Revisiting Financial Literacy and Stock Market Participation:

The Role of Don't Know Responses and Selection Bias

Yoshiyuki Nakazono

Minoru Masujima

Ryoichi Namba

Jun Takahashi

Kento Tango

February 10, 2026

Revisiting Financial Literacy and Stock Market Participation: The Role of Don't Know Responses and Selection Bias*

Yoshiyuki Nakazono[†]   Minoru Masujima[‡]   Ryoichi Namba[§]
Jun Takahashi[¶]   Kento Tango[‖]

Abstract

We study the link between financial literacy and stock market participation using harmonized household survey data from Japan, the United States, Germany, and China, focusing on "Don't Know" (DK) responses. We find that DK responses follow systematic patterns across countries, genders, and question types, and that individuals who frequently answer DK are less engaged in stock markets. We also find that ignoring DK behavior biases empirical estimates: it overstates stock market entry at the extensive margin. We develop a conceptual framework that explains why these biases arise when DK responses are ignored. The results suggest that DK responses contain essential information for assessing the role of financial literacy in stock market participation and portfolio allocation.

JEL Classification:	D12; D14; D84; E21; G11; G41
Keywords:	cross-country analysis; household finance; financial literacy;
	investment behavior; stock market participation; portfolio allocation

1 Introduction

Household participation in stock markets plays a central role in long-run wealth accumulation, risk sharing, and the transmission of macroeconomic shocks. Standard portfolio choice models predict that households should invest in equities to earn the equity premium and to hedge against inflation over the life cycle. Nevertheless, in many countries stock market participation remains surprisingly low, a phenomenon widely referred to as the stock market participation puzzle. This puzzle is particularly salient in cross-country comparisons. For example, Japan exhibits persistently low stock market participation despite prolonged low interest rates and recent inflationary pressures, while participation rates are substantially higher in countries such as the United States and China[1]. Understanding the sources of this limited participation remains a key challenge in household finance.

This paper examines the relationship between financial literacy and stock market participation using large-scale household survey data from Japan, the United States, Germany, and China. First, we document who answers "Don't Know" (DK) in standard financial literacy questions and examine that DK responses are systematic behavioral outcomes rather than random measurement noise. Second, we estimate the relationship between financial literacy and both the probability of stock market participation and the share of stocks held in household portfolios, explicitly accounting for DK-related selection and sample composition.[2] Third, we develop a conceptual framework, motivated by the empirical patterns in the data, that clarifies how ignoring DK responses biases the estimated relationship between financial literacy and stock market participation.

This paper makes three contributions. First, we document who selects DK responses in standard financial literacy questions and show that DK behavior is systematically related to socio-demographic characteristics and individual preferences. In particular, DK responses are more prevalent among women, less educated individuals, and those in lower income groups.

Second, we apply this framework to large-scale cross-country data and show that financial literacy remains associated with stock market participation. After accounting for DK-related selection effects and excluding individuals with past stock market experience, a one-standard-deviation higher level of financial literacy is associated with a 6–7 percentage point higher probability of stock market participation. This magnitude is comparable to recent estimates in the literature and is robust across countries. Our findings do not overturn the conventional view that financially more knowledgeable households are more likely to invest in stocks. Rather, they clarify how measurement and sample selection choices shape the magnitude and interpretation of this relationship.

Third, we offer a conceptual framework that clarifies how ignoring DK responses biases the estimated relationship between financial literacy and stock market participation. The framework is consistent with the empirical patterns documented in the data and shows that omitting DK behavior leads

---

[1]For evidence on the persistently low level of stock market participation in Japan, see Yamori and Ueyama (2022).

[2]Hereafter, we use the term "DK-related selection" to refer to the idea that individuals who systematically opt for DK responses in survey questions are also less likely to engage with financial markets, which can bias the observed relationship between financial literacy and stock market participation.

to an upward bias in estimated effects at the extensive margin (stock market entry), highlighting the importance of accounting for DK-related selection.

Our study contributes to two strands of the literature. First, our study contributes to the emerging literature on DK responses in financial literacy measurement. While DK responses are often coded as incorrect answers or treated as missing data, a growing body of research suggests that they reflect economically meaningful behaviors related to confidence, engagement, and selection into financial decision-making (Lusardi and Mitchell, 2014; van Rooij et al., 2011; Bucher-Koenen et al., 2017; Allgood and Walstad, 2016).[3] Most closely related to our analysis, Bucher-Koenen et al. (2025) show that DK responses often reflect low confidence or under-confidence rather than a pure lack of financial knowledge, and that treating DK mechanically as incorrect answers can distort measured financial literacy. Our analysis treats DK responses as informative outcomes in a large, harmonized cross-country dataset. We show that classical measures of financial literacy can be systematically distorted when DK behavior is ignored.

Second, our study contributes to the literature on household finance and stock market participation. A large body of work documents that low financial literacy is strongly associated with limited participation in equity markets, highlighting informational barriers to stock market participation (van Rooij et al., 2011; Lusardi and Mitchell, 2014; Grinblatt et al., 2011; Campbell, 2006). A related strand emphasizes the role of cultural and institutional factors—such as trust in financial markets and investor protection—in explaining substantial cross-country variation in household stockholding (Guiso et al., 2003, 2008; Giannetti and Koskinen, 2010; Giannetti and Wang, 2016). A third line of research points to informational frictions, behavioral responses, and selection mechanisms, including social interactions, past experiences, and heterogeneous beliefs, as key determinants of market participation (Hong et al., 2004; Malmendier and Nagel, 2011; Calvet et al., 2007; Vissing-Jorgensen, 2003; Gennaioli et al., 2018). Our analysis complements recent experimental evidence, such as Bucher-Koenen et al. (2025), which documents gender differences in DK responses and their implications for measured financial literacy. While that study focuses on a targeted survey experiment, we extend the analysis to the general population across multiple countries and show that DK behavior is systematically related to stock market participation for both men and women.

While the broad link between financial literacy and stock market participation is well established, our contribution lies in clarifying how measurement and sample selection shape the empirical magnitude and interpretation of this relationship. Using large, harmonized cross-country data, we show that accounting for DK behavior leads to economically meaningful differences in estimated associations, even when the qualitative conclusions remain unchanged.

The remainder of the paper is organized as follows. Section 2 describes the data sources and survey design. Section 3 outlines the empirical strategy and discusses how DK responses affect the measurement of financial literacy. Section 4 analyzes stock market participation at the extensive margin and

---

[3]Evidence from experimental and cognitive studies further shows that DK and guessing behavior are closely linked to confidence and willingness to guess, with important implications for measurement and inference (Baldiga, 2014).

portfolio allocation at the intensive margin. Section 5 presents a conceptual framework. Section 6 concludes.

2 Survey Design

We rely on a large-scale household survey conducted by INTAGE Inc. to investigate stock market participation and portfolio allocation.[4] The survey was conducted in October 2024 and targets approximately 10,000 households in Japan. In addition, comparable surveys were administered to about 4,000 households each in the United States, Germany, and China. In each country, the sample was constructed to be representative of the adult population using population-based quotas by gender and age group, with respondents aged 20–79. The survey collects detailed information on households' demographic characteristics, asset holdings, and financial literacy.

Our analysis focuses on stock market participation to shed light on the stock market participation puzzle. Households report their investment experience across a broad set of financial products, including cash and deposits, domestic and foreign stocks and bonds, gold, and digital financial instruments. In particular, investment experience in domestic stocks is elicited through the following question.

Q10 Please indicate your investment experience in the following financial products (select one for each).

    (1) Never invested

    (2) Have invested but do not currently hold

    (3) Currently hold

The list of financial products presented to respondents includes:

– Domestic stocks
  Stocks issued by domestic companies

– Domestic bonds
  U.S. government bonds, municipal bonds, bonds issued by domestic companies, etc.

– Foreign stocks
  Stocks issued by foreign companies

– Foreign bonds
  Bonds issued by foreign governments, companies, etc.

– Foreign exchange futures/options such as FX
  Derivatives related to foreign exchange [Financial derivatives]

Based on this question, we construct a stock-market-participation dummy that equals one if the respondent currently holds domestic or foreign stocks and zero otherwise. In addition, we measure

---

[4]See Appendix A for the complete survey questionnaire.

the intensive margin of stock market participation by eliciting the share of stocks in households' financial portfolios. Respondents report the percentage composition of their financial assets, excluding life insurance and real assets, with the total summing to 100%.[5]

To measure financial literacy, we include four standard multiple-choice questions that assess respondents' basic financial knowledge: simple interest, compound interest, the real interest rate, and risk diversification. Each question explicitly includes a DK option. This design allows us to distinguish incorrect answers from DK responses and to retain information on DK responses, which are informative about respondents' knowledge and confidence and play a central role in our empirical analysis. Financial literacy is measured as the number of correct answers, while DK responses are used separately to capture selection and confidence-related behavior.

Overall, the survey design allows us to analyze both the extensive margin (whether households participate in the stock market) and the intensive margin (how much they invest in stocks), while explicitly accounting for the role of financial literacy and DK responses. This structure provides a suitable framework for examining how financial knowledge and selection mechanisms contribute to the stock market participation puzzle.

3 Determinants of DK Responses

3.1 Systematic Patterns in DK Responses

Before analyzing the relationship between financial literacy and stock market participation, we first document who selects the DK option in financial literacy questions. Table 1 and Figure 1 summarize DK response rates by country and gender.

Two robust patterns emerge. First, DK responses are substantially more frequent among women than men across all four countries. For example, in Japan, approximately 63% of women respond DK to the risk diversification question, compared with about 45% of men. This gender gap is notably larger than in the United States, where the corresponding figures are around 36% for women and 19% for men. Second, DK responses are particularly prevalent for conceptually demanding questions, such as inflation and risk diversification.

Across countries, women consistently display higher DK response rates than men, as illustrated in Figure 1. The gender gap is especially pronounced for more abstract financial concepts, including the real interest rate and portfolio diversification (see Panels C and D of Table 1). In addition, cross-country differences are sizable: Japan exhibits systematically higher DK rates than the United States and Germany, while China shows relatively lower DK rates for several questions.

These patterns indicate that DK responses are not random measurement errors. Rather, they reflect systematic differences in individuals' engagement with financial concepts. Because participation in eq-

---

[5]Table B.1 in Appendix B reports summary statistics for the main variables used in the analysis. Consistent with the stock market participation puzzle, Japanese households exhibit substantially lower stock market participation and lower average stock shares than households in the United States and China, while participation and stock shares in Germany are also relatively low.

uity markets involves non-trivial informational and psychological costs, individuals who are more prone to answering DK may face higher barriers to processing financial information and making investment decisions. Motivated by these descriptive patterns, we treat DK responses as an outcome of interest and examine their determinants directly in the next subsection, providing a foundation for the subsequent analysis of stock market participation that explicitly accounts for DK-related selection mechanisms.

3.2 Empirical specification

To characterize who answers DK, we estimate regression models in which DK behavior is treated as the dependent variable. While Table 1 documents heterogeneity in DK response rates, it does not isolate the individual-level correlates of DK behavior. Table 2 addresses this by estimating the following specification:

$$DK_i = \alpha + X_i'\beta + \sum_c \gamma_c D_{c,i} + \varepsilon_i,$$

where DK_i denotes the number of DK responses across the financial literacy questions. The vector X_i includes a rich set of individual characteristics: a gender indicator, age-class indicators defined in ten-year intervals, marital status, an education indicator for college completion or higher, and an ordinal measure of household income based on Q6 of the survey, coded as increasing integers from lower to higher income categories. In addition, X_i includes standardized indices of time preference and risk tolerance, constructed so that higher values indicate greater patience and greater willingness to take risks, respectively.[6] The terms $D_{c,i}$ are country indicator variables that capture cross-country differences in baseline DK response rates, and ε_i is an idiosyncratic error term. In the pooled specification, country fixed effects are included with Japan as the reference group, and we also estimate country-specific regressions. Financial literacy itself is not included as a regressor, as the objective is to identify systematic correlates of DK behavior rather than to explain variation in knowledge levels.

In our framework, DK responses are informative because standard survey-based literacy scores combine true financial knowledge with confidence in that knowledge. This interpretation is in line with Bucher-Koenen et al. (2025), who show that DK answers often reflect low confidence (including under-confidence) rather than a pure lack of knowledge, and that treating DK mechanically as incorrect answers can understate true knowledge. Because our data do not include experimental variation in the availability of question-specific confidence measures, we do not separately identify knowledge and confidence. Instead, we use the number of DK responses as a parsimonious proxy for confidence-related response behavior, following the practical recommendation in Bucher-Koenen et al. (2025).

---

[6]Here, *Time preference$_i$* is constructed from responses to an intertemporal choice question (Question 31 in Appendix A), with higher values indicating greater patience. *Risk tolerance$_i$* is derived from responses to hypothetical lottery-choice questions (Questions 29 and 30 in Appendix A), with higher values indicating greater willingness to take risks. Both measures are standardized to have mean zero and unit variance.

3.3 Main results: from pooled to country-specific evidence

We begin with the pooled results reported in Column (5) of Table 2, which summarize the determinants of DK responses across all four countries. Three patterns emerge. First, women are significantly more likely to answer DK. Second, individuals with higher education and higher income are substantially less likely to select DK. Third, preference-related traits matter: respondents who are more patient and more risk-tolerant exhibit systematically lower DK response rates. Taken together, these results indicate that DK responses are closely related to both socio-demographic characteristics and economically meaningful preferences, rather than reflecting random uncertainty or inattentive survey behavior.

The pooled estimates also reveal sizable cross-country differences captured by the country fixed effects. Relative to Japan, DK responses are significantly less frequent in the United States and Germany, while China displays a distinct pattern. These differences persist after controlling for a rich set of individual characteristics and preferences, suggesting that country-level institutional, cultural, or educational factors shape respondents' willingness to engage with financial questions. This interpretation is consistent with the descriptive patterns documented in Table 1.[7]

Turning to the country-specific results in Columns (1)–(4) of Table 2, we find that the qualitative patterns observed in the pooled analysis are largely preserved within each country. In Japan, the likelihood of answering DK declines significantly with education and income, and women are significantly more prone to DK responses. Similar patterns hold in the United States and Germany, where both socioeconomic status and preference measures are important predictors of DK behavior. In these countries, the coefficients on time preference and risk tolerance are statistically significant, underscoring the role of attitudes toward uncertainty and delayed rewards in shaping engagement with financial concepts.

China exhibits a somewhat different configuration. While income remains an important determinant of DK responses, the gender gap is smaller and, in some specifications, statistically insignificant. Moreover, the association between DK responses and preference measures is weaker than in the other countries. Nonetheless, even in China, DK responses are far from random and remain systematically related to observable individual characteristics.

Overall, the results in Table 2 demonstrate that DK responses capture behavioral heterogeneity across and within countries. The consistent associations with education, income, and preferences imply that DK responses are informative indicators of cognitive and psychological frictions. These findings provide a crucial empirical foundation for the subsequent analysis, where we examine how financial literacy affects stock market participation and portfolio allocation after explicitly accounting for DK-related selection mechanisms.[8]

---

[7]Our contribution is primarily empirical and methodological: by explicitly accounting for DK responses in a large cross-country dataset, we clarify how measurement choices affect the estimated relationship between financial literacy and stock market participation.

[8]As a robustness check, we estimate analogous specifications using question-specific DK indicators for each of the four financial literacy questions (simple interest, compound interest, real interest rate, and risk diversification). The results, reported in Tables B.2–B.5, confirm that the determinants of DK behavior are stable across question types.

4   Financial Literacy and Stock Market Participation

This section examines the relationship between financial literacy and stock market participation.

4.1   Baseline relationship between financial literacy and stock market participation (extensive margin)

This subsection presents the baseline relationship between financial literacy and stock market participation at the extensive margin.

We begin by estimating the following baseline specification:

$$D_i = \alpha + \beta L_i + X_i'\gamma + \varepsilon_i, \tag{1}$$

where D_i is an indicator equal to one if household i currently owns stocks and zero otherwise, L_i denotes the standardized financial literacy index constructed as the sum of correct answers to the four basic financial literacy questions, and X_i is a vector of control variables including demographic characteristics and preference measures.

We report the results in Table 3. Financial literacy is positively associated with stock market participation across all specifications. Columns (1) and (2) show that this positive association is robust to alternative treatments of DK responses, whether DK is coded as -1 or as 0.

When constructing the financial literacy index, we code DK responses as -1 in one set of specifications. This coding reflects the idea that DK responses convey information that is distinct from both correct and incorrect answers. In particular, choosing DK indicates lower confidence in one's financial knowledge than providing an incorrect answer, which requires actively committing to a specific option. This interpretation is consistent with Bucher-Koenen et al. (2025), who show that DK responses often reflect low confidence or under-confidence rather than a pure lack of knowledge. Coding DK responses as -1 therefore imposes a natural ordinal structure on the literacy index, with correct answers coded as 1, incorrect answers as 0, and DK responses as -1. Importantly, our main results are robust to alternative treatments of DK responses, including coding DK as 0 and controlling directly for the number of DK responses.

4.2   The role of DK responses in the extensive margin

We next examine how accounting for DK responses affects the estimated relationship between financial literacy and stock market participation. Columns (2) and (3) of Table 3 code DK responses as zero when constructing the financial literacy index, but Column (3) additionally controls for the number of DK responses.

Two patterns emerge. First, the estimated coefficient on financial literacy is smaller in Column (3) than in Column (2). Second, the coefficient on the number of DK responses is negative and statistically

significant. Together, these results indicate that DK responses capture confidence-related variation that is directly related to stock market participation, over and above their role in the construction of the literacy index. Consistent with Bucher-Koenen et al. (2025), DK responses may therefore reflect, at least in part, under-confidence in financial knowledge rather than a pure lack of knowledge.

When DK responses are not controlled for, the financial literacy coefficient partly reflects differences in confidence in financial knowledge, rather than knowledge alone. Once DK behavior is explicitly accounted for, the estimated relationship between financial literacy and stock market participation more closely reflects the association between true financial knowledge and the decision to participate in the stock market.[9]

4.3 Past investment experience, confidence, and sample selection

We next investigate how the inclusion of individuals with past stock market experience influences the estimated relationship between financial literacy and current stock market participation. Column (4) of Table 3 expands the sample to include individuals who have previously invested in stocks but do not currently hold any. Relative to the baseline specification in Column (3), the estimated coefficient on financial literacy becomes smaller.

A natural interpretation is that past investment experience affects measured financial literacy not only through the accumulation of financial knowledge, but also through confidence in one's financial understanding. Individuals who have invested in stocks in the past are likely to have acquired familiarity with financial concepts, terminology, and market mechanisms, which can persist even after they exit the stock market. As a result, measured financial literacy partly reflects learning-by-doing and confidence gained through prior exposure to financial markets, rather than an exogenous determinant of current participation.

This channel is closely related to the role of DK responses emphasized throughout the paper. As discussed in Sections 3 and 4.2, DK behavior captures confidence-related response patterns that are systematically linked to stock market participation. Past investment experience may reduce the propensity to answer DK by increasing confidence, even if individuals no longer actively participate in the stock market. Consequently, excluding past investors mitigates, but does not fully eliminate, the endogeneity arising from the joint determination of investment experience, confidence, and measured financial literacy.

Importantly, this concern implies that the attenuation observed in Column (4) should be interpreted as a sample composition effect rather than evidence against the relevance of financial literacy. Including former investors introduces observations with relatively high measured literacy and confidence but zero current stockholdings. These observations weaken the contemporaneous covariance between financial literacy and current participation, mechanically attenuating the estimated coefficient.

---

[9]The same pattern is observed in the country-specific results reported in Tables B.6–B.9 in Appendix B, confirming that the role of DK responses is not driven by a particular country.

For this reason, our preferred specifications exclude households with past stock market experience and focus on a cleaner comparison between current participants and never-participants. This restriction allows us to interpret the financial literacy coefficient in Column (3) as reflecting a more direct association between financial knowledge, confidence-related response behavior, and the decision to enter the stock market, abstracting from legacy effects of prior market exposure.

4.4 Robustness check: financial literacy and portfolio stock shares (intensive margin)

As a robustness check, we examine the intensive margin of stock market participation by analyzing how financial literacy relates to the share of stocks held in household portfolios. Unlike the extensive margin, which captures the decision to enter the stock market, the intensive margin reflects portfolio allocation conditional on participation.

We estimate the following specification for the intensive margin:

$$S_i = \alpha + \beta L_i + X_i'\gamma + \varepsilon_i, \tag{2}$$

where S_i denotes the share of stocks in household i's financial portfolio.

Table 4 shows that financial literacy is positively associated with portfolio stock shares across all specifications.[10] This positive association is robust to alternative treatments of DK responses. When DK responses are coded as zero, explicitly controlling for the number of DK responses does not alter the qualitative relationship between financial literacy and portfolio stock shares. Across specifications, the estimated association remains positive and statistically significant, indicating that financial literacy is systematically related not only to stock market participation but also to portfolio allocation conditional on participation.[11] These results suggest that financial literacy is robustly related to portfolio allocation conditional on participation, even when accounting for DK responses as an additional control variable.

Our findings do not overturn the conventional view that financially more knowledgeable households are more likely to invest in stocks. Rather, they clarify how measurement and sample selection choices shape the magnitude and interpretation of this relationship. In particular, while the sign of the financial literacy coefficient remains positive across all specifications, accounting for DK responses materially changes the magnitude of the estimated effect. This distinction matters for interpretation: without DK controls, part of the estimated literacy effect reflects confidence-related response behavior, rather than financial knowledge per se.

---

[10] A natural alternative would be to estimate a two-equation selection model that jointly models stock market participation and portfolio shares. Given the cross-country nature of our data and the focus on DK behavior, we leave such an extension to future research.

[11] Similar patterns emerge in the country-specific results when controlling for the number of DK responses (Tables B.10–B.13 in Appendix B).

5 Why Ignoring DK Responses Biases Estimates

This section provides a simple theoretical argument for why omitting DK behavior biases the estimated relationship between financial literacy and stock market participation. Throughout, let L_i^* denote an individual's latent (true) financial literacy, and let DK_i denote an individual's propensity to answer DK in literacy questions.[12]

5.1 Setup: measured literacy and DK-related behavior

A key feature of survey-based financial literacy measures is that they conflate true knowledge with confidence in that knowledge. We capture this by allowing the observed literacy index L_i to depend on both latent (true) financial literacy and a confidence-related response component:

$$L_i = L_i^* + \gamma C_i + u_i, \tag{3}$$

where C_i represents confidence in financial knowledge that affects whether respondents attempt an answer rather than choosing DK, $\gamma > 0$, and u_i is classical measurement error. The propensity to answer DK is negatively related to confidence, so that (conceptually) higher DK_i corresponds to lower C_i.[13]

5.2 Omitting DK induces upward bias

Suppose the causal relationship between true financial literacy and stock market participation is

$$D_i = \beta_E L_i^* + \delta_E C_i + \varepsilon_i^E, \qquad E[\varepsilon_i^E \mid L_i^*, C_i] = 0, \tag{4}$$

where $D_i \in \{0, 1\}$ is an indicator that equals one if the household currently participates in the stock market and zero otherwise. The parameter $\beta_E > 0$ captures the causal effect of true financial literacy on market entry, while $\delta_E > 0$ reflects that higher confidence in financial knowledge lowers perceived participation costs and increases the likelihood of entering the stock market (e.g., by reducing hesitation or avoidance in financial decision-making).

Researchers often regress D_i on the observed literacy index L_i without controlling for DK behavior (hence without controlling for C_i). Substituting Equation (3) into Equation (4) and applying the

---

[12]While Bucher-Koenen et al. (2025) develop a latent class model with experimental variation to separate knowledge and confidence, our approach differs conceptually. We provide a stylized econometric framework that illustrates how omitting DK-related behavior can bias coefficient estimates in standard regressions of stock market participation or portfolio allocation. Our focus is not on identifying latent traits but on demonstrating the empirical implications of failing to control for DK-related selection using standard survey data.

[13]In the empirical analysis, DK_i is proxied by the number of DK responses. The role of C_i is to capture confidence-related response behavior that links DK choices to both measured literacy and financial decisions, consistent with the interpretation in Bucher-Koenen et al. (2025).

standard omitted-variable bias formula yields

$$\text{plim}\, \hat{\beta}_E^{\text{no DK}} = \beta_E + \delta_E \frac{Cov(L_i, C_i)}{Var(L_i)}, \tag{5}$$

where $\hat{\beta}_E^{\text{no DK}}$ is denoted as the OLS estimator of the effect of L_i on the extensive margin of stock market participation, obtained from a regression that omits C_i. Because $\gamma > 0$ in Equation (3), we have $Cov(L_i, C_i) > 0$ (higher confidence raises measured literacy when DK is not modeled separately). With $\delta_E > 0$, the second term in Equation (5) is positive, implying

$$\hat{\beta}_E^{\text{no DK}} > \beta_E.$$

Thus, ignoring DK behavior generates an upward bias in the estimated effect of financial literacy on the probability of stock market participation. Intuitively, when DK responses are not controlled for, the financial literacy coefficient absorbs both true financial knowledge and confidence-related response behavior, overstating the effect of knowledge on market entry.

5.3 Sample Composition Bias from Past Stock Market Experience

In addition to the estimation bias arising from ignoring DK responses, this subsection explains why the estimated coefficient on financial literacy becomes smaller when the sample includes households with past stock market experience. This pattern appears in Column (4) of Table 3, where such households are added to the baseline specification.

The mechanism is sample composition, not a change in the underlying relationship between financial literacy and stock market participation. Households that previously invested in stocks but no longer hold them differ systematically from those that have never participated. Past participation is likely to be positively correlated with latent financial literacy through learning-by-doing, exposure to financial information, and experience-based belief updating.

Including former participants changes the interpretation of measured financial literacy. In this expanded sample, financial literacy partly reflects past market exposure rather than an exogenous determinant of current participation. As a result, variation in financial literacy is less informative about contemporaneous stock market participation. This weakens the covariance between literacy and participation and mechanically attenuates the estimated coefficient.

Formally, former participants introduce observations with relatively high measured financial literacy but a zero current share of stocks in the household portfolio. These observations increase dispersion in financial literacy without a corresponding increase in participation, biasing the estimated relationship toward zero.

This pattern is particularly relevant for older households who decumulate assets after retirement. In Japan, for example, many households aged 65 and above rely on pension income and gradually draw down financial assets, including stocks, despite having accumulated financial knowledge through

past market participation. Such households therefore combine high measured financial literacy with a zero current share of stocks in the household portfolio. As a result, the financial literacy coefficient in Column (4) of Table 3 is smaller than in specifications that exclude households with past investment experience.

For this reason, our preferred specifications exclude households with past stock market experience. Focusing on this restricted sample yields a cleaner interpretation of the financial literacy coefficient and avoids attenuation driven by sample composition.

6 Conclusion

This paper examines the relationship between financial literacy and stock market participation using large-scale cross-country household survey data from Japan, the United States, Germany, and China, with a particular focus on the role of DK responses in financial literacy measurement. By exploiting harmonized survey data collected from a large number of households across diverse institutional and cultural settings, we pursue three related objectives and document three corresponding sets of findings.

First, we document who selects DK responses in standard financial literacy questions across countries, genders, and question types. We find that DK responses are systematic behavioral outcomes rather than random measurement error, and that individuals who frequently answer DK are substantially less engaged in stock markets.

Second, we examine how accounting for DK-related selection and sample composition affects the estimated relationship between financial literacy and stock market participation and portfolio allocation. Once DK-related selection and sample composition are properly accounted for, financial literacy remains robustly and economically meaningfully associated with stock market participation. In our benchmark specifications, a one-standard-deviation higher level of financial literacy is associated with an approximately 6–7 percentage point higher probability of stock market participation and with higher stock shares among participating households.

Third, we develop a conceptual framework that clarifies how ignoring DK responses distorts empirical estimates of the relationship between financial literacy and stock market behavior. This framework provides a unified interpretation of why empirical estimates in the existing literature may vary across studies, even when they rely on similar survey-based measures of financial literacy. By explicitly accounting for DK behavior and sample composition, our analysis shows how measurement choices can lead to differences in estimated effects.

We emphasize that our analysis is based on cross-sectional survey data and therefore documents conditional associations rather than causal effects. While we address important measurement and selection issues related to DK behavior, establishing causal effects of financial literacy—such as through experimental or quasi-experimental variation—remains an important direction for future research.

References

Allgood, Sam, and William B. Walstad (2016). "The Effects of Perceived and Actual Financial Literacy on Financial Behaviors." *Economic Inquiry* 54(1), 675–697.

Baldiga, Katherine (2014). "Gender Differences in Willingness to Guess." *Management Science* 60(2), 434–448

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

Bucher-Koenen, Tabea, Rob Alessie, Annamaria Lusardi, and Maarten van Rooij (2025). "Fearless Woman: Financial Literacy, Confidence, and Stock Market Participation." *Management Science* 71(9), 7414–7430.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini (2007). "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes." *Journal of Political Economy* 115(5), 707–747.

Campbell, John Y. (2006). "Household Finance." *Journal of Finance* 61(4), 1553–1604.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny (2018). "A Crisis of Beliefs: Investor Psychology and Financial Fragility." *Princeton University Press*.

Giannetti, Mariassunta, and Yrjö Koskinen (2010). "Investor Protection, Equity Returns, and Financial Globalization." *Journal of Financial and Quantitative Analysis* 45(1), 135–168.

Giannetti, Mariassunta, and Tracy Yue Wang (2016). "Corporate Scandals and Household Stock Market Participation." *Journal of Finance* 71(6), 2591–2636.

Grinblatt, Mark, Matti Keloharju, and Juhani T. Linnainmaa (2011). "IQ and Stock Market Participation." *Journal of Finance* 66(6), 2121–2164.

Guiso, Luigi, Michael Haliassos, and Tullio Jappelli (2003). "Household Stockholding in Europe: Where Do We Stand and Where Do We Go?" *Economic Policy* 18(36), 123–170.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2008). "Trusting the Stock Market." *The Journal of Finance* 63(6), 2557–2600.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein (2004). "Social Interaction and Stock-Market Participation." *Journal of Finance* 59(1), 137–163.

Lusardi, Annamaria, and Olivia S. Mitchell (2014). "The Economic Importance of Financial Literacy: Theory and Evidence." *Journal of Economic Literature* 52(1), 5–44.

Malmendier, Ulrike, and Stefan Nagel (2011). "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?" *Quarterly Journal of Economics* 126(1), 373–416.

van Rooij, Maarten, Annamaria Lusardi, and Rob Alessie (2011). "Financial Literacy and Stock Market Participation." *Journal of Financial Economics* 101(2), 449–472.

Vissing-Jorgensen, Annette (2003). "Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth?" *Journal of Finance* 58(1), 97–142.

Yamori, Nobuyoshi, and Hitoe Ueyama (2022). "Financial Literacy and Low Stock Market Participation of Japanese Households." *Finance Research Letters* 44, 102074.

Table 1: "Don't Know" response rates by question, country, and gender

	Japan (%)	U.S. (%)	Germany (%)	China (%)	All countries (%)
Panel A: Question 1 (simple interest calculation)					
Female	29.33	15.38	14.07	14.89	21.42
Male	19.71	8.39	7.11	17.58	14.98
Panel B: Question 2 (compound interest calculation)					
Female	30.98	17.26	14.47	7.04	21.18
Male	20.68	8.85	8.02	7.75	13.85
Panel C: Question 3 (real interest rate calculation)					
Female	66.96	37.14	33.08	10.08	45.16
Male	41.74	20.53	19.04	8.99	27.73
Panel D: Question 4 (risk diversification)					
Female	62.77	35.81	28.39	12.00	42.50
Male	44.51	18.50	16.58	11.46	28.63

Notes: This table shows the proportion of respondents who answered "Don't Know" (DK) to each of the four financial literacy questions, by country and gender. Question 1: simple interest calculation; Question 2: compound interest calculation; Question 3: real interest rate calculation; Question 4: risk diversification.

Table 2: Determinants of "Don't Know" Responses

Variables	(1) Japan	(2) United States	(3) Germany	(4) China	(5) All
			Number of DK responses		
Male	−0.252***	−0.230***	−0.243***	0.028	−0.199***
	(0.032)	(0.034)	(0.031)	(0.026)	(0.016)
Higher education	−0.399***	−0.210***	−0.134***	0.078***	−0.240***
	(0.029)	(0.035)	(0.035)	(0.029)	(0.016)
Age class	−0.094***	−0.052***	−0.061***	0.001	−0.077***
	(0.009)	(0.010)	(0.010)	(0.011)	(0.005)
Married	0.035	−0.049	−0.012	−0.050	0.006
	(0.031)	(0.036)	(0.033)	(0.043)	(0.018)
Income class	−0.023***	−0.024***	−0.035***	−0.027***	−0.028***
	(0.006)	(0.005)	(0.005)	(0.003)	(0.002)
Time preference (std.)	−0.196***	−0.137***	−0.076***	−0.179***	−0.166***
	(0.013)	(0.017)	(0.017)	(0.014)	(0.008)
Risk tolerance (std.)	−0.445***	−0.311***	−0.273***	−0.125***	−0.355***
	(0.012)	(0.018)	(0.016)	(0.018)	(0.008)
Country dummy: U.S.					−0.467***
					(0.022)
Country dummy: Germany					−0.576***
					(0.022)
Country dummy: China					−0.394***
					(0.023)
Constant	2.149***	1.400***	1.251***	0.699***	1.886***
	(0.044)	(0.051)	(0.051)	(0.055)	(0.027)
Observations	7,645	3,954	3,845	3,991	19,435
R^2	0.258	0.195	0.145	0.109	0.264

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Variable definitions: The dependent variable is the **Number of DK responses** across basic financial literacy questions. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices. Column (5) pools all observations with country-specific intercepts (Japan as the baseline), while Columns (1)–(4) provide country-specific estimates.

Table 3: Determinants of Stock Market Participation: Pooled Analysis

	(1)	(2)	(3)	(4)
		$D^{Stock\ Market\ Participation}$		
Financial literacy (DK = −1, std.)	0.083***			
	(0.004)			
Financial literacy (DK = 0, std.)		0.080***	0.067***	0.056***
		(0.004)	(0.005)	(0.005)
Number of DK responses			−0.022***	−0.023***
			(0.005)	(0.005)
Control Variables				
Male	0.034***	0.032***	0.032***	0.025***
	(0.007)	(0.007)	(0.007)	(0.006)
Higher education	0.088***	0.085***	0.085***	0.081***
	(0.008)	(0.008)	(0.008)	(0.007)
Age class	0.020***	0.019***	0.019***	0.012***
	(0.002)	(0.002)	(0.002)	(0.002)
Married	0.039***	0.040***	0.040***	0.032***
	(0.008)	(0.008)	(0.008)	(0.007)
Time preference (std.)	0.009**	0.014***	0.011***	0.005
	(0.004)	(0.004)	(0.004)	(0.003)
Risk tolerance (std.)	0.083***	0.088***	0.085***	0.068***
	(0.004)	(0.004)	(0.004)	(0.004)
Income class	0.027***	0.027***	0.027***	0.024***
	(0.001)	(0.001)	(0.001)	(0.001)
Country dummy: United States	−0.049***	−0.030***	−0.039***	−0.070***
	(0.010)	(0.010)	(0.010)	(0.009)
Country dummy: Germany	−0.100***	−0.086***	−0.094***	−0.102***
	(0.009)	(0.009)	(0.010)	(0.009)
Country dummy: China	−0.056***	−0.026**	−0.038***	−0.122***
	(0.012)	(0.012)	(0.013)	(0.010)
Constant	0.057***	0.047***	0.054***	0.081***
	(0.011)	(0.011)	(0.011)	(0.010)
Observations	14,255	14,255	14,255	19,435
R^2	0.232	0.232	0.233	0.162

Notes: Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$. The dependent variable is a dummy equal to one if the household currently owns domestic or foreign stocks.
Variable definitions: **Financial literacy** is a standardized index constructed from basic financial literacy questions. In Column (1), DK responses are coded as −1; in Columns (2)–(4), DK responses are coded as 0. **Number of DK responses** is the count of DK answers across the literacy questions. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices. All columns report pooled regressions with country fixed effects (Japan as the baseline).

Table 4: Financial Literacy and Portfolio Stock Shares (Intensive Margin): Baseline Analysis

Variables	(1) StockShare DK=−1	(2) StockShare DK=0	(3) StockShare DK=0
Financial literacy (DK = −1, std.)	3.468***		
	(0.157)		
Financial literacy (DK = 0, std.)		3.603***	3.756***
		(0.158)	(0.204)
Number of DK responses			0.279
			(0.191)
Control Variables			
Male	1.791***	1.671***	1.676***
	(0.270)	(0.269)	(0.270)
Higher education	3.271***	3.001***	2.997***
	(0.326)	(0.325)	(0.325)
Age class	1.110***	1.018***	1.016***
	(0.093)	(0.093)	(0.093)
Married	0.545*	0.610*	0.616**
	(0.313)	(0.312)	(0.312)
Time preference (std.)	−0.392***	−0.158	−0.124
	(0.143)	(0.142)	(0.145)
Risk tolerance (std.)	2.833***	2.959***	2.999***
	(0.180)	(0.177)	(0.181)
Income class	0.806***	0.831***	0.835***
	(0.041)	(0.041)	(0.041)
Country dummy: U.S.	−2.557***	−1.681***	−1.567***
	(0.408)	(0.404)	(0.417)
Country dummy: Germany	−3.680***	−3.065***	−2.965***
	(0.402)	(0.400)	(0.405)
Country dummy: China	−7.326***	−5.951***	−5.800***
	(0.417)	(0.413)	(0.429)
Constant	0.610	0.419	0.336
	(0.462)	(0.453)	(0.461)
Observations	19,435	19,435	19,435
R^2	0.130	0.135	0.135

Notes: Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable is the share of stocks in the household portfolio.

Variable definitions: **Financial literacy** is a standardized index constructed from basic financial literacy questions. In Column (1), DK responses are coded as −1; in Columns (2) and (3), DK responses are coded as 0. **Number of DK responses** is the count of DK answers across the literacy questions. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices. All columns report pooled regressions with country fixed effects (Japan as the baseline).
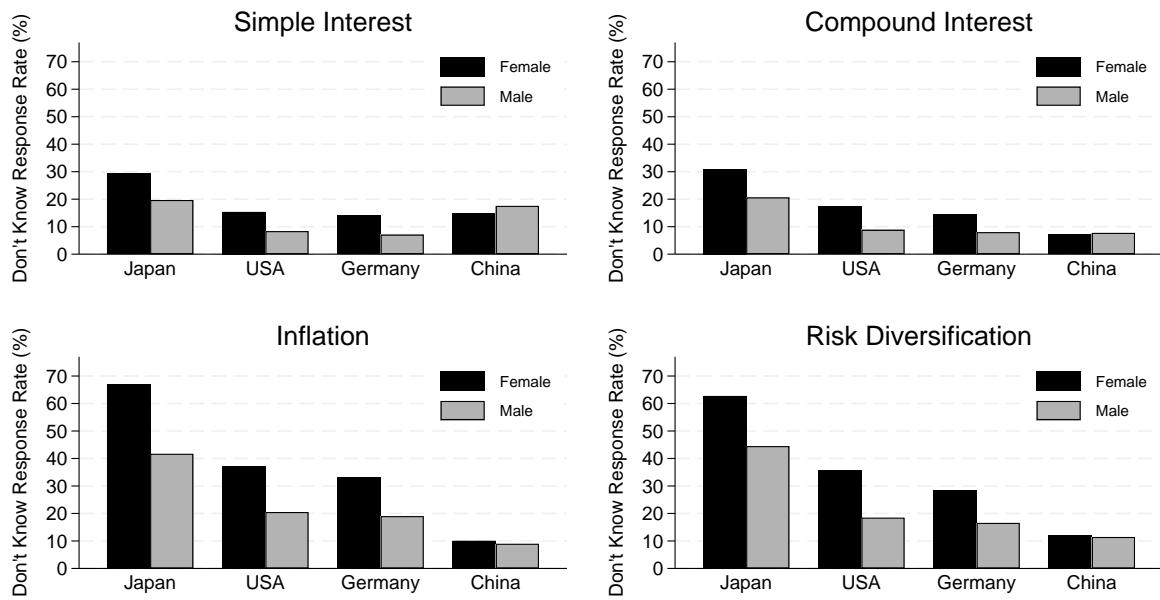
Figure 1: "Don't Know" (DK) Response Rates to Financial Literacy Questions by Country and Gender
Notes: Bars show the percentage of respondents who answered "Don't Know" to each financial literacy question, separately for females and males, in Japan, the United States, Germany, and China. The four questions correspond to standard measures of financial knowledge covering (i) simple interest calculation, (ii) compound interest calculation, (iii) real interest rate (inflation), and (iv) risk diversification. DK responses are interpreted as reflecting both knowledge gaps and confidence-related behavior. The figure is based on the full survey sample used in the analysis.

Appendix A   Questionnaire

Q1. Please select your gender (select only one).

    (1) Male

    (2) Female

Q2. Please choose your current age (select only one).

    (1) 20 – 29 years old

    (2) 30 – 39 years old

    (3) 40 – 49 years old

    (4) 50 – 59 years old

    (5) 60 – 69 years old

    (6) 70 – 79 years old

    (7) 80 years old and over

Q3. Please choose your educational background. If you are still in school, please choose the school you last attended (select only one).

    (1) Junior High School Graduate

    (2) High School Graduate

    (3) Vocational school graduate

    (4) Junior college graduate

    (5) University graduate

    (6) Graduate degree

    (7) Other

Q4. Please indicate your current marital status.

    (1) Unmarried

    (2) Married

Q5. Do you own a residence (including apartments, etc.) for you and your family to live in? Please do not include residences owned for investment purposes that are not used by you or your family.

*Please include cases where the property has not been registered or where the mortgage or other payments have not been completed.

(1) I own a home

(2) I purchased it myself

(3) I acquired it through inheritance, etc.

(4) I don't own a home

(5) I don't own a home (e.g., rent)

Q6. Please indicate your total annual income (before taxes) for the last year.

   (1) No income

   (2) $1 – $9,999

   (3) $10,000 – $19,999

   (4) $20,000 – $29,999

   (5) $30,000 – $39,999

   (6) $40,000 – $49,999

   (7) $50,000 – $59,999

   (8) $60,000 – $69,999

   (9) $70,000 – $79,999

   (10) $80,000 – $89,999

   (11) $90,000 – $99,999

   (12) $100,000 – $119,999

   (13) $120,000 – $139,999

   (14) $140,000 – $159,999

   (15) $160,000 – $179,999

   (16) $180,000 – $199,999

   (17) $200,000 or more

   (18) I don't know

   (19) Prefer not to say

Q7. How much in financial assets do you own? Please select the amount that you believe is closest.

*Including crypto-assets and other digital financial products. However, please exclude life insurance and physical assets such as real estate and cars.

*Even if you have any loans, you do not have to subtract the amount of those loans.

   (1) $0 – $9,999

   (2) $10,000 – $29,999

   (3) $30,000 – $49,999

   (4) $50,000 – $99,999

   (5) $100,000 – $149,999

   (6) $150,000 – $199,999

   (7) $200,000 – $249,999

   (8) $250,000 – $299,999

   (9) $300,000 – $399,999

  (10) $400,000 – $499,999

  (11) $500,000 – $599,999

  (12) $600,000 – $699,999

  (13) $700,000 – $799,999

  (14) $800,000 – $899,999

  (15) $900,000 – $999,999

  (16) $1,000,000 or more

  (17) I don't know

  (18) Prefer not to say

Q8. If you don't mind answering this question, what is the balance of loans/debt you have? Please select the amount that you believe is closest.

   (1) $0 – $9,999

   (2) $10,000 – $29,999

   (3) $30,000 – $49,999

   (4) $50,000 – $99,999

   (5) $100,000 – $149,999

   (6) $150,000 – $199,999

   (7) $200,000 – $249,999

   (8) $250,000 – $299,999

   (9) $300,000 – $399,999

  (10) $400,000 – $499,999

  (11) $500,000 – $599,999

  (12) $600,000 – $699,999

  (13) $700,000 – $799,999

  (14) $800,000 – $899,999

  (15) $900,000 – $999,999

(16) $1,000,000 or more

(17) I don't know

(18) Prefer not to say

Q9. How familiar are you with each of the following financial products?

- **Domestic stocks**
  Stocks issued by domestic companies
- **Domestic bonds**
  U.S. government bonds, municipal bonds, bonds issued by domestic companies, etc.
- **Foreign stocks**
  Stocks issued by foreign companies
- **Foreign bonds**
  Bonds issued by foreign governments, companies, etc.
- **Foreign exchange futures/options such as FX**
  Derivatives related to foreign exchange [Financial derivatives]

(1) Don't know/Have never heard of them

(2) Have heard of them, but don't know much about them

(3) Not enough to teach others, but have some knowledge

(4) I know enough about it to teach others

Q10. Please indicate your investment experience in the following financial products (Select only one of each)

- **Domestic stocks**
  Stocks issued by domestic companies
- **Domestic bonds**
  U.S. government bonds, municipal bonds, bonds issued by domestic companies, etc.
- **Foreign stocks**
  Stocks issued by foreign companies
- **Foreign bonds**
  Bonds issued by foreign governments, companies, etc.
- **Foreign exchange futures/options such as FX**
  Derivatives related to foreign exchange [Financial derivatives]

(1) Never invested

(2) Have invested but do not currently hold

(3) Currently hold

Q11. Do your parents or siblings have experience investing in stocks, bonds, etc.?

(1) Yes

(2) No

(3) I don't know

Q12. Have you ever received financial education at home, school, or work?

*Financial education is an educational program that teaches students about the overall workings of money, including economic policy, asset management, and financial troubles.

(1) Yes

(2) No

(3) I don't remember

Q13. What are your thoughts on domestic stocks and investment trusts that invest mainly in domestic stocks (all denominated in domestic currency)? (Multiple selections)

(1) I would expect profits

(2) Diversifying investments across financial products with different price fluctuations reduces investment risk

(3) Convenience

(4) Interest in the product

(5) I am concerned about large price fluctuations

(6) I am concerned about losses caused by conflicts of interest of brokers

(7) I am concerned about losses resulting from fraudulent accounting practices of the issuer (e.g., window dressing)

(8) I am concerned about losses caused by unfair trading by investors (e.g., insider trading)

(9) I don't understand the system, risks, etc., of the product

(10) Worry about the investment method/utilization method is difficult and cannot be mastered

(11) There is a risk that investor/consumer protection may not be sufficient

(12) I do not trust the issuer or intermediary

(13) Worry about leakage of personal information

(14) Worry about misuse of identity authentication, identity theft fraud, etc.

(15) Disadvantageous tax treatment

(16) Difficult to use due to strict regulations

(17) Other

(18) No particular impressions

Q14. What are your thoughts on foreign stocks, and investment trusts that invest mainly in foreign stocks (all denominated in foreign currencies)? (Multiple selections)

   (1) I would expect profits

   (2) Diversifying investments across financial products with different price fluctuations reduces investment risk

   (3) Convenience

   (4) Interest in the product

   (5) I am concerned about large price fluctuations

   (6) I am concerned about losses caused by conflicts of interest of brokers

   (7) I am concerned about losses resulting from fraudulent accounting practices of the issuer (e.g., window dressing)

   (8) I am concerned about losses caused by unfair trading by investors (e.g., insider trading)

   (9) I don't understand the system, risks, etc., of the product

   (10) Worry about the investment method/utilization method is difficult and cannot be mastered

   (11) There is a risk that investor/consumer protection may not be sufficient

   (12) I do not trust the issuer or intermediary

   (13) Worry about leakage of personal information

   (14) Worry about misuse of identity authentication, identity theft fraud, etc.

   (15) Disadvantageous tax treatment

   (16) Difficult to use due to strict regulations

   (17) Other

   (18) No particular impressions

Q15. What are your thoughts on foreign exchange futures/options such as FX? (Multiple selections)

   (1) I would expect profits

   (2) Diversifying investments across financial products with different price fluctuations reduces investment risk

   (3) Convenience

   (4) Interest in the product

   (5) I am concerned about large price fluctuations

(6) I am concerned about losses caused by conflicts of interest of brokers

(7) I am concerned about losses resulting from fraudulent accounting practices of the issuer (e.g., window dressing)

(8) I am concerned about losses caused by unfair trading by investors (e.g., insider trading)

(9) I don't understand the system, risks, etc., of the product

(10) Worry about the investment method/utilization method is difficult and cannot be mastered

(11) There is a risk that investor/consumer protection may not be sufficient

(12) I do not trust the issuer or intermediary

(13) Worry about leakage of personal information

(14) Worry about misuse of identity authentication, identity theft fraud, etc.

(15) Disadvantageous tax treatment

(16) Difficult to use due to strict regulations

(17) Other

(18) No particular impressions

Q16. For each of the following financial products, select the response that best describes your investment gains/losses (yield) since you started investing.

*For instance, if your initial investment of $10,000 has now grown to $15,000, respond "+50%."

- Domestic stocks and investment trusts that invest mainly in domestic stocks (all denominated in domestic currency)
- Foreign stocks, and investment trusts that invest mainly in foreign stocks (all denominated in foreign currencies)
- Foreign exchange futures/options such as FX

(1) +100% or more of the investment principal

(2) +50% or more of the investment principal

(3) +30% or more of the investment principal

(4) +10% or more of the investment principal

(5) +1% or more of the investment principal

(6) Virtually no gain/loss

(7) −1% or less of the investment principal

(8) −10% or less of the investment principal

(9) −30% or less of the investment principal

(10) −50% or less of the investment principal

Q17. How familiar are you with each of the following (crypto assets, etc.)?

- **Crypto assets**
  Property value that can be exchanged over the Internet using distributed ledger technology called blockchain and know-how of tokens (coins), virtual currencies: e.g. Bitcoin, Ethereum, Litecoin, Ripple, Stellar

- **Stablecoin**
  Virtual currencies designed to be linked (pecked) to a stable asset such as the US dollar or gold: e.g. USDT, USDC, BUSD

- **ST**
  Security token is a digital security that applies distributed ledger technology called blockchain and know-how of tokens (coins) to securities

- **NFTs**
  Non-fungible token, digital tokens without substitutability issued using distributed ledger technology called blockchain or know-how of tokens (coins): e.g. Cryptokitties (games), CryptoPunks (digital art)

- **CBDC**
  Digital currency issued by a central bank

(1) Don't know/Have never heard of it

(2) Have heard of them, but don't know much about them

(3) Not enough to teach others, but have some knowledge

(4) I know enough about it to teach others. I know a lot about it

Q18. We would like to ask those who are aware of [new digital financial instruments (crypto-assets, stable coins, STs, NFTs)]. Please tell us about your experience investing in and using [new digital financial instruments (crypto-assets, stable coins, STs, NFTs)]."

- **Crypto assets**
  Property value that can be exchanged over the Internet using distributed ledger technology called blockchain and know-how of tokens (coins), virtual currencies: e.g. Bitcoin, Ethereum, Litecoin, Ripple, Stellar

- **Stablecoin**
  Virtual currencies designed to be linked (pecked) to a stable asset such as the US dollar or gold: e.g. USDT, USDC, BUSD

- **ST**
  Security token is a digital security that applies distributed ledger technology called blockchain and know-how of tokens (coins) to securities

- **NFTs**

    Non-fungible token, digital tokens without substitutability issued using distributed ledger technology called blockchain or know-how of tokens (coins): e.g. Cryptokitties (games), CryptoPunks (digital art)

(1) Never invested

(2) Have invested but do not currently hold

(3) Currently hold

Q19. We would like to ask this question to those who currently own [new digital financial instruments (crypto-assets, stable coins, STs, NFTs)]. How much of each of the following [new digital financial instruments (crypto-assets, stable coins, STs, NFTs)] do you own? Please base your answer on the current appraised value, not the amount at the time of purchase.

- Crypto assets
- Stablecoin
- ST
- NFTs

(1) Less than $100

(2) $100 – $499

(3) $500 – $999

(4) $1,000 – $4,999

(5) $5,000 – $9,999

(6) $10,000 – $49,999

(7) $50,000 – $99,999

(8) $100,000 – $499,999

(9) $500,000 – $999,999

(10) $1,000,000 or more

Q20. Please indicate the percentage of financial assets you own.

*Please exclude life insurance and physical assets such as real estate and cars.

*Please answer the percentage of each so that the total is 100%.

*Please enter "0" for those that you do not own.

(1) Cash and deposits

(2) Domestic stocks and mutual funds managed primarily in domestic stocks (all denominated in the local currency)

(3) Domestic bonds

(4) Foreign stocks and mutual funds managed primarily in foreign stocks (all denominated in foreign currencies)

(5) Foreign bonds

(6) Gold

(7) New digital financial instruments (crypto-assets, stable coins, STs, NFTs)

(8) Other

Q21. We would like to ask this question to those who currently own [new digital financial instruments (crypto-assets, stable coins, STs, NFTs)]. How has your investment in the following new digital financial instruments (crypto-assets, stable coins, STs, NFTs) changed over the past year? Please answer about changes in your investments without considering changes in market prices.

- Crypto assets
- Stablecoin
- ST
- NFTs

(1) I did not own this product a year ago

(2) Started investing in this product over the past year

(3) I owned this product a year ago and currently own it as well

(4) Increased the amount I invested

(5) Slightly increased the amount I invested

(6) Almost no change in the amount I invested

(7) Slightly decreased the amount I invested

(8) Decreased the amount I invested

Q22. You said you have owned crypto assets. What was your purpose in acquiring/owning crypto assets? (Multiple selections)

(1) Short-term (period of less than 1 year) investment purpose

(2) Long-term (period of more than 1 year) investment purpose

(3) To purchase other crypto assets, stablecoins, security tokens, or NFTs

(4) To use as a means of payment for the purchase of goods or services (except for the purchase of crypto assets or stablecoins)

(5) For remittance purposes

(6) Studying crypto assets, blockchain, etc.

(7) Received it as a gift/for free

(8) Participating in the crypto asset community (including for emotional support)

(9) Other

Q23. What do you think about [new digital financial instruments (crypto-assets, stable coins, STs, NFTs)]? (Multiple answers allowed)

    (1) There is an expectation of profit and price appreciation

    (2) Diversifying investments across financial products with different price fluctuations reduces investment risk

    (3) Convenience

    (4) Interest about new technology and products, and expectation of future potential

    (5) I am concerned about large price fluctuations

    (6) I am concerned about losses caused by conflicts of interest of brokers

    (7) I am concerned about losses resulting from fraudulent accounting practices of the issuer (e.g., window dressing)

    (8) I am concerned about losses caused by unfair trading by investors (e.g., insider trading)

    (9) I do not understand the contents of the product

    (10) Investment/utilization methods are difficult to understand and there is worry about not being able to use the products.

    (11) Insufficient investor/consumer protection

    (12) I do not trust the issuer or intermediary

    (13) Expensive fees

    (14) Worry about leakage of personal information

    (15) Worry about misuse of identity authentication, identity theft fraud, etc.

    (16) Disadvantageous tax treatment

    (17) Difficult to use due to strict regulations

    (18) Concern about whether crypto asset exchanges are engaged in segregated management (segregated management refers to a system in which the assets of the exchange and the assets of investors are managed separately, meaning investor assets will be protected even if the exchange goes bankrupt)

    (19) Few investment trusts, ETFs, etc., that invest in crypto assets makes investing in them inconvenient

    (20) Other

    (21) No particular impressions

Q24. Do you think the activities of your country's central bank help stabilize prices and the financial system?

    (1) I think so

    (2) If I had to choose, I'd say yes

    (3) If I had to choose, I'd say no

    (4) I don't think so

Q25. You will not be evaluated on whether you get this question right or wrong. Respond with what you personally think is correct. Suppose you deposited $10,000 into a savings account with an annual interest rate of 2%. How much do you think the balance for this account will be a year later, if there are no other deposits to or withdrawals from this account? Ignore any taxes on the interest.Select the response that is closest to your answer.

    (1) $10,000

    (2) $10,200

    (3) $12,000

    (4) I don't know

Q26. How much do you think the balance for the account will be five years later? Ignore any taxes on the interest.

    (1) More than $11,000

    (2) $11,000

    (3) Less than $11,000

    (4) I don't know

Q27. Which of the two options below do you think would have the higher actual yield?

    (1) 2% interest rate on deposits, with a 0% future inflation rate

    (2) 5% interest rate on deposits, with a 4% future inflation rate

    (3) They would have the same actual yield

    (4) I don't know

Q28. Please indicate how you feel about the following statement. "Buying shares of one company is usually a safer investment than buying mutual funds."

(1) I think this is correct

(2) I think this is incorrect

(3) I don't know

Q29. Suppose there is an investment product that offers a 50% chance of yielding a profit of $1,200, and a 50% chance of yielding a profit of $800. How much would you be willing to pay for this investment product? Select the maximum amount you would be willing to pay.

(1) Buy even if it costs more than $1,000

(2) $1,000

(3) $980

(4) $960

(5) $940

(6) $920

(7) $900

(8) $880

(9) $860

(10) $840

(11) $820

(12) $800

(13) I wouldn't buy it

Q30. Suppose there is an investment product that offers a 10% probability of yielding a profit of $10,000, and a 90% chance of yielding a profit of $1,000. How much would you be willing to pay for this investment product? Select the maximum amount you would be willing to pay.

(1) Buy even if it costs more than $2,500

(2) $2,500

(3) $2,000

(4) $1,900

(5) $1,800

(6) $1,700

(7) $1,600

(8) $1,500

(9) $1,400

(10) $1,300

(11) $1,200

(12) $1,100

(13) $1,000

(14) I wouldn't buy it

Q31. Imagine you were blessed with the opportunity to either receive $1,000 immediately, or receive more than $1,000 a year from now. How much money would you have to receive to want to choose the option to receive the money a year later?

(1) $1,005

(2) $1,010

(3) $1,020

(4) $1,030

(5) $1,040

(6) $1,050

(7) $1,060

(8) $1,070

(9) $1,080

(10) $1,090

(11) $1,100

(12) $1,150

(13) $1,200

(14) $1,300

(15) $1,500 or more

(16) I would choose to receive the money immediately, regardless of how much I could have received a year later

Q32. What do you think about the following statement? Please choose the one that best reflects your feelings. "Men should have paid jobs to keep their families financially stable."

(1) I think so

(2) If I had to choose, I'd say yes

(3) If I had to choose, I'd say no

(4) I don't think so

(5) I prefer not to answer

Q33. What do you think about the following statement? Please choose the one that best reflects your feelings. "Men are better suited to be leaders of organizations."

   (1) I think so

   (2) If I had to choose, I'd say yes

   (3) If I had to choose, I'd say no

   (4) I don't think so

   (5) I prefer not to answer

Q34 We would like to ask you about U.S.'s economic growth rate over the next year. U.S.'s economic growth rate over the next year is the percentage change in the total value of new goods and services produced in the country over the next year compared to the total value over the past year. A positive growth rate means that the economy has grown, while a negative growth rate means that the economy has contracted.

What do you expect U.S.'s economic growth rate to be over the next year? Since there are many possibilities for the future, the economy may grow (with a positive growth rate) or shrink (with a negative growth rate).

What level of economic growth would you consider to be "high"?
What level of economic growth would you consider to be "medium"?
What level of economic growth would you consider to be "low"?

If you think the economy will grow, slide the bar to a positive number, and if you think the economy will shrink, slide the bar to a negative number. If you think the economy will remain the same (neither growing nor shrinking), leave the bar at "0".

*Drag the semi-transparent button and slide it to the position you think applies to the extent of your feelings. The first state with the button in the middle will be "the economy remains the same (neither growing nor shrinking)".

   (1) I feel that the economic growth rate is "high" if: Economic growth rate $-50\% \sim +50\%$

   (2) I feel that the economic growth rate is "medium" if: Economic growth rate $-50\% \sim +50\%$

   (3) I feel that the economic growth rate is "low" if: Economic growth rate $-50\% \sim +50\%$

Q34.1. For each of the three cases you answered in the previous question, how likely do you think they are to occur in the future? Please enter so that the total of the three cases is "100%".

   (1) The likelihood of achieving a "high" rate of economic growth

   (2) The likelihood of achieving a "medium" rate of economic growth

(3) The likelihood of achieving a "low" rate of economic growth

Q35. How do you expect prices to change over the next year? Since there are many possibilities for the future, prices may rise or fall.

What level of price increase would you consider "high"?
What level of price increase would you consider "medium"?
What level of price increase would you consider to be "low"?

*If you think prices will rise, slide the bar to a positive number, and if you think prices will fall, slide the bar to a negative number. If you believe that prices will not change (neither rise nor fall), leave the value at "0".

*Drag the semi-transparent button and slide it to the position you think applies to the extent of your feelings. The first state with the button in the middle will be "the price remains the same (neither growing nor shrinking)".

(1) I feel that the rate of price increases is "high" if: Price increase rate $-50\% \sim +50\%$

(2) I feel that the rate of price increases is "medium" if: Price increase rate $-50\% \sim +50\%$

(3) I feel that the rate of price increases is "low" if: Price increase rate $-50\% \sim +50\%$

Q35.1. For each of the three cases you answered in the previous question, how likely do you think they are to occur in the future? Please enter so that the total of the three cases is "100%".

(1) The likelihood of a "high" price increase rate

(2) The likelihood of a "medium" price increase rate

(3) The likelihood of a "low" price increase rate

Q36. We would like to know your outlook on the exchange rate of your country's currency. The exchange rate of a currency is the ratio of one currency to another. A rising currency means an increase in value against another currency, while a falling currency means a decrease in value against another currency. For example, if the exchange rate goes from 0.9 euro per dollar to 1.0 euro per dollar, the value of the dollar against the euro rises; if the exchange rate goes from 0.9 euro per dollar to 0.8 euro per dollar, the value of the dollar against the euro falls. For this question, you are asked to consider the probability that something will happen in the future. On a scale of 0 to 100, 0 means not likely at all and 100 means it will definitely happen.

For example, the following figures.

- 2-5% : "Almost unlikely"

- Around 18%: "Not very likely"
- 47-52%: "Very likely"
- Around 83%: "Extremely likely"
- 95-98%: "Almost certain"

How do you think the exchange rate of the U.S. dollar will change in the next year? There are various possibilities in the future, and the U.S. dollar exchange rate may rise or fall. Please estimate the likelihood of each of the following 7 cases occurring. Make sure the total for all 7 cases adds up to 100%.

(1) Rise of 25% or more

(2) Rise of more than 15% but less than 25%

(3) Rise of more than 5% but less than 15%

(4) Fall of less than 5% or rise of less than 5%

(5) Fall of more than 5% but less than 15%

(6) Fall of more than 15% but less than 25%

(7) Fall of more than 25%

Q37. Please indicate your ideal percentage of financial assets held in one year.
*Please exclude life insurance and physical assets such as real estate and cars.
*Please answer the percentage of each so that the total is 100%.

(1) Cash and deposits

(2) Domestic stocks and mutual funds managed primarily in domestic stocks (all denominated in the local currency)

(3) Domestic bonds

(4) Foreign stocks and mutual funds managed primarily in foreign stocks (all denominated in foreign currencies)

(5) Foreign bonds

(6) Gold

(7) New digital financial instruments (crypto-assets, stable coins, STs, NFTs)

(8) Other

Q37.1. Dow Jones Industrial Average has risen +57% over the past 5 years and +141% over the past 10 years. Please enter the number that best describes your ideal percentage of your personal financial holdings at the end of next year. Make sure that the numbers you enter add up to 100.
*Please exclude life insurance and physical assets such as real estate and cars.
*Please answer the percentage of each so that the total is 100%.

(1) Cash and deposits

(2) Domestic stocks and mutual funds managed primarily in domestic stocks (all denominated in the local currency)

(3) Domestic bonds

(4) Foreign stocks and mutual funds managed primarily in foreign stocks (all denominated in foreign currencies)

(5) Foreign bonds

(6) Gold

(7) New digital financial instruments (crypto-assets, stable coins, STs, NFTs)

(8) Other

Q37.2. The price of Bitcoin has increased by more than 6 times in the past five years and by more than 100 times in the past decade. Please enter the number that best describes your ideal percentage of your personal financial holdings at the end of next year. Make sure that the numbers entered total 100.
*Please exclude life insurance and physical assets such as real estate and cars.
*Please answer the percentage of each so that the total is 100%.

(1) Cash and deposits

(2) Domestic stocks and mutual funds managed primarily in domestic stocks (all denominated in the local currency)

(3) Domestic bonds

(4) Foreign stocks and mutual funds managed primarily in foreign stocks (all denominated in foreign currencies)

(5) Foreign bonds

(6) Gold

(7) New digital financial instruments (crypto-assets, stable coins, STs, NFTs)

(8) Other

Q37.3. The International Monetary Fund (IMF) expects prices in the U.S. to rise 2.0% next year compared to last year. Please enter the number that best describes your ideal percentage of your personal financial holdings at the end of next year. Make sure that the numbers entered total 100.
*Please exclude life insurance and physical assets such as real estate and cars.
*Please answer the percentage of each so that the total is 100%.

(1) Cash and deposits

(2) Domestic stocks and mutual funds managed primarily in domestic stocks (all denominated in the local currency)

(3) Domestic bonds

(4) Foreign stocks and mutual funds managed primarily in foreign stocks (all denominated in foreign currencies)

(5) Foreign bonds

(6) Gold

(7) New digital financial instruments (crypto-assets, stable coins, STs, NFTs)

(8) Other

Q38. Please answer the following questions regarding your investment in real estate for investment purposes (real estate acquired by individuals for investment purposes).

(1) I have never invested and do not intend to invest in the next year

(2) I have never invested, but intend to do so in the next year

(3) I currently have investments, but do not intend to make additional investments in the next year

(4) I currently have investments, and intend to invest further over the next year

(5) I currently have investments, and intend to sell more over the next year

Q38.1. You have just answered "cite answer to Q38" regarding real estate for investment purposes, but could you please provide specific amounts for the following?

(1) Amount currently held: (   ) dollars

(2) Estimated investment amount for the next year: (   ) dollars

(3)   Estimated amount of sales for the next year: (   ) dollars

Appendix B   Tables

Table B.1: Summary statistics

	Japan	United States	Germany	China	All countries
Age	51.45	47.55	49.41	46.69	49.51
Male (%)	49.67	49.45	49.90	50.62	49.85
High education (%)	42.74	35.88	21.00	59.03	40.50
Married (%)	59.10	41.53	48.08	84.60	58.54
$D^{Stock\ Market\ Participation}$ (%)	28.08	30.02	24.11	49.44	30.86
Stocks in portfolio (%)	9.51	9.92	8.61	10.37	9.58
Observations	10,000	4,000	4,000	4,000	22,000

Notes: High education is a dummy variable that takes 1 if the respondent has a university degree or higher and otherwise 0. Stocks in portfolio represents the proportion of domestic and foreign stocks in the portfolio. $D^{Stock\ Market\ Participation}$ is a dummy variable that takes the value of 1 if domestic or foreign stocks are currently held, and 0 if they have never been held.

Table B.2: Determinants of "Don't Know" Response (Question 1: Simple Interest)

Variables	(1) Japan	(2) United States	(3) Germany DK response	(4) China	(5) All countries
Male	−0.014***	−0.024**	−0.042***	0.024**	−0.014***
	(0.009)	(0.010)	(0.010)	(0.011)	(0.005)
Higher education	−0.077***	−0.033***	−0.024***	0.064***	−0.035***
	(0.008)	(0.010)	(0.010)	(0.013)	(0.006)
Age class	−0.030***	−0.026***	−0.017***	−0.004	−0.027***
	(0.003)	(0.004)	(0.004)	(0.005)	(0.002)
Married	0.006	−0.012	−0.001	0.002	0.004
	(0.010)	(0.011)	(0.010)	(0.020)	(0.006)
Income class	−0.004	−0.007***	−0.006***	−0.010***	−0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Time preference (std.)	−0.045***	−0.020***	−0.012**	−0.061***	−0.039***
	(0.004)	(0.005)	(0.005)	(0.006)	(0.003)
Risk tolerance (std.)	−0.083***	−0.054***	−0.050***	−0.008	−0.061***
	(0.003)	(0.005)	(0.005)	(0.007)	(0.003)
Country dummy: U.S.					−0.047***
					(0.007)
Country dummy: Germany					−0.059***
					(0.007)
Country dummy: China					0.080***
					(0.008)
Constant	0.354***	0.281***	0.220***	0.222***	0.314***
	(0.014)	(0.016)	(0.016)	(0.025)	(0.009)
Observations	7,645	3,954	3,845	3,991	19,435
R^2	0.123	0.081	0.057	0.051	0.084

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Variable definitions: The dependent variable is a dummy variable for **DK response**, which takes the value of 1 if the respondent chose the DK option in Question 1, and 0 otherwise. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices. Column (5) pools all observations with country-specific intercepts (Japan as the baseline), while Columns (1)–(4) provide country-specific estimates.

Table B.3: Determinants of "Don't Know" Response (Question 2: Compound Interest)

Variables	(1) Japan	(2) United States	(3) Germany DK response	(4) China	(5) All countries
Male	−0.021**	−0.033***	−0.038***	0.011	−0.021***
	(0.010)	(0.011)	(0.010)	(0.008)	(0.005)
Higher education	−0.074***	−0.043***	−0.010	0.011	−0.046***
	(0.009)	(0.011)	(0.011)	(0.009)	(0.005)
Age class	−0.029***	−0.018***	−0.016***	0.001	−0.021***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.002)
Married	0.002	−0.011	−0.004	0.001	−0.001
	(0.010)	(0.011)	(0.010)	(0.014)	(0.006)
Income class	−0.005**	−0.007***	−0.006***	−0.005***	−0.005***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Time preference (std.)	−0.044***	−0.027***	−0.015***	−0.035***	−0.034***
	(0.004)	(0.005)	(0.005)	(0.005)	(0.002)
Risk tolerance (std.)	−0.070***	−0.055***	−0.054***	−0.033***	−0.061***
	(0.004)	(0.006)	(0.005)	(0.005)	(0.003)
Country dummy: U.S.					−0.052***
					(0.007)
Country dummy: Germany					−0.069***
					(0.007)
Country dummy: China					−0.022***
					(0.007)
Constant	0.369***	0.271***	0.222***	0.106***	0.310***
	(0.014)	(0.016)	(0.017)	(0.018)	(0.009)
Observations	7,645	3,954	3,845	3,991	19,435
R^2	0.121	0.080	0.056	0.053	0.103

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Variable definitions: The dependent variable is a dummy variable for **DK response**, which takes the value of 1 if the respondent chose the DK option in Question 2, and 0 otherwise. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices. Column (5) pools all observations with country-specific intercepts (Japan as the baseline), while Columns (1)–(4) provide country-specific estimates.

Table B.4: Determinants of "Don't Know" Response (Question 3: Inflation)

Variables	(1) Japan	(2) United States	(3) Germany	(4) China	(5) All countries
			DK response		
Male	−0.152***	−0.086***	−0.099***	−0.008	−0.102***
	(0.012)	(0.014)	(0.014)	(0.009)	(0.006)
Higher education	−0.131***	−0.058***	−0.049***	0.011	−0.082***
	(0.011)	(0.015)	(0.016)	(0.011)	(0.007)
Age class	−0.024***	−0.004	−0.024***	0.009**	−0.019***
	(0.004)	(0.004)	(0.005)	(0.004)	(0.002)
Married	0.013	0.001	−0.008	−0.015	0.008
	(0.012)	(0.015)	(0.014)	(0.016)	(0.007)
Income class	−0.062***	−0.008***	−0.012***	−0.008***	−0.009***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Time preference (std.)	−0.047***	−0.037***	−0.028***	−0.040***	−0.043***
	(0.005)	(0.007)	(0.007)	(0.005)	(0.003)
Risk tolerance (std.)	−0.136***	−0.106***	−0.071***	−0.025***	−0.106***
	(0.005)	(0.007)	(0.007)	(0.006)	(0.004)
Country dummy: U.S.					−0.182***
					(0.009)
Country dummy: Germany					−0.215***
					(0.009)
Country dummy: China					−0.248***
					(0.009)
Constant	0.754***	0.420***	0.473***	0.154***	0.661***
	(0.015)	(0.020)	(0.023)	(0.020)	(0.010)
Observations	7,645	3,954	3,845	3,991	19,435
R^2	0.203	0.124	0.081	0.057	0.225

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Variable definitions: The dependent variable is a dummy variable for **DK response**, which takes the value of 1 if the respondent chose the DK option in Question 3, and 0 otherwise. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices. Column (5) pools all observations with country-specific intercepts (Japan as the baseline), while Columns (1)–(4) provide country-specific estimates.

Table B.5: Determinants of "Don't Know" Response (Question 4: Risk Diversification)

Variables	(1) Japan	(2) United States	(3) Germany DK response	(4) China	(5) All countries
Male	−0.065***	−0.087***	−0.064***	0.001	−0.063***
	(0.012)	(0.014)	(0.013)	(0.010)	(0.006)
Higher education	−0.117***	−0.067***	−0.052***	−0.008	−0.078***
	(0.011)	(0.014)	(0.015)	(0.012)	(0.007)
Age class	−0.011***	−0.004	−0.004	−0.005	−0.011***
	(0.002)	(0.004)	(0.004)	(0.004)	(0.002)
Married	0.012	−0.027**	0.018	−0.037**	−0.004
	(0.012)	(0.014)	(0.013)	(0.018)	(0.007)
Income class	−0.010***	−0.010***	−0.010***	−0.005***	−0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Time preference (std.)	−0.060***	−0.054***	−0.021***	−0.042***	−0.050***
	(0.005)	(0.007)	(0.007)	(0.005)	(0.003)
Risk tolerance (std.)	−0.140***	−0.097***	−0.097***	−0.058***	−0.118***
	(0.005)	(0.007)	(0.007)	(0.006)	(0.004)
Country dummy: U.S.					−0.187***
					(0.009)
Country dummy: Germany					−0.235***
					(0.009)
Country dummy: China					−0.203***
					(0.009)
Constant	0.669***	0.428***	0.336***	0.217***	0.601***
	(0.016)	(0.019)	(0.020)	(0.021)	(0.010)
Observations	7,645	3,954	3,845	3,991	19,435
R^2	0.183	0.146	0.099	0.072	0.224

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Variable definitions: The dependent variable is a dummy variable for **DK response**, which takes the value of 1 if the respondent chose the DK option in Question 4, and 0 otherwise. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices. Column (5) pools all observations with country-specific intercepts (Japan as the baseline), while Columns (1)–(4) provide country-specific estimates.

Table B.6: Determinants of Stock Market Participation: Japan

Variables	(1)	(2)	(3)	(4)
		$D^{Stock\ Market\ Participation}$		
Financial literacy (DK = −1, std.)	0.085***			
	(0.006)			
Financial literacy (DK = 0, std.)		0.085***	0.078***	0.065***
		(0.006)	(0.010)	(0.009)
Number of DK responses			−0.009	−0.013
			(0.009)	(0.008)
Control Variables				
Male	−0.005	−0.007	−0.007	−0.010
	(0.012)	(0.012)	(0.012)	(0.011)
Higher education	0.074***	0.071***	0.071***	0.067***
	(0.012)	(0.012)	(0.012)	(0.011)
Age class	0.026***	0.025***	0.025***	0.016***
	(0.004)	(0.004)	(0.004)	(0.003)
Married	0.024**	0.024**	0.024**	0.021**
	(0.011)	(0.011)	(0.011)	(0.011)
Time preference (std.)	0.006	0.008	0.007	0.003
	(0.006)	(0.006)	(0.006)	(0.005)
Risk tolerance (std.)	0.096***	0.098***	0.097***	0.090***
	(0.007)	(0.006)	(0.006)	(0.006)
Income class	0.023***	0.023***	0.023***	0.022***
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.046***	0.053***	0.056***	0.081***
	(0.016)	(0.016)	(0.016)	(0.015)
Observations	6,551	6,551	6,551	7,645

Notes: Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$. The dependent variable ($D^{Stock\ Market\ Participation}$) is a dummy equal to one if the household currently owns domestic or foreign stocks.

Variable definitions: **Financial literacy** is a standardized index constructed from basic financial literacy questions. In Column (1), DK responses are coded as −1; in Columns (2)–(4), DK responses are coded as 0. **Number of DK responses** is the count of DK answers across the literacy questions. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices.

Table B.7: Determinants of Stock Market Participation: United States

Variables	(1)	(2)	(3)	(4)
		$D^{Stock\ Market\ Participation}$		
Financial literacy (DK = -1, std.)	0.083***			
	(0.007)			
Financial literacy (DK = 0, std.)		0.084***	0.068***	0.043***
		(0.008)	(0.010)	(0.009)
Number of DK responses			-0.031***	-0.037***
			(0.010)	(0.009)
Control Variables				
Male	0.106***	0.106***	0.105***	0.082***
	(0.016)	(0.016)	(0.016)	(0.014)
Higher education	0.157***	0.154***	0.153***	0.119***
	(0.020)	(0.020)	(0.020)	(0.016)
Age class	0.023***	0.020***	0.021***	0.016***
	(0.005)	(0.005)	(0.005)	(0.004)
Married	-0.001	-0.001	-0.001	-0.004
	(0.017)	(0.017)	(0.017)	(0.014)
Time preference (std.)	0.028***	0.035***	0.032***	0.013*
	(0.009)	(0.009)	(0.009)	(0.007)
Risk tolerance (std.)	0.041***	0.049***	0.044***	0.035***
	(0.009)	(0.009)	(0.009)	(0.008)
Income class	0.033***	0.033***	0.033***	0.029***
	(0.003)	(0.003)	(0.003)	(0.002)
Constant	-0.069***	-0.065***	-0.068***	-0.059***
	(0.022)	(0.022)	(0.022)	(0.019)
Observations	2,690	2,690	2,690	3,954

Notes: Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$. The dependent variable ($D^{Stock\ Market\ Participation}$) is a dummy equal to one if the household currently owns domestic or foreign stocks.

Variable definitions: **Financial literacy** is a standardized index constructed from basic financial literacy questions. In Column (1), DK responses are coded as -1; in Columns (2)–(4), DK responses are coded as 0. **Number of DK responses** is the count of DK answers across the literacy questions. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices.

Table B.8: Determinants of Stock Market Participation: Germany

Variables	(1)	(2)	(3)	(4)
		$D^{Stock\ Market\ Participation}$		
Financial literacy (DK = −1, std.)	0.077***			
	(0.007)			
Financial literacy (DK = 0, std.)		0.075***	0.056***	0.057***
		(0.008)	(0.011)	(0.009)
Number of DK responses			−0.038***	−0.029***
			(0.011)	(0.010)
Control Variables				
Male	0.087***	0.087***	0.086***	0.068***
	(0.015)	(0.015)	(0.015)	(0.013)
Higher education	0.133***	0.131***	0.132***	0.114***
	(0.023)	(0.023)	(0.023)	(0.019)
Age class	0.008*	0.006	0.007	0.007
	(0.005)	(0.005)	(0.005)	(0.004)
Married	−0.006	−0.003	−0.005	0.025*
	(0.016)	(0.017)	(0.016)	(0.014)
Time preference (std.)	0.012	0.016**	0.013*	0.012*
	(0.008)	(0.008)	(0.008)	(0.007)
Risk tolerance (std.)	0.061***	0.069***	0.063***	0.042***
	(0.009)	(0.008)	(0.009)	(0.007)
Income class	0.042***	0.042***	0.042***	0.028***
	(0.003)	(0.003)	(0.003)	(0.002)
Constant	−0.067***	−0.066***	−0.075***	−0.036*
	(0.024)	(0.024)	(0.024)	(0.021)
Observations	2,617	2,617	2,617	3,845

Notes: Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$. The dependent variable ($D^{Stock\ Market\ Participation}$) is a dummy equal to one if the household currently owns domestic or foreign stocks.

Variable definitions: **Financial literacy** is a standardized index constructed from basic financial literacy questions. In Column (1), DK responses are coded as −1; in Columns (2)–(4), DK responses are coded as 0. **Number of DK responses** is the count of DK answers across the literacy questions. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices.

Table B.9: Determinants of Stock Market Participation: China

Variables	(1)	(2)	(3)	(4)
		$D^{Stock\ Market\ Participation}$		
Financial literacy (DK = -1, std.)	0.058***			
	(0.008)			
Financial literacy (DK = 0, std.)		0.056***	0.046***	0.040***
		(0.009)	(0.010)	(0.008)
Number of DK responses			-0.037***	-0.014
			(0.014)	(0.012)
Control Variables				
Male	0.012	0.008	0.011	-0.003
	(0.019)	(0.019)	(0.019)	(0.014)
Higher education	0.014	0.009	0.013	0.035**
	(0.023)	(0.023)	(0.023)	(0.017)
Age class	-0.018**	-0.017**	-0.017**	-0.016***
	(0.008)	(0.008)	(0.008)	(0.006)
Married	0.209***	0.209***	0.209***	0.133***
	(0.033)	(0.033)	(0.033)	(0.024)
Time preference (std.)	-0.013	-0.005	-0.011	-0.005
	(0.009)	(0.009)	(0.010)	(0.008)
Risk tolerance (std.)	0.083***	0.086***	0.083***	0.053***
	(0.009)	(0.009)	(0.009)	(0.008)
Income class	0.021***	0.022***	0.021***	0.022***
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.160***	0.147***	0.139***	0.070***
	(0.036)	(0.036)	(0.036)	(0.026)
Observations	2,397	2,397	2,397	3,991

Notes: Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$. The dependent variable ($D^{Stock\ Market\ Participation}$) is a dummy equal to one if the household currently owns domestic or foreign stocks.

Variable definitions: **Financial literacy** is a standardized index constructed from basic financial literacy questions. In Column (1), DK responses are coded as -1; in Columns (2)–(4), DK responses are coded as 0. **Number of DK responses** is the count of DK answers across the literacy questions. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices.

Table B.10: Financial Literacy and Portfolio Stock Shares (Intensive Margin): Japan

Variables	(1) $StockShare$ DK=−1	(2) $StockShare$ DK=0	(3) $StockShare$ DK=0
Financial literacy (DK = −1, std.)	3.677***		
	(0.275)		
Financial literacy (DK = 0, std.)		3.877***	4.342***
		(0.278)	(0.434)
Number of DK responses			0.579
			(0.363)
Control Variables			
Male	1.111**	0.955*	0.951*
	(0.505)	(0.504)	(0.504)
Higher education	2.681***	2.459***	2.459***
	(0.518)	(0.519)	(0.519)
Age class	0.696***	0.628***	0.624***
	(0.153)	(0.153)	(0.153)
Married	0.940*	0.925*	0.917*
	(0.501)	(0.499)	(0.500)
Time preference (std.)	−0.456*	−0.396*	−0.357
	(0.241)	(0.238)	(0.240)
Risk tolerance (std.)	3.924***	3.944***	4.004***
	(0.302)	(0.297)	(0.302)
Income class	0.724***	0.720***	0.722***
	(0.105)	(0.105)	(0.105)
Constant	1.312*	1.807**	1.632**
	(0.716)	(0.719)	(0.730)
Observations	7,645	7,645	7,645
R^2	0.137	0.141	0.141

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable ($StockShare$) is the share of stocks in the household portfolio.
Variable definitions: **Financial literacy** is a standardized index constructed from basic financial literacy questions. In Column (1), DK responses are coded as −1; in Columns (2) and (3), DK responses are coded as 0. **Number of DK responses** is the count of DK answers across the literacy questions. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices.

Table B.11: Financial Literacy and Portfolio Stock Shares (Intensive Margin): United States

Variables	(1) $StockShare$ DK=−1	(2) $StockShare$ DK=0	(3) $StockShare$ DK=0
Financial literacy (DK = −1, std.)	3.161***		
	(0.314)		
Financial literacy (DK = 0, std.)		3.593***	3.702***
		(0.363)	(0.453)
Number of DK responses			0.238
			(0.390)
Control Variables			
Male	1.869***	1.826***	1.842***
	(0.640)	(0.638)	(0.639)
Higher education	4.593***	4.406***	4.411***
	(0.780)	(0.776)	(0.777)
Age class	2.050***	1.828***	1.819***
	(0.210)	(0.210)	(0.212)
Married	−0.856	−0.782	−0.773
	(0.692)	(0.689)	(0.689)
Time preference (std.)	0.181	0.508	0.539
	(0.342)	(0.342)	(0.350)
Risk tolerance (std.)	1.772***	2.012***	2.052***
	(0.399)	(0.391)	(0.402)
Income class	1.075***	1.083***	1.086***
	(0.105)	(0.104)	(0.105)
Constant	−6.322***	−5.582***	−5.539***
	(0.856)	(0.860)	(0.861)
Observations	3,954	3,954	3,954
R^2	0.163	0.169	0.169

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable ($StockShare$) is the share of stocks in the household portfolio.
Variable definitions: **Financial literacy** is a standardized index constructed from basic financial literacy questions. In Column (1), DK responses are coded as −1; in Columns (2) and (3), DK responses are coded as 0. **Number of DK responses** is the count of DK answers across the literacy questions. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices.

Table B.12: Financial Literacy and Portfolio Stock Shares (Intensive Margin): Germany

Variables	(1) $StockShare$ DK=−1	(2) $StockShare$ DK=0	(3) $StockShare$ DK=0
Financial literacy (DK = −1, std.)	3.592***		
	(0.289)		
Financial literacy (DK = 0, std.)		3.918***	3.985***
		(0.332)	(0.431)
Number of DK responses			0.153
			(0.407)
Control Variables			
Male	4.138***	4.036***	4.044***
	(0.585)	(0.579)	(0.582)
Higher education	3.058***	2.784***	2.780***
	(0.886)	(0.888)	(0.889)
Age class	1.206***	1.023***	1.019***
	(0.198)	(0.199)	(0.201)
Married	0.053	0.130	0.133
	(0.615)	(0.613)	(0.613)
Time preference (std.)	−0.152	0.107	0.120
	(0.307)	(0.306)	(0.308)
Risk tolerance (std.)	1.652***	1.904***	1.926***
	(0.356)	(0.349)	(0.362)
Income class	1.077***	1.107***	1.109***
	(0.111)	(0.110)	(0.110)
Constant	−4.787***	−4.239***	−4.201***
	(0.906)	(0.907)	(0.912)
Observations	3,845	3,845	3,845
R^2	0.139	0.145	0.145

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable ($StockShare$) is the share of stocks in the household portfolio.
Variable definitions: **Financial literacy** is a standardized index constructed from basic financial literacy questions. In Column (1), DK responses are coded as −1; in Columns (2) and (3), DK responses are coded as 0. **Number of DK responses** is the count of DK answers across the literacy questions. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices.

Table B.13: Financial Literacy and Portfolio Stock Shares (Intensive Margin): China

Variables	(1) $StockShare$ DK=−1	(2) $StockShare$ DK=0	(3) $StockShare$ DK=0
Financial literacy (DK = −1, std.)	1.144***		
	(0.263)		
Financial literacy (DK = 0, std.)		1.511***	1.670***
		(0.269)	(0.287)
Number of DK responses			0.707*
			(0.408)
Control Variables			
Male	0.669	0.565	0.532
	(0.491)	(0.490)	(0.491)
Higher education	1.569***	1.440**	1.384**
	(0.584)	(0.582)	(0.581)
Age class	−0.282	−0.231	−0.221
	(0.217)	(0.217)	(0.217)
Married	4.064***	4.006***	4.012***
	(0.823)	(0.823)	(0.823)
Time preference (std.)	−0.482*	−0.341	−0.245
	(0.255)	(0.254)	(0.260)
Risk tolerance (std.)	1.852***	1.844***	1.889***
	(0.272)	(0.269)	(0.272)
Income class	0.610***	0.636***	0.652***
	(0.054)	(0.053)	(0.055)
Constant	0.447	0.200	0.385
	(0.845)	(0.842)	(0.841)
Observations	3,991	3,991	3,991
R^2	0.092	0.095	0.096

Notes: Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable ($StockShare$) is the share of stocks in the household portfolio.
Variable definitions: **Financial literacy** is a standardized index constructed from basic financial literacy questions. In Column (1), DK responses are coded as −1; in Columns (2) and (3), DK responses are coded as 0. **Number of DK responses** is the count of DK answers across the literacy questions. **Male** and **Married** are dummy variables. **Higher education** is a dummy for holding a university degree or higher. **Age class** is treated as an ordinal variable, categorized into 10-year intervals with the 20s set as the reference group. **Income class** is also an ordinal variable; however, observations where respondents chose "I don't know" or "Prefer not to say" for their income are excluded from the analysis. **Time preference** and **Risk tolerance** are standardized indices.