



Public Trust in Banks and Financial Institutions

A Knowledge Mining Approach

Presented By: Soha Arian, Kasra Akbari, Eman Ajmal, Joseph Martinez

Meet The Team

Soha Arian Kasra Akbari Eman Ajmal Joseph Martinez

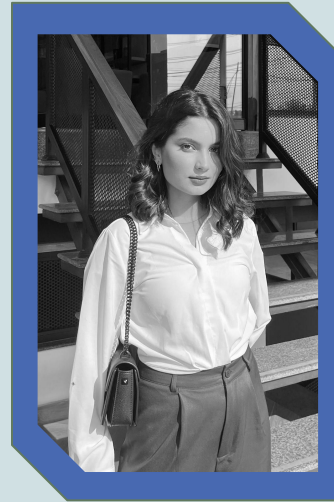


Table Of Contents



01 Problem Statement

04 Methodology

02 Research Question

05 Preliminary Findings

03 Solution

06 Progress & Delivery





Problem Statement

- Public trust in banks is an important part of financial behavior and institutional legitimacy
- Confidence in banks may differ across income, education, and age groups
- These differences may also shift across historical periods
- A data-driven analysis is needed to show whether these demographic patterns are stable, changing, or uneven over time

Research Question

How does confidence in banks vary by income, education, and age in the United States, and do these relationships remain consistent over time? Does an AI-generated summary accurately reflect the statistical findings?

Methodology Overview

Data Source: General Social Survey (1972–2024)

Outcome Variable: Confidence in banks

Predictors: Income, education, age, survey year

Step 1 Through 5 Approach:

- Step 1: Select and clean the required GSS variables
- Step 2: Recode confidence in banks as an ordered outcome
- Step 3: Group income within survey year for fair comparison over time
- Step 4: Apply GSS survey weights
- Step 5: Run weighted descriptive statistics by income, education, age, and periods



Methodology Overview

Data Source: General Social Survey (1972–2024)

Outcome Variable: Confidence in banks

Predictors: Income, education, age, survey year

Steps 6 Through 10 Approach:

- Step 6: Estimate ordered logistic regression
- Step 7: Test whether demographic relationships vary over time using interactions with survey year
- Step 8: Check model fit and sensitivity, including alternative income constructions
- Step 9: Generate predicted probabilities
- Step 10: Evaluate AI summary accuracy using a fixed prompt and scoring rubric



SHINY APP METHODOLOGY

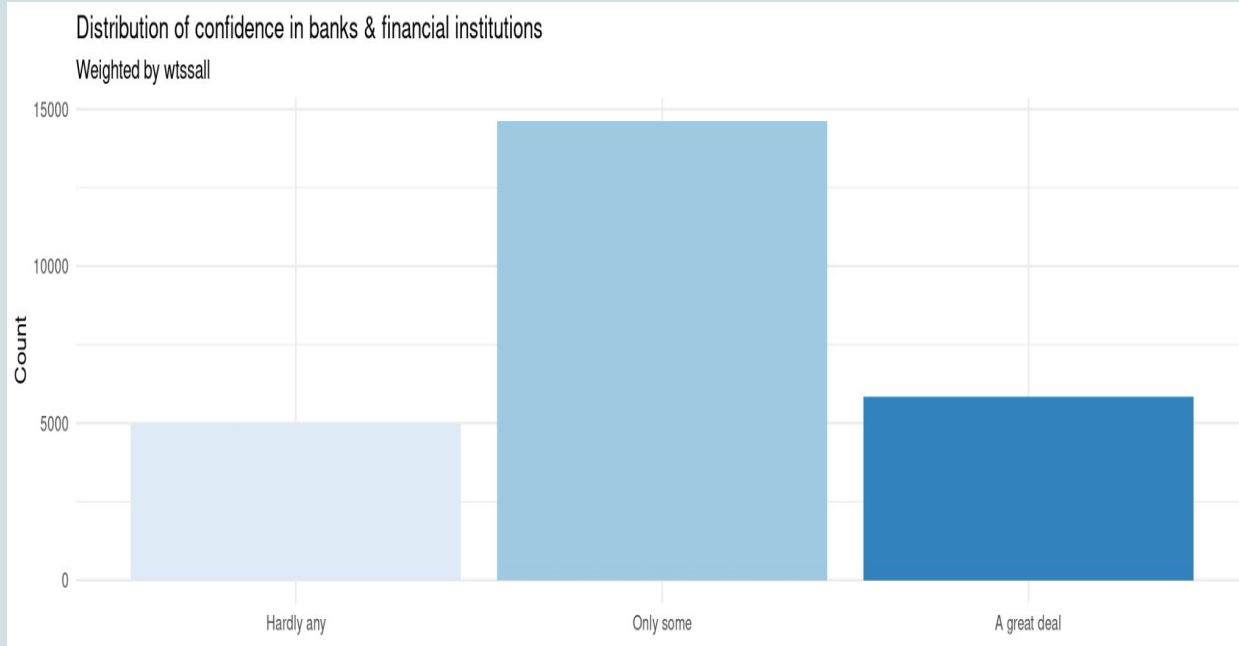
- Interactive deliverable for reproducing and inspecting GSS analysis
- Processes cleaned data files (`gss_banks.rds`, `cpi_annual.rds`)
- Filter by year, age, sex, race, and education degree
- Select income measure and survey weighting options
- Choose period mode and click "Fit model"

Outputs

- Bank confidence overview with bar and trend charts
- Confidence breakdown by income group
- Ordered-logit model results (odds ratios & confidence intervals)
- Predicted probabilities by income group
- Model results export to AI prompt (bonus feature)

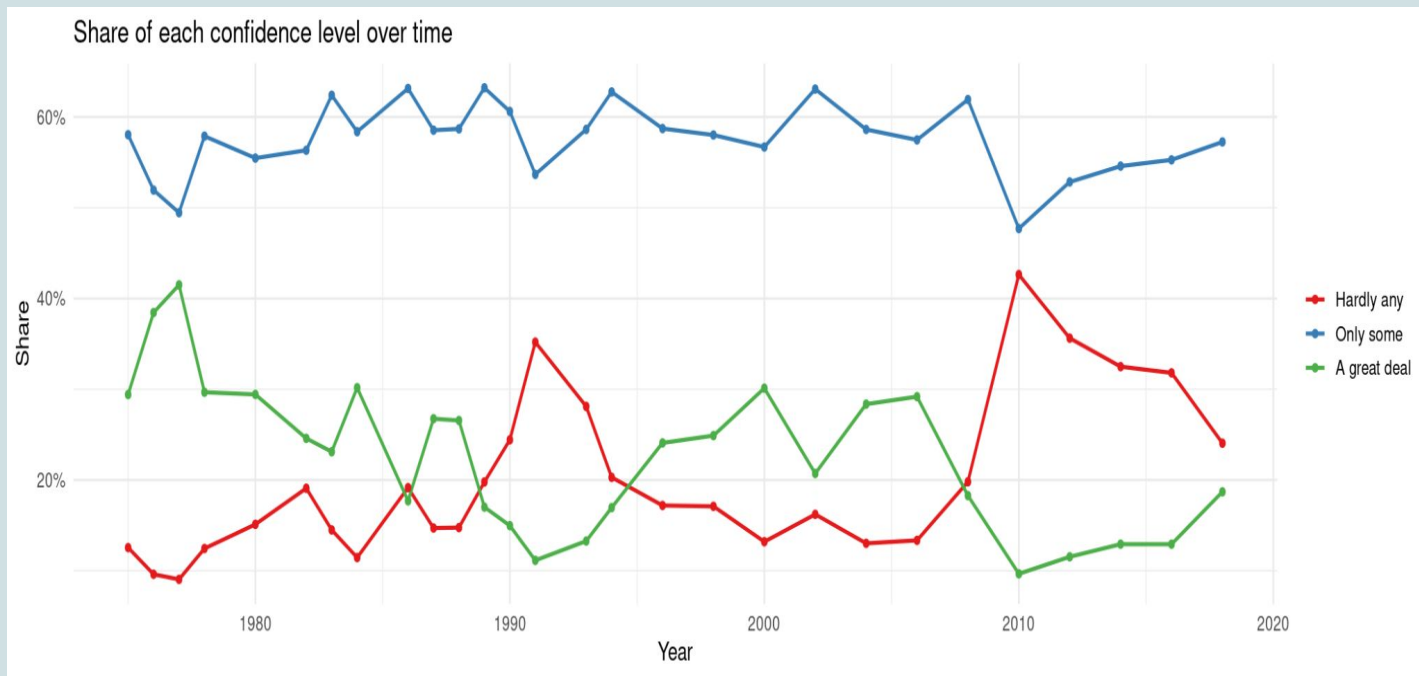


Figure 1. Distribution of confidence in banks and financial institutions



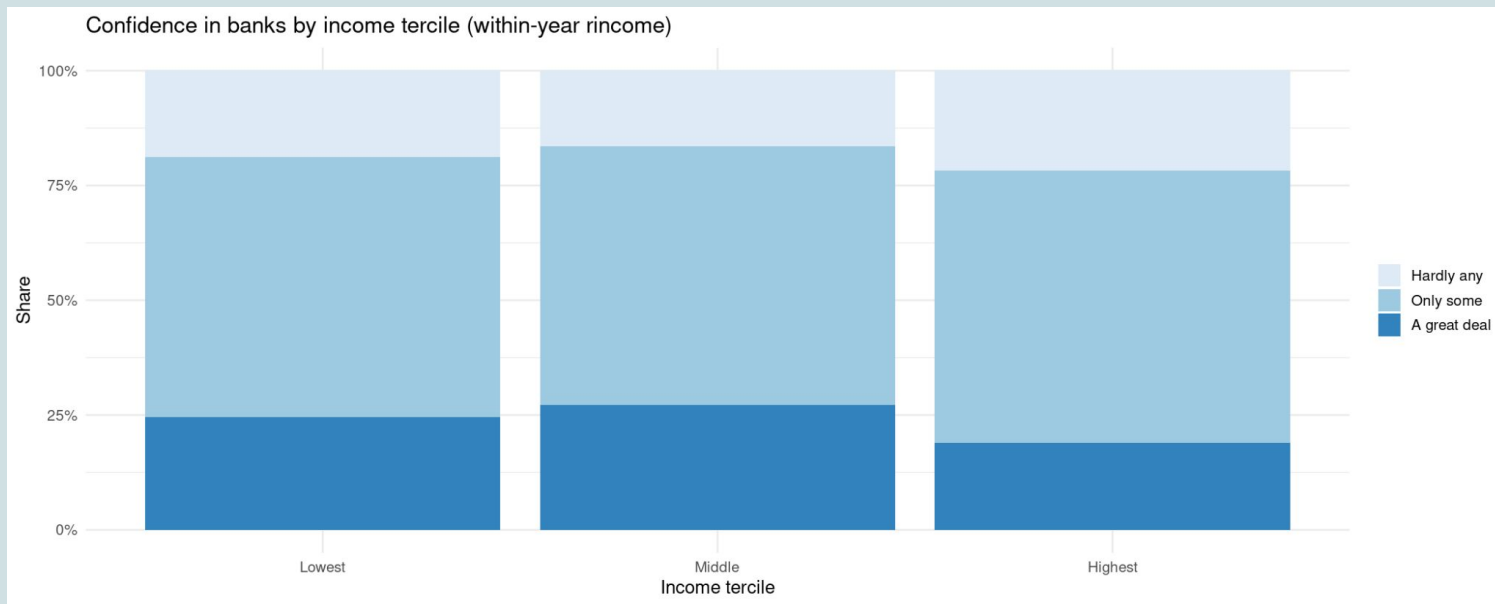
Source: Shiny app Overview tab. This figure shows the weighted distribution of responses to the GSS confidence-in-banks variable. The chart shows that “Only some” confidence is the most common response category.

Figure 2. Share of each confidence level over time



Source: Shiny app Overview tab. This figure shows the weighted share of respondents reporting “Hardly any,” “Only some,” and “A great deal” of confidence across survey years. The graph shows that public confidence in banks changes over time rather than staying flat.

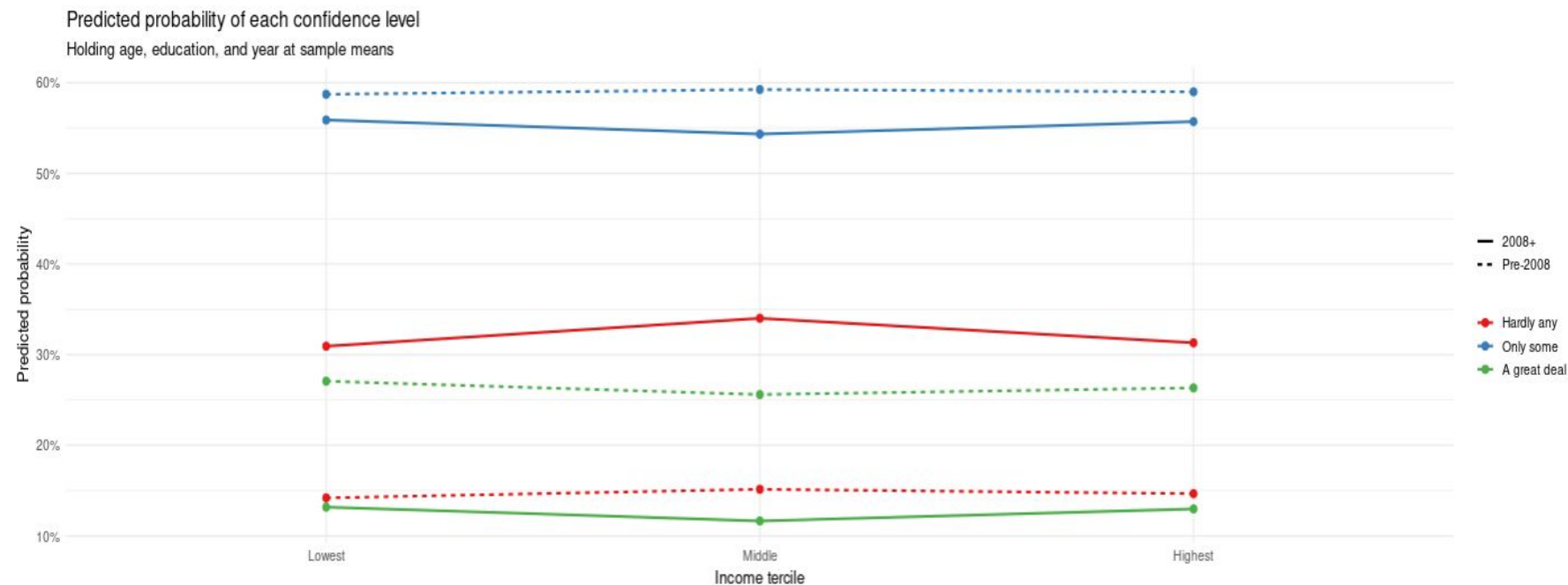
Figure 3. Predicted probability of each confidence level by income group



Source: Shiny app Predicted Probabilities tab. This figure shows predicted probabilities for “Hardly any,” “Only some,” and “A great deal” confidence across income terciles while holding age, education, and year at sample values.

Figure 4. Predicted probability of each confidence level Pre and Post 2008 Financial Crisis by income groups

Predicted probabilities at mean age, mean education, and median year within each sample.



Source: Shiny app Predicted Probabilities tab. This figure shows predicted probabilities for “Hardly any,” “Only some,” and “A great deal” confidence across income terciles while holding age, education, and year at sample values.

Preliminary Findings

- We successfully prepared the GSS data for analysis on the Shiny app and resolved the Stata variable-limit issue during data setup
- We constructed the main analysis variables, including the ordered confidence variable and within-year income groups
- The initial ordered logistic regression has been completed
- The current results shown here are preliminary and focus on income patterns



Preliminary Findings

- Predicted probabilities show that “Only some” confidence is the most common response across all income groups
- Differences by income appear modest rather than large
- Higher-income respondents are slightly more likely to report “Hardly any” confidence
- Higher-income respondents are slightly less likely to report “A great deal” of confidence



KEY ACTIONS

- We will generate an AI summary using a fixed prompt
- We will compare the AI summary directly to the model results
- We will score whether the summary correctly reports direction, size, and significance of the findings
- We will also record omissions, exaggerations, and factual errors



NEXT STEPS

- Keep testing AI summaries of the findings using a fixed prompt
- Score whether the AI summary accurately reflects the statistical results
- Finalize the report and website posting



References

Grable, John E., Eun Jin Kwak, and Kristy L. Archuleta. “Distrust of Banks among the Unbanked and Banked.” *International Journal of Bank Marketing*, vol. 41, no. 6, 2023, pp. 1498–1520.

<https://doi.org/10.1108/IJBM-10-2022-0441>.

Croignani, Matteo, et al. “Financial Inclusion in the United States: Measurement, Determinants, and Recent Developments.” *SSRN Electronic Journal*, 2025, <https://doi.org/10.2139/ssrn.5534239>.

General Social Survey. “Get the Data in Stata Format.” *GSS – NORC at the University of Chicago*, <https://gss.norc.oregonstate.edu/get-the-data/stata.html>. Accessed 26 Feb. 2026.