financial Well Being Analysis*

Xinxiang Gao

December 3, 2024

This should be an abstract. The report is still being finalized.

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) Lusardi and Mitchell (2008) provides a foundational three-item measure of basic financial concepts, while the Knoll and Houts financial knowledge scale (KHscore) Knoll and Houts (2012) offers a more comprehensive 10-item assessment using item response theory. Additionally, the CFPB Financial Skill Scale (FSscore) 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.

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

This paper investigates how demographic characteristics (age, education, income, marital status) and psychosocial factors (health status, psychological distress) influence individuals' financial well-being as measured by the FWBscore, while accounting for their financial knowledge and skills as measured by the LMscore, KHscore, and FSscore. By examining these relationships comprehensively, we aim to identify which population segments may be more vulnerable to financial challenges and how personal circumstances interact with financial literacy to shape financial well-being outcomes.

Our findings indicate [brief summary of key results], suggesting important implications for financial education programs and policy interventions. These results highlight the need for tailored approaches that consider both financial knowledge and individual circumstances in promoting financial well-being.

The rest of the paper is organized as follows: Section 2 reviews relevant literature on financial well-being and its determinants. Section 3 describes our data and methodology. Section 4 presents our empirical results. Section 5 discusses the implications of our findings and concludes.

2 Data

2.1 Overview of the Dataset

The dataset for this study is the Consumer Financial Protection Bureau's (CFPB) National Financial Well-Being Survey Public Use File (PUF), collected in 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).

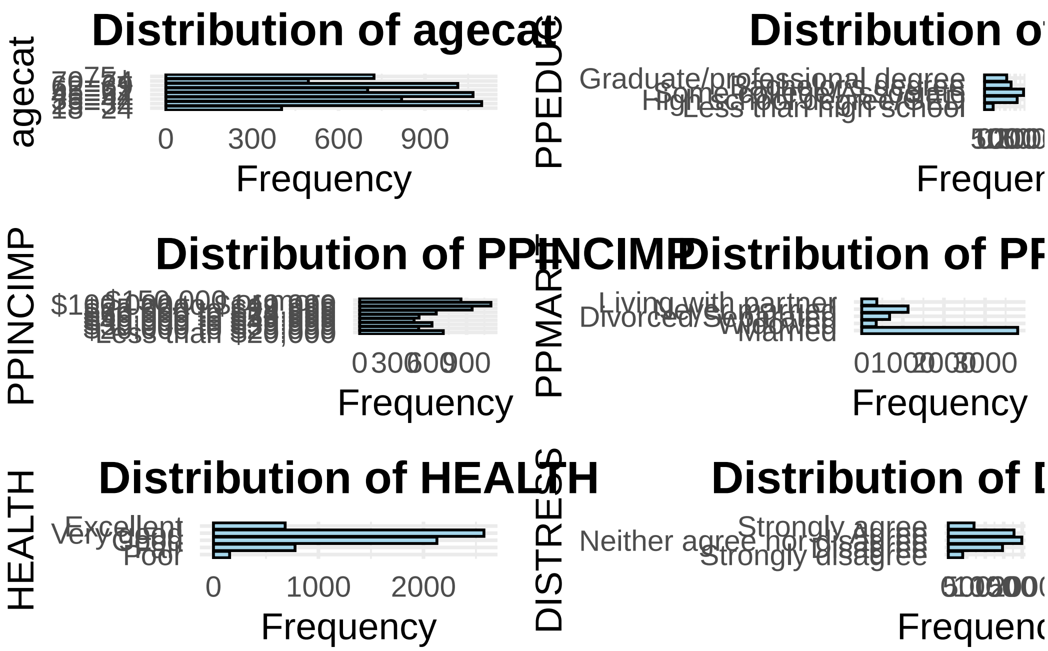
Stress (DISTRESS): Perceived level of stress, rated on a 5-point Likert scale from 1 (Not at all) to 5 (Extremely).

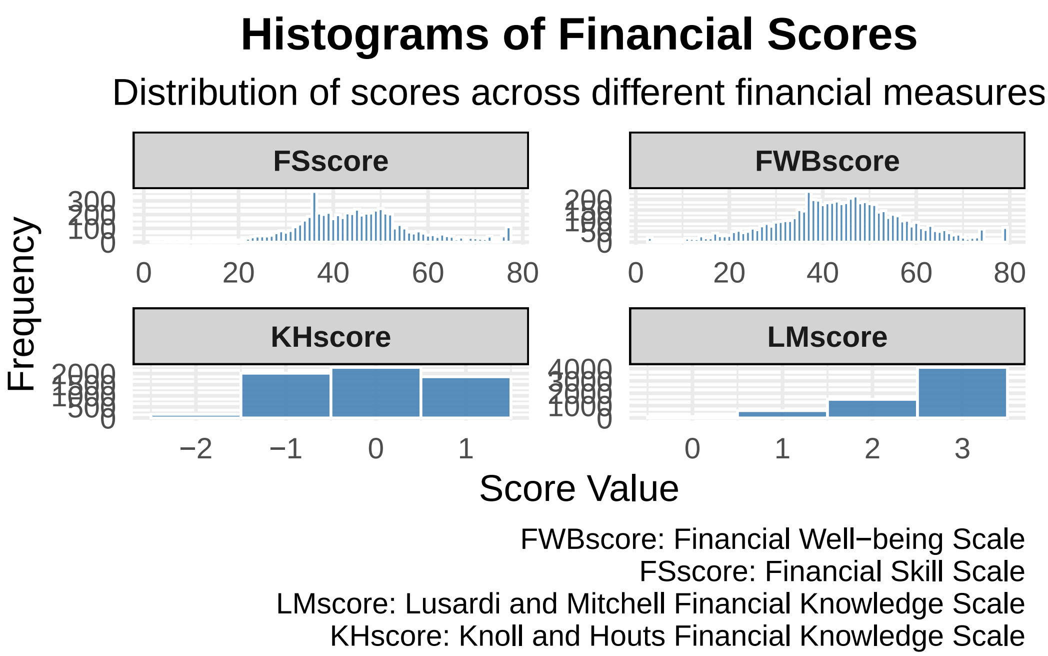
2.2.3 Additional Variables

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.





3 Model

The linear regression model can be expressed mathematically as:

$$y_i = \beta_0 + \beta_1 \cdot \text{LMscore}_i + \beta_2 \cdot \text{KHscore}_i + \beta_3 \cdot \text{FSscore}_i + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma^2)$$

3.0.1 Model Components

- y_i: The dependent variable, FWBscore_i, for observation i.
- β_0: The intercept term (constant in the model).
- $\beta_1, \beta_2, \beta_3, \beta_4$: The slopes (coefficients) of the respective predictor variables:
  - LMscore_i: Lusardi and Mitchell financial knowledge scale score.
  - KHscore_i: Knoll and Houts financial knowledge scale score.
  - FWBscore_i: Financial well-being scale score.
  - FSscore_i: Financial skill scale score.

4 Results

Table 1: Summary of Linear Regression Model

	Term	Estimate	Std_Error	t_value	p_value
(Intercept)	(Intercept)	19.618	0.844	23.248	0
LMscore	LMscore	0.874	0.248	3.527	0
KHscore	KHscore	3.748	0.230	16.300	0
FSscore	FSscore	0.501	0.012	41.524	0

Table 2: 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

Variable	Value	estimate	std.error	statistic	p.value
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

```r
# Calculate VIF
library(car)
vif(lm_model)
```

```
             GVIF Df GVIF^(1/(2*Df))
agecat   2.000516  7        1.050776
PPEDUC   1.559984  4        1.057158
PPINCIMP 1.773351  8        1.036453
PPMARIT  1.885586  4        1.082507
HEALTH   1.313284  4        1.034653
DISTRESS 1.275537  4        1.030888
```

5 Discussion

6 Appendix

6.1 Data cleaning

6.2 Surveys, sampling, and observational data

6.3 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 provide useful functions for data tests.

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.

References

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/.

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.

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.

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/.