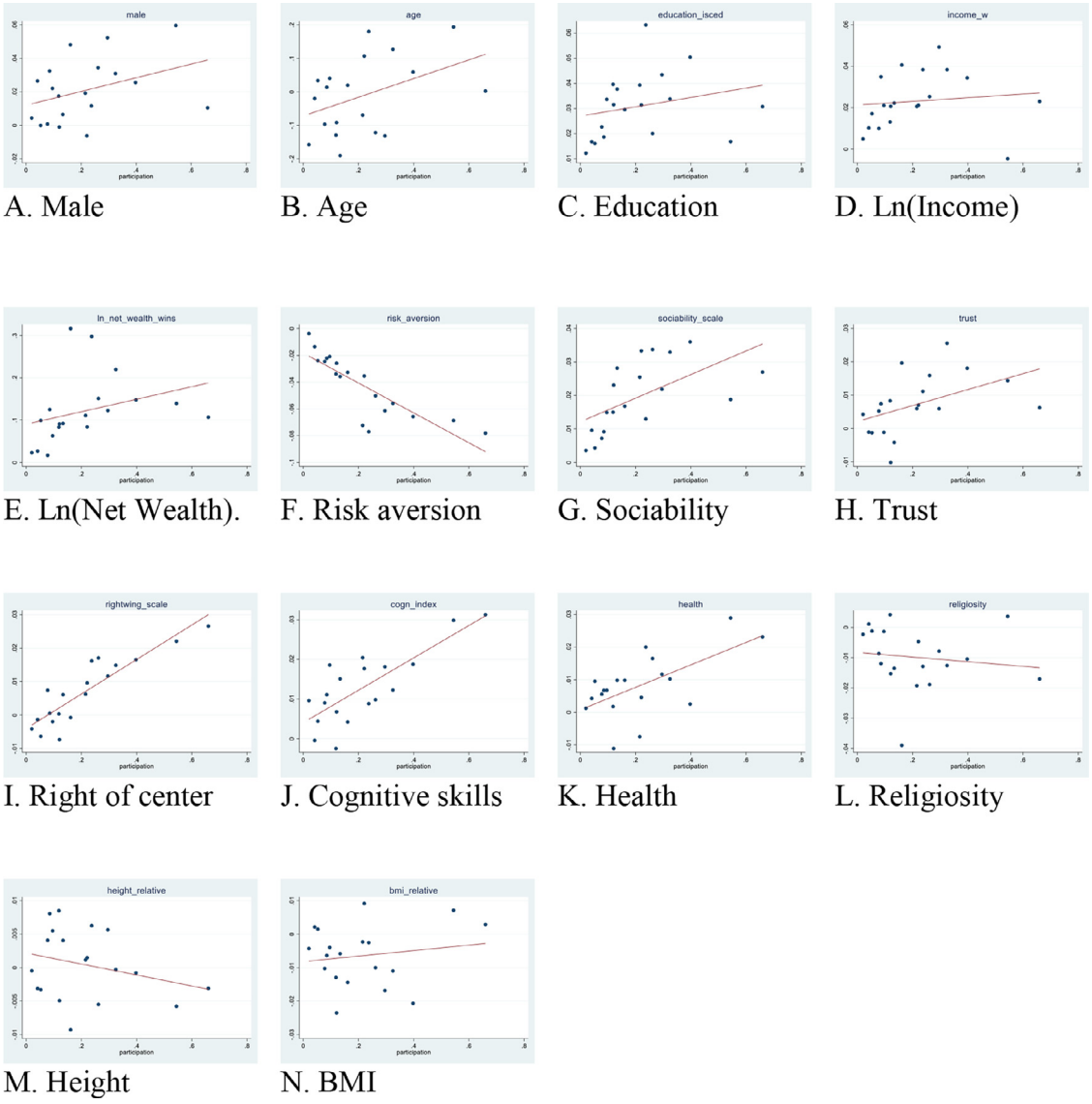


Fig. A2.1. Marginal effects of variables across countries.
We run our baseline probit model (Table 4, col.3-BL) separately for each country. For each variable, we plot the marginal effect of that variable against the country's average stock market participation rate. The marginal effects are on the vertical axis, and the participation rate on the horizontal. The order of countries from lowest to highest participation rate is: Poland, Hungary, Greece, Estonia, Italy, Spain, Portugal, Slovenia, Austria, Israel, Netherlands, Czech Republic, Luxembourg, Germany, France, Belgium, Switzerland, Denmark, Sweden. The red line in each figure shows the fitted values from the OLS regression of the marginal effect on the participation rate.

References

Addoum, J.M., Korniotis, G., Kumar, A., 2017. Stature, obesity, and portfolio choice. Manag. Sci. 63 (10), 3147–3529.

Ameriks, J., Caplin, A., Laufer, S., Van Nieuwerburgh, S., 2011. The joy of giving or assisted living? Using strategic surveys to separate public care aversion from bequest motives. J. Financ. 66 (2), 519–561.

Ameriks, J., and Zeldes, S.P., 2004. How do household portfolio shares vary with age? Working paper, Columbia Business School, September 2004.

Ang, A., Bekaert, G., Liu, J., 2005. Why stocks may disappoint. J. Financ. Econ. 76, 471–508.

Atella, V., Brunetti, M., Maestas, N., 2012. Household portfolio choices, health status and health care systems: a cross-country analysis based on SHARE. J. Bank. Financ. 36, 1320–1335.

Barber, B., Odean, T., 2001. Boys will be boys: gender, overconfidence and common stock investment. Q. J. Econ. 116 (1), 261–292.

Barro, R., McCleary, R., 2003. Religion and economic growth across countries. Am. Sociol. Rev. 68, 760–781.

Bender, S., Choi, J.J., Dyson, D., Robertson, A.Z., 2022. Millionaires speak: what drives their personal investment decisions? J financ econ forthcoming.

Benjamin, D.J., Brown, S.A., Shapiro, J.M., 2013. Who is behavioral? Cognitive ability and anomalous preferences. J. Eur. Econ. Assoc. 11 (6), 1231–1255.

Bergmann, M. Kneip, T., De Luca, G., & Scherpenzeel, A., 2019. Survey participation in the Survey of Health, Aging and Retirement in Europe (SHARE), Wave 1-7. Based on release 7.0.0. SHARE Working Paper Series 41-2019. Munich: SHARE-ERIC.

Bertaut, C., 1998. Stockholding behavior of U.S. households: evidence from the 1983-1989 survey of consumer finances. Rev. Econ. Stat. 80 (2), 263–275.

Borghans, L., Golsteyn, B.H.H., 2006. Time discounting and the body mass index: evidence from the Netherlands. Econ. Hum. Biol. 4, 39–61.

Börsch-Supan, A, Brandt, M, Hunkler, C, Kneip, T, Korbmacher, J, Malter, F, Schaan, B, Stuck, S, Zuber, S, 2013. Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). Int. J. Epidemiol. doi:10.1093/ije/dyt088.

Bressan, S., Pace, N., Pelizzon, L., 2014. Health status and portfolio choice: is their relationship economically relevant? Int. Rev. Financ. Anal. 32, 109–122.

Brown, J., Ivković, Z., Smith, P., Weisbenner, S., 2008. Neighbors matter: causal community effects and stock market participation. J. Financ. 63, 1509–1531.

Calvet, L.E., Sodini, P., 2014. Twin picks: disentangling the determinants of risk-taking in household portfolios. J. Financ. 69 (2), 867–906.

Campante, F., Yanagizawa-Drott, D., 2015. Does religion affect economic growth and happiness? Evidence from Ramadan. Q. J. Econ. 130 (2), 615–658.

Campbell, J., 2006. Household finance. J. Financ. 61 (4), 1553–1604.

Changwony, F.K., Campbell, K., Tabner, I.T., 2021. Savings goals and wealth allocation in household financial portfolios. J. Bank. Financ. 124.

Changwony, F.K., Campbell, K., Tabner, I.T., 2015. Social engagement and stock market participation. Rev. Financ. 19 (1), 317–366.

Choi, J.J., Robertson, A., 2020. What matters to individual investors? Evidence from the horse's mouth. J. Financ. 75 (4), 1965–2020.

Christelis, D, 2011. Imputation of missing data in Waves 1 and 2 of SHARE. CSEF Working Papers 278, Centre for Studies in Economics and Finance (CSEF). University of Naples, Italy.
Christelis, D., Georgarakos, D., Haliassos, M., 2013. Differences in portfolios across countries: economic environment versus household characteristics. Rev. Econ. Stat. 95 (1), 220–236.
Christelis, D., Jappelli, T., Padula, M., 2010. Cognitive abilities and portfolio choice. Eur. Econ. Rev. 54, 18–38.
Cocco, J., Gomes, F., Maenhout, P., 2005. Consumption and portfolio choice over the life cycle. Rev. Financ. Studies 18, 490–533.
Cohn, R., Lewellen, W., Lease, R., Schlarbaum, G., 1975. Individual investor risk aversion and investment portfolio composition. J. Finance 30, 605–620.
Cole, S., Paulson, A., Shastry, G.K., 2014. Smart money: the effect of education on financial behavior. Rev. Financ. Stud. 27 (7), 2022–2051.
Curcuru, S., Heaton, J., Lucas, D., Moore, D., ait-Sahalia, Y., Hansen, L.P., 2010. Heterogeneity and portfolio choice: theory and evidence. In: Handbook of Financial Econometrics, I. North-Holland, Amsterdam, pp. 337–382.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2010. Are risk aversion and impatience related to cognitive ability? Am. Econ. Rev. 100, 1238–1260.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G., 2011. Individual risk attitudes: measurement, determinants and behavioral consequences. J. Eur. Econ. Assoc. 9 (3), 522–550.
Dominitz, J., Manski, C.F., 2007. Expected equity returns and portfolio choice: evidence from the health and retirement study. J. Eur. Econ. Assoc. 5, 369–379.
Duflo, E., Saez, E., 2002. Participation and investment decisions in a retirement plan: the influence of colleagues' choices. J. Public Econ. 85 (1), 121–148.
Dynan, K., Maki, D., 2001. Does Stock Market Wealth Matter for Consumption? Finance and Economics Discussion Series 2001–23. Washington: Board of Governors of the Federal Reserve System..
Edwards, R., 2008. Health risk and portfolio choice. J. Bus. Econ. Stat. 26 (4), 472–485.
Elliott, F., Zhao, R., 2009. Health status and portfolio choice: causality or heterogeneity? J. Bank. Financ. 33, 1079–1088.
Fey, J.-C., Lerbs, O., Schmidt, C., Weber, M., 2021. Risk attitude and capital market participation: is there a gender gap in Germany? (May 17, 2021). Available at SSRN: https://ssrn.com/abstract=3847935.
Frantonini, M.C., 1998. Homeownership and investment in risky assets. J. Urban Econ. 44, 27–42.
Frederick, S., 2005. Cognitive reflection and decision making. J. Econ. Perspect. 19, 25–42.
Gennaioli, N., Shleifer, A., Vishny, R., 2015. Money doctors. J. Financ. 70 (1), 91–114.
Georgarakos, D., Pasini, G., 2011. Trust, sociability, and stock market participation. Rev. Financ. 15 (4), 693–725.
Giannetti, M., Koskinen, Y., 2010. Investor protection, equity returns, and financial globalization. J. Financ. Quant. Anal. 45 (1), 135–168.
Giannetti, M., Wang, T., 2016. Corporate scandals and household stock market participation. J. Financ. 71, 2591–2636.
Gomes, F.J., Michaelides, A., 2005. Optimal life-cycle asset allocation: understanding the empirical evidence. J. Financ. 60, 869–904.
Grinblatt, M., Keloharju, M., Linnainmaa, J., 2011. IQ and stock market participation. J. Financ. 66 (6), 2121–2164.
Guiso, L., Haliassos, M., Jappelli, T, 2003a. Household Portfolios. MIT Press, Boston.
Guiso, L., Haliassos, M., Jappelli, T., 2003b. Household stockholding in Europe: where do we stand and where do we go? Econ. Policy 18 (36), 123–170.
Guiso, L., Jappelli, T., Terlizzese, D., 1996. Income risk, borrowing constraints, and portfolio choice. Am. Econ. Rev. 86, 158–172.
Guiso, L., Sapienza, P., Zingales, L., 2004. The role of social capital in financial development. Am. Econ. Rev. 94, 526–556.
Guiso, L., Sapienza, P., Zingales, L., 2008. Trusting the stock market. J. Financ. 63, 2557–2600.
Guiso, L., Sapienza, P., Zingales, L., 2018. Time varying risk aversion. J. Financ. Econ. 128 (3), 403–421.
Guiso, L., Sodini, P., Harris, M., Stulz, R., 2013. Household finance: an emerging field. Ch. 21 in in George constantinides. Handbook of the Economics of Finance (eds.). Amsterdam, North-Holland.
Guvenen, F., 2006. Reconciling conflicting evidence on the elasticity of intertemporal substitution: a macroeconomic perspective. J. Monet. Econ. 53, 1451–1472.
Halek, M., Eisenhauer, J.G., 2001. Demography of risk aversion. J. Risk Insur. 68, 1–24.
Haliassos, M., Bertaut, C., 1995. Why do so few hold stocks? Econ. J. 105, 1110–1129.
Halko, M.-L., Kaustia, M., Alanko, E., 2012. The gender effect in risky asset holdings. J. Econ. Behav. Organ. 83, 66–81.
Heaton, J., Lucas, D., 2000a. Portfolio choice in the presence of background risk. Econ. J. 110, 1–26.
Heaton, J., Lucas, D., 2000b. Portfolio choice and asset prices: the importance of entrepreneurial risk. J. Financ. 55, 1163–1198.
Hong, H., Kubik, J., Stein, J., 2004. Social interaction and stock-market participation. J. Financ. 59, 137–163.
Hosmer, D.W., Lemeshow, S., 2000. Applied Logistic Regression. John Wiley & Sons, Inc.
Ikeda, S., Kang, M.-I., Ohtake, F., 2010. Hyperbolic discounting, the sign effect, and the body mass index. J. Health Econ. 29, 268–284.
Jianakoplos, N., Bernasek, A., 1998. Are women more risk averse? Econ. Inq. 36, 620–630.
Kaustia, M., Knüpfer, S., 2012. Peer performance and stock market entry. J. Financ. Econ. 104 (2), 321–338.
Kaustia, M., Torstila, S., 2011. Stock market aversion? Political preferences and stock market participation. J. Financ. Econ. 100 (1), 98–112.
Knüpfer, S., Rantapuska, E., Sarvimäki, M., 2017. Formative experiences and portfolio choice: evidence from the finnish great depression. J. Financ. 72 (1), 133–166.
Kullmann, C. Siegel, S., 2005. Real estate and its role in household portfolio choice (November 1, 2005). EFA 2003 Annual Conference Paper No. 918, Sauder School of Business Working Paper, Available at SSRN: https://ssrn.com/abstract=424322.
Kumar, A., Page, J.P., Spalt, O.G., 2011. Religious beliefs, gambling attitudes, and financial market outcomes. J. Financ. Econ. 102, 671–708.
Laudenbach, C., Malmendier, U., Niessen-Ruenzi, A., 2020. The long-lasting effects of living under communism on attitudes towards financial markets. NBER Working Paper No. 26818, March 2020.
Li, Y., Baldassi, M., Johnson, E.J., Weber, E.U., 2013. Complementary cognitive capabilities, economic decision making, and aging. Psychol. Aging 28 (3), 595.
Love, D.A., Smith, P.A., 2010. Does health affect portfolio choice? Health Econ. 19, 1441–1460.
Madrian, B., Shea, D., 2001. The power of suggestion: inertia in 401(k) participation and savings behavior. Q. J. Econ. 116, 1149–1187.
Malmendier, U., Nagel, S., 2011. Depression babies: do macroeconomic experiences affect risk-taking? Q. J. Econ. 126, 373–416.
Mankiw, N., Zeldes, S., 1991. The consumption of stockholders and non-stockholders. J. Financ. Econ. 29, 97–112.
Maslow, A., 1943. A theory of human motivation. Psychol. Rev. 50 (4), 370–396.
Mehra, R., Prescott, E., 1985. The equity premium - a puzzle. J. Monet. Econ. 15, 145–161.
Merton, R., 1969. Lifetime portfolio selection under uncertainty: the continuous time case. Rev. Econ. Stat. 50, 247–257.
Merton, R., 1971. Optimum consumption and portfolio rules in a continuous-time model. J. Econ. Theory 3, 373–413.
Packiam Alloway, T., Alloway, R.G., 2010. Investigating the predictive roles of working memory and IQ in academic attainment. J. Exp. Child Psychol. 106 (1), 20–29.
Pagano, M., Volpin, P., 2006. Alfred Marshall lecture: shareholder protection, stock market development, and politics. J. Eur. Econ. Assoc. 4 (2–3), 315–341.
In Poterba, J., Samwick, A., Ogura, S., Tachibanaki, T., David Wise, A., 2001. Portfolio allocations over the life cycle. In: Aging Issues in the United States and Japan. University of Chicago Press, Chicago, pp. 65–103.
Puri, M., Robinson, D., 2007. Optimism and economic choice. J. Financ. Econ. 86 (1), 71–99.
Rosen, H., Wu, S., 2004. Portfolio choice and health status. J. Financ. Econ. 72 (3), 457–484.
Riley, W., Chow, K., 1992. Asset allocation and individual risk aversion. Financ. Anal. J. 48 (6), 32–37.
Sapienza, P., Toldra-Simats, A., Zingales, L., 2013. Understanding trust. Econ. J. 123 (573), 1313–1332.
Solon, G., Haider, S.J., Wooldridge, J.M., 2015. What are we weighting for? J. Hum. Resour. 50 (2), 301–316.
Stulz, R., Williamson, R., 2003. Culture, openness and finance. J. Financ. Econ. 70, 313–349.
Sundén, A., Surette, B., 1998. Gender differences in the allocation of assets in retirement savings plans. Am. Econ. Rev. 88 (2), 207–211.
Tabachnick, B.G., Fidell, L.S., Ullman, J.B., 2007. In: Using Multivariate Statistics, 5. Pearson, Boston, MA, pp. 481–498.
Tanner, H., 1981. Italy suspends all stock trading. The New York Times, 9 July 1981, https://www.nytimes.com/1981/07/09/business/italy-suspends-all-stock-trading.html.
Van Rooij, M., Lusardi, A., Alessie, R., 2007. Financial literacy and stock market participation. J. Financ. Econ. 101 (2), 449–472.
Vestman, R., 2019. Limited stock market participation among renters and homeowners. Rev. Financ. Stud. 32 (4), 1494–1535.
Vissing-Jørgensen, A., 2002. Limited asset market participation and the elasticity of intertemporal substitution. J. Polit. Econ. 110, 825–853.
Vissing-Jørgensen, A., 2004. Perspectives on behavioral finance: does irrationality disappear with wealth? Evidence from expectations and actions. NBER Macroeconomics Annual 2003. MIT Press, Cambridge.