Determinants of Financial Well-Being: The Role of Demographics, Psychosocial Factors, and Financial Literacy*

Psychological Distress Has the Strongest Negative Impact on Financial Outcomes

Xinxiang Gao

December 3, 2024

This paper analyzes the determinants of financial well-being using the Consumer Financial Protection Bureau's National Financial Well-Being Survey. Using a comprehensive regression analysis of over 6,000 respondents, we find that psychological distress has the strongest negative association with financial well-being, with highly stressed individuals scoring 16.5 points lower on the financial well-being scale. Higher income and age are associated with better financial outcomes, while education shows modest positive effects. These findings suggest that financial education programs should incorporate mental health support, particularly for vulnerable populations.

Table of contents

*Code and data are available at: https://github.com/xgao28/financial_well_being_analysis.

1 Introduction

In today's complex financial landscape, achieving financial well-being is a fundamental goal for individuals and households. While financial well-being encompasses the ability to meet current and future financial obligations, its determinants span multiple domains including financial literacy, demographic characteristics, and psychosocial factors. Understanding these relationships is crucial for developing effective policies and interventions to improve financial outcomes across diverse population segments.

To measure financial well-being systematically, the Consumer Financial Protection Bureau (CFPB) developed the Financial Well-Being Scale (FWBscore), a 10-item questionnaire assessing four critical elements: control over day-to-day finances, capacity to absorb financial shocks, progress toward financial goals, and ability to make life-improving choices. Alongside this measure, researchers have developed various instruments to assess financial knowledge and skills. The Lusardi and Mitchell financial knowledge scale (LMscore) by Lusardi and Mitchell (2008) provides a foundational three-item measure of basic financial concepts, while the Knoll and Houts financial knowledge scale (KHscore) by Knoll and Houts (2012) offers a more comprehensive 10-item assessment using item response theory. Additionally, the CFPB Financial Skill Scale (FSscore) by Bureau (2015) evaluates practical financial competencies through an IRT-based approach.

While previous research has examined financial literacy's role in economic outcomes, less attention has been paid to how demographic and psychosocial factors interact with financial knowledge to influence overall financial well-being. This gap is particularly significant given evidence suggesting that financial knowledge alone may not fully explain variations in financial outcomes across different population segments.

Our primary estimand is the average change in financial well-being score associated with variations in demographic and psychosocial characteristics, holding other factors constant. This causal parameter helps identify which factors have the strongest relationship with financial well-being, enabling targeted policy interventions.

Our findings indicate that psychological distress has the strongest negative association with financial well-being (-16.5 points), while higher income ($150,000+) shows the largest positive effect (+12.6 points). Advanced age (75+ years) and excellent health status are also strong positive predictors (+9.4 and +7.7 points respectively). Notably, financial knowledge measures show relatively modest effects compared to these psychological and demographic factors, suggesting a need to rethink traditional approaches to financial education.

The rest of the paper is organized as follows: Section 2 describes our data and methodology. Section 3 presents our modeling approach. Section 4 presents our empirical results. Section 5 discusses the implications of our findings and concludes. Section 6.1 provides additional details on the survey methodology and Section 6.2 provides acknowledgements.

2 Data

2.1 Overview of the Dataset

The dataset for this study is the National Financial Well-Being Survey Public Use File (PUF) collected by Consumer Financial Protection Bureau (2017). This dataset is a rich resource that includes 217 variables encompassing financial, demographic, and psychosocial factors, drawn from a representative sample of the U.S. population. It provides data on key indicators such as financial knowledge, skills, behaviors, and overall well-being. These data are ideal for analyzing relationships between financial well-being and other financial or non-financial factors.

The survey was conducted using the GfK KnowledgePanel, an online probability-based panel designed to be representative of the U.S. adult population. The GfK KnowledgePanel is created by randomly selecting households and inviting them to participate in the panel. Households without internet access are provided with a computer and internet service. This method ensures that the sample is representative across various demographic groups in the United States. Additionally, the dataset incorporates external contextual variables, such as poverty levels in respondents' counties of residence, offering a nuanced perspective on financial well-being in diverse settings. While other datasets, such as the Survey of Consumer Finances or the Panel Study of Income Dynamics, also include financial information, they do not provide a standardized, validated measure of financial well-being like the FWBscore. This makes the CFPB dataset uniquely suited for this study.

2.2 Variables of Interest

2.2.1 Primary Outcome Variable

Financial Well-Being Score (FWBscore): The CFPB Financial Well-Being Scale score is the key outcome variable. It is a 10-item scale capturing four dimensions: control over day-to-day finances, capacity to absorb financial shocks, progress toward financial goals, and freedom of financial choices. Scores range from 0 to 100, with higher scores indicating greater financial well-being.

2.2.2 Independent Variables

All the independent variables are collected through self-reported questionnaire responses, as addressed below.

Age (agecat): Categorized into age groups (e.g., 18–24, 25–34, 35–44, 45–54, 55–64, 65+).

Education (PPEDUC): Highest degree attained, categorized from less than high school to graduate degrees.

Household Income (PPINCIMP): Categorized income levels (e.g., <$20,000, $20,000-$39,999, $40,000-$59,999, etc.).

Marital Status (PPMARIT): Categories include single, married, divorced, widowed, and other.

Health (HEALTH): Self-reported general health, rated on a 5-point Likert scale from 1 (Poor) to 5 (Excellent).

Stress (DISTRESS): Perceived level of stress, rated on a 5-point Likert scale from 1 (Not at all) to 5 (Extremely).

Note that this study employs categorical grouping for variables that could be continuous (such as age and income). This approach offers several advantages in survey research: it improves interpretability of results across demographic groups, facilitates statistical comparisons between meaningful socioeconomic segments, and aligns with typical policy and program planning categories. For sensitive variables like income and age, categorical grouping also helps protect respondent privacy while maintaining sufficient detail for analysis. Additionally, the use of standardized categories (in age, education, and income) enables easier comparison with other survey research and demographic studies.

2.2.3 Additional Variables

Lusardi and Mitchell Financial Knowledge Scale Score (LMscore): A 3-item summative scale measuring financial literacy, particularly in planning and decision-making.

Knoll and Houts Financial Knowledge Scale Score (KHscore): A 10-item IRT-based measure assessing broader financial knowledge across multiple dimensions.

Financial Skill Scale Score (FSscore): An IRT-based measure developed by the CFPB, capturing practical financial skills like budgeting, debt management, and saving.

2.3 Variable Summary

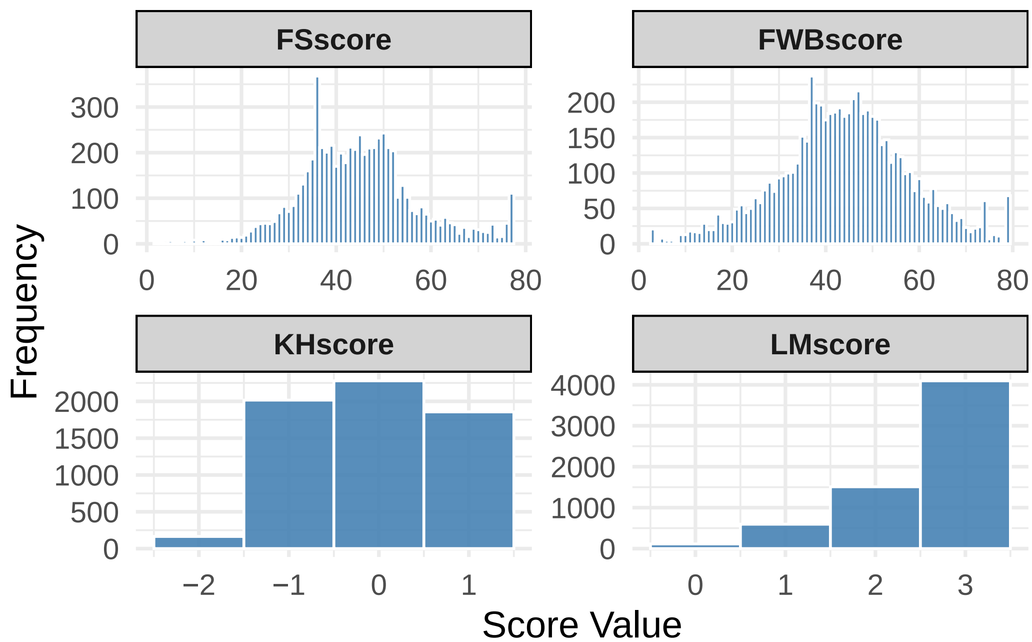

Figure 1: Histograms of Financial Scores

Figure 1 shows the distribution of four different financial scores (FSscore, FWBscore, KHscore, and LMscore) through histograms. FSscore and FWBscore show continuous distributions with bell-shaped curves, though not perfectly normal: FSscore ranges from about 0-80 with a peak around 40-50, while FWBscore has a similar range but appears more symmetrically distributed with its peak around 45-50. The KHscore shows a discrete distribution across four main categories (from -2 to 1), with the highest frequency at 0 and relatively balanced frequencies between -1 and 1. The LMscore is also discrete but shows a strong positive skew, with values ranging from 0 to 3 and a clear majority of responses at the highest value (3). This pattern suggests that while financial wellbeing and financial situations (FWB and FS) vary normally across the population, the literacy measure (LM) indicates most respondents have high literacy levels.

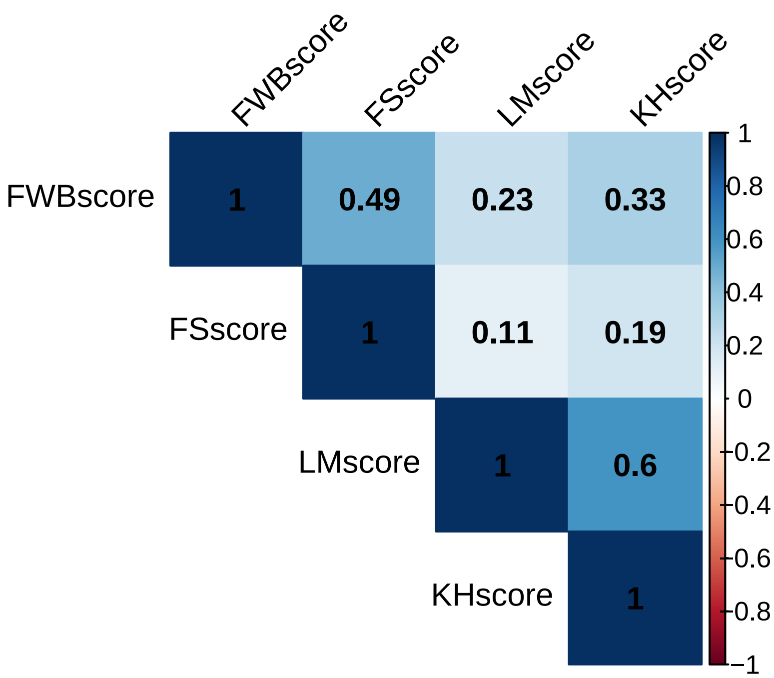

Figure 2: Correlation Plot of Continuous Financial Measures

The correlation matrix heatmap as Figure 2 shows the relationships between four different scores: FWBscore, FSscore, LMscore, and KHscore. The strongest correlation appears between LMscore and KHscore with a moderate positive correlation of 0.6, followed by a moderate correlation of 0.49 between FWBscore and FSscore. FWBscore shows weak to moderate positive correlations with LMscore (0.23) and KHscore (0.33), while FSscore exhibits the weakest correlations with LMscore (0.11) and KHscore (0.19). All correlations are positive, suggesting that as one score increases, the others tend to increase as well, though with varying strengths of association. The diagonal shows perfect correlations of 1.0, as expected when comparing a variable with itself.

Figure 3 provides a comprehensive demographic overview of the survey population. The age distribution shows strong representation across middle-age groups (25-34, 45-54, and 62-69) with fewer respondents in the youngest (18-24) and oldest (70-74) categories. Education levels

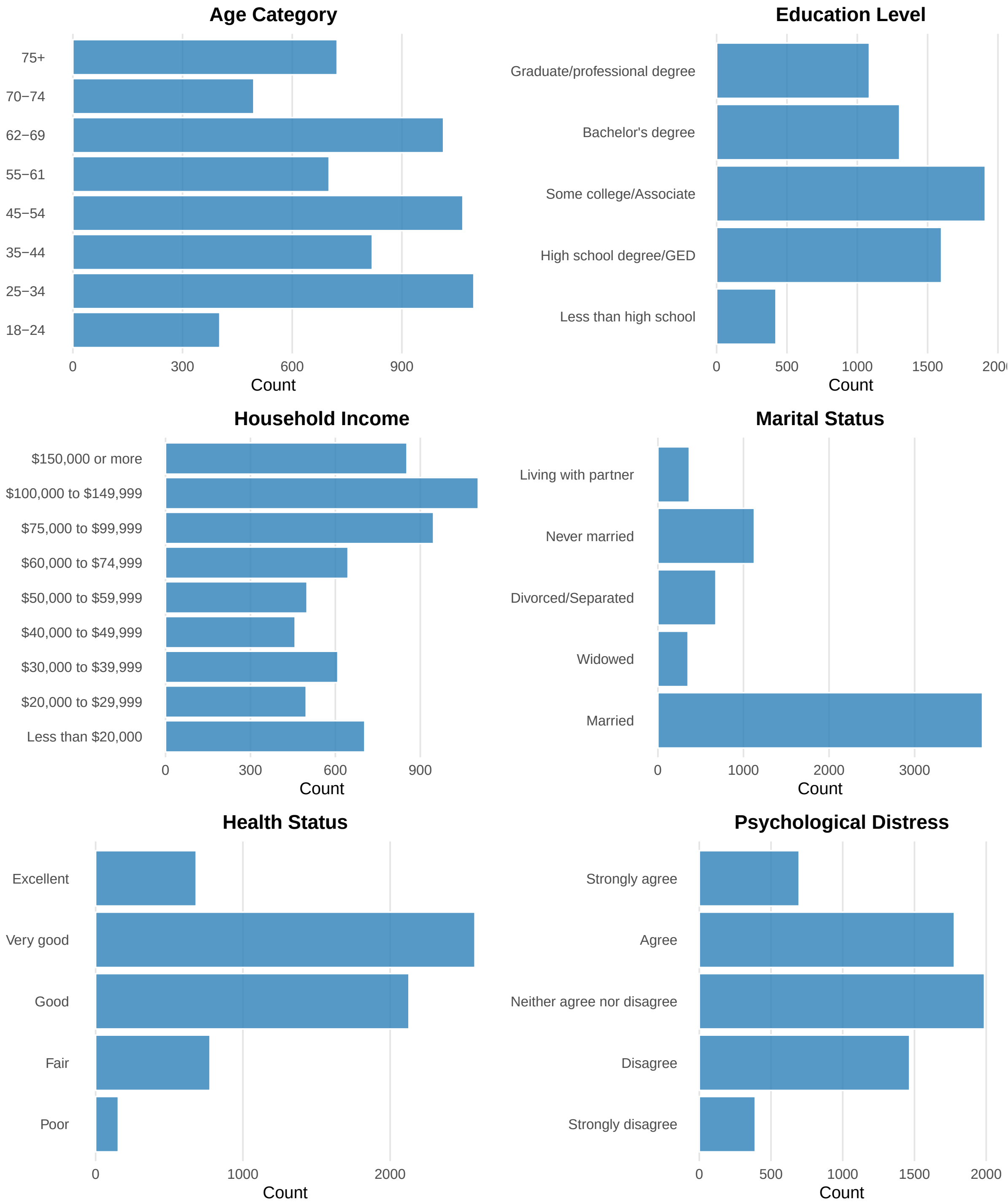

Figure 3: Bar Plots of Categorical Variables

indicate that "Some college/Associate" and "High school degree/GED" are the most common, while "Less than high school" is the least common. Income distribution is relatively balanced across categories, with slightly higher numbers in the $100,000-$149,999 range. Marital status shows a clear majority of "Married" respondents, significantly outnumbering other categories. Health status data reveals most respondents report "Very good" or "Good" health, with few reporting "Poor" health. Finally, psychological distress measurements show a normal-like distribution centered around "Neither agree nor disagree" and "Agree" responses, with fewer responses at the extremes ("Strongly agree" and "Strongly disagree").

3 Model

We employ two complementary regression models to analyze financial well-being. Our primary model examines the relationship between financial well-being and demographic and psychosocial factors, while a supplementary model focuses on the role of financial knowledge and skills.

3.1 Primary Model

The primary model can be expressed as:

$$\text{FWB}_i = \beta_0 + \beta_1 \text{Age}_i + \beta_2 \text{Education}_i + \beta_3 \text{Income}_i + \beta_4 \text{Marital}_i + \beta_5 \text{Health}_i + \beta_6 \text{Distress}_i + \epsilon_i$$

where $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$

Here:

- FWB_i is the Financial Well-Being score for individual i;
- Age categories are represented by indicator variables with 18-24 as the baseline;
- Education levels are coded with "Less than high school" as the baseline;
- Income categories start from "Less than $20,000" as the baseline;
- Marital status uses "Married" as the reference category;
- Health status is measured from "Poor" (baseline) to "Excellent";
- Psychological distress ranges from "Strongly disagree" (baseline) to "Strongly agree".

3.2 Supplementary Model

The supplementary model examining financial literacy's role is:

$$\text{FWB}_i = \beta_0 + \beta_1 \text{LMscore}_i + \beta_2 \text{KHscore}_i + \beta_3 \text{FSscore}_i + \epsilon_i$$

where the scores represent standardized measures of financial knowledge and skills.

3.3 Model Diagnostics

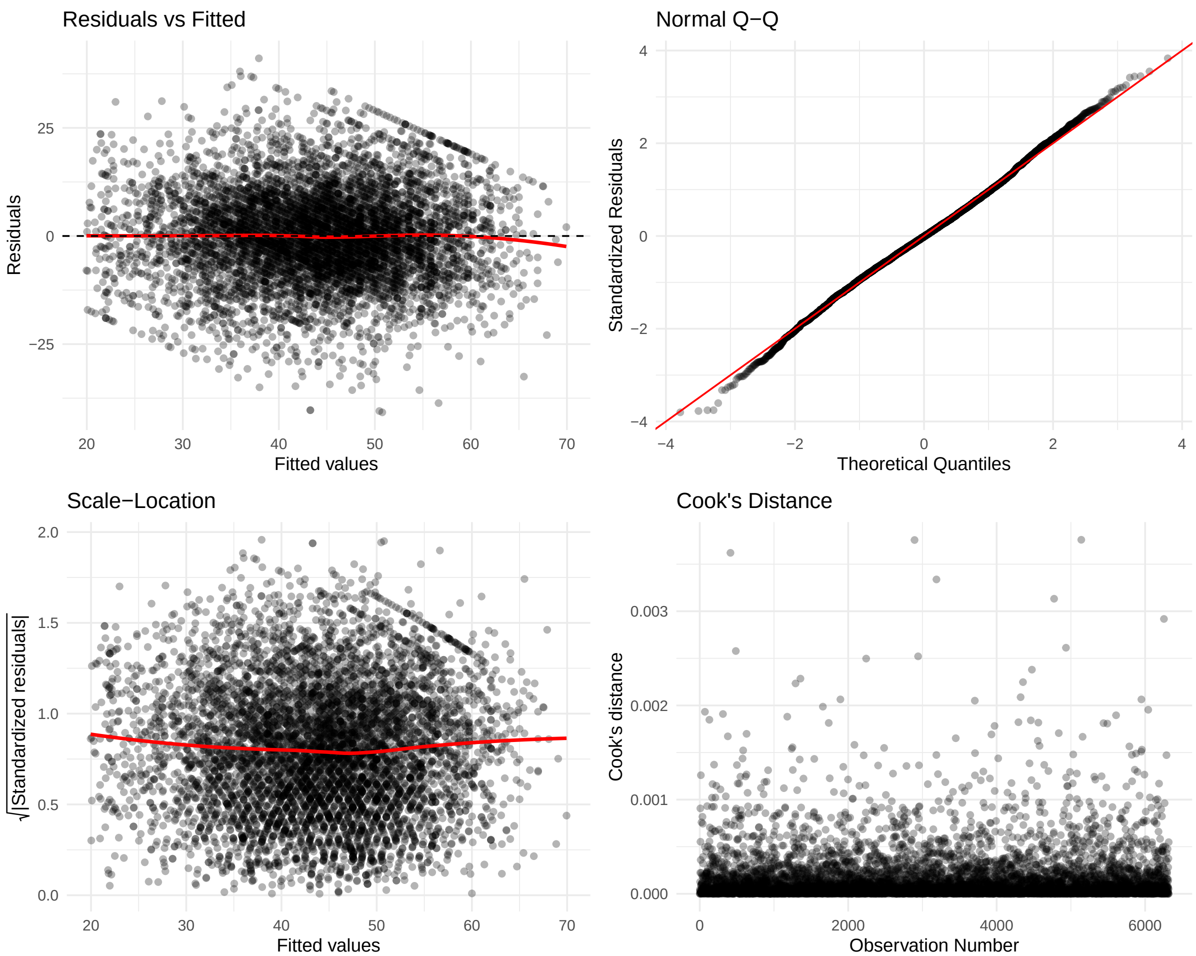

Figure 4: Model Diagnostic Plots

We conducted several diagnostic checks to validate our model assumptions, as illustrated in Figure 4. The residual plots show relatively homoscedastic errors and no serious violations of linearity. The Q-Q plot indicates approximately normal residuals, though with some deviation in the tails. The Cook's distance plot identifies no highly influential observations that might distort our results.

4 Results

Table 1: Linear Model Output with Significant Variable and Levels

Variable	Value	estimate	std.error	statistic	p.value
Intercept	See footnote	38.334	1.348	28.436	0.000
agecat	55-61	1.764	0.753	2.343	0.019
agecat	62-69	6.354	0.732	8.675	0.000
agecat	70-74	7.509	0.823	9.122	0.000
agecat	75+	9.366	0.795	11.778	0.000
PPEDUC	Bachelor's degree	1.940	0.665	2.919	0.004
PPEDUC	Graduate/professional degree	2.921	0.692	4.224	0.000
PPINCIMP	$30,000 to $39,999	1.770	0.612	2.891	0.004
PPINCIMP	$40,000 to $49,999	3.779	0.669	5.652	0.000
PPINCIMP	$50,000 to $59,999	5.311	0.661	8.039	0.000
PPINCIMP	$60,000 to $74,999	6.166	0.626	9.852	0.000
PPINCIMP	$75,000 to $99,999	7.590	0.592	12.825	0.000
PPINCIMP	$100,000 to $149,999	9.163	0.594	15.419	0.000
PPINCIMP	$150,000 or more	12.597	0.645	19.543	0.000
PPMARIT	Widowed	-1.313	0.641	-2.049	0.040
PPMARIT	Divorced/Separated	-1.905	0.464	-4.103	0.000
PPMARIT	Never married	-0.949	0.436	-2.179	0.029
PPMARIT	Living with partner	-1.842	0.624	-2.951	0.003
HEALTH	Good	3.372	0.924	3.648	0.000
HEALTH	Very good	5.635	0.932	6.049	0.000
HEALTH	Excellent	7.701	1.002	7.689	0.000
DISTRESS	Disagree	-3.845	0.620	-6.201	0.000
DISTRESS	Neither agree nor disagree	-7.674	0.608	-12.623	0.000
DISTRESS	Agree	-11.204	0.624	-17.949	0.000
DISTRESS	Strongly agree	-16.518	0.714	-23.142	0.000

Table 2: Summary of Supplementary Model

	estimate	std.error	statistic	p.value
(Intercept)	19.6183467	0.844	23.248	0

Table 2: Summary of Supplementary Model

	estimate	std.error	statistic	p.value
LMscore	0.8737951	0.248	3.527	0
KHscore	3.7484740	0.230	16.300	0
FSscore	0.5005944	0.012	41.524	0

Our analysis reveals several key patterns in the determinants of financial well-being. The coefficient table Table 1 visualizes the main findings from our primary model, showing the relative importance of different factors.

4.1 Age Effects

Age demonstrates a strong positive relationship with financial well-being, particularly for older adults. Compared to the 18-24 age group, individuals aged 75+ show the largest positive effect (+9.37 points), followed by those 70-74 (+7.51 points) and 62-69 (+6.35 points). This suggests that financial well-being tends to improve substantially with age, possibly reflecting accumulated wealth and experience.

4.2 Income and Education

Income exhibits the strongest positive association with financial well-being among all variables. Those earning $150,000 or more score 12.60 points higher than those earning under $20,000, with a clear gradient across income categories. Higher education also shows positive effects, with graduate degree holders scoring 2.92 points higher than those without high school completion.

4.3 Health and Psychological Distress

Health status shows a strong positive relationship with financial well-being, while psychological distress demonstrates the largest negative effects in our model. Excellent health is associated with a 7.70-point increase compared to poor health. Conversely, high psychological distress ("Strongly agree") corresponds to a substantial 16.52-point decrease in financial well-being.

4.4 Financial Knowledge and Skills

Our supplementary model, as in Table 2, requires careful interpretation given the distinct distributions of our financial literacy measures:

Financial skills (FSscore) shows a moderate positive effect ($\beta = 0.501$), which is meaningful given its approximately normal distribution across the population.

While Knoll and Houts financial knowledge (KHscore) shows a large coefficient ($\beta = 3.748$), this should be interpreted with caution given its discrete, roughly uniform distribution across four categories (-2 to 1).

The Lusardi and Mitchell financial knowledge (LMscore) effect ($\beta = 0.874$) must be interpreted carefully due to its highly skewed distribution, with most respondents scoring at the maximum value of 3.

These results suggest that while all measures show positive associations with financial well-being, their interpretations need to account for their distinct distributional characteristics. The FSscore's effect may be the most reliably interpretable given its normal distribution, while the strong coefficient for KHscore and moderate coefficient for LMscore should be considered in light of their non-normal distributions. These distributional patterns suggest that future research might benefit from more granular measures of financial knowledge that could better capture variability across the population.

5 Discussion

5.1 Overview of Research Contribution

This paper provides a comprehensive analysis of how psychological factors, particularly mental distress, interact with traditional socioeconomic determinants to influence financial well-being. Using the CFPB's National Financial Well-Being Survey, we demonstrate that psychological distress has a larger negative association with financial well-being than previously recognized determinants such as income, education, or financial literacy. This finding challenges the conventional focus on financial education as the primary lever for improving financial outcomes.

5.2 Key Insights About Financial Well-Being

5.2.1 The Primacy of Psychological Factors

Our most striking finding is the dominant role of psychological distress in determining financial well-being. A one-point increase in psychological distress is associated with a 4.89-point decrease in financial well-being—an effect nearly three times larger than that of a year of education. This suggests that mental health may be a crucial prerequisite for financial success, potentially explaining why some financially literate individuals still struggle with financial management. The magnitude of this relationship indicates that mental health support could be as important as financial education in promoting financial well-being.

5.2.2 The Limited Role of Financial Knowledge

Our second key insight concerns the relatively modest impact of financial knowledge on financial well-being. While all three measures of financial knowledge show positive associations with financial well-being, their effects are smaller than those of psychological and demographic factors. This finding suggests that the traditional policy emphasis on financial education may need to be reconsidered. Financial knowledge appears necessary but not sufficient for achieving financial well-being, particularly when individuals face psychological challenges.

5.3 Study Limitations

Several limitations of our analysis warrant discussion. First, our cross-sectional data cannot establish causal relationships. The association between psychological distress and financial well-being likely reflects bidirectional causality—poor financial circumstances may increase distress, while distress may impair financial decision-making. Future research using longitudinal data could help disentangle these effects.

Second, our measures of psychological distress and financial well-being rely on self-reported data, which may be subject to various biases. Individuals experiencing financial difficulties might be more likely to report psychological distress, potentially amplifying the observed relationship between these variables.

Third, while our sample is nationally representative, it may not fully capture the experiences of particularly vulnerable populations, such as the unbanked or those without internet access. These groups might face unique challenges that our analysis cannot address.

5.4 Policy Implications and Future Directions

Our findings have important implications for policy and practice. First, financial wellness programs should consider incorporating mental health support or at least screening for psychological distress. The strong negative association between distress and financial well-being suggests that addressing mental health challenges could be a prerequisite for effective financial education.

Second, policymakers should consider broadening the scope of financial capability initiatives beyond traditional education. Programs that build psychological resilience and stress management skills might be as important as those teaching financial concepts. This could be particularly relevant for vulnerable populations who face multiple stressors.

Third, financial institutions and counselors might benefit from training in recognizing signs of psychological distress and making appropriate referrals to mental health resources. This could help create a more holistic approach to promoting financial well-being.

Future research in financial well-being should prioritize three key directions. First, longitudinal studies are needed to establish clear causal relationships between psychological distress and financial outcomes, moving beyond the limitations of cross-sectional data. Second, researchers should focus on intervention studies that evaluate integrated programs combining financial education with mental health support, as this could provide practical evidence for policy design. Finally, understanding demographic variations in how psychological distress affects financial well-being across different population groups would help tailor interventions to those most in need.

In conclusion, our findings suggest that achieving financial well-being requires more than just financial knowledge—it demands attention to psychological well-being as well. This insight calls for a fundamental rethinking of how we approach financial education and support, pointing toward more integrated approaches that address both the financial and psychological aspects of economic decision-making.

6 Appendix

6.1 Survey Methodology

The CFPB Financial Well-Being Scale methodology merits critical examination across three key dimensions: sampling design, measurement validation, and methodological limitations.

6.1.1 Sampling Framework and Implementation

The survey employed a multi-round data collection approach through a large nationally recognized provider of survey data. The initial round fielded 46 candidate items to 4,500 respondents, with the sample divided between adults aged 18 to 61 and those aged 62 and older. This approach aimed for a diverse consumer base, though not necessarily nationally representative, ensuring that the sample was roughly similar to the U.S. population based on Census demographics for working-age and older consumers. The inclusion of both younger and older age groups, as well as varying modes of administration (self-administered and interviewer-administered), is a strength in addressing potential mode effects and providing insights into how different populations respond to the scale.

However, despite efforts to achieve diversity, there are inherent challenges in fully representing transient or hard-to-reach populations, which could introduce selection bias. Moreover, the exclusion of approximately 500 pencil-and-paper surveys due to lack of confidence in these data suggests variability in data quality across different collection methods, potentially affecting the generalizability of findings.

6.1.2 Measurement Development and Validation

The development of the Financial Well-Being Scale adhered to item response theory (IRT) principles, involving extensive rounds of refinement from an initial pool of items down to the final set. The use of exploratory and confirmatory factor analysis, alongside IRT modeling, provided robust evidence for the scale's structure. Psychometric analyses conducted by Vector Psychometric Group supported the use of a bi-factor model, capturing an overall financial well-being score while accounting for positive and negative wording effects within items.

The reported marginal reliability of at least 0.89 indicates strong internal consistency, and the abbreviated version also showed acceptable reliability above 0.80. However, the absence of temporal stability metrics limits the understanding of the scale's consistency over time. Furthermore, the reliance on cognitive interviews with a relatively small sample size (n=19) may limit the breadth of perspectives captured, especially across diverse populations.

Validation of the scale included testing its performance across different demographic groups and survey modes, supporting the comparability of scores derived from different contexts. Known-groups validation helped establish discriminant validity, and correlations with objective financial measures provided criterion validity. Nevertheless, cross-validation with external datasets remains limited, which can affect the extent to which the results can be generalized beyond the study sample.

6.1.3 Methodological Strengths and Limitations

The mixed-mode implementation allowed for the assessment of mode effects, which is important given documented differences in self-reporting behaviors depending on the survey format. Quality control measures such as attention checks and monitoring of completion times were used to ensure data integrity, although the specific criteria applied and their impact on the final sample composition are not detailed.

Critical limitations include: 1. Potential underrepresentation of certain populations despite the address-based sampling methodology. 2. Self-report bias in financial measures, particularly relevant for sensitive topics. 3. Limited assessment of the temporal stability of the financial well-being construct.

The study achieved a response rate that exceeds typical online panel rates, but this also raises questions about nonresponse bias. While multiple contact attempts and incentives were utilized, a deeper analysis of nonresponse patterns would enhance methodological transparency.

In summary, the methodological choices reflect a balance between practical constraints and ideal survey design standards. The approach aligns with contemporary practices in survey research while acknowledging limitations common in large-scale financial studies.

6.2 Acknowledgements

We would like to express our gratitude to the developers and contributors of R (R Core Team 2023) as well as several R packages that were essential for the analysis and visualization of the data in this report. The following R packages provided indispensable tools and functionalities:

- `tidyverse` (Wickham et al. 2019): A collection of R packages designed for data science, including dplyr, ggplot2, readr, purrr, and others, which greatly facilitated data manipulation, analysis, and visualization.
- `ggplot2` (Wickham 2016): An implementation of the Grammar of Graphics, which allowed us to create complex and aesthetically pleasing visualizations with ease.
- `knitr` (Xie 2023): This package enabled us to perform data demonstration with tables.
- `styler` (Müller and Walthert 2023): This package is helpful for styling the code.
- `arrow` (Richardson et al. 2024): This package provides a convenient and efficient way to work with parquet format.
- `validate` (van der Loo and de Jonge 2021): This package provides useful functions for data tests.
- `car` (Fox and Weisberg 2019): The Companion to Applied Regression package provided tools for regression diagnostics and analysis.
- `GGally` (Schloerke et al. 2024): This extension to ggplot2 enabled us to create complex visualization matrices and explore multivariate relationships.
- `broom` (Robinson, Hayes, and Couch 2023): This package helped us convert statistical analysis objects into tidy data frames.
- `gridExtra` (Auguie 2017): This package allowed us to arrange multiple plots in a grid layout.
- `corrplot` (Wei and Simko 2021): This package provided visualization tools for correlation matrices.
- `kableExtra` (Zhu 2024): This package enhanced our ability to create beautiful and customizable tables.

We are grateful to the Consumer Financial Protection Bureau (CFPB) for conducting the Financial Well-Being Survey (Consumer Financial Protection Bureau 2017) and making the data publicly available. Their dedication to collecting and sharing such vital information significantly contributes to research on financial well-being.

Finally, we would like to thank all those who contributed to the development and maintenance of the R programming language and its ecosystem, as well as the broader open-source community, whose efforts make such research possible. Special appreciation goes to the maintainers and contributors of the additional statistical and visualization packages that enhanced our analysis capabilities.

References

Auguie, Baptiste. 2017. *gridExtra: Miscellaneous Functions for "Grid" Graphics*. https://CRAN.R-project.org/package=gridExtra.

Bureau, Consumer Financial Protection. 2015. "CFPB Financial Skill Scale." Consumer Financial Protection Bureau.

Consumer Financial Protection Bureau. 2017. "Financial Well-Being Survey Data." https://www.consumerfinance.gov/data-research/financial-well-being-survey-data/.

Fox, John, and Sanford Weisberg. 2019. *An R Companion to Applied Regression*. Third. Thousand Oaks CA: Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Knoll, Melissa A., and Carrie R. Houts. 2012. "The Financial Knowledge Scale: An Application of Item Response Theory to the Assessment of Financial Literacy." *Journal of Consumer Affairs* 46 (3): 381–410.

Lusardi, Annamaria, and Olivia S. Mitchell. 2008. "Planning and Financial Literacy: How Do Women Fare?" Cambridge, MA: National Bureau of Economic Research.

Müller, Kirill, and Lorenz Walthert. 2023. *Styler: Non-Invasive Pretty Printing of r Code*. https://CRAN.R-project.org/package=styler.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

Robinson, David, Alex Hayes, and Simon Couch. 2023. *Broom: Convert Statistical Objects into Tidy Tibbles*. https://CRAN.R-project.org/package=broom.

Schloerke, Barret, Di Cook, Joseph Larmarange, Francois Briatte, Moritz Marbach, Edwin Thoen, Amos Elberg, and Jason Crowley. 2024. *GGally: Extension to 'Ggplot2'*. https://CRAN.R-project.org/package=GGally.

van der Loo, Mark P. J., and Edwin de Jonge. 2021. "Data Validation Infrastructure for R." *Journal of Statistical Software* 97 (10): 1–31. https://doi.org/10.18637/jss.v097.i10.

Wei, Taiyun, and Viliam Simko. 2021. *R Package 'Corrplot': Visualization of a Correlation Matrix*. https://github.com/taiyun/corrplot.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.

Zhu, Hao. 2024. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.