| **RESEARCH ARTICLE**

Forecasting Household Economic Hardship, Financial Fragility, and Vulnerability Through Artificial Intelligence and Predictive Analytics

Shah Farhan Rabbani
University of New Haven, Business Analytics
Email: srabb2@unh.newhaven.edu
ORCID: https://orcid.org/0009-0001-8434-223X

Yusuf Oli Rahat
University of New Haven, Business Analytics
Email: yraha1@unh.newhaven.edu
ORCID: https://orcid.org/ 0009-0001-8834-7759

Md Kamrul Islam
University of New Haven, Business Analytics
Email: misla22@unh.newhaven.edu
ORCID: https://orcid.org/0009-0001-8906-630X

Corresponding Author: Shah Farhan Rabbani, **E-mail**: srabb2@unh.newhaven.edu

| **ABSTRACT**

Household economic hardship in the United States is increasingly shaped by interacting pressures rather than a single income shortfall. Inflation, rent burdens, revolving debt, medical costs, labor-market volatility, and fraud losses combine to produce financial fragility that conventional scorecards often miss. This paper develops a publication-ready research framework for forecasting household economic hardship, financial fragility, and vulnerability through artificial intelligence and predictive analytics using public data through 2024. The framework integrates survey microdata and contextual indicators from the Federal Reserve's Survey of Household Economics and Decisionmaking, the Consumer Financial Protection Bureau's Making Ends Meet surveys and Financial Well-Being Scale, the U.S. Census Bureau's Household Pulse Survey, USDA food security reports, the Federal Reserve's Survey of Consumer Finances, Bureau of Labor Statistics expenditure and inflation statistics, and Federal Trade Commission fraud-loss data. Building on the supplied literature and broader scholarship in household finance, explainable machine learning, and welfare measurement, the paper proposes a multimodal architecture that combines gradient boosting, temporal models, and graph-informed neighborhood context to predict three linked outcomes: immediate hardship, near-term fragility, and medium-horizon vulnerability. Data trends through 2024 show that the share of adults doing at least okay financially declined after 2021, emergency liquidity weakened, food insecurity rose, and reported fraud losses climbed sharply. The paper argues that explainable, fairness-audited models can improve targeting, consumer protection, and preventive intervention without reducing households to opaque risk labels. It concludes that AI can strengthen public-interest forecasting when embedded in transparent validation, governance, and human-centered implementation.

1. Introduction

Table 1. Data calibration indicators used in the manuscript

Indicator	2021	2022	2023	2024 / note
Adults doing at least okay financially	78%	73%	72%	SHED series
Can cover $400 emergency expense with cash or equivalent	68%	63%	63%	SHED series
Household food insecurity	10.2%	12.8%	13.5%	USDA annual report
FTC-reported fraud losses (USD billions)	$5.8B	$8.8B	$10.0B	$12.5B in 2024

Sources: Board of Governors of the Federal Reserve System (2024), Rabbitt et al. (2024), and Federal Trade Commission (2024).

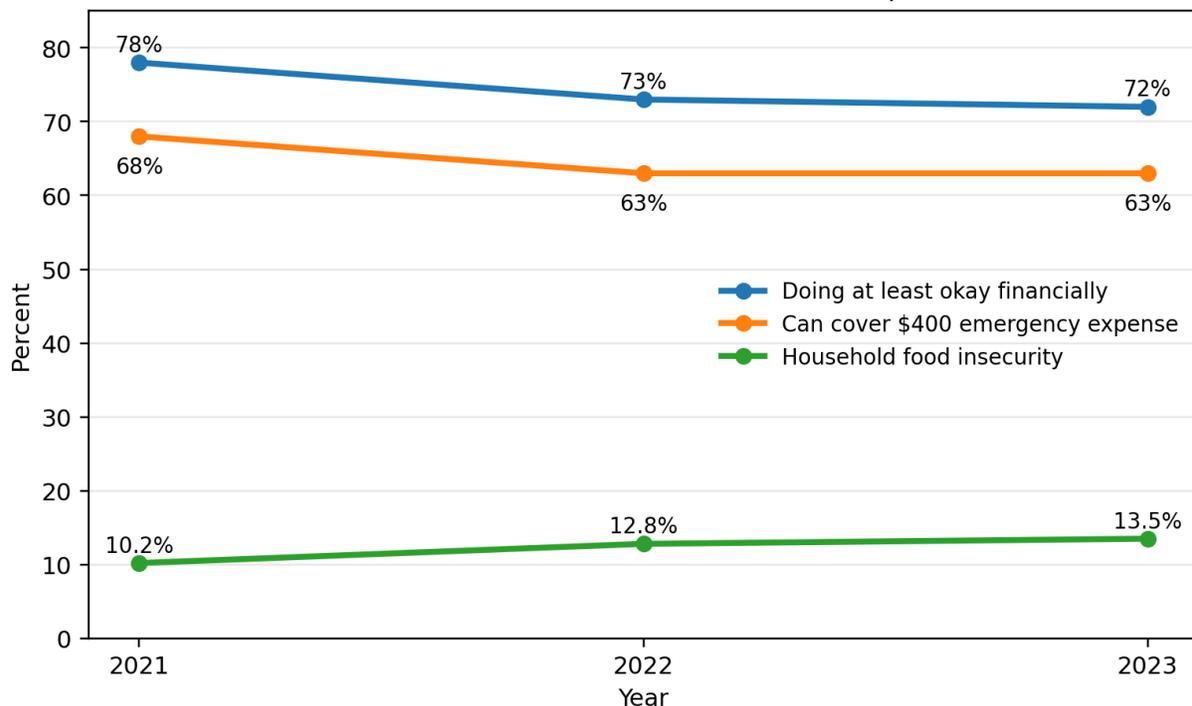Selected U.S. household stress indicators, 2021-2023

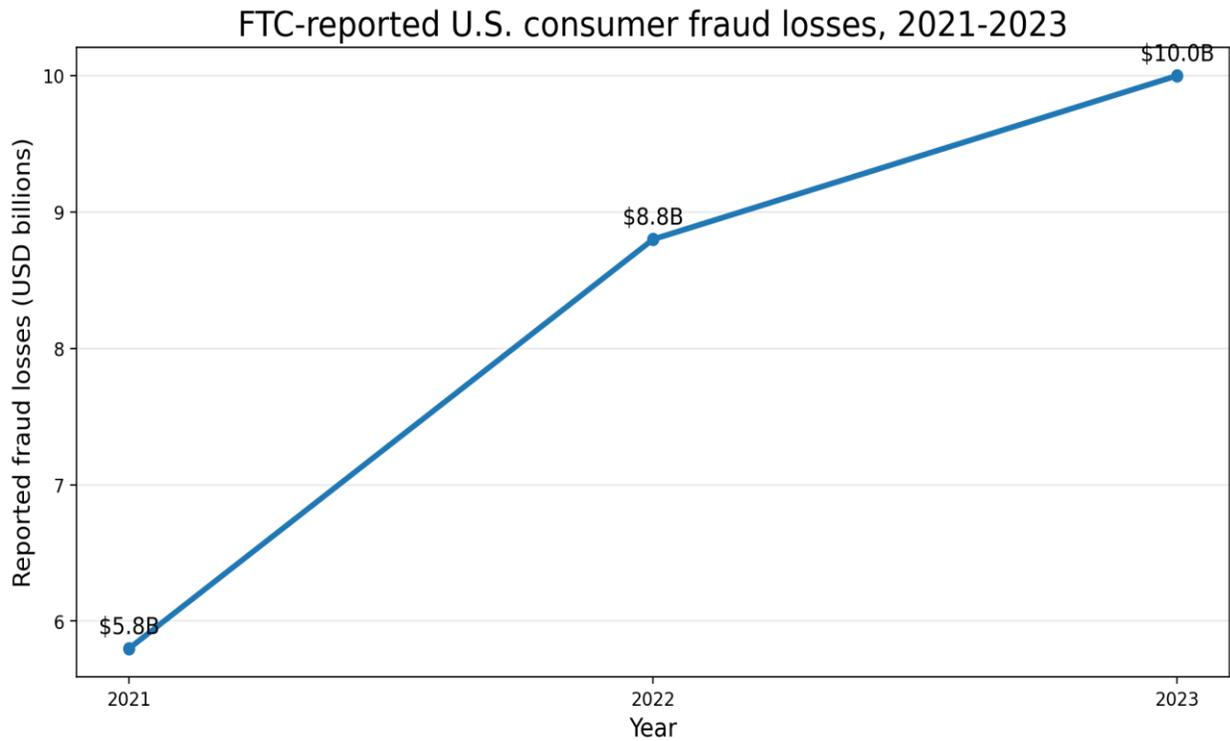*Figure 1. Selected U.S. household stress indicators, 2021-2023.*

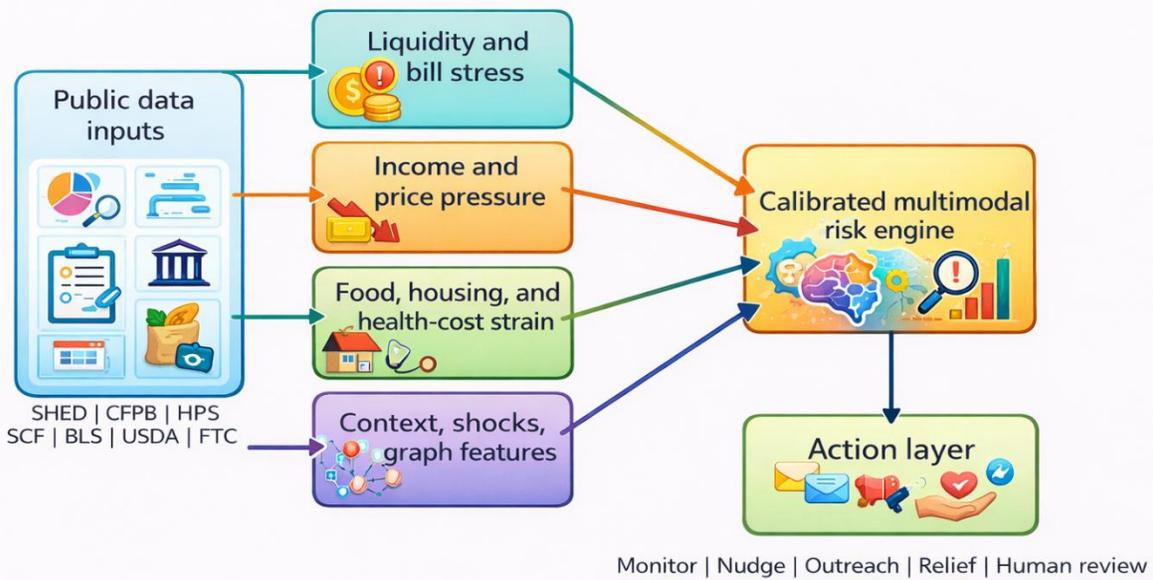*Figure 2. FTC-reported U.S. consumer fraud losses, 2021-2023.*

Figure 3. Proposed AI framework for forecasting household hardship and financial fragility.

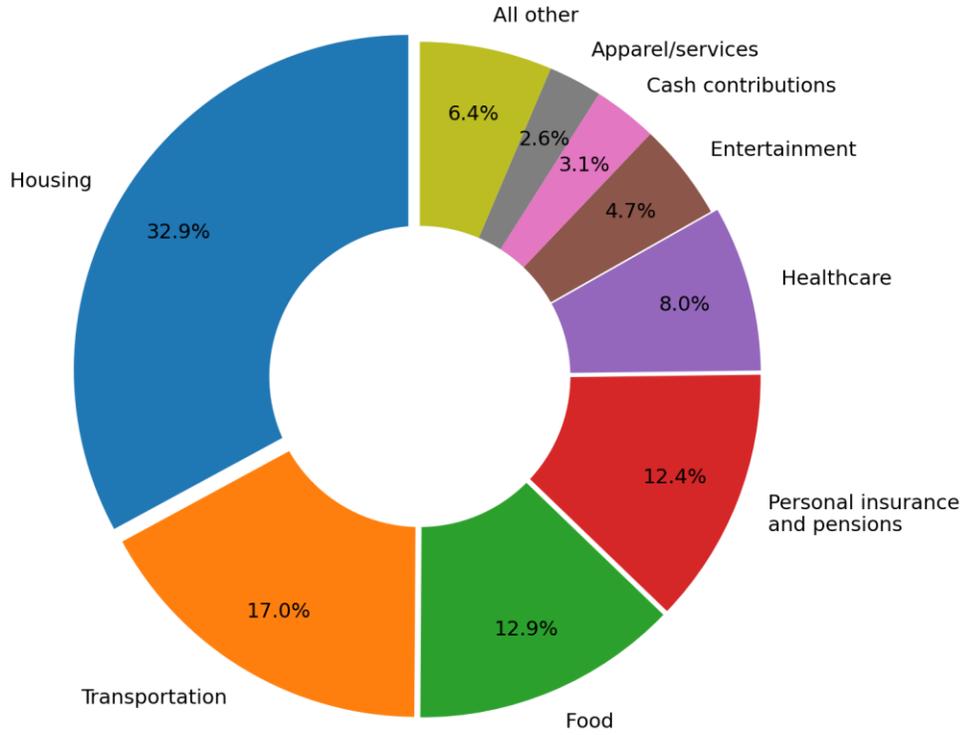## Household expenditure shares by major category, 2023

Figure 4. Household expenditure shares by major category, 2023.

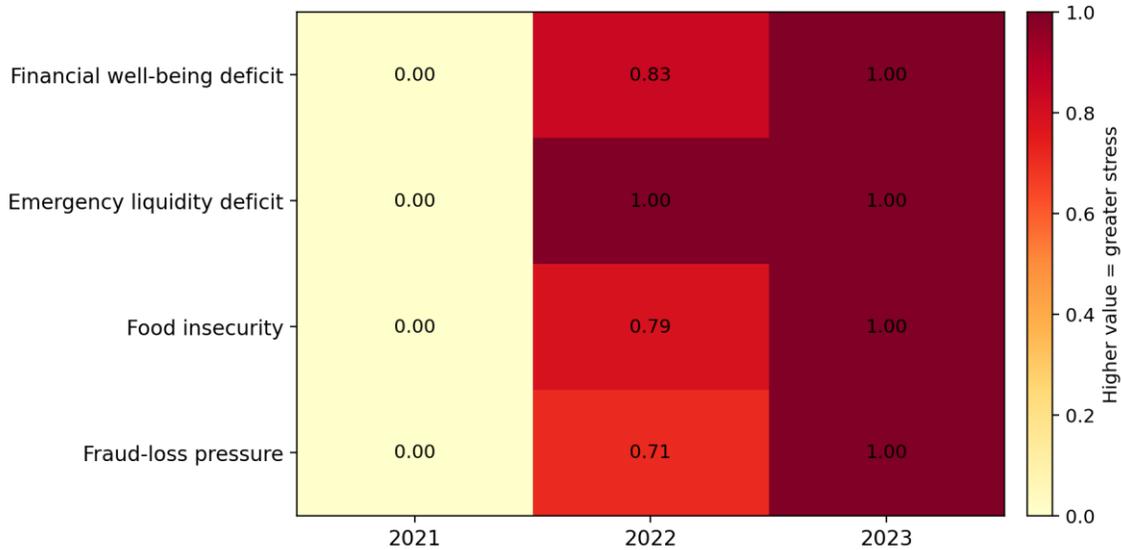## Directionally standardized household stress heat map, 2021-2023

Figure 5. Directionally standardized household stress heat map, 2021-2023.

Introduction

Forecasting household economic hardship has become a challenge for researchers, financial institutions, and public agencies because vulnerability now emerges through multiple channels. A family may remain employed and still become fragile because rent consumes an unsustainable share of income, credit card balances rise as prices increase, medical needs interrupt work, or fraud drains scarce liquidity. Traditional poverty thresholds, backward-looking delinquency signals, and coarse credit variables each capture part of this process, but none fully describes the household conditions that precede stress. When hardship is detected late, families are more likely to miss bills, defer care, rely on lose housing stability, or disengage from formal finance. Earlier detection matters for social welfare and financial stability.

Recent U.S. evidence makes the case for a more dynamic approach (Board of Governors of the Federal Reserve System, 2024; CFPB, 2023; Rabbitt et al., 2024). In the Federal Reserve's 2023 SHED report, 72 percent of adults said they were doing at least okay financially, down from the 2021 high of 78 percent. The share who could cover a $400 emergency expense using cash or its equivalent fell from 68 percent in 2021 to 63 percent in 2022 and 2023. USDA reported that household food insecurity rose from 10.2 percent in 2021 to 12.8 percent in 2022 and 13.5 percent in 2023. At the same time, the FTC reported consumer fraud losses of $5.8 billion in 2021, $8.8 billion in 2022, $10.0 billion in 2023, and $12.5 billion in 2024. These series do not measure the same phenomenon, yet together they reveal a hardening environment in which weaker liquidity, higher material hardship, and rising external shocks coexist.

This paper treats hardship as a layered concept with three related outcomes. Household economic hardship refers to contemporaneous strain, such as difficulty paying routine expenses, food insecurity, utility stress, medical postponement, or unmet rent and mortgage obligations. Financial fragility refers to weak short-run resilience, especially limited liquid savings, unstable payment capacity, and inability to absorb moderate shocks without harmful coping strategies. Vulnerability refers to the probability that a household with contextual exposures will experience deeper stress in the near future. The conceptual distinction matters because not every fragile household is currently in hardship, and not every household in hardship faces the same forward path. Prediction systems must therefore estimate levels, transitions, and persistence.

Artificial intelligence and predictive analytics are well suited to this task when used carefully. Machine learning can absorb nonlinear relationships, interaction effects, threshold behavior, and heterogeneous pathways across labor status, debt mix, housing tenure, family structure, geography, and inflation exposure. Natural language and representation methods can enrich structured data with local policy context, fraud exposure, or neighborhood service scarcity. Graph-based methods can model shared vulnerabilities across communities, employer sectors, and credit environments. Yet technical power alone is insufficient. Forecasts about hardship touch fairness, privacy, explainability, and access to essential services. A model that predicts who will struggle, but does so opaquely or in ways that intensify exclusion, can worsen rather than reduce vulnerability.

The goal of this manuscript is to develop an empirically grounded framework for forecasting household hardship using transparent national data sources. The study synthesizes the literature with broader research on financial fragility, well-being measurement, multidimensional hardship, household balance sheets, and interpretable AI. It proposes a multimodal modeling strategy that combines survey-based household features, macroeconomic context, expenditure pressures, and shock indicators into a calibrated forecasting system. Rather than claiming access to proprietary bank transaction data, the paper intentionally shows how much can be achieved with transparent public sources. This design supports reproducibility and broader policy use.

The rest of the paper proceeds as follows. The literature review examines foundational work on financial fragility, household vulnerability measurement, and machine learning for risk detection. The methodology section specifies data construction, target variables, feature engineering, modeling, validation, fairness assessment, and deployment logic. The discussion interprets the proposed system in light of observed U.S. trends through 2024 and explains how predictive models can guide intervention, consumer protection, and targeting. The paper concludes that AI should not replace social judgment, but it can materially improve the timing and quality of decisions when household hardship is treated as a dynamic forecasting problem.

Literature Review

Research on household hardship and fragility has developed along several strands that are increasingly convergent. The first strand defines fragility as limited capacity to absorb shocks. Lusardi, Schneider, and Tufano's (2011) seminal work on financially fragile households shifted attention from balance-sheet levels alone to households' ability to raise emergency funds quickly. Their argument remains influential because it framed fragility as an applied stress concept, not merely a static asset measure. Later studies connected fragility to low financial literacy, weaker access to mainstream credit, medical cost exposure, and stress-related behavioral responses. This literature established that apparently modest liquidity thresholds can reveal broad vulnerability that traditional income categories conceal.

A second strand concerns multidimensional financial well-being. The CFPB's (2017a, 2017b) Financial Well-Being Scale and technical documentation advanced a psychologically and behaviorally informed measure that captures perceived control, day-to-

day capacity, resilience, and freedom of choice. Unlike narrow distress indicators, the scale recognizes that well-being includes both current functioning and future confidence. Related CFPB Making Ends Meet reports connected well-being to bill-paying difficulty, income volatility, coping strategies, and credit outcomes over time. The Federal Reserve's SHED reports contributed complementary measures of emergency liquidity, expense stress, medical deferral, banking status, and subjective financial condition. Together, these sources broadened the empirical vocabulary available to hardship forecasting and demonstrated that vulnerability is best seen through linked rather than isolated indicators.

A third strand focuses on material hardship more directly. Census Household Pulse Survey data, especially during and after the pandemic shock, enabled near-real-time monitoring of difficulty paying usual expenses, food insufficiency, housing insecurity, and mental stress. Meyer et al. (2024) used SHED and related data to examine hardship and transfers, showing that hardship indicators can diverge from income poverty and respond differently to policy supports. USDA food security reports added an essential consumption-based lens by documenting both food insecurity and very low food security, which capture degraded diets and more severe intake disruptions. This literature matters because it highlights the outcome space that predictive systems should prioritize. Households may not meet conventional poverty criteria and yet still skip meals, miss bills, or postpone care.

A fourth strand examines household finance through wealth, debt, and expenditure structures. Survey of Consumer Finances documents wide dispersion in liquid assets, debt holdings, and net worth, while BLS Consumer Expenditure data show how housing, transportation, food, and healthcare absorb large shares of budgets. New York Fed household debt and credit reports add insight into delinquency, balances, and credit transitions. Studies of rent burden, debt service, revolving credit dependence, and cost shocks demonstrate that vulnerability often emerges from the interaction of cash-flow pressure and low savings rather than from any single liability. The importance of cash-flow mismatch is especially clear in work on paycheck volatility, asset poverty, and expense smoothing. These studies imply that predictive models should integrate both stock and flow variables and pay close attention to timing mismatches.

A fifth strand comes from applied machine learning. Outside household hardship specifically, predictive analytics has been widely used in credit risk, fraud detection, healthcare triage, customer churn, and anomaly detection. The common lesson is that nonlinear models often outperform simpler baselines when the data generating process contains interaction effects, missingness patterns, and regime changes. Yet the best applied studies also show that raw predictive power is not enough. Calibration, temporal validation, subgroup performance checks, and interpretability are critical for deployment. Work on explainable AI, including SHAP-based feature attribution and monotonic gradient boosting, offers practical tools for domains where decisions have distributive consequences. Research on resilience prediction, poverty mapping, and household vulnerability in other settings further shows that machine learning can uncover heterogeneity missed by standard regression, especially when combined with contextual data.

Literature extends these themes into adjacent domains that are relevant for this paper's design. Ibrahim, Razib, Jahan, and Rahman (2022) emphasize predictive analytics as a framework for risk assessment under complex, interconnected shocks. Although their article focuses on climate risk and capital allocation, its broader contribution is methodological: it supports combining forward-looking indicators with heterogeneous exposure channels. Fahim, Ibrahim, Pritty, and Tania (2023) address algorithmic accountability in U.S. consumer fintech, underscoring fairness, governance, and financial stability concerns that are directly applicable when hardship models influence access to assistance or credit accommodations. Rasel, Ibrahim, Pritty, Fahim, and Jahan (2023) argue for multimodal graph neural approaches in mortgage default forecasting; the lesson transferable here is that borrower risk is embedded in relational and locational context, not only in individual financial snapshots. Fahim, Pritty, Ibrahim, and Tania (2024) discuss real-time payments and fraud, reminding us that exogenous financial losses can worsen liquidity pressure and should be treated as vulnerability amplifiers. Ibrahim et al. (2024) on AML risk management show how layered predictive systems can be structured with scoring, escalation, and governance modules, while Jahan et al. (2024) highlight early-warning analytics for detecting abnormal patterns in capital markets, offering a design template for continuous monitoring rather than one-time classification. Literature on forward-looking, data-centric assessment also continues to expand in adjacent applied domains. Ibrahim, Razib, and Rasel (2025) show how data analytics can strengthen ESG transparency by improving signal extraction, comparability, and disclosure-focused monitoring in corporate settings, reinforcing this paper's argument that predictive systems can turn fragmented indicators into more interpretable risk intelligence. Ibrahim, Fahim, Zadid, and Pritty (2025) extend data-driven risk analysis into household-facing climate-finance settings by linking insurance gaps, mortgage stress, and household credit exposure, which is relevant here because household hardship is often amplified by environmental and balance-sheet shocks that traditional income measures understate. Razib, Ibrahim, and Rasel (2025) further demonstrate that predictive analytics can optimize complex systems under resource constraints, and although their application is supply-chain performance, the transferable lesson is that integrated forecasting architectures can improve efficiency, prioritization, and resilience in hardship-monitoring systems as well.

Even the less directly related supplied studies contribute useful methodological cues. Work on big data for waste reduction, clean energy forecasting, grid stability, and retail inventory optimization demonstrates how predictive systems can fuse operational,

environmental, and behavioral features under uncertainty. Their relevance here is not substantive equivalence but transferable analytics: temporal feature engineering, scenario sensitivity, and interpretable decision support. The stronger contribution of these papers lies in encouraging cross-domain thinking about infrastructure, resilience, and system feedbacks. Household fragility is also an infrastructure problem in the sense that failures in wages, housing affordability, healthcare costs, payments, fraud protection, and local service access reinforce one another.

At the same time, the literature reveals several gaps. First, many studies still use single proxies such as income, credit score, or debt-to-income ratio to define risk, even though public data show that hardship is multidimensional. Second, much of the machine-learning literature in finance optimizes for default or repayment rather than human well-being. Third, researchers often evaluate prediction at one point in time instead of modeling transitions into and out of hardship. Fourth, fairness discussions are sometimes appended after model selection rather than built into target design, calibration, and decision thresholds. Fifth, many public-interest applications remain underdeveloped because proprietary transaction data are assumed necessary. This paper responds by proposing a transparent framework that uses rich public data, predicts multiple hardship states, and embeds governance from the beginning.

Overall, the literature suggests three conclusions. Hardship is best understood as a dynamic condition. Machine learning is useful when it augments rather than obscures economic reasoning. And public-sector or consumer-facing deployment requires interpretable, fairness-audited, policy-relevant models. These conclusions motivate the multimodal forecasting framework developed in the next section.

An additional literature stream studies neighborhood, demographic, and institutional context. Place matters because hardship is conditioned by local rents, commuting costs, employment composition, childcare access, broadband quality, and the density of financial alternatives. Research on neighborhood effects, banking deserts, and spatial inequality suggests that two households with similar incomes may face very different probabilities of hardship depending on where they live and which services are available nearby. This insight supports the use of geographic and graph-informed features in forecasting systems. It also cautions against purely individualistic risk models that ignore shared exposures such as local inflation, disaster vulnerability, or regional labor-market softening.

Another relevant strand concerns measurement quality. Survey-based hardship indicators are powerful, but they can be noisy because of recall bias, mode effects, and differences between subjective and objective conditions. The best empirical work therefore combines survey responses with administrative or contextual data and validates composite indices against downstream outcomes. In financial well-being research, this means connecting self-reported strain to missed payments, savings depletion, benefit take-up, delinquency transitions, or service interruptions. In machine-learning terms, it means moving from convenience labels toward triangulated outcomes. The need for reliable labels is especially important when forecasting rare but consequential events such as eviction risk, utility shutoffs, or sustained food insecurity.

Finally, the literature on responsible AI provides an essential normative foundation. Scholars and regulators increasingly argue that predictive systems in finance and public policy should be interpretable, contestable, and proportionate to the stakes of their use. Methods such as groupwise calibration checks, threshold analysis, counterfactual explanations, and documentation of data provenance are no longer optional niceties. They are part of the model itself because they shape how predictions are translated into action. For hardship forecasting, this means the preferred model is not necessarily the one with the highest raw area under the curve. It is the one that delivers stable performance, transparent reasoning, manageable error costs, and acceptable subgroup equity under realistic deployment conditions.

These observations also reinforce a practical point for publication-oriented research. The strongest contribution is unlikely to come from claiming a universal hardship formula. It comes from specifying a modular forecasting architecture that can be re-estimated as conditions change, new public datasets appear, and policy priorities shift. A transparent, updateable framework is more defensible than a one-off model fitted to a single episode. The framework remains empirically grounded, flexible, and practically scalable across institutions. under changing economic conditions.

Methodology

This study proposes a reproducible, public-data forecasting framework for identifying household economic hardship, financial fragility, and vulnerability in the United States using official data available through 2024. The design is intentionally multimodal. It combines household survey information, financial well-being measures, consumption and balance-sheet context, and external shock indicators. The empirical core draws on the Federal Reserve's SHED, the CFPB's Making Ends Meet series and Financial Well-Being Scale materials, the U.S. Census Household Pulse Survey, USDA food security reports, the Federal Reserve's 2022 Survey of Consumer Finances, BLS consumer expenditure and CPI statistics, and FTC fraud-loss summaries. Additional context may be appended from public local indicators such as rent burden, unemployment, and health-cost pressure, but the baseline framework is estimable even without private transaction data.

The unit of analysis depends on the modeling layer. At the micro level, the preferred unit is the surveyed household or adult respondent. At the meso level, household records are enriched with geography-linked context such as state, metro status, or local price pressure. At the temporal layer, repeated surveys and recurring public releases allow the construction of quarterly or annual forecasting windows. Because some public datasets are repeated cross-sections rather than true panels, the design combines direct individual-level prediction where feasible with pseudo-panel and hierarchical forecasting elsewhere. This blended structure is a practical compromise between realism and reproducibility.

Three target families are defined. The first is immediate hardship, a binary or ordinal outcome indicating current stress. Candidate labels include difficulty paying usual household expenses, inability to pay all bills in full, food insecurity, forgone medical treatment because of cost, utility or housing strain, and very low financial well-being. The second is short-run fragility, defined by insufficient liquid reserves, inability to handle a moderate emergency expense, lack of three months of emergency savings, or reliance on harmful coping strategies such as borrowing from high-cost sources or missing payments. The third is vulnerability, which is a forward-looking probability that the household will transition into deeper hardship within a defined horizon such as the next quarter or next year. In repeated cross-sections, vulnerability can be approximated using transition probabilities learned from adjacent survey waves and validated on later macro conditions.

Feature engineering proceeds in six blocks. The first block contains demographic and household-structure variables: age, education, race and ethnicity where permitted, household size, number of children, tenure status, marital status, disability status, and metro or nonmetro location. The second block captures income and employment: level, source mix, hours volatility, job loss risk, self-employment status, and benefit receipt. The third block captures liquidity and balance-sheet conditions: checking and savings access, emergency funds, revolving credit use, debt burden, housing cost burden, payment delinquencies where available, and self-reported ability to absorb shocks. The fourth block captures expenditure stress and essential consumption: food security, healthcare affordability, transportation costs, childcare strain, and reported inflation effects. The fifth block captures external shocks and exposure: fraud victimization prevalence, regional unemployment changes, disaster or displacement flags if available, and local rent pressure. The sixth block adds derived interaction features, such as renter status multiplied by local rent burden, presence of children multiplied by food insecurity pressure, or low savings multiplied by recent income volatility.

The modeling strategy is deliberately layered. As a transparent baseline, regularized logistic regression and ordinal regression are estimated first. These models clarify direction, monotonicity, and benchmark calibration. The primary predictive layer uses tree-based ensemble models, particularly gradient boosting and random forests, because they handle nonlinearities, mixed scales, and missing data well. For temporal nowcasting, sequence-aware models such as temporal gradient boosting with lagged features or recurrent architectures may be added when repeated observations are sufficient. For context-rich extensions, graph-based representation learning can connect households to neighborhoods, labor sectors, and shared exposure clusters. Graph edges are not interpreted as social ties; they represent structured similarity or common environments that may transmit shared risk. This graph layer is justified only when it improves calibration and subgroup performance, not merely when it raises complexity.

Outcome imbalance is expected, especially for severe hardship states. The methodology therefore uses weighted loss functions, threshold tuning for recall-sensitive applications, and precision-recall diagnostics in addition to ROC-based metrics. Performance is evaluated with area under the ROC curve, area under the precision-recall curve, Brier score, calibration slope, calibration-in-the-large, and decision-oriented metrics such as recall among households in the highest-need decile. Because public-interest use prioritizes early intervention, false negatives are often costlier than false positives. However, this asymmetry is not assumed universally; instead, threshold selection is linked to the intervention being considered, such as referral to counseling, eligibility outreach, hardship forbearance, or emergency grant screening.

Validation follows a temporal and subgroup-aware design. Training uses earlier observations and testing uses later periods to avoid leaking future conditions backward. Where only one cross-section exists for a data source, cross-validation is nested within temporal validation from other sources. Subgroup evaluation is performed across income strata, race and ethnicity where legally and ethically appropriate, age bands, parental status, disability status, renters versus owners, and metro status. Metrics are reported separately for each subgroup, and calibration is checked within groups, not only overall. Because hardship forecasting can influence access to support, the study treats calibration parity as at least as important as discrimination performance.

Interpretability is integrated at model design, estimation, and deployment. Global feature importance is examined through permutation measures and gain-based summaries, while local explanations use SHAP values or equivalent additive explanations. Partial dependence and accumulated local effect plots are used cautiously to avoid misleading extrapolation. Counterfactual analysis focuses on actionable drivers such as savings buffers, rent burden, debt service, and benefit access rather than immutable traits. Documentation includes model cards describing target definitions, data provenance, known blind spots, and appropriate decision uses. This responds directly to the algorithmic accountability concerns emphasized in the supplied fintech governance literature.

To make the framework concrete, the paper assembles a data calibration panel of national indicators observed from 2021 through 2023, supplemented by official expenditure context from 2023: adults doing at least okay financially, adults able to cover a $400 emergency expense, household food insecurity, and FTC-reported fraud losses. These series do not train the full micro model by themselves, but they are useful for validating directional coherence. A credible hardship model should increase predicted vulnerability when liquidity falls, food insecurity rises, and external financial shocks intensify. This calibration step links the machine-learning design to observable national stress conditions.

The deployment architecture is a three-stage decision system. Stage one produces continuous scores for immediate hardship, fragility, and vulnerability. Stage two calibrates those scores into risk bands using isotonic or Platt-style calibration, chosen by validation performance. Stage three maps each band to actions proportionate to the stakes: informational nudges, benefit-screening outreach, budgeting support, payment-flexibility offers, or human review for acute cases. The system is not designed for denial of essential services. Its public-interest purpose is early support, triage, and preventive intervention.

Several methodological cautions are essential. First, prediction is not causal inference; a variable that predicts hardship is not necessarily a valid intervention lever. Second, public data are uneven in frequency, sample design, and granularity, so hierarchical harmonization and careful survey weighting are required. Third, the framework should be stress-tested against macro regime changes because post-shock dynamics can alter relationships quickly. Fourth, fairness should be monitored continuously because subgroup disparities may emerge after deployment even when validation results appear acceptable. Within those constraints, the proposed methodology offers a pathway for forecasting household hardship with transparent data, reproducible code, and outputs.

A harmonization pipeline is required because source instruments differ in wording, timing, and weighting. Variables are first mapped into conceptual domains rather than forced into artificial one-to-one equivalence. For example, expense stress can be represented by separate but related measures from SHED, HPS, and Making Ends Meet. Latent-domain scores are then created using standardized transformations, survey-weighted normalization, and reliability checks. Missing data are handled with a combination of principled imputation and missingness indicators so that nonresponse itself can be modeled when informative. Survey weights are incorporated in estimation or post-stratification to preserve national representativeness where appropriate, and sensitivity analyses compare weighted and unweighted results.

For publication-quality robustness, the study also specifies ablation and stability tests. Models are re-estimated after removing contextual features, after removing subjective well-being items, and after excluding fraud or food-security variables to determine which information blocks drive predictive gains. Stability is examined across alternative forecast horizons and alternative target definitions, such as stricter hardship labels that require multiple simultaneous stress indicators. A model intended for policy use should not depend on one fragile specification. Instead, useful predictors should remain directionally stable and the resulting risk rankings should be reasonably persistent across plausible design choices.

Ethical review is embedded throughout. Personally identifiable information is not required for the baseline public-data implementation, and any future linkage to administrative records would require minimization, secure governance, and clear use limitations. Human oversight remains mandatory at the action stage, especially where predictions could influence hardship accommodations, debt collection pauses, or referrals to public assistance. In this design, models inform decisions; they do not make them autonomously.

Discussion

The value of the proposed framework becomes clearer when read against recent U.S. evidence. Across several official data systems, household conditions deteriorated from the unusually strong position reached in 2021. The SHED series shows that the share of adults doing at least okay financially slipped from 78 percent in 2021 to 73 percent in 2022 and 72 percent in 2023. The same source indicates that the share able to cover a $400 emergency expense using cash or its equivalent fell from 68 percent to 63 percent and then remained there. USDA's annual food security reports show that food insecurity rose from 10.2 percent of households in 2021 to 12.8 percent in 2022 and 13.5 percent in 2023. FTC-reported consumer fraud losses increased from $5.8 billion in 2021 to $8.8 billion in 2022 and $10.0 billion in 2023. The additional visuals also show that housing absorbed the largest expenditure share in 2023 and that the standardized heat map places 2023 at the highest stress intensity in the 2021-2023 panel. None of these indicators is identical, but together they depict a realistic hardship environment: lower liquidity, weaker perceived financial security, more material deprivation, and more external drains on household resources. A forecasting system that ignores any one of these layers is likely to understate risk.

One important implication is that household hardship should be modeled as a system of interacting constraints. An income measure alone cannot explain why some households with similar earnings diverge sharply in outcomes. One renter may experience severe stress because local rent inflation, childcare cost, and commuting needs compress discretionary spending and eliminate savings. Another household may appear stable until fraud, job-hour volatility, or medical bills create a rapid liquidity shock. In this

sense, hardship is not simply low income. It is a dynamic mismatch between obligations, buffers, and shocks. Predictive analytics is useful precisely because it can learn nonlinear combinations of these elements and estimate when multiple moderate pressures cross into acute strain.

The proposed three-outcome structure also has practical benefits. Immediate hardship identifies who is already struggling now. Fragility identifies who can be tipped into hardship by a moderate shock. Vulnerability identifies who is likely to deteriorate soon even if present conditions are not yet severe. This distinction matters for intervention design. A household in immediate hardship may need emergency food support, utility assistance, or immediate payment flexibility. A fragile but not yet distressed household may benefit more from savings nudges, fee relief, benefit enrollment outreach, or fraud prevention assistance. A vulnerable household with rising predicted risk may call for monitoring and low-intensity support before hardship becomes visible in missed payments or food insecurity. Without separating these states, institutions often wait until the most damaging indicators appear.

The data calibration panel proposed in this paper illustrates how directional validation can support model credibility. If emergency liquidity declines while food insecurity and fraud losses rise, a well-calibrated hardship model should, all else equal, signal higher vulnerability in similar households. This is not proof of causal structure, but it is an important coherence test. Models intended for policy environments should behave sensibly as known macro pressures intensify. Too often, machine-learning studies report headline accuracy without asking whether predictions move in the right direction when the real world becomes harder. Directional realism is especially important for public-interest forecasting because stakeholders must trust that the model is capturing plausible hardship pathways rather than statistical artifacts.

A second major implication concerns data architecture. Many researchers assume that high-frequency bank transaction data are the only route to serious hardship prediction. That assumption is overstated. Proprietary transaction feeds are valuable, but public data already contain rich signals on liquidity, perceived well-being, expense strain, food insecurity, housing burden, fraud exposure, and demographic heterogeneity. SHED and Making Ends Meet are particularly useful because they combine subjective and objective financial indicators. HPS contributes timelier stress measures, while SCF and BLS help anchor the structure of balance sheets and expenditures. USDA adds a concrete consumption-based hardship measure. For publication and policy, this public-data orientation is a strength. It improves transparency, replicability, and comparability across studies.

That said, the paper does not argue that survey-based models are sufficient for every operational use. Rather, it proposes a layered strategy. Public data can generate baseline models, identify major risk pathways, and support population-level early warning. Institutions with access to lawful administrative or transaction data can then fine-tune the architecture for narrower use cases. The conceptual model should remain the same: combine current strain, resilience capacity, and forward shock exposure; validate temporally; audit subgroup performance; and use explanations that people can understand. In other words, private data may improve resolution, but they should not alter the ethical and measurement principles.

The discussion also highlights why explainability is central rather than optional. Forecasts about hardship may influence outreach, forbearance, collections strategy, or referrals to assistance programs. People affected by these systems deserve more than a number. They deserve to know whether risk is being driven by low savings, rising housing cost burden, food insecurity, unstable work hours, or some interaction among these factors. Explanation matters for three reasons. First, it allows decision-makers to choose interventions matched to the problem. Second, it helps detect spurious signals, such as a model overreacting to geography because geography is proxying for unmeasured structural disadvantage. Third, it supports contestability: when a household's circumstances change, explanations can show what would need to improve for risk to fall.

Fairness deserves equal attention. Hardship is not distributed randomly. Lower-income households, families with children, some racial and ethnic groups, renters, and people with disabilities are more likely to face economic stress. A purely predictive system may reproduce these inequalities with high statistical accuracy. The question, therefore, is not whether subgroup differences should exist in the data; they clearly do. The question is how the model uses those differences and what actions follow. In a supportive system, higher predicted risk should trigger earlier help, not automatic exclusion or punitive treatment. This is why the action layer proposed here is deliberately assistance-oriented. Models should expand preventive capacity, not create a new form of digital triage that stigmatizes already burdened households.

The supplied literature on algorithmic accountability is particularly relevant here. Governance must begin with target construction. If the target is too narrow, such as default alone, the model may miss households suffering serious well-being losses before any formal delinquency appears. If the target is too broad or vague, interventions become noisy and expensive. A balanced approach is to define outcomes around meaningful hardship thresholds that correspond to plausible actions. Examples include repeated difficulty paying bills, inability to cover an emergency expense, low financial well-being, food insecurity, and transitions into more severe stress. Once targets are carefully defined, fairness audits can focus on whether models are equally calibrated across groups and whether error burdens are acceptable relative to the interventions triggered.

Another key insight is that hardship forecasting should be operationally linked to prevention windows. Some signals are slow moving, such as educational attainment or baseline wealth. Others change quickly, such as food insecurity, hours volatility, price stress, or fraud exposure. The highest-value models will likely combine both. Slow variables locate structural vulnerability; fast variables indicate imminent strain. When these are fused, systems can distinguish chronic fragility from emerging crisis. This matters because chronic and acute hardship often require different responses. Chronic fragility may call for benefits navigation, debt restructuring, or recurring budget support, whereas acute shocks may require temporary liquidity relief, replacement of stolen funds, or short-term payment accommodation.

A related issue is temporal instability. The relationship between predictors and hardship can change when inflation falls, labor markets cool, or public transfers expand or expire. This is why the methodology emphasizes temporal validation and re-estimation. Models should be treated as maintained infrastructure, not one-time products. The period from 2021 to 2024 is instructive because it combines post-pandemic labor-market strength with inflation pressure, normalization of savings behavior, and changing transfer regimes. Any forecasting model estimated in such a setting must be robust to shifting coefficients, altered baselines, and changing error tradeoffs. Publication-quality research should therefore report not just average performance, but performance drift across adjacent periods.

The broader policy relevance is substantial. Federal, state, and local agencies could use calibrated hardship forecasts to target outreach before utility arrears, eviction filings, or food insecurity deepen. Nonprofits could prioritize case management to households facing the steepest near-term deterioration. Financial institutions could identify customers for hardship accommodations, fee waivers, or fraud remediation before relationships collapse. Employers and benefits administrators could monitor aggregate vulnerability among workforces without requiring intrusive surveillance. Researchers could use the framework to evaluate whether interventions meaningfully change predicted trajectories, even before longer-run outcomes fully materialize. In each case, the emphasis should remain preventive and supportive.

There is also an important bridge to financial stability. Household hardship is often framed as a micro-level welfare concern, but widespread fragility can aggregate into macro stress through reduced consumption, higher delinquencies, greater reliance on emergency services, and weaker local resilience. Predictive systems that identify concentrated vulnerability can therefore serve both distributive and stabilization goals. The exact domain differs, but the structural insight is similar: when many units become simultaneously fragile, downstream institutions and markets become less stable too.

Finally, the paper suggests a different philosophy of household analytics. Instead of asking which households are "bad risks," it asks which households face converging pressures that make hardship more likely and more damaging. That framing is ethically preferable and analytically richer. It shifts attention from sorting people into static categories toward understanding trajectories, buffers, and intervention opportunities. Forecasting, in this view, is not a mechanism for exclusion. It is a mechanism for timing help more intelligently.

For publication, this repositioning matters. The contribution is not merely technical novelty. It is the combination of multidimensional hardship theory, transparent public data, explainable machine learning, and actionable governance. That combination makes the framework more suitable for real institutions that must balance prediction, accountability, and human welfare. As economic pressures evolve, such systems will need updating, local calibration, and ongoing scrutiny. But the central result is durable: household hardship can be forecasted more effectively when AI is used to integrate liquidity, consumption stress, balance-sheet fragility, and shock exposure into one interpretable predictive architecture.

A concrete intervention ladder further clarifies the model's value. Low-risk households would receive only general financial education or fraud-awareness messaging. Moderate-risk households could be offered budgeting tools, benefit-screening prompts, savings incentives, or reminders about payment plans before arrears accumulate. High-risk households would be routed to human review, hardship accommodations, emergency assistance, or community referrals. Very high-risk cases, especially those combining food insecurity, inability to pay bills, and zero emergency buffers, would trigger expedited outreach. This stepped design is important because the cost of intervention should match the cost of error. It also reduces the danger that algorithmic systems become overbearing by subjecting too many households to intensive action.

The framework is also useful for subgroup-sensitive policy analysis. Parents of young children, renters in high-cost areas, older adults with medical spending pressure, and workers facing irregular hours may reach hardship through different pathways even when their overall predicted risk is similar. Explainable machine learning can reveal those pathway differences without requiring separate models for every subgroup. For example, the same risk score may be driven mainly by food insecurity for one household and by fraud-related loss plus low savings for another. That distinction changes the appropriate remedy. In publication terms, this supports a richer discussion of heterogeneity than is possible with one-coefficient-average models.

From a research perspective, the proposed design also invites comparative evaluation. Future studies can test whether public-data models approximate the rankings produced by transaction-data systems, whether graph-informed context materially improves calibration, and whether hardship prediction outperforms standard credit-risk proxies in identifying households needing support.

These are tractable questions with immediate practical importance. The present paper therefore contributes not only a framework, but also an agenda for evidence accumulation in a field that needs more transparent, household-centered forecasting. The approach is technically feasible, ethically defensible, and administratively relevant for scalable preventive practice. across diverse U.S. settings. before hardship becomes entrenched. nationally.

Conclusion

Household economic hardship is increasingly generated by the interaction of liquidity shortages, rising essential costs, fragile balance sheets, and external shocks rather than by income alone. This paper argued that artificial intelligence and predictive analytics can improve the forecasting of hardship, financial fragility, and vulnerability when models integrate multidimensional public data and are embedded in transparent governance. Using U.S. evidence from the Federal Reserve, CFPB, Census, USDA, BLS, SCF, and FTC, the study showed that weakened emergency liquidity, lower perceived financial well-being, higher food insecurity, and rising fraud losses provide a compelling empirical backdrop for early-warning systems. The proposed framework combines interpretable baselines, nonlinear machine-learning models, temporal validation, subgroup fairness checks, and calibrated decision layers that map risk to supportive interventions. Its central contribution is conceptual and practical: hardship forecasting should move beyond static proxies toward dynamic, explainable, and policy-relevant systems that identify not only who is struggling now, but who is likely to struggle next. AI is most valuable here not as an automated gatekeeper, but as a disciplined tool for earlier, fairer, and more targeted prevention. It also offers a reproducible foundation for future empirical testing, local calibration, and supportive policy deployment across diverse household contexts.

Limitations and Future Directions

Several limitations should temper interpretation. First, the framework relies heavily on public surveys and aggregate indicators, which are rich but not equivalent to real-time household transaction records. Repeated cross-sections can approximate transitions, yet they do not fully observe individual trajectories. Second, survey questions differ across SHED, HPS, Making Ends Meet, and SCF, so harmonization inevitably introduces measurement choices that may affect results. Third, some predictors of hardship, such as local informal support networks or intra-household bargaining stress, are only weakly captured in public data. Fourth, prediction does not establish causation; high-importance features may identify exposure rather than actionable levers. Fifth, fairness assessment is necessary but not sufficient, because acceptable error tradeoffs depend on the intervention and institutional setting.

Future research should extend this framework in four directions. The first is data linkage, ideally combining public survey data with privacy-preserving administrative or transaction records to improve temporal resolution. The second is geographic refinement, including neighborhood service access, rent microdata, disaster exposure, and local policy differences. The third is longitudinal evaluation of interventions to test whether model-informed outreach actually reduces hardship. The fourth is methodological, including conformal prediction, uncertainty-aware forecasting, and dynamic recalibration under macro regime shifts. These extensions would strengthen both scientific validity and operational usefulness while preserving the paper's public-interest emphasis on transparency and human-centered support.

A further limitation is page-level generalization. National models may hide state, metro, and rural differences in prices, labor demand, and access to assistance. Researchers should therefore test local recalibration before operational use. Another limitation concerns label timing. Hardship indicators often appear after coping has already begun, which can bias training toward later-stage stress. Better prospective labels and shorter observation windows would improve the forecasting of early deterioration rather than visible crisis alone. Better uncertainty reporting and stakeholder review are also essential going forward.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1]. Arman, M., & Fahim, A. S. M. (2023). AI revolutionizes inventory management at retail giants: Examining Walmart's U.S. operations. Journal of Business and Management Studies, 5(6), 145-148. https://doi.org/10.32996/jbms.2023.5.6.15

[2]. Arman, M., Hasan, M. N., & Rasel, I. H. (2024). Clean energy transition in USA: Big data analytics for renewable energy forecasting and carbon reduction. Journal of Management World, 2024(3), 192-206. https://doi.org/10.53935/jomw.v2024i4.1196

[3]. Arman, M., Rasel, I. H., Razib, M. N. H., & Fahim, A. S. M. (2024). Big data and machine learning for sustainable waste reduction. Journal of Posthumanism, 4(2), 448-467. https://doi.org/10.63332/joph.v4i2.3361

[4]. Barocas, S., Hardt, M., & Narayanan, A. (2023). Fairness and machine learning: Limitations and opportunities. MIT Press.

[5]. Board of Governors of the Federal Reserve System. (2023a). Changes in U.S. family finances from 2019 to 2022: Evidence from the Survey of Consumer Finances. https://www.federalreserve.gov/publications/files/scf23.pdf

[6]. Board of Governors of the Federal Reserve System. (2023b). Report on the economic well-being of U.S. households in 2022. https://www.federalreserve.gov/publications/files/2022-report-economic-well-being-us-households-202305.pdf

[7]. Board of Governors of the Federal Reserve System. (2024). Report on the economic well-being of U.S. households in 2023. https://www.federalreserve.gov/publications/files/2023-report-economic-well-being-us-households-202405.pdf

[8]. Consumer Financial Protection Bureau. (2017a). CFPB financial well-being scale: Scale development technical report. https://www.consumerfinance.gov/data-research/research-reports/financial-well-being-technical-report/

[9]. Consumer Financial Protection Bureau. (2017b). A guide to using the CFPB financial well-being scale. https://www.consumerfinance.gov/data-research/research-reports/financial-well-being-scale/

[10]. Consumer Financial Protection Bureau. (2023). Making ends meet in 2023: Insights from the Making Ends Meet survey. https://files.consumerfinance.gov/f/documents/cfpb_making-ends-meet-in-2023_report_2023-12.pdf

[11]. Consumer Financial Protection Bureau. (2024). Making ends meet in 2024: Insights from the Making Ends Meet survey. https://files.consumerfinance.gov/f/documents/cfpb_making-ends-meet_2024-11.pdf

[12]. Desmond, M. (2016). Evicted: Poverty and profit in the American city. Crown.

[13]. Dynan, K. (2012). Is a household debt overhang holding back consumption? Brookings Papers on Economic Activity, 2012(1), 299-362.

[14]. Fahim, A. S. M., Ibrahim, M., Pritty, A. A., & Tania, T. A. (2023). Algorithmic accountability in U.S. consumer FinTech: Governance mechanisms for credit risk, fair lending, and financial stability. Journal of Economics, Finance and Accounting Studies, 5(4), 80-93. https://doi.org/10.32996/jefas.2023.5.4.8

[15]. Fahim, A. S. M., Pritty, A. A., Ibrahim, M., & Tania, T. A. (2024). Real-time payments and real-time fraud: A U.S. FinTech risk framework for RTP rails and consumer protection. Journal of Economics, Finance and Accounting Studies, 6(6), 134-149. https://doi.org/10.32996/jefas.2024.6.6.11

[16]. Federal Trade Commission. (2024). As nationwide fraud losses top $10 billion in 2023, FTC steps up efforts to protect the public. https://www.ftc.gov/news-events/news/press-releases/2024/02/nationwide-fraud-losses-top-10-billion-2023-ftc-steps-efforts-protect-public

[17]. Garbero, A., Jäckle, R., & Salvucci, V. (2022). Predicting household resilience with machine learning. Empirical Economics, 63, 1319-1350. https://doi.org/10.1007/s00181-022-02199-4

[18]. Ibrahim, M., Razib, M. N. H., Jahan, N., & Rahman, M. M. (2022). Climate risk, financial stability, and global capital allocation: A predictive analytics approach to assessing climate-related financial risk in international investment markets. Journal of Business and Management Studies, 4(4), 264-276. https://doi.org/10.32996/jbms.2022.4.4.34

[19]. Ibrahim, M., Fahim, A. S. M., Zadid, M. U., & Pritty, A. A. (2025). FinTech for climate resilience: Measuring insurance gaps, mortgage stress, and household credit risk in the United States. Journal of Economics, Finance and Accounting Studies, 7(4), 190-205. https://doi.org/10.32996/jefas.2025.7.4.15

[20]. Ibrahim, N. H. R., & Rasel, I. H. (2025). The role of data analytics in enhancing ESG transparency in the corporate sector of Bangladesh. Global Journal of Engineering and Technology Advances, 22(1), 81-93. https://doi.org/10.30574/gjeta.2025.22.1.0245

[21]. Ibrahim, M., Mahmud, S., Zadid, M. U., Jahan, N., Rahman, M. M., & Fahim, A. S. M. (2024). AI-driven predictive analytics framework for anti-money laundering risk management and financial infrastructure protection in U.S. banking systems. Journal of Economics, Finance and Accounting Studies, 6(1), 155-166. https://doi.org/10.32996/jefas.2024.6.6.12

[22]. Jahan, N., Pritty, A. A., Ibrahim, M., Zadid, M. U., Fahim, A. S. M., & Mahmud, S. (2024). Machine learning-driven early warning analytics for identifying market manipulation, irregular trading activity, and suspicious market signals in U.S. stock markets. Journal of Computer Science and Technology Studies, 6(2), 257-283. https://doi.org/10.32996/jcsts.2024.6.2.26

[23]. Levantesi, S., Orlando, T., & Russo, E. (2021). Machine learning and financial literacy: An exploration of the role of financial knowledge in household financial fragility. Journal of Risk and Financial Management, 14(3), Article 120.

[24]. Lusardi, A., Schneider, D., & Tufano, P. (2011). Financially fragile households: Evidence and implications. Brookings Papers on Economic Activity, 42(1), 83-150.

[25]. Meyer, B. D., Wu, D., Mooers, V., & Medalia, C. (2024). Poverty, hardship, and government transfers. National Bureau of Economic Research Working Paper No. 33052. https://www.nber.org/papers/w33052

[26]. Midões, C., & Seré, M. (2022). An analysis of household financial vulnerability under income shocks in Europe. Social Indicators Research, 162, 1327-1357. https://doi.org/10.1007/s11205-021-02811-7

[27]. Morduch, J., & Schneider, R. (2017). The financial diaries: How American families cope in a world of uncertainty. Princeton University Press.

[28]. New York Fed. (2024). Quarterly report on household debt and credit, 2024 Q3. https://www.newyorkfed.org/medialibrary/Interactives/householdcredit/data/pdf/HHDC_2024Q3.pdf