Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Credit-reporting beliefs, affordability-related use and self-reported late payment among U.S. buy-now-pay-later users: a cross-sectional analysis of the 2025 SHED

Kaibo Tang

U300863g@ecs.osaka-u.ac.jp

The University of Osaka

Research Article

Keywords: Buy now pay later, household finance, consumer credit, affordability, credit reporting, late payment

Posted Date: August 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10409150/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Buy now, pay later (BNPL) credit can relax immediate payment constraints but creates future obligations shaped by information, product use and household resources. This study compares credit-reporting beliefs and affordability-related use as correlates of self-reported late payment. We analyzed the public 2025 U.S. Survey of Household Economics and Decisionmaking and its analytical domain of 2,004 past-year BNPL users, 487 of whom reported any late payment. Survey-weighted logistic models were summarized using standardized probabilities and risk differences. Only 14.5% of users rejected both credit-reporting statements in accordance with the survey-time benchmark; 31.4% were uncertain and 54.1% accepted at least one statement. In a common expanded specification, benchmark-discordant agreement differed by + 0.2 percentage points from the benchmark-consistent category (95% confidence interval − 5.3 to + 5.7). Relative to their reference categories, standardized risk differences were + 16.5 points for the affordability motive (95% confidence interval + 12.0 to + 20.9), + 10.9 for groceries-or-medical use (+ 6.2 to + 15.6), + 8.9 for three or more versus zero or one other-purpose categories (+ 2.8 to + 15.1), and + 8.1 for financial strain (+ 3.7 to + 12.5). Because the affordability motive concerns the most recent purchase whereas the outcome records any late payment in the preceding year, these are respondent-level cross-sectional contrasts, not transaction-matched or causal effects. Taken together, the estimates suggest that survey-benchmark reporting beliefs and the affordability context of BNPL use provide different descriptive information about self-reported late payment.

JEL Classification: D14; D83; G41; G51

1 Introduction

Buy now, pay later (BNPL) credit has moved short-term borrowing into the ordinary retail checkout process. A typical pay-in-four arrangement divides a purchase into an initial payment and three later installments, often without stated interest. This structure can help households align expenditure with expected income and may extend payment flexibility to consumers with limited access to conventional credit. It also creates multiple future obligations that can overlap with other claims on household cash flow. In the 2025 Survey of Household Economics and Decisionmaking (SHED), 16% of U.S. adults reported using BNPL during the previous year, and slightly more than one quarter of those users reported at least one late payment (Board of Governors of the Federal Reserve System 2026a). Administrative evidence further indicates that many BNPL borrowers hold other unsecured debt and use more than one provider, while some installments are paid with credit cards rather than current account balances (Consumer Financial Protection Bureau 2025; Guttman-Kenney et al. 2023).

From a household-finance perspective, late payment may arise through several economically distinct margins. Consumers may differ in what they believe the product offers, including whether using BNPL or paying on time produces benefits through the credit-reporting system. They may also differ in how frequently and for what purposes they use the product, which determines the number and timing of payment obligations. Finally, consumers facing tight budgets may be unable to meet an installment even

when they understand its terms and intend to pay. Information, usage exposure and payment capacity are therefore related but not interchangeable dimensions of BNPL use. Distinguishing them matters because a common information pattern does not necessarily identify the users most likely to report a late payment, and a conditional payment gradient does not by itself identify the mechanism that generated it.

Economic research on consumer financial protection provides a basis for examining the information margin. Credit contracts contain attributes that differ in salience, timing and the effort required to understand them. Consumers may devote limited attention to attributes that are remote from the purchase decision, while firms and product interfaces can make some costs or benefits more visible than others (Campbell et al. 2011; Gabaix and Laibson 2006). Deferred and installment payment formats can also alter perceived affordability and purchasing behavior by separating immediate consumption from the full payment stream (Maesen and Ang 2025; Siemens 2007). These arguments make product-specific beliefs economically relevant, but they do not imply that a particular survey response must be associated with late payment in a predetermined direction. Believing that on-time BNPL payment improves a credit score could increase the perceived return to paying on time. The same belief could also increase demand for BNPL and hence the number of opportunities to be late.

Product-specific beliefs should also be distinguished from general financial literacy. Debt-literacy research links knowledge of interest, borrowing costs and debt accumulation to credit portfolios and overindebtedness, although the strength and interpretation of these relationships vary across measures and institutional settings (Disney and Gathergood 2013; Lusardi and Mitchell 2014; Lusardi and Tufano 2015). Recent BNPL studies have examined financial literacy, responsible financial behavior, perceived risk and product adoption (Gerrans et al. 2022; Powell et al. 2023; Relja et al. 2024; Wang and Wong 2026). The two SHED items used here are narrower. They ask whether BNPL helps a user build a credit history and whether making payments on time helps a credit score. They do not measure knowledge of fees, installment totals, concurrent obligations, numeracy or general debt management. Their association with late payment must therefore be evaluated as an empirical question about specific credit-reporting beliefs rather than as a general test of financial capability.

Credit reporting is a useful setting for this question because the institutional environment has been both consequential and changing. BNPL obligations historically appeared only incompletely in U.S. credit files, limiting lender visibility into concurrent borrowing and reducing the extent to which on-time BNPL payment could contribute to a consumer's credit record (Consumer Financial Protection Bureau 2025; Zeballos Doubinko and Akana 2023). Evidence from a large digital platform outside the United States indicates that introducing credit reporting changed BNPL use and repayment among borrowers with previous defaults, showing that actual credit-record consequences can alter behavior (Dong et al. 2025). A policy that changes those consequences is not equivalent, however, to a consumer's contemporaneously measured belief about them. Provider practices may vary, reporting arrangements may evolve, and prior product experience may shape survey responses. These features require a time-specific and provider-agnostic interpretation of the SHED items.

The BNPL literature has documented both demand-side and resource-related patterns. Product demand varies with consumer and transaction characteristics (Chen et al. 2024), while BNPL adoption has been associated with subsequent overdraft and credit-card interest charges in transaction data (deHaan et al. 2024). U.S. survey and administrative evidence consistently places BNPL use within broader household balance sheets: users often have lower liquid resources, carry other unsecured obligations and report signs of financial fragility (Consumer Financial Protection Bureau 2025; Donou-Adonsou and Leslie-Piper 2025). Using the 2023 SHED, Hayashi and Routh (2025) showed that financial constraints were markedly more common among BNPL users who reported paying late. Existing research has therefore established that information, product demand and household resources are all relevant to BNPL. What remains less clear is how a newly measured, product-specific information variable compares with affordability-related use and contemporaneous financial strain when each is related to the same payment outcome among current users.

This study addresses that question using the public-use 2025 SHED. The analysis is restricted to the BNPL-user analytical domain, comprising 2,004 adults who reported using BNPL during the previous year; 487 reported any late BNPL payment. Four questions organize the analysis. First, how were responses to the two credit-reporting statements jointly distributed? Second, were the resulting survey-benchmark categories associated with any late payment after socioeconomic adjustment? Third, within one expanded specification, how did those belief contrasts compare descriptively with an affordability motive, groceries-or-medical use, breadth across other purchase purposes and financial strain? Fourth, did the belief association differ according to strain? Survey-weighted models are reported with standardized probabilities and risk differences on an absolute scale.

The findings separate prevalence from association. Only 14.5% of users were in the benchmark-consistent category, while 54.1% showed benchmark-discordant agreement with at least one statement. In the common expanded model, the standardized risk difference for benchmark-discordant agreement was + 0.2 percentage points (95% confidence interval − 5.3 to + 5.7). The affordability-motive, groceries-or-medical, three-or-more versus zero-or-one breadth, and strain contrasts ranged from + 8.1 to + 16.5 points; the two-category breadth contrast was + 1.6 points. Exploratory within-model differences between risk differences for the resource-context indicators—affordability-related use measures and contemporaneous financial strain—and the benchmark-discordant-agreement risk difference are reported below with joint delta-method uncertainty. These model-dependent comparisons remain descriptive rather than causal. Together, the estimates suggest that information consistency and the resource context of BNPL use are empirically distinct correlates of any late payment.

The study contributes by linking newly available product-specific reporting beliefs to a payment outcome, making the belief construction auditable through the complete item distribution, and comparing belief and resource-context gradients on a common probability scale. The analysis remains descriptive: it does not identify a behavioral mechanism, a transaction-level late-payment probability or the effects of regulation.

2 Institutional setting and economic framework

2.1 BNPL credit and the survey-time credit-reporting benchmark

BNPL combines a retail payment interface with a short-duration credit contract. The consumer receives the good immediately, while the provider and merchant allocate payment across several dates. Many pay-in-four products advertise no interest when installments are paid as scheduled, but late payments can generate provider charges, overdraft or non-sufficient-funds fees, interruption of product access and other costs (Consumer Financial Protection Bureau 2022). The product can consequently serve different economic purposes. It may provide convenience to a consumer who could otherwise pay immediately, smooth a temporary mismatch between income and expenditure, or finance a purchase that the consumer reports being unable to afford through another available payment method.

Credit reporting adds a potential intertemporal benefit or cost to this contract. If on-time payments contribute to a credit record, timely repayment may carry value beyond avoiding an immediate fee. If missed payments are reported, the expected future cost of lateness may also increase. U.S. BNPL reporting practices, however, have varied across providers and have been changing. At the time of the October 2025 SHED, the Federal Reserve described the two positive credit-reporting statements in the questionnaire as inconsistent with prevailing arrangements at the three nationwide credit bureaus (Board of Governors of the Federal Reserve System 2026a). The present study therefore uses disagreement with both statements as the benchmark-consistent category for that survey time. It does not treat this category as a universal or permanent statement about every provider, contract or later reporting regime. Agreement is correspondingly described as benchmark-discordant agreement, not as a validated measure of general financial knowledge.

2.2 Perceived credit returns and usage exposure

A simple payment decision illustrates why the belief association need not have a unique sign. When an installment becomes due, a consumer weighs the current value of retaining liquid resources against the immediate and expected future consequences of paying late. A belief that on-time payment improves a credit score may increase the perceived future return to timely payment and could therefore be associated with a lower per-payment probability of lateness. Conversely, believing that BNPL helps build a credit history may increase the perceived value of using the product. If that belief is associated with more plans or more frequent purchases, it may increase the number of payment occasions over which at least one late payment can occur.

The outcome available in SHED is an extensive-margin event rather than a transaction-level rate. In a simplified setting with N_i conditionally independent payment opportunities and a constant per-opportunity probability q_i of lateness for consumer i,

Pr(Y_i = 1 | q_i, N_i) = 1 - (1 - q_i)^(N_i).

The same observed outcome can therefore result from a high per-payment probability, a large number of payment opportunities or both. SHED does not report the number of transactions, active plans, providers, installments or total BNPL amount. Other-purpose breadth records distinct categories rather than transaction frequency, so the analysis cannot separate within-transaction payment behavior from usage exposure.

Beliefs are also potentially endogenous to product experience. Provider-specific communication, marketing exposure and contract terms may influence both responses to the credit-reporting items and BNPL use. A prior late payment may lead a consumer to seek information or revise a belief, creating reverse updating. Moreover, the analysis conditions on having used BNPL, although adoption itself may depend on beliefs, financial resources and unmeasured preferences. If unmeasured credit demand, credit access, risk tolerance or prior product experience affects both entry into BNPL use and late payment, conditioning on use can attenuate, amplify or reverse the association observed within the user domain; its direction cannot be inferred from these data. Limited attention and shrouded product attributes remain possible explanations for how some beliefs arise (Campbell et al. 2011; Gabaix and Laibson 2006), but neither construct is measured in SHED. The reported coefficients are therefore not tests of limited attention or another behavioral mechanism.

2.3 Resource conditions, purchase motives and payment capacity

Payment capacity supplies a separate economic margin. Even when a consumer perceives the contract accurately and values future credit access, an installment may arrive when available resources are insufficient relative to other obligations. The SHED self-reported financial strain indicator distinguishes respondents who described themselves as “just getting by” or “finding it difficult to get by” from those who were “doing okay” or “living comfortably.” This item captures contemporaneous subjective resource pressure. It is not a direct measure of cash on hand, income volatility or a binding liquidity constraint, and its timing does not establish whether strain preceded BNPL use or followed payment problems.

The survey also provides measures closer to the context in which BNPL was used. The affordability motive identifies respondents who selected “Only way I could afford it” for their most recent BNPL purchase. Groceries-or-medical use covers groceries or food delivery and medical or veterinary expenses during the preceding year. To avoid mechanically counting those same categories twice, other-purpose breadth counts only clothing, travel, furniture or appliances, electronics and other purchases, grouped as zero or one, two, and three or more categories. These measures may reflect household resources, preferences, purchase size, access to other credit and prior BNPL experience; they do not establish necessity, a binding liquidity constraint or transaction frequency.

The expanded descriptive model places these measures alongside credit-reporting beliefs, socioeconomic characteristics and contemporaneous financial strain. Its purpose is comparative rather than structural. Reporting all focal contrasts from this common specification permits an economically

interpretable within-model comparison, but does not make the motive, purpose or strain gradients causal.

2.4 Empirical questions and interpretive boundaries

The empirical strategy is organized around distinct estimands rather than directional causal hypotheses. The principal socioeconomic-adjusted Model 2 estimates the cross-sectional association between the survey-benchmark belief categories and self-reported any late BNPL payment after adjustment for age, gender, race and ethnicity, education, household income and employment. Because a belief may be associated with both the perceived return to timely payment and usage exposure, no directional sign is imposed. Model-standardized predicted probabilities and risk differences are the principal absolute-scale outputs, with odds ratios reported as complementary relative measures.

The strain-conditioned Model 3 asks how the belief association appears after additionally conditioning on contemporaneous financial strain. The expanded model then estimates belief categories, other-purpose breadth, the affordability motive, groceries-or-medical use and strain in one specification; its standardized contrasts form the basis of Table 4 and Fig. 1. Finally, the belief-by-strain model evaluates exploratory heterogeneity using within-strain risk differences and differences between those risk differences.

All estimands are conditional on reported BNPL use during the previous year. They do not describe adoption, nonusers, individual contracts or per-transaction payment behavior. They also do not estimate what would happen if beliefs, purchase motives or financial strain were changed by an intervention. Within those boundaries, the framework permits a focused household-finance assessment: whether alignment with a time-specific information benchmark and observable indicators of the resource context of use provide similar or distinct cross-sectional information about self-reported any late payment.

Table 1

Measurement reference periods, specifications and interpretive boundaries **Panel A. Measurement reference periods and timing boundaries**

Measure	Questionnaire reference	Analytical unit	Timing boundary
Self-reported any late BNPL payment	Any late BNPL payment in the preceding year	Respondent-level extensive-margin outcome	No transaction date, count, amount or link to a reported purchase motive or purpose
Credit-reporting beliefs	Survey-time responses in October 2025	Respondent-level beliefs about BNPL generally	Not linked to a provider, transaction or the timing of any late payment
Affordability motive	Reason selected for the most recent BNPL purchase	Purchase-specific motive reported by the respondent	The purchase cannot be linked to the late-payment report or placed before or after it
Groceries-or-medical use and other-purpose breadth	BNPL purchase purposes during the preceding year	Respondent-level indicators and category count	Shares the outcome's broad window but is not transaction-linked; breadth is not frequency
Self-reported financial strain	How the respondent was managing financially ‘these days’	Respondent-level survey-time status	May precede, follow or share causes with a late payment during the preceding year
Panel B. Empirical questions and estimands			

Empirical question	Specification	Principal output	Permitted interpretation
Distribution of survey-benchmark beliefs	Weighted descriptive analysis	Full 3 by 3 item distribution and three-category shares	Belief distribution in the 2025 BNPL-user analytical domain
Belief category and any late payment	Model 2: socioeconomic-adjusted principal belief model	Model-standardized probabilities, risk differences, odds ratios and global Wald test	Cross-sectional adjusted association among BNPL users
Belief association after contemporaneous strain conditioning	Model 3: secondary strain-conditioned model	Odds ratios and financial-strain standardized contrast	Conditional association; not mediation or a strain effect
Affordability-related use and any late payment	Expanded descriptive resource-use model	Standardized probabilities, risk differences, odds ratios and exploratory within-model differences	Common-specification association comparison; not an identified liquidity constraint or causal effect
Belief association by financial strain	Exploratory interaction model	Within-stratum risk differences and difference-in-risk-differences	Exploratory risk-scale effect modification with uncertainty

Note: All estimands concern respondents who reported BNPL use in the preceding year. The affordability motive refers to the most recent purchase and cannot be linked to the late-payment event. Models containing strain, purchase motives or purposes are conditional descriptive specifications, not causal decompositions.

3 Data and empirical strategy

3.1 Data source and analytical domain

The analysis used the public-use 2025 SHED (Board of Governors of the Federal Reserve System 2026b). Ipsos administered the survey through its probability-based KnowledgePanel from 17 to 28 October 2025. The public file contains 12,934 U.S. adults aged 18 years or older and a cross-sectional post-stratification weight calibrated to Current Population Survey benchmarks (Board of Governors of the Federal Reserve System 2026a). The analytical domain comprised respondents who reported using BNPL credit to make a purchase in the preceding year. This restriction yielded 2,004 respondents, including 487 who reported at least one late BNPL payment. The respondent was the independent analytical unit.

SHED imputes item nonresponse for non-text variables and provides item-specific imputation flags. After those official imputations, all 2,004 BNPL users had values for the primary outcome, belief classification, self-reported financial strain indicator, selected covariates and final weight. The primary analysis

retained the supplied values and treated the sample as the full BNPL-user analytical domain after public-file imputation, rather than as an unimputed complete-case sample. Sensitivity analyses excluded records flagged as imputed for focal variables.

3.2 Outcomes and credit-reporting beliefs

The primary outcome was self-reported any late BNPL payment in the preceding year (BNPL3; yes = 1, no = 0). It is an extensive-margin respondent-level outcome: the survey does not report the number, timing, amount or duration of late payments, or the number of transactions over which each user was at risk. It therefore cannot support a per-transaction delinquency estimand. Table 1, Panel A records the reference period and timing boundary for each focal measure.

Two items measured product-specific credit-reporting beliefs. Respondents indicated whether they agreed, disagreed or were not sure that BNPL “helps me build a credit history” (item A, BNPL7_a) and that “making payments on time helps my credit score” (item B, BNPL7_b). The Federal Reserve treated rejection of both statements as consistent with its survey-time benchmark. A hierarchical three-category variable classified respondents as: (1) both statements rejected, the benchmark-consistent reference category; (2) uncertain, defined as no agreement and at least one “Not sure” response; or (3) at least one statement accepted, termed benchmark-discordant agreement. Agreement took precedence if one item was agreed with and the other was uncertain or rejected. This classification is time bounded and product specific; it is neither a general financial-literacy scale nor a claim about every provider's subsequent reporting practice. Item A and item B were also modeled separately, with “Disagree” as the reference.

3.3 Resource-related measures and covariates

The self-reported financial strain indicator classified “just getting by” or “finding it difficult to get by” as financially strained and “doing okay” or “living comfortably” as not financially strained. Because this measure is contemporaneous with the outcome window, it may be an antecedent, a consequence or a correlate of self-reported any late payment. Models containing strain are therefore conditional descriptions, not mediation analyses.

The expanded model added three BNPL-use measures. Other-purpose breadth counted five past-year categories—clothing, travel, furniture or appliances, electronics and other—and grouped the count as zero or one, two, or three or more. Groceries or food delivery and medical or veterinary expenses were deliberately excluded because they define the separate groceries-or-medical indicator. Breadth is not transaction frequency. The affordability-motive indicator identified respondents selecting “Only way I could afford it” as a reason for their most recent BNPL purchase. A sensitivity specification replaced the combined groceries-or-medical indicator with simultaneous indicators for groceries or food delivery and for medical or veterinary expenses. The two indicators were not mutually exclusive, and the questionnaire did not permit groceries to be separated from food delivery or medical expenses from veterinary expenses.

The affordability-motive item referred to the respondent's most recent BNPL purchase, whereas the outcome captured any late BNPL payment during the preceding year. The survey does not identify whether the motive and late-payment report referred to the same transaction or even to the same temporal sequence. Purpose measures covered the preceding year, while financial strain described how respondents were managing “these days”; neither was transaction-linked.

Socioeconomic covariates were age group, gender as supplied in the binary public-use variable, race and ethnicity, education, annual household-income category and employment status. The adjustment set summarized demographic and economic differences that may structure consumer-credit access, use and repayment; it was not used to claim conditional exchangeability. Sensitivity analyses used alternative belief, strain, fragility, credit-card, imputation, weighting and purchase-use specifications. Pairwise overlap among the four resource-context indicators was summarized with final-weighted Pearson correlations in the 2,004-user domain. The affordability-motive, groceries-or-medical-use and financial-strain indicators were coded 0/1, whereas other-purpose breadth entered as its 0–5 count; correlations between two binary indicators therefore equal weighted phi coefficients.

3.4 Models and estimands

Survey-weighted logistic regressions with a quasibinomial working variance reported odds ratios, 95% confidence intervals and design-based Wald tests. Model 1 contained only the three-category belief measure. Model 2, the principal belief model, added the socioeconomic covariates. Model 3 additionally conditioned on contemporaneous financial strain. The expanded descriptive model simultaneously entered belief category, socioeconomic covariates, other-purpose breadth, affordability motive, groceries-or-medical use and strain. Separate socioeconomic-adjusted models retained the original three responses to each belief item. Item A was also substituted for the composite belief category in the expanded specification. Four single-resource specifications, E1–E4, each added one of the affordability motive, groceries-or-medical use, other-purpose breadth or financial strain to the belief category and socioeconomic covariates; E5 was the common expanded model containing all four resource-context indicators. An exploratory model added the belief-by-strain interaction, evaluated jointly with a two-degree-of-freedom Wald test.

The estimands were kept distinct. Model 2 was the principal estimate of the belief and late-payment association. Model 3 described that association conditional on strain and separately quantified the strain gradient; it did not decompose direct and indirect pathways. The expanded model compared conditional resource-context gradients and the belief association within one descriptive specification. Exploratory within-model comparisons subtracted the benchmark-discordant-agreement risk difference from each resource-context risk difference in the expanded model. The split-purpose sensitivity similarly compared the grocery-or-food-delivery and medical-or-veterinary risk differences. These differences compare model-standardized associations and do not place the underlying constructs on a causal-effect scale. Because standardized risk differences depend on each measure's coding and reference categories, between-contrast differences do not provide a universal ranking of predictive, economic or

policy importance. The interaction model assessed whether belief risk differences varied by strain, but did not identify causal effect modification.

Because late payment was common, focal results were also expressed as model-standardized predicted probabilities and risk differences. For each focal classification, observed covariates were retained, predictions were averaged over the observed final-weight distribution, and uncertainty was computed by the delta method. Item-level probabilities and risk differences were standardized over the same observed final-weight distribution. For differences between two risk differences from one model, standard errors used the difference between their coefficient gradients and the full within-model covariance matrix. Table 3 retains Model 2 standardization for the principal belief estimand. Table 4, Panel A and Fig. 1 use only standardized belief and resource-context contrasts from the common expanded model. The interaction analysis compared belief risk differences between strained and not-strained users. These quantities are model-dependent descriptive contrasts, not treatment effects.

3.5 Weights, uncertainty and robustness

The full public file was declared as a survey design with one observation per respondent and the supplied cross-sectional final weight; estimation then used the BNPL-user domain. The public file has no primary-sampling-unit, stratum or replicate-weight variables. Taylor-linearized uncertainty was therefore estimated under a weighted independent-observation design and does not reconstruct all uncertainty from sampling and calibration. The analytic-domain weights summed to 2,132.70; the survey regression routine rescaled them to the analytic n of 2,004 for model fitting. The Kish effective sample size was approximately 1,757.

Robustness analyses changed belief coding, strain measurement, weighting, credit-card adjustment, imputation restrictions and purchase-use definitions. They also split the combined purpose indicator, compared E1–E4 single-resource specifications with E5, and re-estimated the item-A “Not sure” contrast in expanded, unweighted, non-imputed and credit-card-adjusted specifications. Exploratory internal cross-validation is reported in Online Resource 1. No observations were removed as outliers, and focal tests were two sided. Two post hoc exploratory Benjamini–Hochberg families were defined: the four non-reference coefficient tests across the two socioeconomic-adjusted item models and the five expanded-model resource-minus-belief risk-difference tests. The corresponding q values are reported alongside the unadjusted coefficient or risk-difference p values, and 95% confidence intervals were not multiplicity adjusted. Other exploratory and sensitivity contrasts were not adjusted for multiplicity and are interpreted through magnitude and uncertainty rather than isolated p values.

Analyses used R 4.5.3 and the survey package 4.5. All reported model matrices were full rank and all models converged. Full numerical and diagnostic outputs are supplied as Online Resource 1; the complete public-data workflow, model code, table inputs and figure source data are supplied as Online Resource 2. The analysis was not preregistered.

3.6 Ethics, reproducibility and generative artificial intelligence disclosure

Ethics approval

and consent were not applicable because this study analyzed a public, deidentified secondary dataset and involved no participant contact or intervention. The SHED public data were downloaded from the Federal Reserve data repository on 18 July 2026. The analyzed CSV had SHA-256 checksum c728d566008309c9abd87a026069ee70c2b58e7aef2d5c701884d371101d2a97. OpenAI Codex, using the GPT-5 model in the desktop application and accessed on 18 and 19 July 2026, assisted with literature discovery, initial prose drafting, R-code development and language editing. The author independently verified source relevance, variable definitions, analysis code, numerical results and the final text. All reported statistics were generated locally in R; no output from the language model was treated as empirical evidence.

4 Results

4.1 Analytical-domain characteristics and joint beliefs

Among the 2,004 BNPL users, 487 reported a late payment; the weighted prevalence was 25.8% (95% confidence interval 23.7 to 27.9%). The benchmark-consistent category contained 285 respondents and represented 14.5% (12.8 to 16.1%) of the weighted domain. The uncertain category contained 605 respondents and represented 31.4% (29.2 to 33.6%), while 1,114 respondents showed benchmark-discordant agreement, representing 54.1% (51.8 to 56.5%).

The complete joint distribution shows how the hierarchical classification was formed (Table 2, Panel A). Agreement with both items was the largest cell, containing 734 respondents and 36.2% of the weighted domain. Another 530 respondents, or 27.7% weighted, selected “Not sure” for both. The agree-agree and not-sure-not-sure cells together accounted for 63.9% of the weighted domain. The benchmark-consistent disagree-disagree cell contained 285 respondents. The remaining six mixed-response cells were retained under the stated hierarchical rule rather than discarded.

Financial strain was reported by 952 users, or 47.0% weighted (44.7 to 49.3%). The affordability motive was selected by 1,161 users, or 58.7% (56.4 to 61.0%), and 479 reported groceries or medical use, or 25.0% (22.9 to 27.1%) (Table 2, Panel B). Adverse-payment consequence indicators are reported in Online Resource 1.

Table 2

Joint credit-reporting beliefs and selected characteristics of BNPL users **Panel A. Complete three-by-three response distribution**

Item A response / Item B response	Disagree	Not sure	Agree
Disagree	285; 14.5% (12.8–16.1)	55; 2.6% (1.9–3.3)	102; 4.9% (3.9–5.9)
Not sure	20; 1.1% (0.6–1.6)	530; 27.7% (25.6–29.8)	247; 11.6% (10.2–13.1)
Agree	15; 0.7% (0.3–1.1)	16; 0.7% (0.3–1.1)	734; 36.2% (33.9–38.4)

Note: Rows are responses to item A, 'helps me build a credit history'; columns are responses to item B, 'making payments on time helps my credit score.' Each cell reports unweighted n; weighted percentage (95% confidence interval).

Panel B. Selected analytical-domain characteristics

Characteristic	Category	Unweighted n	Weighted %	95% CI
Survey-benchmark belief category	Both statements rejected: benchmark-consistent	285	14.5	12.8–16.1
	Uncertain	605	31.4	29.2–33.6
	At least one accepted: benchmark-discordant agreement	1114	54.1	51.8–56.5
Self-reported any late BNPL payment	Yes	487	25.8	23.7–27.9
Self-reported financial strain	Yes	952	47.0	44.7–49.3
Lacks a three-month emergency fund	Yes	1309	66.5	64.3–68.7
Unable to cover a USD 400 emergency immediately	Yes	403	19.9	18.1–21.8
Only-way-to-afford motive selected	Yes	1161	58.7	56.4–61.0
Groceries or medical use	Yes	479	25.0	22.9–27.1

Note: Percentages and confidence intervals use the supplied SHED final weight. Financial strain denotes 'just getting by' or 'finding it difficult to get by.'

4.2 Principal belief model and item-level results

Unadjusted late-payment prevalence was 27.9% (22.3 to 33.5%) in the benchmark-consistent category, 24.1% (20.4 to 27.7%) among uncertain respondents and 26.2% (23.4 to 29.1%) under benchmark-discordant agreement. The belief term was imprecise both without adjustment ($F(2, 2001) = 0.73$, $p = 0.482$) and in Model 2 ($F(2, 1978) = 1.28$, $p = 0.278$). Relative to the benchmark-consistent category, Model 2 estimated an odds ratio of 0.79 (95% confidence interval 0.55 to 1.13; $p = 0.191$) for uncertainty and 0.95 (0.68 to 1.33; $p = 0.768$) for benchmark-discordant agreement.

Model-standardized predicted probabilities were 27.6% (22.3 to 33.0%) for the benchmark-consistent category, 23.4% (20.1 to 26.8%) for uncertainty and 26.7% (23.9 to 29.5%) for benchmark-discordant agreement. Relative to the benchmark, the risk differences were − 4.2 percentage points (95% confidence interval − 10.5 to 2.2; $p = 0.198$) and − 0.9 points (-7.0 to 5.2; $p = 0.769$), respectively. The interval remained compatible with modest differences in either direction and did not establish equivalence.

Item-level models did not provide clear evidence of an elevated late-payment association for agreement. For item A, agreement versus disagreement had an odds ratio of 0.93 (95% confidence interval 0.68 to 1.26; $p = 0.620$), whereas “Not sure” had an odds ratio of 0.64 (0.47 to 0.87; $p = 0.005$); the global test was $F(2, 1978) = 5.34$, $p = 0.005$. After Benjamini–Hochberg adjustment across the four non-reference item coefficients, the item-A “Not sure” contrast had $q = 0.018$. Its model-standardized probabilities were 29.4% for “Disagree” and 21.7% for “Not sure”, a risk difference of -7.7 percentage points (95% confidence interval − 13.1 to -2.3). For item B, agreement had an odds ratio of 0.90 (0.65 to 1.25) and “Not sure” had an odds ratio of 0.76 (0.54 to 1.08), with an imprecise global test ($F(2, 1978) = 1.36$, $p = 0.257$). The remaining three item-coefficient q values were 0.620, 0.620 and 0.252. The isolated item-A pattern remains secondary and exploratory and does not establish that uncertainty protects against late payment (Table 3; Online Resource 1).

Table 3

Credit-reporting beliefs and self-reported any late BNPL payment **Panel A. Three-category belief models and Model 2 standardization**

Belief category	Model 1 OR	Model 2 OR	Model 3 OR	Model 2 standardized probability	Model 2 risk difference
Both statements rejected: benchmark-consistent	Reference	Reference	Reference	27.6% (22.3–33.0)	Reference
Uncertain	0.82 (0.58–1.15); p = 0.255	0.79 (0.55–1.13); p = 0.191	0.82 (0.57–1.17); p = 0.275	23.4% (20.1–26.8)	-4.2 pp (-10.5 to + 2.2); p = 0.198
At least one statement accepted: benchmark-discordant agreement	0.92 (0.67–1.26); p = 0.595	0.95 (0.68–1.33); p = 0.768	1.06 (0.75–1.48); p = 0.748	26.7% (23.9–29.5)	-0.9 pp (-7.0 to + 5.2); p = 0.769
Global Wald test for belief category	0.482	0.278	0.168	—	—

Note: OR odds ratio. Model 1 contains belief category only. Model 2 adjusts for age group, gender, race and ethnicity, education, household income and employment. Model 3 additionally conditions on contemporaneous financial strain. Probabilities and risk differences are standardized from Model 2 over the observed final-weight distribution. The global-test rows report p values.

Panel B. Socioeconomic-adjusted item-level models

Survey item	Contrast	OR (95% CI); p	BH q	Standardized probability: reference to comparison	Risk difference	Global p
Item A: helps build a credit history	Agree versus Disagree	0.93 (0.68–1.26); p = 0.620	0.620	29.4% to 28.0%	-1.4 pp (-7.1 to + 4.2)	0.005
	Not sure versus Disagree	0.64 (0.47–0.87); p = 0.005	0.018	29.4% to 21.7%	-7.7 pp (-13.1 to -2.3)	
Item B: on-time payment helps credit score	Agree versus Disagree	0.90 (0.65–1.25); p = 0.543	0.620	28.2% to 26.4%	-1.8 pp (-7.7 to + 4.1)	0.257
	Not sure versus Disagree	0.76 (0.54–1.08); p = 0.126	0.252	28.2% to 23.5%	-4.7 pp (-10.9 to + 1.4)	

Note: Disagree is the reference category. Each global p value applies to both contrasts for the corresponding survey item. Probabilities and risk differences are standardized over the observed final-weight distribution. BH q values adjust the four non-reference item-coefficient p values as one post hoc exploratory Benjamini–Hochberg family; risk-difference confidence intervals are unadjusted. Item-level tests are secondary and do not establish that uncertainty is protective. All models used n = 2,004 respondents and 487 late-payment events.

4.3 Financial strain and belief-by-strain contrasts

Late-payment prevalence was 36.3% (33.0 to 39.6%) among financially strained users and 16.5% (14.0 to 18.9%) among other users. In Model 3, financial strain was associated with higher late-payment odds (odds ratio 2.44, 95% confidence interval 1.90 to 3.13; $F(1, 1977) = 48.36$, $p < 0.001$). The model-standardized predicted probabilities were 18.1% (15.5 to 20.7%) for not-strained users and 33.4% (30.2 to 36.6%) for strained users, a risk difference of 15.3 percentage points (11.1 to 19.6). This is a conditional disparity, not an estimated effect of changing strain. The belief term remained imprecise after strain was added ($F(2, 1977) = 1.78$, $p = 0.168$): odds ratios were 0.82 (0.57 to 1.17) for uncertainty and 1.06 (0.75 to 1.48) for benchmark-discordant agreement.

The exploratory interaction did not provide clear evidence that belief associations varied by strain ($F(2, 1975) = 1.00$, $p = 0.368$). Within-strain risk differences and differences between them were imprecise (Table 4, Panel B; Online Resource 1).

4.4 Common expanded descriptive model

Unadjusted late-payment prevalence was 36.4% (33.4 to 39.4%) when the affordability motive was selected and 10.7% (8.4 to 13.0%) otherwise. It was 42.1% (37.4 to 46.8%) for groceries-or-medical use

and 20.3% (18.1 to 22.6%) for other users. Across other-purpose breadth, prevalence rose from 22.8% (20.3 to 25.3%) for zero or one categories to 28.8% (24.4 to 33.3%) for two and 36.2% (29.4 to 43.0%) for three or more.

In the expanded model, the odds ratios were 0.88 (95% confidence interval 0.61 to 1.27) for uncertainty and 1.01 (0.72 to 1.43) for benchmark-discordant agreement; the global belief test was F(2, 1973) = 0.53, p = 0.589. The affordability motive had an odds ratio of 2.91 (2.13 to 3.97), groceries-or-medical use 1.90 (1.46 to 2.47), and strain 1.66 (1.26 to 2.18). For other-purpose breadth, the odds ratios were 1.10 (0.83 to 1.47) for two categories and 1.70 (1.20 to 2.42) for three or more relative to zero or one; the global test was F(2, 1973) = 4.39, p = 0.013.

In that same model, benchmark-discordant agreement had a standardized risk difference of + 0.2 percentage points (-5.3 to + 5.7) and uncertainty − 2.0 points (-7.9 to + 3.8). The affordability-motive contrast was + 16.5 points (+ 12.0 to + 20.9), groceries-or-medical use + 10.9 (+ 6.2 to + 15.6), financial strain + 8.1 (+ 3.7 to + 12.5), and three or more other-purpose categories + 8.9 (+ 2.8 to + 15.1); two categories had a + 1.6-point difference (-3.0 to + 6.1).

Exploratory subtraction of the benchmark-discordant-agreement risk difference from each resource-context risk difference yielded differences of + 16.2 percentage points for the affordability motive (95% confidence interval + 9.1 to + 23.4; Benjamini–Hochberg q < 0.001), + 10.7 for groceries-or-medical use (+ 3.6 to + 17.8; q = 0.007), + 8.7 for three or more other-purpose categories (+ 0.3 to + 17.1; q = 0.054), and + 7.9 for financial strain (+ 1.0 to + 14.8; q = 0.041). The difference for two other-purpose categories was + 1.3 points (-5.8 to + 8.5; q = 0.711). These are exploratory differences between model-standardized associations, not causal comparisons or universal rankings of variable importance. Weighted pairwise correlations among the four resource-context indicators ranged from − 0.041 to 0.407. The largest was between the affordability motive and strain (r = 0.407); all others had absolute values no greater than 0.192, providing no evidence of severe pairwise collinearity while not establishing that the constructs are independent (Online Resource 1). The component gradients likewise remain conditional respondent-level associations rather than distinct causal contributions (Table 4, Panel A; Fig. 1).

Table 4

Common-specification gradients and exploratory belief-by-strain contrasts **Panel A. Common expanded descriptive model**

Measure	Category	OR (95% CI); p	Standardized probability	Risk difference
Survey-benchmark belief	Both statements rejected: benchmark-consistent	Reference	26.3% (21.5–31.1)	Reference
	Uncertain	0.88 (0.61–1.27); p = 0.497	24.3% (21.0–27.6)	-2.0 pp (-7.9 to + 3.8); p = 0.500
	At least one accepted: benchmark-discordant agreement	1.01 (0.72–1.43); p = 0.941	26.5% (23.8–29.2)	+ 0.2 pp (-5.3 to + 5.7); p = 0.941
Other-purpose breadth	0–1 other categories	Reference	24.3% (21.9–26.7)	Reference
	2 other categories	1.10 (0.83–1.47); p = 0.500	25.8% (22.0–29.6)	+ 1.6 pp (-3.0 to + 6.1); p = 0.503
	3 or more other categories	1.70 (1.20–2.42); p = 0.003	33.2% (27.6–38.8)	+ 8.9 pp (+ 2.8 to + 15.1); p = 0.005
Affordability motive	Other reasons	Reference	14.9% (11.8–18.0)	Reference
	Only-way-to-afford motive selected	2.91 (2.13–3.97); p < 0.001	31.3% (28.6–34.1)	+ 16.5 pp (+ 12.0 to + 20.9); p < 0.001
Groceries-or-medical use	No groceries-or-medical use	Reference	22.6% (20.3–24.8)	Reference
	Any groceries-or-medical use	1.90 (1.46–2.47); p < 0.001	33.5% (29.5–37.4)	+ 10.9 pp (+ 6.2 to + 15.6); p < 0.001
Self-reported financial strain	Not financially strained	Reference	21.3% (18.4–24.2)	Reference
	Financially strained	1.66 (1.26–2.18); p < 0.001	29.4% (26.5–32.3)	+ 8.1 pp (+ 3.7 to + 12.5); p < 0.001

Note: All Panel A estimates come from one model containing belief category, socioeconomic covariates, other-purpose breadth, affordability motive, groceries-or-medical use and financial strain. Standardized results average over the observed final-weight distribution. Other-purpose breadth excludes the grocery and medical categories and is not transaction frequency. Global p values were 0.589 for belief, 0.013 for other-purpose breadth, and < 0.001 for affordability, groceries-or-medical use and strain.

Panel B. Exploratory risk-scale belief contrasts by financial strain

Belief contrast	Risk difference: not strained	Risk difference: strained	Difference in risk differences
Uncertain versus benchmark-consistent	-0.9 pp	-5.3 pp	-4.4 pp (-16.5 to + 7.7); p = 0.478
Benchmark-discordant agreement versus benchmark-consistent	+ 4.8 pp	-2.3 pp	-7.0 pp (-18.5 to + 4.4); p = 0.228

Note: Risk differences are percentage points. The global two-degree-of-freedom interaction test was F(2, 1975) = 1.00, p = 0.368. Difference-in-risk-differences intervals use the delta method. These exploratory contrasts do not identify causal effect modification.

4.5 Robustness and incremental descriptive information

Alternative belief encodings did not provide clear evidence of an adverse association for benchmark-discordant responses. Combining all non-benchmark responses gave an odds ratio of 0.96 (0.70 to 1.32), and the count of accepted items gave an odds ratio of 0.99 per item (0.84 to 1.17). Estimates were similar in an unweighted model, after adjustment for credit-card behavior and after excluding imputed focal records. Excluding imputed BNPL1, BNPL3 and belief items left n = 1,989 and 480 events; the odds ratios were 0.81 (0.57 to 1.17) for uncertainty and 1.03 (0.74 to 1.45) for benchmark-discordant agreement. Additionally excluding imputed strain left n = 1,986 and produced nearly identical estimates. The item-A “Not sure” contrast was also similar in the expanded, unweighted, item-A-specific non-imputed and credit-card-adjusted specifications: odds ratios ranged from 0.62 to 0.66 and standardized risk differences from − 8.0 to -6.6 percentage points. The expanded item-A global test was F(2, 1973) = 3.97, p = 0.019 (Online Resource 1).

Resource-related results were stable to alternative specifications. Relative to living comfortably, the four-level strain model yielded odds ratios of 2.06 (1.32 to 3.20) for just getting by and 4.03 (2.52 to 6.47) for finding it difficult to get by. Lacking a three-month emergency fund had an odds ratio of 1.96 (1.49 to 2.59), and inability to cover a USD 400 expense immediately had an odds ratio of 2.42 (1.84 to 3.18). A sensitivity model using continuous seven-category purchase-purpose breadth gave odds ratios of 1.37 (1.22 to 1.55) per category and 2.89 (2.12 to 3.94) for the affordability motive. Requiring that motive to be the most important reason gave an odds ratio of 1.90 (1.46 to 2.49).

In the simultaneous split-purpose model, the standardized risk difference was + 10.4 percentage points (95% confidence interval + 5.4 to + 15.5) for groceries or food delivery and + 8.6 points (+ 1.0 to + 16.2) for medical or veterinary expenses. The two were not clearly distinguishable: grocery or food delivery minus medical or veterinary use was + 1.9 points (-7.7 to + 11.4; p = 0.703). Component prevalence, odds ratios and standardized probabilities appear in Online Resource 1.

The risk differences for all four resource-context indicators, including both breadth contrasts, attenuated from their single-resource E1–E4 specifications to the mutually adjusted E5 model. This pattern describes shared descriptive information among the indicators and is not a mediation decomposition; the complete estimates appear in Online Resource 1.

Exploratory internal cross-validation did not indicate meaningful incremental descriptive information from the belief categories alone. Imputation exclusions and model diagnostics did not materially change the focal conclusions. Full results appear in Online Resource 1.

5 Discussion

5.1 Main findings and contribution to household finance

This study distinguishes an information margin, represented by survey-benchmark beliefs about credit reporting, from a resource-context margin represented by affordability-motivated use, groceries-or-medical use, other-purpose breadth and financial strain. Benchmark-discordant agreement was common, yet the belief categories did not clearly differentiate self-reported late payment after adjustment. In the common expanded model, the affordability-motive, groceries-or-medical, three-or-more versus zero-or-one breadth, and strain risk differences were farther from zero than the benchmark-discordant-agreement contrast; the two-category breadth contrast was not. After Benjamini–Hochberg adjustment across the five expanded-model resource-minus-belief risk-difference tests, the between-contrast evidence was clearest for the affordability motive, groceries-or-medical use and financial strain, while the three-or-more breadth comparison was less precise. These remain model-dependent descriptive comparisons rather than causal-effect contrasts or universal rankings of variable importance.

The household-finance contribution is to compare product-specific information consistency with observed resource context on a common probability scale. Imperfect information, limited attention, liquidity and repayment behavior may be connected, but need not identify the same households. The results do not imply that information is unimportant; within the reported specifications, the composite survey-benchmark belief category added little incremental descriptive information for distinguishing users who reported any late payment.

5.2 Credit-reporting beliefs as an information and attention margin

Behavioral models of consumer finance emphasize that product complexity, limited attention and incomplete understanding can weaken consumers' responses to future costs (Campbell et al. 2011; Gabaix and Laibson 2006). Experimental and quasi-experimental studies further show that appropriately framed information or increased fee salience can alter high-cost borrowing and overdraft behavior (Bertrand and Morse 2011; Stango and Zinman 2014). The present results do not contradict that

literature. Instead, they indicate a boundary condition: benchmark-discordant agreement was not clearly associated with self-reported late payment among people who had already selected into BNPL use.

Several features could produce this pattern. Credit reporting is only one component of the perceived cost of lateness. Users may respond more strongly to immediate consequences, including late fees, account suspension, overdraft charges or the need to preserve cash for other bills. Reporting practices may also differ across providers, products and stages of delinquency. A respondent's answer can therefore combine general knowledge, experience with a particular provider, uncertainty about changing rules and interpretation of the survey wording. The belief categories should consequently be understood as survey-benchmark consistency, not as a comprehensive measure of financial literacy or attentiveness.

Selection into BNPL use also matters. Beliefs and resource conditions can affect adoption, provider choice, the number of plans opened and repayment effort. Conditioning on use may therefore conceal associations operating at these earlier margins and may induce selection through unmeasured credit demand, access or risk tolerance. Attenuation of the belief gradient is one possible result, but the bias need not be toward zero and could instead amplify or reverse the within-user association. Moreover, the outcome records whether the respondent had any late payment over the past year. It combines the probability of lateness on a given obligation with the number of opportunities to become late. A more informed user could exercise greater repayment effort yet undertake more BNPL transactions, leaving the association with any late payment ambiguous. Because product-specific beliefs are observed only for users and the public survey does not report transaction-level exposure, the present analysis cannot identify or separate these channels.

The uncertainty category should not be treated as an intermediate score on a knowledge scale. Its adjusted composite estimate was below the benchmark-consistent estimate, but the risk-difference confidence interval crossed zero. Separately, item-A “Not sure” was associated with a lower standardized late-payment probability than “Disagree” and survived adjustment within the four non-reference item-coefficient family; the item-B estimate was imprecise. One possible but unmeasured selection account is that ambiguity-averse or cautious respondents use BNPL less intensively and face fewer late-payment opportunities; uncertainty could instead mark weaker knowledge. SHED measures neither these traits nor transaction exposure. Item interpretation, provider experience, selection and residual confounding are also possible. The isolated, non-preregistered pattern does not establish protection or a monotonic knowledge gradient. Future studies should link disaggregated beliefs about fees, credit reporting, account restrictions, automatic payment and collections to provider-verified exposure and repayment records.

5.3 Affordability, groceries-or-medical use and other-purpose breadth

The positive resource-context gradients are consistent with an extensive household-finance literature showing that liquid resources and short-term financial fragility shape credit use and repayment behavior (Gross and Souleles 2002; Lusardi et al. 2011). They also accord with evidence that consumers with

lower income or creditworthiness exhibit greater demand for BNPL and that financing can relax an immediate payment constraint (Aidala et al. 2025). Federal Reserve evidence similarly shows that consumers experiencing financial difficulty are especially likely to report that BNPL was the only way they could afford a purchase (Larrimore et al. 2024).

Users selecting “Only way I could afford it” had a standardized late-payment probability of 31.3%, compared with 14.9% among other users in the expanded model. The item may capture a high shadow value of current liquidity: paying the full price immediately could require reducing other consumption, drawing down scarce balances or missing another obligation. It is nevertheless subjective and may also reflect payment-smoothing preferences, purchase size or retrospective justification; it is not an identified binding liquidity constraint.

Groceries-or-medical use was associated with a 10.9-percentage-point higher standardized probability. Such use may mark attempts to bridge expenditure and cash-flow timing, but the categories do not establish necessity and cannot reveal whether BNPL relieved hardship, increased repayment pressure or both. When the combined measure was split, both component gradients remained positive and their difference was imprecise. The survey nevertheless cannot distinguish groceries from food delivery or medical from veterinary expenses, so neither component identifies a narrowly defined necessity category.

Other-purpose breadth excludes the grocery and medical categories used in the separate binary measure. Users selecting three or more of the remaining categories had an 8.9-percentage-point higher standardized probability than those selecting zero or one. Breadth is not transaction frequency, loan count, balance or plan stacking, but may mark broader integration of BNPL into household spending and more opportunities for at least one late payment.

Financial strain remained associated with lateness after the use measures were included. Its 8.1-point risk difference should not be read as a mediated or direct effect because strain and the payment outcome have unresolved temporal ordering. More modestly, the common model shows conditional descriptive gradients across resource-context indicators that were related but not highly correlated. The attenuation from the E1–E4 single-resource specifications to the mutually adjusted E5 model is consistent with shared descriptive information among the resource indicators, but it neither partitions mechanisms nor establishes mediation.

5.4 Policy implications and relevance to SDG 10

The study does not evaluate a disclosure intervention. Its descriptive pattern instead motivates testing clear disclosure together with exposure-management and repayment-support measures. Clear reporting information remains valuable for informed choice and procedural fairness, while candidate measures for evaluation include consolidated displays of obligations across plans, reminders showing total upcoming payments, transparent overdraft and fee warnings, and accessible hardship or rescheduling options. Evidence from credit-card regulation demonstrates that consumer-finance rules can reduce avoidable

fees, but policy effects depend on product design and consumer responses (Agarwal et al. 2015). The weak belief gradient in this user domain is not an estimate that disclosure policy is ineffective.

Affordability assessment presents a genuine trade-off. Stronger checks may reduce borrowing that is unlikely to be repaid, but they may also exclude households using BNPL to manage time-sensitive expenditure. A proportionate design could test more intensive review for larger or unusually stacked commitments while preserving lower-value payment flexibility, but the present estimates identify neither an appropriate threshold nor the welfare effects of broader access or tighter screening. They should not justify automated denial based on survey characteristics. Policy evaluations should jointly measure approval and exclusion, repayment, fees, consumption smoothing, substitution toward costlier credit and access to beneficial purchases across resource groups rather than treating reduced credit use alone as success.

The relevance to Sustainable Development Goal 10 lies in evaluating access and repayment burden jointly, consistent with its emphasis on economic inclusion and effective oversight of financial markets and institutions (United Nations 2015). Higher model-standardized late-payment probabilities among respondents reporting affordability-motivated use or financial strain suggest that regulators and providers should test whether safeguards reduce late fees, overdraft exposure and repayment burden for resource-constrained households without shifting them toward costlier credit or excluding beneficial purchases. Audits stratified by household resources could also test whether automated affordability rules create disparate exclusion. This study does not measure an official SDG indicator or establish that any particular intervention would reduce inequality; its contribution is to identify household-resource margins along which access, benefits and burdens may be unequally distributed.

5.5 Strengths and limitations

The study uses a recent public national survey, reports weighted absolute and relative estimates, preserves the complete belief-response structure and examines alternative indicators of financial fragility. These features improve transparency, but do not overcome the design's inferential limits.

Several limitations constrain inference. First, the design is cross-sectional and cannot establish temporal ordering or causal effects. The affordability-motive item referred to the respondent's most recent BNPL purchase, whereas the outcome captured any late BNPL payment during the preceding year. The survey does not identify whether the motive and late-payment report referred to the same transaction or even to the same temporal sequence. Consequently, the affordability-motive contrast should be interpreted as a respondent-level cross-sectional correlate rather than as the repayment outcome of an affordability-motivated transaction. Financial strain may likewise precede, follow or share causes with late payment.

Second, all variables are self-reported and may be affected by recall and interpretation; “any late payment” is not lender-recorded delinquency or default. Third, the sample is restricted to past-year BNPL users and does not represent all adults, former users, rejected applicants or deterred nonusers. Conditioning on use may induce selection when beliefs, resources and unmeasured determinants of

adoption also relate to late payment. The direction is not identified: selection could attenuate, amplify or reverse associations within the user domain, and product-specific beliefs are unavailable for a comparable nonuser analysis.

Fourth, the public data do not include transaction counts, amounts, schedules, outstanding obligations, provider identities or event dates. Other-purpose breadth cannot be converted into an exposure rate, and provider-specific reporting rules cannot be matched to beliefs. Fifth, the two reporting items are not a validated psychometric scale. The supplied gender variable is binary. Sixth, the final weight is not accompanied by primary-sampling-unit, stratum or replicate-weight variables, and official single imputation does not propagate imputation uncertainty.

Finally, simultaneous conditioning does not separate mechanisms. Only the two post hoc exploratory families—the four non-reference item-coefficient family and five expanded-model resource-minus-belief risk-difference family—received Benjamini–Hochberg adjustment; the remaining exploratory and sensitivity tests did not. Internal cross-validation does not establish transportability. The results are unsuitable for individual underwriting, causal policy evaluation or provider-level inference.

6 Conclusion

Among U.S. adults who used BNPL, benchmark-discordant credit-reporting agreement was common but did not clearly differentiate self-reported any late payment. In the common expanded model, point estimates for affordability-motivated use, groceries-or-medical use, three-or-more versus zero-or-one other-purpose categories and financial strain were farther from zero than the belief contrasts; the two-category breadth estimate was not. Exploratory within-model tests likewise indicated larger risk differences for several resource-context contrasts than for benchmark-discordant agreement. Because the contrasts depend on coding and reference categories, they do not provide a universal ranking of variable importance; nor do they prove the absence of a belief association or identify causal effects.

The household-finance implication is that information consistency and resource-context indicators should not be treated as interchangeable correlates of self-reported late payment. Better disclosure may improve informed decision-making, but policies intended to reduce late payment should also consider the resource conditions under which BNPL is used. The largest affordability contrast links a motive reported for the most recent purchase to any late payment reported over the preceding year; it cannot be attributed to the same transaction or read as a causal repayment pathway. Transaction-level longitudinal research is needed to determine whether these patterns arise from liquidity relief, greater credit exposure, repayment burden or selection into BNPL, and to evaluate interventions that protect financially vulnerable users without unnecessarily restricting access to useful payment flexibility.

Declarations

Funding

The author received no financial support for the research, authorship or publication of this article.

Competing interests

The author declares no competing interests.

Author contributions

Kaibo Tang performed the conceptualization, methodology, software, formal analysis, investigation, data curation, visualization, writing of the original draft, review and editing, and project administration. The author read and approved the final manuscript.

Ethics approval

Not applicable. This study used a public, deidentified secondary dataset and involved no interaction with participants.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Data availability

The public 2025 SHED dataset and documentation are available from the Board of Governors of the Federal Reserve System at https://doi.org/10.17016/datasets.002. The specific public CSV analyzed in this study had SHA-256 checksum c728d566008309c9abd87a026069ee70c2b58e7aef2d5c701884d371101d2a97 and was accessed on 18 July 2026.

Code availability

The complete R scripts used to reproduce the analysis, tables and figure are provided with the submission as Online Resource 2.

Materials availability

The questionnaire, codebook and analytical variable definitions are available with the public SHED documentation and in Online Resource 2.

References

1. Agarwal S, Chomsisengphet S, Mahoney N, Stroebel J (2015) Regulating consumer financial products: evidence from credit cards. Q J Econ 130(1):111–164. https://doi.org/10.1093/qje/qju037
2. Aidala F, Koşar G, Mangrum D, van der Klaauw W (2025) Understanding consumer demand for Buy Now, Pay Later. Federal Reserve Bank of New York Staff Reports No. 1167, revised February 2026. https://doi.org/10.59576/sr.1167
3. Bertrand M, Morse A (2011) Information disclosure, cognitive biases, and payday borrowing. J Financ 66(6):1865–1893. https://doi.org/10.1111/j.1540-6261.2011.01698.x
4. Board of Governors of the Federal Reserve System (2026a) Economic well-being of U.S. households in 2025. Board of Governors of the Federal Reserve System, Washington, DC. https://doi.org/10.17016/8960.2
5. Board of Governors of the Federal Reserve System (2026b) Survey of Household Economics and Decisionmaking [dataset]. Board of Governors of the Federal Reserve System, Washington, DC. https://doi.org/10.17016/datasets.002
6. Campbell JY, Jackson HE, Madrian BC, Tufano P (2011) Consumer financial protection. J Econ Perspect 25(1):91–114. https://doi.org/10.1257/jep.25.1.91
7. Chen T, Marshall BR, Nguyen NH, Visaltanachoti N (2024) What influences demand for Buy Now. Pay Later credit? Econ Lett 242:111857. https://doi.org/10.1016/j.econlet.2024.111857
8. Consumer Financial Protection Bureau (2022) Buy Now, Pay Later: Market trends and consumer impacts. Consumer Financial Protection Bureau, Washington, DC. https://www.consumerfinance.gov/data-research/research-reports/buy-now-pay-later-market-trends-and-consumer-impacts/. Accessed 18 July 2026
9. Consumer Financial Protection Bureau (2025) Consumer use of Buy Now, Pay Later and other unsecured debt. Consumer Financial Protection Bureau, Washington, DC. https://www.consumerfinance.gov/data-research/research-reports/consumer-use-of-buy-now-pay-later-and-other-unsecured-debt/. Accessed 18 July 2026
10. deHaan E, Kim J, Lourie B, Zhu C (2024) Buy Now Pay (Pain?) Later. Manag Sci 70(8):5586–5598. https://doi.org/10.1287/mnsc.2022.03266
11. Disney R, Gathergood J (2013) Financial literacy and consumer credit portfolios. J Bank Financ 37(7):2246–2254. https://doi.org/10.1016/j.jbankfin.2013.01.013
12. Dong Y, Hu J, Huang Y, Qiu H, Zhang Y (2025) The use and disuse of FinTech credit: When buy-now-pay-later meets credit reporting. BIS Working Papers No. 1239, Bank for International Settlements, Basel. https://www.bis.org/publ/work1239.htm. Accessed 18 July 2026
13. Donou-Adonsou F, Leslie-Piper N (2025) BNPL and financial fragility in U.S. households. Financ Res Lett 86:108423. https://doi.org/10.1016/j.frl.2025.108423
14. Gabaix X, Laibson D (2006) Shrouded attributes, consumer myopia, and information suppression in competitive markets. Q J Econ 121(2):505–540. https://doi.org/10.1162/qjec.2006.121.2.505
15. Gerrans P, Baur DG, Lavagna-Slater S (2022) Fintech and responsibility: Buy-now-pay-later arrangements. Aust J Manag 47(3):474–502. https://doi.org/10.1177/03128962211032448

16. Gross DB, Souleles NS (2002) Do liquidity constraints and interest rates matter for consumer behavior? Evidence from credit card data. Q J Econ 117(1):149–185. https://doi.org/10.1162/003355302753399472
17. Guttman-Kenney B, Firth C, Gathergood J (2023) Buy now, pay later (BNPL) … on your credit card. J Behav Exp Financ 37:100788. https://doi.org/10.1016/j.jbef.2023.100788
18. Hayashi F, Routh A (2025) Financial constraints among Buy Now, Pay Later users. Fed Reserve Bank Kans City Econ Rev 110(4):1–18. https://doi.org/10.18651/ER/v110n4HayashiRouth
19. Larrimore J, Lloro A, Merchant Z, Tranfaglia A (2024) The Only Way I Could Afford It: Who Uses Buy Now, Pay Later and Why. FEDS Notes, Board of Governors of the Federal Reserve System, Washington, DC, 20 December 2024. https://www.federalreserve.gov/econres/notes/feds-notes/the-only-way-i-could-afford-it-who-uses-bnpl-and-why-20241220.html. Accessed 18 July 2026
20. Lusardi A, Mitchell OS (2014) The economic importance of financial literacy: theory and evidence. J Econ Lit 52(1):5–44. https://doi.org/10.1257/jel.52.1.5
21. Lusardi A, Schneider DJ, Tufano P (2011) Financially fragile households: evidence and implications. Brookings Pap Econ Act 2011(1):83–134. https://doi.org/10.1353/eca.2011.0002
22. Lusardi A, Tufano P (2015) Debt literacy, financial experiences, and overindebtedness. J Pension Econ Financ 14(4):332–368. https://doi.org/10.1017/S1474747215000232
23. Maesen S, Ang D (2025) Buy Now, Pay Later: impact of installment payments on customer purchases. J Mark 89(3):13–35. https://doi.org/10.1177/00222429241282414
24. Powell R, Do A, Gengatharen D, Yong J, Gengatharen R (2023) The relationship between responsible financial behaviours and financial wellbeing: the case of buy-now-pay-later. Acc Financ 63(4):4431–4451. https://doi.org/10.1111/acfi.13100
25. Relja R, Ward P, Zhao AL (2024) Understanding the psychological determinants of buy-now-pay-later in the UK: a user perspective. Int J Bank Mark 42(1):7–37. https://doi.org/10.1108/IJBM-07-2022-0324
26. Siemens JC (2007) When consumption benefits precede costs: towards an understanding of buy now, pay later transactions. J Behav Decis Mak 20(5):521–531. https://doi.org/10.1002/bdm.564
27. Stango V, Zinman J (2014) Limited and varying consumer attention: evidence from shocks to the salience of bank overdraft fees. Rev Financ Stud 27(4):990–1030. https://doi.org/10.1093/rfs/hhu008
28. United Nations (2015) Transforming our world: the 2030 Agenda for Sustainable Development. A/RES/70/1. United Nations, New York. https://sdgs.un.org/2030agenda. Accessed 18 July 2026
29. Wang M, Wong MCS (2026) Financial literacy, financial vulnerability, and buy now, pay later use. Econ Lett 267:113080. https://doi.org/10.1016/j.econlet.2026.113080
30. Zeballos Doubinko V, Akana T (2023) How does Buy Now, Pay Later affect customers' credit? Federal Reserve Bank of Philadelphia Consumer Finance Institute Discussion Paper DP 23 – 01. https://doi.org/10.21799/frbp.dp.2023.01

Figures

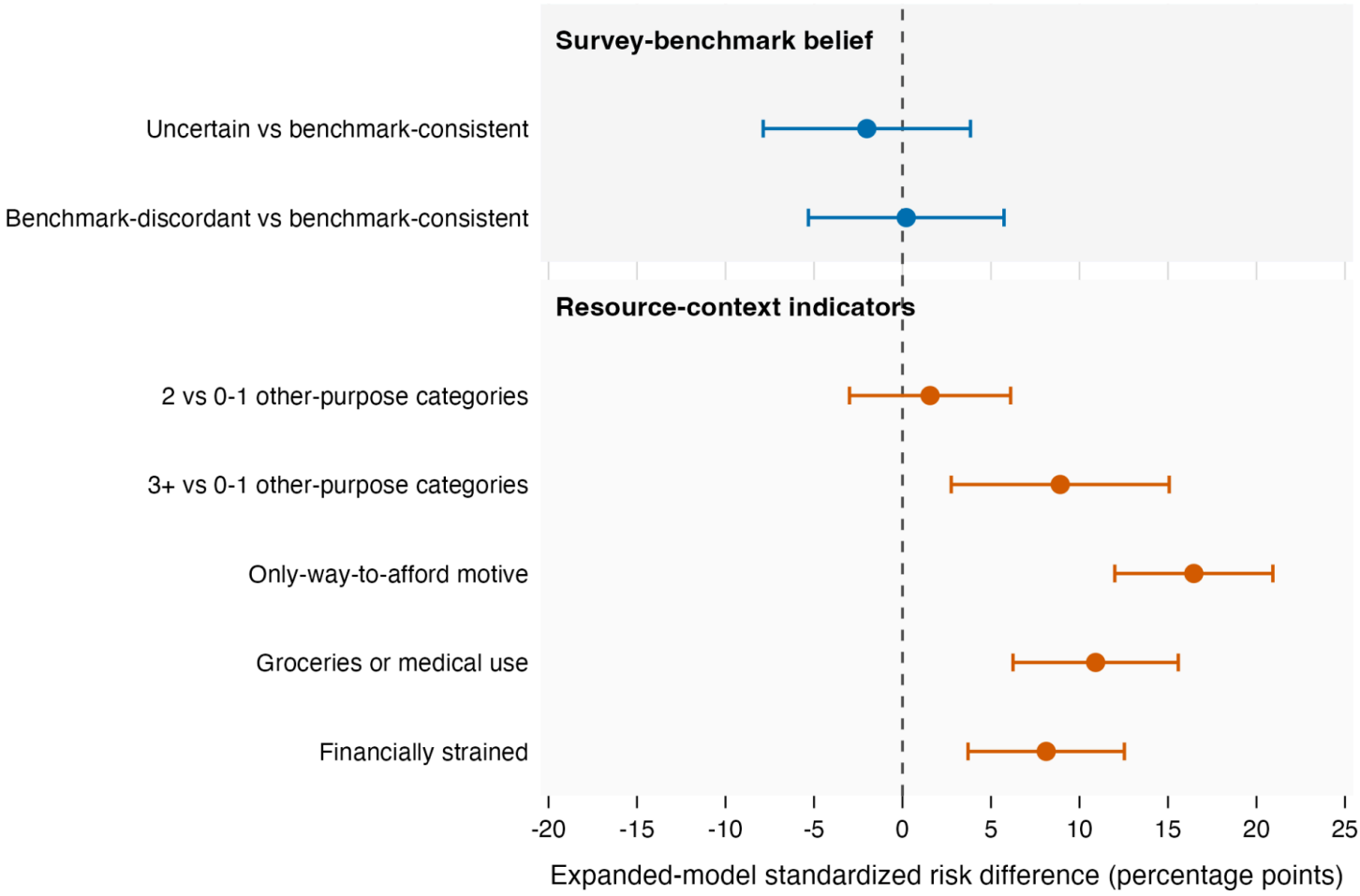


Figure 1

Expanded-model standardized risk differences in self-reported any late BNPL payment. Every estimate comes from the same descriptive specification, which simultaneously included belief category, age group, gender, race and ethnicity, education, household income, employment, other-purpose breadth, affordability motive, groceries-or-medical use and financial strain. Blue denotes belief contrasts and orange denotes resource-context indicators. Points show risk differences in percentage points; horizontal bars show delta-method 95% confidence intervals. Each contrast changes the focal classification while averaging predictions over the observed final-weight distribution of the 2,004 BNPL users. The dashed line marks zero. Other-purpose breadth excludes groceries or food delivery and medical or veterinary expenses and is not transaction frequency. The affordability motive refers to the most recent purchase, whereas the outcome covers any late payment in the preceding year; that contrast is not transaction matched. All contrasts are descriptive and model dependent.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SNBusinessEconomicsOnlineResource1.pdf
- SNBusinessEconomicsOnlineResource2.zip