**Title:** Can Machine Learning Predict Stimulus Check Usage? Evidence from the 2021 NFCS Dataset

**Abstract**

Understanding how individuals used their COVID-19 stimulus checks is essential for evaluating the policy's effectiveness and designing future economic relief programs. Using data from the 2021 National Financial Capability Study (NFCS), we classify recipients into three behavioral categories: spenders, savers, and debtors. We train multiple machine learning models to predict stimulus usage behavior from demographic and financial features, comparing performance across classifiers. Our best model, a class-weighted XGBoost, achieves balanced performance across all classes (macro F1 > 0.55), with interpretability enabled via SHAP analysis. Key predictors include income volatility, emergency savings, and financial fragility. Our findings reveal that stimulus use is strongly tied to financial vulnerability and highlight the need for targeted policy design. These insights not only improve our understanding of public financial behavior but also offer practical implications for how aid policies could be structured in the future.

**1. Introduction**

The COVID-19 pandemic triggered an economic crisis that led to the implementation of several relief measures by the U.S. government, including three rounds of stimulus checks. These Economic Impact Payments totaled over $800 billion and aimed to support consumer spending, reduce debt burdens, and prevent household-level financial collapse. While aggregate effects such as GDP growth and poverty reduction have been evaluated, less attention has been paid to how individual recipients used the funds. This behavioral question is critical for designing future relief efforts that are both efficient and equitable.

Conventional wisdom suggests that low-income and financially fragile individuals are more likely to spend their stimulus money, while higher-income and financially stable households are more likely to save it. However, survey evidence often lacks the resolution or predictive power needed to analyze these trends in depth. Machine learning (ML) offers an opportunity to classify individual-level behavior based on observable features and to interpret those patterns using modern tools like SHAP (SHapley Additive Explanations).

This study uses publicly available microdata from the 2021 National Financial Capability Study (NFCS) to predict and interpret how stimulus recipients used their checks. We define three behavioral classes—spender, saver, debtor—and train a series of machine learning models to classify individuals based on their demographic and financial attributes. We further use SHAP to interpret feature influence and error analysis to understand model limitations. Our findings show strong associations between financial vulnerability indicators and spending or debt usage, providing useful guidance for policymakers. Can interpretable machine learning models accurately predict how individuals used their COVID-19 stimulus payments — whether to spend, save, or repay debt — based on financial vulnerability indicators such as income volatility, emergency savings, and job/income shock?

**2. Related Work**

Most existing literature on stimulus checks has focused on macroeconomic outcomes, such as changes in household consumption (Ganong et al., 2020) or poverty rates (Cox et al., 2021). Survey-based studies from the Federal Reserve and U.S. Census Bureau report self-declared usage categories, but few incorporate predictive modeling to anticipate behavior based on pre-existing attributes. Our study bridges this gap by integrating machine learning with financial survey data, enabling both prediction and interpretation at the individual level. While some prior work has used logistic regression or decision trees to analyze consumer finance, we contribute by applying state-of-the-art models like XGBoost and visual explanations with SHAP.

**3. Data and Labeling**

We use the 2021 wave of the NFCS dataset, a national survey covering household finances, credit use, and economic resilience. From the full dataset of over 27,000 responses, we filtered for individuals who confirmed receiving a COVID-19 stimulus check (QJ51 = 1).

To classify behavior, we used the following questions:

* QJ52: "What did you primarily use your stimulus check for?"
* QJ53: "What else did you use it for, if anything?"

Based on these, we defined a hierarchy of labeling:

1. **Debtor**: Paid down debt
2. **Saver**: Saved or invested
3. **Spender**: Used primarily for consumption

Each respondent was assigned exactly one label using this priority order to avoid ambiguity.

**4. Feature Engineering and Preprocessing**

We engineered features based on prior economic theory and exploratory data analysis. Selected variables include:

* **Income Stability**:
  + A6: Frequency of income changes
  + J1: Recent job loss or income shock
* **Financial Fragility**:
  + J2: Ability to cover a $2,000 emergency expense
  + J20: Emergency savings presence
* **Credit and Demographics**:
  + A50A\_2: Self-reported credit rating
  + A9\_\*: Age brackets
  + A3Ar\_\*: Education level
  + A4A\_new\_\*: Gender

Categorical variables were one-hot encoded. Features with non-responses (codes 98/99) were removed or imputed. All data was split into an 80/20 train-test split using stratified sampling to preserve label proportions. We avoided feature scaling due to the tree-based nature of our best models.

**5. Modeling and Evaluation**

We trained the following classifiers:

1. **Logistic Regression**: Baseline model
2. **Random Forest**: Nonlinear ensemble method
3. **XGBoost**: Gradient-boosted trees

Hyperparameters for XGBoost were tuned via grid search using 5-fold cross-validation, optimizing for weighted F1-score.

To address class imbalance, we applied sample weights using compute\_sample\_weight(class\_weight='balanced').

The final weighted XGBoost model achieved:

* Accuracy: ~58%
* Macro F1-score: ~0.55
* Class-specific F1:
  + Debtor: 0.63
  + Spender: 0.59
  + Saver: 0.41

**[Insert classification report table here]**  
**[Insert confusion matrix heatmap here]**

Weighted XGBoost clearly outperformed other models, particularly in improving saver prediction, which is typically underrepresented.

**6. Interpretation with SHAP**

To interpret model behavior, we used SHAP to generate both global and class-level importance rankings.

**Global Insights:**

* A6 (Income Volatility) was the most influential feature across all classes
* J2 (Financial Fragility) and J20 (Emergency Savings) ranked consistently high

**Class-Level Insights:**

* **Spenders**: Low savings, high volatility, lower credit scores
* **Debtors**: Moderate savings but recent job shocks or unstable income
* **Savers**: Higher education, higher emergency fund access, and better credit

**[Insert SHAP global bar chart here]**  
**[Insert SHAP beeswarm plots for 3 classes here]**

These visualizations confirm that the model captures economically meaningful patterns.

**7. Error Analysis**

We analyzed misclassified test cases to understand failure modes.

Example 1:

* **True Label**: Saver
* **Prediction**: Spender
* **Attributes**: Has some emergency savings but recently experienced job loss

Example 2:

* **True Label**: Debtor
* **Prediction**: Saver
* **Attributes**: Low income volatility, no emergency savings

These examples suggest that borderline cases with mixed financial traits often confuse the model.

**[Insert sample from misclassified\_examples.csv]**

**8. Discussion and Policy Implications**

Our model and SHAP interpretation suggest that stimulus behavior is largely driven by pre-existing financial conditions.

* Spenders are often under severe financial stress and use the funds for immediate consumption needs
* Debtors prioritize stability and view the stimulus as a chance to reduce liabilities
* Savers are generally more secure and see the check as discretionary income

This classification can inform targeted communication strategies:

* Encourage saving among those with some savings buffer
* Offer debt forgiveness or matched repayment programs to debtors
* Ensure spenders have access to flexible support mechanisms (e.g., food or rent aid)

**9. Conclusion**

We developed and interpreted a machine learning model to predict individual stimulus check behavior using financial survey data. Our class-weighted XGBoost model demonstrated that this behavior is predictable and largely explained by financial fragility, income volatility, and emergency savings status.

This approach bridges data science and public policy, offering a way to design more personalized economic relief efforts in the future. We hope this work contributes both methodologically to interpretable ML and practically to social impact analysis.

**Appendix (Optional)**

* Full confusion matrices for all models
* SHAP class-specific plots (saver, spender, debtor)
* Misclassified example table
* Feature correlation heatmap