The Effects of Deleting Medical Debt from Consumer Credit Reports*

Victor Duarte
University of Illinois

Julia Fonseca
University of Illinois and NBER

Divij Kohli
University of Illinois

Julian Reif
University of Illinois and NBER

May 2026

Abstract

In April 2023, credit bureaus stopped reporting medical debt collections below \$500. We study the effects of this information deletion on credit access and financial health. Regression discontinuity estimates comparing individuals just above and below the \$500 threshold show that deletion reduced reported medical collections by 61 percent. We find no evidence of benefits over the subsequent two years. To interpret these findings, we build credit scoring models and show that medical debts, regardless of size, add minimal information for default prediction. Our results suggest that eliminating medical collections entirely from credit reports would be unlikely to affect credit outcomes.

JEL Codes: D18, G51, H75, I18
Keywords: medical debt, credit reports, consumer finance

*We thank Anthony DeFusco, Tatyana Deryugina, Martin Kanz (discussant), Ray Kluender (discussant), Neale Mahoney, Scott Nelson, David Sovich (discussant), and seminar participants at the American Finance Association, ASU, Dartmouth, Georgetown, Georgia State University, Maryland, the Philadelphia Federal Reserve, University of Illinois, and the Wash U Finance Conference for helpful comments and discussions. Hyong-gu Hwang provided excellent research assistance.

Conflict-of-Interest Disclosure Statement

Victor Duarte

I have nothing to disclose.

Julia Fonseca

I have nothing to disclose.

Divij Kohli

I have nothing to disclose.

Julian Reif

I have nothing to disclose.

1 Introduction

One in seven Americans carry medical debt (U.S. Census Bureau, 2022). Unpaid medical bills are often sent to collection agencies and subsequently reported to credit bureaus, resulting in $69 billion in medical debt appearing on consumer credit reports as of 2022.[1] Nearly one-fifth of this total comes from balances under $500. While credit reporting can reduce information frictions and adverse selection, policymakers increasingly worry that the visibility of medical debt on credit reports—often reflecting health shocks—could impair access to credit (The White House, 2025). In response, the three major U.S. credit bureaus announced in April 2023 that they were no longer including medical debt collections below $500 in credit reports. Building on this reform, the Consumer Financial Protection Bureau (CFPB) finalized a rule in January 2025 to eliminate all remaining medical debt collections above $500 from credit reports. However, a federal court struck down the rule in July 2025, leaving the regulatory future of medical-debt reporting in doubt. Thirty U.S. senators criticized the policy reversal, arguing that keeping medical debt on credit reports "blocks working families from access to credit" (Warnock et al., 2025).

Lenders routinely rely on credit-report variables and commercial credit scores in screening and pricing decisions. Basic economic theory therefore suggests that removing negative information from credit reports should improve credit access for affected individuals by raising commercial credit scores such as FICO or reducing predicted default risk in lenders' internal risk models. However, medical debt may be a poor signal of credit risk. These debts are often inaccurate, frequently reflecting bills that were already paid or should have been covered by insurance (CFPB, 2022). If medical debt contains little predictive information beyond standard credit-report variables, its removal may have limited effects on lending outcomes. Thus, the ultimate effect of removing medical debt information on lending is uncertain and remains an open empirical question.

This paper uses a regression discontinuity (RD) design to estimate the effects of deleting medical debt collections from credit reports on a comprehensive set of credit outcomes. Our empirical strategy exploits the $500 threshold introduced by the April 2023 policy change,

[1] Unless otherwise specified, our use of the term "medical debt" refers to medical debt collections that appear on consumer credit reports, which represent a subset of consumers' total medical liabilities.

comparing individuals whose medical debts in 2022 were just above or below the reporting cutoff prior to deletion. Using data from the Gies Consumer and Small Business Credit Panel (GCCP), we show that the April 2023 deletion reduced the number of medical debt collections per consumer by 0.30 (61%) by 2024. However, we find no evidence that removing these debts affected credit scores, credit limits or utilization, repayment behavior, payday borrowing, or other related outcomes in the year following removal. The 95% confidence intervals rule out increases in credit scores greater than 5.48 points (0.89%) and decreases in the balance-to-limit ratio of revolving credit exceeding 2.7 percentage points (5.1%). More broadly, we are well-powered to reject effects comparable to those found in the bankruptcy flag removal literature (e.g., Musto, 2004; Gross et al., 2020). We find no evidence of strategic manipulation of the running variable around the $500 threshold, and our placebo and falsification tests show that outcomes exhibited no discontinuities at the cutoff prior to the April 2023 deletion, supporting the validity of our RD design. We obtain similarly null results when extending the analysis to 2025, two years after the deletion.

To help interpret these null findings, we investigate whether medical debt collections contained incremental predictive information for default risk prior to 2023. We first show, in simple parametric specifications, that medical debt predicts default in isolation, but this signal diminishes rapidly once standard credit-report variables are included. We then evaluate incremental predictive content using a flexible classification model, comparing otherwise identical models trained on rich credit histories with and without medical debt variables. The models perform nearly identically, implying that small medical debts contribute minimally to risk assessment. This conclusion is reinforced by placebo tests showing that these variables behave similarly to a randomly generated noise predictor. Despite their limited informational content, we show that medical debt variables can still affect predicted risk by injecting noise into credit risk models, particularly for low-income borrowers with short credit histories.

Extending this analysis, we show that larger medical debts ($\geq$ \$500) are similarly uninformative: excluding all medical debt from the model produces only negligible changes in classification performance, with no meaningful deterioration in the model's ability to identify defaulters. These results suggest that eliminating *all* medical debts from credit reports would be unlikely to materially affect lending decisions.

Our findings are directly relevant to ongoing policy debates over the treatment of medical debt under today's reporting and scoring environment. Over the past decade, reporting rules and model specifications have evolved. These developments include the removal of paid medical collections from credit reports and the exclusion of medical debt collections from VantageScore, a widely used credit scoring model. Although these changes reduced the prominence of medical collections in credit reporting and scoring, they did not eliminate the role of medical debt in risk assessment. During the period we study, FICO scores—the most widely used credit scores in U.S. lending—continued to incorporate medical collections, and these variables remained available as inputs into lenders' internal risk models. Consequently, whether medical debt appears on credit reports has remained central to policy debate, with advocates arguing that removing such information could expand credit access (CFPB, 2022; Warnock et al., 2025; The White House, 2025). In addition, because the visibility of medical collections may shape repayment incentives, removing medical collections from credit reports could affect not only credit access but also subsequent credit performance.

Our study contributes to research on debt relief and financial outcomes, which has examined contexts including residential mortgages (Agarwal et al., 2017; Cherry et al., 2021; Ganong and Noel, 2020; Cespedes et al., 2025), credit cards (Dobbie and Song, 2020), bankruptcy (Dobbie and Song, 2015), student debt (Di Maggio et al., 2020; Dinerstein et al., 2025), and medical debt (Adams et al., 2022). The work most closely related to ours is Kluender et al. (2024), who conduct two large-scale randomized experiments to evaluate the effects of medical debt forgiveness—a policy often discussed alongside information deletion. They find that debt forgiveness modestly improves credit scores and credit card limits when medical debts are reported to credit bureaus, but has no effect otherwise.[2] Their results highlight the central role of information visibility—specifically, whether a debt appears on a credit report—in determining the effectiveness of debt relief. Our study tests this mechanism directly by isolating the information channel. Using a much larger sample of consumers with

[2] Kluender et al. (2024) find no average effect of debt relief on credit scores in their two main samples, which include many consumers whose forgiven medical debts never appeared on credit reports. However, they emphasize that effects emerge precisely when the debt is visible to lenders: in their credit reporting sub-experiment—which focuses on consumers whose medical debts would otherwise have appeared on their credit reports—debt relief increases credit scores and limits.

reported medical debts and exploiting the 2023 reporting cutoff, we show that deleting medical debts from credit reports has no detectable effects across a broad set of credit outcomes, including payday borrowing, an important margin rarely examined alongside mainstream credit outcomes. This finding runs counter to expectations expressed by policy advocates and by the CFPB in its 2025 rulemaking, both of which anticipated that removing medical debt from credit reports would expand credit access.

We also extend the literature on information deletion by studying a novel context: medical debt. Prior work has studied the effects of removing bankruptcy flags and unpaid debts from credit reports on borrowing or labor market outcomes (Musto, 2004; Bos et al., 2018; Liberman et al., 2019; Dobbie et al., 2020; Gross et al., 2020; Herkenhoff et al., 2021; Jansen et al., 2024), as well as the effect of expanding lenders' access to borrower credit histories (Miller, 2015). Beyond credit markets, related studies examine the removal of criminal history or credit-report information in employer screening, highlighting how information deletion can affect screening incentives and produce unintended labor market effects (Agan and Starr, 2017; Bartik and Nelson, 2025; Agan et al., 2026). A distinctive feature of our setting is that lending decisions are based on predicting default using an observable set of credit-report variables, allowing us to test directly whether the deleted information improves risk prediction. More broadly, our results suggest that in settings where researchers can observe the underlying information set, the predictive content of deleted information can help identify the likely effects of information removal.

Third, we relate to the literature on machine learning in credit markets. We provide rigorous, out-of-sample evidence that medical debt collections have negligible incremental predictive value for default risk in modern credit scoring models.[3] However, we also show that even unreliable information like medical debt can still influence predicted default probabilities, particularly for consumers with short credit histories—not because it contains meaningful signal, but because it injects noise into prediction models when reliable information is scarce—a phenomenon previously documented by Blattner and Nelson (2022). Prior studies

[3] Brevoort and Kambara (2014) show that medical debt collections are less predictive of future credit performance than non-medical debt collections, but do not evaluate the incremental contribution of medical debt to overall model performance. FICO (2023) evaluates the effect of mechanically removing paid and small medical debt collections from existing score formulas and finds little change in predictive performance of FICO models.

have used machine learning to study information deletion (Liberman et al., 2019) and to develop credit scoring models (e.g. Khandani et al., 2010; Frost et al., 2020; Sadhwani et al., 2020; Fuster et al., 2022; Meursault et al., 2022; Agarwal et al., 2023; Blattner et al., 2024; Chioda et al., 2024). Building on this body of work, we construct a credit scoring model using XGBoost, a state-of-the-art prediction algorithm, and achieve high predictive accuracy across multiple standard metrics.

Finally, we contribute to a growing literature on debts in collections and the debt collection industry (e.g., Fedaseyeu and Hunt, 2018; Fedaseyeu, 2020; Cheng et al., 2020; Kluender et al., 2021; Guttman-Kenney et al., 2022; Keys et al., 2022; Fonseca, 2023; Lin, 2024). Consistent with our RD estimates, Batty et al. (2022) and Goldsmith-Pinkham et al. (2023) find that expanded health insurance coverage reduces medical debt collections but does not improve other financial outcomes. Our results help explain this pattern by showing that medical debt collections contain little incremental information about default risk relative to standard credit-report variables, so changes to the reporting of medical debt are unlikely to meaningfully affect lending decisions or borrower behavior. Like Fonseca (2023), we study both mainstream and subprime credit outcomes by linking traditional credit reports from a major credit bureau to reports from a bureau specializing in alternative financial services. This linkage provides a more comprehensive set of credit market outcomes, particularly for consumers with limited access to traditional credit.

The remainder of this paper is structured as follows. Section 2 provides background on medical debt, credit reports, and credit scoring models. Section 3 describes the data used in our analysis. Section 4 presents RD estimates of the direct effects of deleting medical debt collections. Section 5 investigates whether medical debt is predictive of default. Section 6 concludes.

2 Background

2.1 Medical debt

Medical debt arises when patients are unable to pay the out-of-pocket portions of their medical bills. Healthcare providers typically first attempt to collect unpaid amounts directly from patients. If these efforts fail, providers often turn to third-party collection agencies. These agencies attempt to recover the debt, typically through calls, letters, or lawsuits, and often report unpaid debts to credit bureaus.[4] Because not all medical debts are reported to credit bureaus, the balances observed on credit reports represent only a subset of total medical debt and should therefore be interpreted as a lower bound on consumers' overall medical liabilities.

The consequences of medical debt are complex and challenging to quantify, in part because repayment rates are exceedingly low: medical debt can be purchased for pennies on the dollar (Kluender et al., 2024). This contrasts sharply with other forms of unsecured debt, such as student loans and credit card debt. Unlike medical debt, student loans are rarely dischargeable in bankruptcy—elimination typically requires proving "undue hardship," a demanding legal standard—while credit card debt has much higher repayment rates, as issuers can threaten to restrict future access to credit for delinquent borrowers. Additionally, many states provide consumer protections specific to medical debt, including limits on wage garnishment and prohibitions on home foreclosure (Robertson et al., 2022).[5]

Popular commentary frequently highlights the relationship between medical debt and personal bankruptcies. While many bankruptcy filers do carry medical debt, this correlation does not necessarily imply causation. To assess causality, Dobkin et al. (2018) examine how hospitalizations affect the likelihood of filing for bankruptcy within four years of admission. They find that hospitalizations account for approximately 4 percent of personal bankruptcies among non-elderly adults and about 6 percent among uninsured non-elderly adults. These results indicate that health shocks can trigger financial distress, but whether medical debt

[4]While hospitals can report unpaid medical bills directly to credit bureaus, this practice is uncommon (Brevoort and Kambara, 2014).

[5]Colorado and New York prohibited reporting of all medical debt collections in August and December 2023, respectively, several months after the nationwide deletion of small medical debts. Our results are robust to excluding these states from the analysis (Table A.14).

itself adds predictive value beyond standard credit variables remains an open question.

2.2 Credit reports

When a consumer applies for credit, the lender's underwriting system typically requests a credit report from one or more of the three national credit bureaus: Equifax, Experian, and TransUnion. The bureau locates the borrower's credit file and returns both the credit report and, often, a credit score such as FICO or VantageScore. The lender then incorporates this information into its internal credit decision process, which may rely primarily on the bureau-provided credit score or, for more sophisticated institutions, combine it with proprietary risk models that draw on the underlying credit-report data. By making borrowers' prior credit histories observable at the point of application, credit reports reduce information frictions and adverse selection.

A credit report summarizes a consumer's credit accounts, including revolving credit lines (e.g., credit cards) and installment loans (auto, student, mortgage). It records balances and limits, credit utilization, payment history, and negative marks such as late payments, debt collections, and bankruptcies. Debt collection entries include information on the number and total balance of medical collection accounts.

Over the past several years, credit bureaus have substantially revised how medical debts are treated in credit reporting. On March 18, 2022, the three national credit bureaus jointly announced several reforms. Beginning July 1, 2022, paid medical debts in collection were removed from consumer credit files, and the waiting period for reporting unpaid medical debt was extended from six months to one year, giving consumers more time to resolve billing disputes or secure insurance coverage.

As part of that same announcement, the bureaus also pledged to delete all medical debt collections with an initial reported balance under $500 by the "first half of 2023" (Business Wire, 2022). On April 11, 2023, they confirmed completion of these deletions. Although the exact timing of the deletions is uncertain, Quinn (2023) reports that they occurred in March and April 2023. By removing this information from lenders' view, the policy also reduced the leverage of debt collectors, since they could no longer use credit reporting as a tool to pressure repayment.[6]

Medical debt collections can influence credit decisions through lenders' internal underwriting models or through commercial credit scores that summarize credit-report information. Section 5 evaluates how removing medical debt information may influence lenders' internal underwriting models; here, we describe its treatment in commercial credit scoring models.

FICO and VantageScore are the two dominant commercial credit scoring frameworks in the United States, each offered in multiple versions that differ in their treatment of specific credit-report variables. Historically, commercial credit scoring models treated medical and nonmedical collections similarly, aggregating them into a single "total collections" variable (Brevoort and Kambara, 2015). This likely reflects the fact that early credit bureau data did not distinguish between medical and non-medical collections, leading early credit scoring models to use a unified collections measure (FICO, 2023). That structure persists in some widely used scoring models today. Notably, FICO 8—the most widely used FICO model during our study period—does not distinguish between medical and non-medical collections and instead incorporates them into that unified collections measure.

Between 2014 and 2017, FICO and VantageScore introduced models that assign a lower weight to medical than to non-medical collections—the FICO 9 and the VantageScore 4.0 (FICO, 2023; VantageScore, 2022). In January 2023, VantageScore went further and removed all medical debt collection data from its 3.0 and 4.0 models, stating that the change would have "minimal" impact on predictive performance (VantageScore, 2022).[7] These changes created differential exposure across scoring models during our study period. Because VantageScore removed medical collections in January 2023, the April 2023 deletion of small medical debts from credit reports could not directly affect VantageScores. By contrast, all versions of the FICO model—including FICO 8, FICO 9, and FICO 10 (introduced in 2020)—continued to incorporate unpaid medical collections through April 2023 and still include medical collections with balances exceeding $500 today (FICO, 2023). Thus, despite

[6]The Fair Debt Collection Practices Act prohibits abusive or deceptive practices by third-party debt collectors.

[7]Because this change affected all consumers, it does not threaten the internal validity of our empirical analysis, which exploits the removal of small medical debts under $500 from credit reports.

reporting reforms, medical debt remained embedded in the dominant commercial credit scoring framework during our study period.

Deleting medical debt collections from credit reports could affect credit outcomes through several distinct channels. First, because widely used scores such as FICO continued to incorporate medical collections during our sample period, removal could mechanically increase scores for affected consumers. Second, deletion changes the information set available to lenders' internal risk models, potentially altering approval decisions or pricing. Third, eliminating the visibility of medical debt collections may weaken debt collectors' leverage, potentially affecting repayment behavior and subsequent credit performance. Whether these channels translate into meaningful changes in credit outcomes is ultimately an empirical question. Deletion may matter if medical collections play a substantive role in underwriting or borrower behavior, but may have limited effects if they are largely redundant with other observable characteristics.

3 Data

Our study uses the Gies Consumer and Small Business Credit Panel (GCCP), a panel dataset of anonymized credit records for consumers and small businesses obtained from a major credit bureau. The GCCP features a one-percent random sample of individuals with a credit report, linked to alternative credit records and business credit records for individuals who own a business.[8] The dataset spans 2004–2025 and contains annual snapshots measured at the close of the first quarter of each year. Sampling is based on the last two digits of Social Security numbers. This sampling method accounts for natural flows into the panel as new Social Security numbers are issued, as well as outflows due to death or prolonged inactivity, ensuring that the sample remains representative of the broader population over time.

Each GCCP record includes all reported debt obligations, or "tradelines," with information on credit type (mortgage, auto, student, or credit card), balances and limits, and

[8] Alternative credit records include information not reported to the major credit bureaus, such as payday loans and title loans. See Fonseca (2023) and Correia et al. (2023) for a discussion of the link between mainstream and alternative credit records in the GCCP, Fonseca and Wang (2023) on the link between consumer and business credit records, and Fonseca and Liu (2024), Howard and Shao (2022), and Fonseca et al. (2024) for other papers using the GCCP.

payment history. The data also include VantageScores, public records such as bankruptcies and judgments, debts in collections, and demographic variables such as age, sex, income, and 5-digit zip code. Demographic variables are administrative or modeled: age is computed from date of birth, sex is a bureau-provided name-based classification, and income is a bureau-provided estimate based on credit-report variables.

We classify a debt collection as medical if the creditor is labeled as Medical/Health Care or if the furnisher operates in a medical or health-related sector.[9] Using this classification, we compute estimates of the prevalence and aggregate balances of medical debt collections in the GCCP. As shown in Tables A.1 and A.2, these GCCP-based estimates closely match corresponding benchmarks from external sources. In 2022, our GCCP data indicate that \$69 billion in medical debt appeared on consumer credit reports, of which \$12 billion came from balances under \$500. As in previous studies, our measure does not capture medical spending paid using credit cards, which cannot be identified in the data.

We restrict the main analysis to the years 2019–2024 and to consumers aged 18 or older. We drop observations with missing age, credit score, or income, and exclude consumers with inconsistent age histories (deviations exceeding 10 years from a consistent birth year). The final sample includes 15,366,574 observations, summarized in Table 1. In the full sample, about half of consumers are female, the average credit score is 702, average annual income is \$51,970, and average total balances across all credit products is \$76,650. About 20 percent of consumers have an alternative credit record, and their average number of medical collections is 0.25. In an extension, we incorporate 2025 data to examine whether longer-run patterns differ from the one-year effects studied in our primary analysis.

Columns (4)–(6) of Table 1 describe consumers with at least one medical debt collection. This group has lower credit scores, lower income, and lower balances than the full sample and is more likely to have subprime credit. Among these consumers, the average number of medical debt collections is 2.44, of which 1.45 involve balances below \$500.

Figure 1 shows that the share of consumers with medical debt collections fell from 16 percent in 2019 to 4 percent in 2024. This decline is driven by the disappearance of medical

[9]Furnisher categories include Dentists, Chiropractors, Doctors, Medical group, Hospitals and clinics, Osteopaths, Pharmacies and drugstore, Optometrists and optical outlets, and Medical and related health-nonspecific.

debts under \$500, which fall from 10 percent in 2022 to zero by 2024 after the credit bureaus stopped reporting them. Because the GCCP snapshots are taken at the end of March each year, the marked decline already evident in the 2023 data implies that many deletions occurred before the April 11, 2023 public announcement, consistent with reports that the deletions occurred in March and April.

For the RD analysis, we further restrict the sample to consumers with at least one medical debt collection in 2022 and a non-missing credit score in 2022–2024. The resulting sample includes 272,490 consumers, totaling 817,470 observations across the three years. Table 2 summarizes this sample in 2022, the year before the deletion of medical collections under \$500. On average, these consumers had 3.53 debts in collections, including 1.56 small medical debts below \$500.

4 Regression Discontinuity Design and Results

4.1 Empirical strategy

We employ an RD design to estimate the direct effect of medical debt deletion on consumer credit outcomes. We first estimate the following model at the account level:

$$Y_{ij}^{2024} = \alpha_1 \mathrm{DEBT}_{ij}^{2022} + \beta_1 \mathrm{ABOVE}_{ij}^{2022} + \gamma_1 (\mathrm{ABOVE}_{ij}^{2022} \times \mathrm{DEBT}_{ij}^{2022}) + \epsilon_{ij} \qquad (1)$$

We refer to this account-level specification as a "first stage," since it documents the effect of deletion on the reporting of individual medical debt collection accounts. The dependent variable, Y_{ij}^{2024}, represents an outcome in 2024 for account j belonging to consumer i. The running variable, $\mathrm{DEBT}_{ij}^{2022}$, is defined as the account's balance relative to the \$500 cutoff in 2022, the year prior to the deletion of small medical debt collections. The indicator variable $\mathrm{ABOVE}_{ij}^{2022}$ is equal to one if $\mathrm{DEBT}_{ij}^{2022} \geq 0$. We approximate the conditional expectation function with a local linear regression, allowing the slope to vary on either side of the cutoff.

We then estimate a separate specification at the consumer level that aggregates the running variable by taking the maximum debt amount across all the consumer's medical debt collection accounts:

$$Y_i^{2024} = \alpha \text{MAXDEBT}_i^{2022} + \beta \text{ABOVE}_i^{2022} + \gamma(\text{ABOVE}_i^{2022} \times \text{MAXDEBT}_i^{2022}) + \epsilon_i \quad (2)$$

The running variable is defined as the largest medical debt for consumer i, relative to the \$500 cutoff: $\text{MAXDEBT}_i^{2022} = \max_j \text{DEBT}_{ij}^{2022}$. Because the April 2023 policy removed all medical collection accounts with balances below \$500, a consumer retains at least one of their 2022 medical collection accounts if any account had a balance of at least \$500; otherwise, all medical collections accounts are deleted. The parameter β therefore captures the intent-to-treat effect of having at least one medical debt collection on the credit report. Equivalently, $-\beta$ measures the effect of having all accounts deleted. Appendix B.1 reports results from alternative specifications that use the minimum debt balance or are estimated at the account level. These approaches define treatment differently—having any account deleted or scaling with the share of deleted accounts—but exploit the same \$500 discontinuity. All yield qualitatively similar results, indicating that our conclusions do not depend on a particular aggregation of accounts.

Our main analysis considers a set of ten consumer credit outcomes spanning credit access, repayment behavior, and payday borrowing. These outcome variables are defined in Table A.4. In addition, our supplementary analysis, reported in the appendix, examines a broader set of outcomes, including credit inquiries, revolving credit limits, alternative credit balances, and mortgage origination.

Our main identifying assumption is that assignment around the \$500 threshold is effectively random. This assumption is plausible because medical charges are typically set by providers using fixed and often opaque pricing, leaving consumers with little opportunity to strategically adjust balances. Additionally, our data come from administrative records, reducing concerns about measurement error or sample selection bias. We assess manipulation by testing for continuity in predetermined covariates and by examining the density of the running variable around the cutoff (Lee, 2008; McCrary, 2008).

The main threat to identification is that other policies or practices might also depend on the \$500 threshold. For instance, if hospitals restrict services once unpaid bills exceed \$500, then any observed discontinuities could reflect hospital practices rather than credit bureau reporting rules. A related concern is that debt collectors may have treated debts below \$500 differently even before the 2023 policy change—for example, by routinely choosing not to report them. To assess these possibilities, we estimate placebo RD specifications with outcomes measured before 2023, prior to the removal of small medical debts from credit reports.

All RD regressions use a triangular kernel. Our preferred specification uses a mean-squared error optimal bandwidth that is symmetric around the cutoff and allowed to vary across outcomes. We report robust bias-corrected confidence intervals to account for potential misspecification of the estimating equation (Calonico et al., 2014).

4.2 Results

We begin by estimating the first-stage effect of the 2023 deletion on medical collections accounts. Panel A of Figure 2 shows that by 2024, nearly all accounts with balances below \$500 in 2022 had been removed from credit reports, whereas more than 20 percent of accounts above the cutoff remained. Panel B shows similar patterns after aggregating to the consumer level: the deletion reduced the number of medical debt collections per person in 2024 by 0.30 (61.4%).[10]

We next turn to credit access. Figure 3 shows no evidence of discontinuities in credit scores, balances, new accounts, or revolving utilization (balance-to-limit ratio). Because VantageScore already excluded medical collections before the April 2023 deletion, the remaining credit-access outcomes—total balances, new accounts, and revolving utilization—provide the most direct test of whether the deletion affected lending decisions.[11] Table 3 presents formal estimates. The 95 percent confidence intervals rule out improvements in credit scores greater than 5.48 points (0.89%), increases in credit balances exceeding \$2,491 (4.98%), increases in

[10] Consumers whose largest 2022 medical collection account was under \$500 may still have medical collections in 2024 if they acquired new accounts above 500 dollars in 2023 or 2024. These new accounts generate the positive values to the left of the cutoff in Panel (b).

[11] Medical debt collections were removed from the VantageScore model in January 2023, so the April 2023 deletion cannot directly affect our credit score measure. Because we measure outcomes one year later, however, indirect effects on the score remain possible: the deletion may influence borrowing behavior or lenders' decisions, which in turn could alter other credit-report variables that enter VantageScore (see Section 2.2).

new credit accounts greater than 0.088 (17.6%), and decreases in revolving utilization larger than 2.7 percentage points (5.1%).

Figure 4 presents analogous results for delinquency, bankruptcy, and alternative credit use (e.g., payday and title loans). The estimates again show no evidence of discontinuities at the cutoff, and for several outcomes the confidence intervals rule out economically meaningful effects. In particular, they rule out decreases in the probability of holding an alternative credit balance greater than 0.31 percentage points (7.86%) and increases in the number of alternative credit accounts exceeding 0.045 (25%). Estimates for delinquent balances and bankruptcy are less precise, but still rule out large effects, including reductions in delinquent balances greater than $761 (38.2%) and decreases in the probability of bankruptcy larger than 1.19 percentage points (37.07%).

Figure A.1 shows results for additional outcomes, including the number of accounts 90+ days past due, the number of new inquiries, revolving limits, total balance in alternative credit accounts, and the number of new mortgage accounts. Again, we find no detectable effects, and the estimates are sufficiently precise to rule out meaningful changes in most outcomes.

Table A.9 focuses on consumers whose debt collections consist solely of medical debts in 2022. For this subgroup, we again find no robust evidence of effects across our main outcomes. The only conventionally significant estimate is a 2.8 percentage point (5.9%) reduction in revolving utilization (balance-to-limit ratio), but this estimate is no longer statistically significant after adjusting for multiple hypothesis testing.[12] We also find no evidence of credit-score gains: the 95 percent confidence interval rules out an increase larger than 9.29 points, or 1.43 percent of the control mean.

These subgroup results differ from Kluender et al. (2024), who found modest benefits of medical debt relief for consumers whose reported collections were entirely medical. In particular, they report a 13.8 point (2.3%) increase in average VantageScores, which falls outside our confidence interval. One important difference is timing. Our outcomes are

[12]We control the family-wise error rate using the Sidak-Holm step-down procedure (Huh and Reif, 2021). The family consists of the 8 outcomes reported in columns (3)–(10) and does not account for the broader set of possible subgroup analyses, which would imply an even more stringent adjustment. For revolving utilization, the conventional (unadjusted) p-value is 0.013 and the Sidak-Holm adjusted p-value is 0.099.

measured after January 2023, when VantageScore removed medical debt collections from its scoring model. The earlier score gains may therefore reflect mechanical effects from scoring models that aggregated medical and non-medical collections into a single collections measure (Brevoort and Kambara, 2014). If non-medical collections are informative about default risk, removing medical collections from an aggregated collections measure can mechanically raise scores even when medical collections contain little independent predictive content or causal relevance for future credit outcomes. Our null effects across a broader set of outcomes are consistent with this interpretation.

Our estimates rule out effects comparable to those from other information deletion settings, such as bankruptcy flag removals. The confidence intervals shown in Table 3 exclude increases in total credit balances (Musto, 2004) and new credit accounts (Gross et al., 2020) similar to those estimated following bankruptcy flag removal. Likewise, Table A.6 rules out similarly sized effects for inquiries (Gross et al., 2020), revolving limits (Gross et al., 2020), and new mortgage accounts (Miller and Soo, 2020). We also reject comparable increases in credit scores, though this comparison should be interpreted cautiously because medical debt was excluded from VantageScore during our sample period, whereas bankruptcies remain included.

Tables A.10 and A.11 extend the baseline 2024 analysis by examining outcomes in 2023 and 2025. The 2023 outcomes should be interpreted cautiously because the GCCP snapshot was measured near the time of deletion and the exact timing of account-level deletions is unobserved. As a result, the 2023 estimates may reflect only partial treatment exposure and provide only a limited window for downstream effects to materialize. Nevertheless, the first-stage effect is already apparent in 2023 and remains visible in 2025. Across all three outcome years, however, the reduced-form estimates for credit access and financial health remain close to zero and statistically insignificant. Overall, the evidence indicates that deleting small medical debt collections had no discernible impact on credit access, repayment behavior, or broader financial outcomes over the two years following deletion.

4.3 Robustness

We evaluate the no-manipulation assumption by testing whether predetermined outcomes are continuous at the cutoff. Figure A.2 shows no statistically significant discontinuities in income, age, or sex. We also examine the density of the running variable. Strategic manipulation would generate excess mass just below the $500 threshold. However, Figure A.3 shows no such pattern. Instead, we observe slight bunching just above the cutoff, consistent with rounding behavior or provider reporting practices rather than strategic manipulation. To assess whether this bunching influences our estimates, we estimate an RD specification that excludes observations in a narrow window above the cutoff ($500–$505), following Barreca et al. (2015). These results, shown in Table A.12, are similar to our main estimates.

We also conduct a series of falsification and placebo tests. Figures A.4, A.5, and A.6 replicate our RD specification using 2022 outcomes instead of 2024 outcomes. As expected, we observe no evidence of discontinuities, including for first-stage outcomes. Figures A.7, A.8, and A.9 repeat our analysis for 2020–2022 in place of 2022–2024 and likewise find no significant effects. Together, these results show that the $500 threshold was not associated with spurious breaks in outcomes prior to the April 2023 deletion, supporting the internal validity of our RD design.

A separate concern is that our sample is restricted to consumers with non-missing credit scores. This restriction could bias our estimates if the deletion of small medical debts affected the probability that a consumer is scorable. However, scoring model design suggests this concern is limited. FICO reports that medical collections do not enter minimum scoring criteria and that their removal has no material impact on the FICO scorable rate (FICO, 2023). Although we use VantageScore, whose minimum scoring criteria are not publicly documented, it removed medical debt collections from its models in January 2023, making it unlikely that medical debts influence scorability during our study period. Nevertheless, to address this concern, Appendix B.2 presents a supplementary analysis showing no evidence that the deletion of small medical debts affected the number of scorable consumers in our data (Table A.13). For completeness, we also estimate difference-in-discontinuities models for our other main outcomes using this unrestricted sample; the resulting coefficients closely match our baseline RD estimates, suggesting that the null findings are not driven by pre-existing discontinuities at the cutoff.

Finally, Figure A.10 presents a complementary event-study difference-in-differences analysis for consumers within \$20 of the \$500 threshold. The event-study estimates show a sharp decline in reported medical collections beginning in 2023 for consumers just below the cutoff, with no corresponding changes in credit access or financial distress. The absence of pre-existing differential trends and the lack of post-deletion responses in the outcome variables reinforce the RD evidence that the reporting change reduced visible medical debt but did not measurably alter consumer credit outcomes.

5 The Predictive Value of Medical Debt

As outlined in Section 2.2, medical debt information could influence lending decisions either through its contribution to commercial credit scores such as FICO or through lenders' proprietary risk models. Yet our RD analysis showed that deleting small medical debt collections from credit reports produced no measurable improvements in credit access or financial health for affected consumers. This absence of effects suggests that information on small medical debts carries little value for underwriting. If these debts contained meaningful information about default risk, their removal would have altered the inputs to lenders' risk assessment models and, in turn, altered loan approval and pricing decisions.

In this section, we test this implication directly by evaluating whether medical debt collections predict default. We train credit scoring models that are identical except for whether they include information on medical collections below \$500, and we repeat the analysis excluding all medical collections. While our model does not replicate any specific lender's proprietary approach, sophisticated lenders deploy cutting-edge risk models and face strong incentives to extract any variable with genuine predictive content (Braunstein, 2010). Thus, if a state-of-the-art model such as ours detects no predictive value in medical debt collections, it is reasonable to infer that lenders' proprietary models are likewise unlikely to extract meaningful signal from them.

We show below that medical debt collections provide no meaningful incremental predictive power beyond standard credit-report variables. We demonstrate this result in several complementary ways: removing medical collections has a negligible effect on different measures of classification error; medical debt variables rank near the bottom of standard measures of variable importance; and excluding them produces only minimal changes in the distribution of predicted default probabilities.

5.1 Credit Scoring With and Without Medical Collections

Credit scoring models estimate the likelihood that a borrower will default based on their financial and credit history.[13] Formally, these models take the form:

$$\Pr(Y = 1 \mid X) = f(X) \tag{3}$$

where Y is a binary default indicator and X denotes a vector of borrower characteristics. The function $f(\cdot)$ represents the mapping from borrower attributes to a predicted default probability, which may be specified parametrically or estimated flexibly using machine learning techniques.

Traditional credit scoring models, such as FICO and VantageScore, are typically logit models estimated on consumer-level data and segmented into groups based on repayment history, commonly referred to as "scorecards" (Federal Reserve Board, 2007; Gibbs et al., 2025). Within each scorecard, a separate logistic regression is estimated, so that model parameters differ across groups. These models generally aim to predict "default," defined as any credit account becoming 90 or more days past due within the next 18–24 months. Predictors typically include variables related to payment history, amounts owed, length of credit history, new credit activity, and credit mix. To capture nonlinear relationships, predictors are binned into discrete categories, which can improve model performance (Federal Reserve Board, 2007).

Rather than relying on binning and group-specific estimation, our baseline model uses

[13] In addition to default probabilities, lenders estimate two other components of expected credit losses: loss given default (LGD) and exposure at default (EAD). Unlike default probabilities, LGD and EAD are rarely estimated at the borrower level; instead, they are typically based on product characteristics or historical averages (Federal Reserve Board, 2013).

XGBoost, a state-of-the-art machine-learning algorithm widely used in classification settings (Chen and Guestrin, 2016). This flexible, tree-based ensemble method captures complex, nonlinear interactions and typically outperforms standard parametric models, making it a natural choice for assessing whether medical collections contain any predictive signal. Its structure also resembles the advanced modeling approaches likely employed by sophisticated lenders. Section 5.3 shows that our conclusions do not hinge on this specification: a traditional logit approach yields similar results.

Our model includes $n = 48$ predictors that capture a broad set of credit-related characteristics. There are four predictors conveying information on medical debts: the number of medical debts below \$500, the total balance on medical debt accounts below \$500, and two analogous variables for medical debts above \$500. In addition, the model includes detailed measures of delinquency (counts and balances of past-due accounts across multiple horizons, derogatory indicators, and recency measures); counts and balances of non-medical collections; indicators of bankruptcies and other public records; balances and account counts across major credit types; credit utilization measures; account-age characteristics; and recent inquiries and new accounts.[14] Consistent with prior work, our model excludes variables prohibited by the Equal Credit Opportunity Act (ECOA)—such as race, sex and marital status—as well as variables that may proxy for them, such as geographic identifiers (Federal Reserve Board, 2007; Blattner and Nelson, 2022; Gibbs et al., 2025). Although age is not prohibited under ECOA, its inclusion requires special documentation and validation under fair-lending regulations (Federal Reserve Board, 2007). We therefore exclude it from our model, consistent with the approach used in FICO scores (FICO, 2025).

We train three person-level credit scoring models. The first includes the number of medical debt collections below \$500 and the total balance on these accounts, as well as analogous variables for medical debts above \$500. The second excludes information on medical collections below \$500 while retaining information on medical debts above \$500. The third excludes all medical collection variables. Using data from 2019 to 2021—prior

[14] "Derogatory" events refer to serious negative credit outcomes, such as repossessions, but do not include collections or bankruptcies, which we model separately. The recency measure captures the time elapsed since the most recent derogatory event. For more information on the predictors included in traditional credit scoring models, see FICO (2025).

to the removal of information on medical collections below \$500—we predict whether a consumer experiences default between 2020 and 2021, based on borrower characteristics measured in 2019. We define default as any account becoming 90 or more days past due, excluding collections. The dataset contains records for over 2.4 million consumers. We split the data into a 10% holdout test set and a 90% training and validation set, then further split the latter into 90% training and 10% validation subsamples (so the validation set accounts for 9% of the full sample). We tune hyperparameters via Bayesian optimization using a tree-structured Parzen estimator (TPE) (Bergstra et al., 2011). At each iteration, we fit the XGBoost classifier on the training set and select parameters that maximize performance in the validation set, measured using the F1 score—a standard metric that balances precision and recall and is described below. We then fit the final model on the combined train and validation data before evaluating model performance out-of-sample on the holdout test set. Predicted default probabilities are converted into binary predictions using a threshold of 50%. Appendix C provides additional details on model fitting and hyperparameter tuning.

We measure model performance using several metrics. We report accuracy—the share of correct predictions—due to its widespread use and interpretability, though it provides limited insight in our setting: a model that predicts no defaults achieves an accuracy score of 86.9%, matching the share of consumers who did not default. We also compute the area under the Receiver Operating Characteristic curve (AUC), which measures the probability that the model assigns a higher default probability to a true defaulter than to a non-defaulter.

Our preferred metrics are precision and recall, which better capture a model's ability to predict rare events than either accuracy or AUC (Davis and Goadrich, 2006). Precision and recall are defined as:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$\text{Recall} = 1 - \text{False Negative Rate} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

Precision is the share of predicted defaulters who actually defaulted, while recall is the share of actual defaulters who were correctly identified. High precision minimizes the misclassification of creditworthy borrowers, helping lenders avoid missed profitable opportunities. High recall ensures that true high-risk borrowers are identified, reducing the likelihood of inadvertently lending to borrowers who are unlikely to repay. To balance these objectives, we also compute the F1 score, which is the harmonic mean of precision and recall:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Finally, for ease of interpretation, we also report the false positive rate, which captures the share of non-defaulters incorrectly classified as defaulters:

$$\text{False Positive Rate} = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}}$$

Appendix C.4 provides formal definitions of these metrics and discusses their relationships.

5.1.1 Results

As a benchmark, we begin by estimating a linear probability model relating the number of medical debt collections to two-year default rates (Table A.15). Column (1) shows a statistically significant association between medical debt and default, but this relationship attenuates sharply as additional credit-report controls are introduced. Across debt categories, the estimated coefficients decline by 60–80 percent once standard credit-report controls are included, and decline further after introducing flexible polynomial terms.[15] This pattern suggests that the predictive content of medical debt largely reflects correlations with other credit-report variables rather than independent signal. Although richer parametric specifications—such as spline-based models or models with extensive interaction terms—could further increase flexibility within this framework, linear probability models are not well suited to classification and are limited in their ability to capture complex nonlinearities. We therefore turn to a flexible, high-dimensional classification approach to assess whether medical debt retains any incremental predictive value once nonlinearities and interactions are allowed for more generally.

[15] Appendix Table A.16 shows a similar attenuation pattern when medical debt is measured by balance amounts rather than counts, using the maximum medical collection balance across accounts.

Table 4 reports out-of-sample model performance for our XGBoost model. Column (1) shows that our baseline model with the full set of predictors performs exceptionally well in predicting default. For instance, we achieve an AUC of 0.905, exceeding the typical range of 0.66 to 0.88 reported in prior work (e.g., Butaru et al., 2016; Fuster et al., 2022; Agarwal et al., 2023; Blattner and Nelson, 2022; Chioda et al., 2024). Our F1 Score of 0.561 also compares favorably to the literature and significantly exceeds the corresponding metric from our linear probability model.[16] To interpret this value, we focus on its components, precision and recall, which are more commonly reported. Our recall of 0.454 ranks among the highest reported values, with most prior studies finding recalls between 0.35 and 0.41 (e.g., Butaru et al. (2016), Agarwal et al. (2023)). Two exceptions are Khandani et al. (2010) and Chioda et al. (2024), who report recalls of 0.654 and 0.749, respectively, but both use substantially shorter prediction horizons of 3 and 6 months.[17] Shorter prediction windows generally yield higher precision and recall, helping explain the stronger performance in these studies. In addition, Chioda et al. (2024) use a 20% threshold—much lower than our 50% threshold—which further boosts recall at the expense of precision.

Our precision score of 0.735 also compares favorably with the literature, where reported values typically range from 0.06 to 0.50 (e.g., Butaru et al. (2016), Fuster et al. (2022), Agarwal et al. (2023), Chioda et al. (2024)). The sole exception is Khandani et al. (2010), who achieve a higher precision of 0.853 but, again, over a much shorter 3-month prediction horizon.

Comparing the first two columns of Table 4 shows that removing information on medical collections below $500 has no meaningful impact on model performance. Our preferred metric, the F1 score, together with accuracy and AUC, remains unchanged up to the third decimal. Decomposing the F1 score down into its two components, we observe a negligible decrease of 0.001 in recall and an *increase* of 0.002 in precision when small medical collections are removed. The third column further shows that deleting information on *all* medical

[16] The comparison should be interpreted with caution, as the linear probability model metrics are computed in-sample, whereas the XGBoost results are evaluated out-of-sample. Out-of-sample evaluation provides the appropriate benchmark for assessing predictive performance. The linear probability model is included solely as a benchmark and is not intended as a competing classification approach.

[17] We compute recall and precision for Khandani et al. (2010) using the confusion matrix for the December 2008 3-month forecast with a 50% classification threshold.

collections, including those exceeding \$500, likewise has a negligible effect. Removing all medical collections *increases* the F1 score by 0.001. These results suggest that the CFPB's 2025 final rule to eliminate all remaining medical collections from credit reports is unlikely to affect the accuracy of credit scoring models.

To shed light on the limited impact of medical collections on model performance, Figure 5 reports average absolute SHAP values, which summarize each predictor's contribution to the model's predictions. The two most important predictors are Credit Amount in Last 6 Months and Number of Accounts Never Past-Due or Derogatory; on average, they shift the model's predicted log-odds of default by 0.337, corresponding to roughly a 4 percentage point change relative to an average default probability of 13 percent.[18] By contrast, the number and balance of medical collections below \$500 rank near the bottom of the importance distribution, with a combined average SHAP value of 0.019 (= 0.007 + 0.012)—nearly 20 times smaller than the top predictors. Medical collections above \$500 are even less important, with the number of large medical collections being the least important predictor in the model. Together, the four medical collection variables account for over 8 percent of the model's 48 predictors but less than 0.7 percent of the total SHAP mass, meaning that excluding all medical collection information would change predicted default risk for the average borrower by a trivial amount relative to standard credit-report information.

Figure A.11 further shows that removing information on medical collections below \$500 has minimal effect on predicted default probabilities: only about 10% of consumers experience a change greater than 2 percentage points.[19] Although the tails of Figure A.11 suggest that small medical debt collections might improve predictive performance for a narrow set of borrowers, these changes are far more likely to reflect estimation noise than meaningful differences in default risk. To investigate this further, we categorize consumers into three groups based on the change in their predicted default probabilities across the two models:

Higher risk: Consumers in the top 5 percent of the distribution, whose predicted

[18]An average default probability of 13 percent implies odds of $0.13/(1-0.13) = 0.149$. A log-odds shift of 0.337 multiplies these odds by $\exp(0.337) \approx 1.40$, corresponding to a change in the predicted default probability of about 4 percentage points.

[19]For context, 2 percentage points corresponds to the difference in 2022Q2–2024Q1 90-day delinquency rates between consumers with a VantageScore of 300–500 and those with scores of 501–520 (https://www.vantagescore.com/lenders/risk-ratio/).

probabilities increase by approximately 2 percentage points or more when small medical debts are removed from the model.

Lower risk: Consumers in the bottom 5 percent of the distribution, whose predicted probabilities decrease by approximately 2 percentage points or more when small medical debts are removed from the model.

Unaffected: Consumers between the 25th and 75th percentiles, whose predicted probabilities change by no more than approximately 0.2 percentage points.

If small medical debts were truly predictive of default, these groups should exhibit clear differences. For instance, consumers reclassified as lower risk should, on average, hold more small medical debts than those reclassified as higher risk.

However, we do not observe this pattern. Table 5 shows summary statistics for all three groups. While both the higher and lower risk groups differ significantly from the unaffected group, they are remarkably similar to each other. Figure 6 illustrates their similarity using balancing regressions. Each dot plots the coefficient from a regression of a standardized variable on an indicator for either the lower risk (blue) or higher risk (red) group indicator. If small medical debts were strongly predictive of default, we would expect a pronounced sorting effect along relevant characteristics—particularly the number of medical debts. However, Figure 6 reveals no such pattern: consumers in the higher and lower risk groups appear statistically indistinguishable across a wide range of characteristics, including the presence of medical debts.

The similarity between the positively and negatively treated groups, despite their substantial divergence from the unaffected group, suggests that the changes in predicted risk are driven by estimation noise rather than meaningful differences in underlying credit risk. Similar to Blattner and Nelson (2022), default probabilities are estimated with considerable noise for low-income consumers with short credit histories. In such settings, even uninformative predictors can receive non-zero weights during model training.[20] Consequently, excluding an uninformative feature can shift predicted probabilities for noisy cases in largely

[20] In theory, machine-learning algorithms such as XGBoost should assign zero weight to uninformative variables. In practice, however, finite sample noise in the training data can lead to spurious associations.

random ways, producing two groups with sizeable—but economically meaningless—changes in predicted risk. This process effectively assigns consumers quasi-randomly to the higher and lower risk groups, while separating them from the more stable, unaffected group. To validate this interpretation, Section 5.2 introduces a randomly generated variable into the model and shows that excluding it produces a nearly identical pattern to the one observed when excluding small medical debt collections.

5.2 Placebo Test Using a Random Noise Predictor

As a placebo test, we compare the effect of removing medical collections below $500 to the effect of removing a purely random predictor. Specifically, we train a version of our model that includes a variable drawn randomly from a uniform distribution and compare its performance to our baseline model, which excludes this noise variable.

Table A.17 reports the performance metrics for this placebo exercise. Column (1) adds a purely random noise variable to the baseline set of predictors. Column (2) reproduces the baseline model; the comparison between columns (1) and (2) thus isolates the effect of removing the noise predictor. Column (3) removes small medical debt collections from this same baseline model, reproducing column (2) of Table 4. Strikingly, the negligible changes in model performance induced by excluding small medical debt collections closely mirror those induced by excluding the random noise variable. Figure A.12 reinforces this comparison, showing that the distribution of changes in predicted default probabilities after removing the random predictor is nearly indistinguishable from the distribution obtained after removing small medical debt collections.

Figure A.13 provides additional evidence. Panel A reproduces the balance plot from Figure 6, showing that removing small medical collections assigns consumers to the positively and negatively treated groups in a way that yields no systematic differences across observable characteristics. Panel B performs the same exercise for the random variable and produces a similar pattern. In both panels, the consumers reclassified as higher or lower "risk" have lower credit scores, lower income, and lower balances, consistent with the conclusion that predicted default probabilities are substantially noisier for these borrowers (Blattner and Nelson, 2022).

One potential concern with this placebo exercise is that removing *any* variable—regardless of its predictive power—might fail to generate systematic differences between the positively and negatively treated groups. To address this concern, Figure A.14 examines the impact of removing a clearly informative predictor: credit history length, as measured by average account age, the age of the oldest account, and the age of the oldest account that was never delinquent or derogatory. Credit history length is widely used in credit scoring models as a predictor of default (Federal Reserve Board, 2007; Gibbs et al., 2025), with longer credit histories generally associated with lower default risk, and average account age is one of the most important features in our model as measured by SHAP values (Figure 5). When we exclude these variables, we observe pronounced and intuitive sorting. Although both tail groups differ from the unaffected group, consumers reclassified as lower risk are systematically younger and have shorter credit histories, as measured by average account age, than consumers reclassified as higher risk. This pattern is consistent with the removal of a predictor that assigns higher risk to borrowers with limited credit histories. In contrast to the effects of removing small medical collections or a random variable, removing a genuinely predictive predictor produces clear and systematic differences across groups. This contrast confirms the validity of our placebo test: removing uninformative variables produces largely symmetric or quasi-random sorting, whereas removing valuable predictors yields meaningful differences.

5.3 Robustness Using a Traditional Logit Model

As a robustness check, we show that our conclusions are unchanged when we replace the XGBoost model with a traditional logit scorecard approach, described briefly in Section 5.1 and Gibbs et al. (2025), and in detail in Federal Reserve Board (2007). Following industry practice, we estimate separate models (scorecards) for three groups of consumers: major derogatory files, thin files, and clean files. The major derogatory scorecard includes consumers with at least one account 90 or more days past due, an account in collections, or a public record such as a bankruptcy or foreclosure. The thin scorecard covers consumers not in the major derogatory group who have fewer than three accounts. The clean scorecard applies to consumers with three or more accounts who are not classified as major derogatory.

We use 2019 data to construct the same 48 predictors as in our baseline model, along with the variables needed to segment consumers into the three scorecards. In 2019, 32.4% of consumers fell into the major derogatory scorecard, 12.7% into the thin scorecard, and 54.9% into the clean scorecard. Within each scorecard, we bin predictors using the OptBinning Python library, which applies a mixed-integer programming model to create monotonic and statistically meaningful bins—an approach consistent with traditional industry scorecard construction (Federal Reserve Board, 2007).

Table A.18 reports the performance of this logit scorecard model. Column (1) shows results including medical collections, Column (2) excludes the number and balance amounts of medical collections below $500, and Column (3) excludes all medical collections. Although the logit model underperforms our baseline model on all metrics except the False Positive Rate (see Table 4), its predictive accuracy still compares favorably to the prior literature summarized in Section 5.1.

Consistent with our baseline results (Table 4), removing small medical collections has no meaningful effect: all metrics remain unchanged up to the third decimal place (column 1 vs. 2 of Table A.18). Likewise, dropping all medical collections produces virtually no change relative to the full model: the only differences are a 0.001 decline in the F1 score, driven by a 0.001 decline in recall, and a 0.001 decline in accuracy. These results reinforce the conclusion that medical debt collections—whether below or above $500—offer little, if any, incremental value for predicting default.

6 Conclusion

This paper studies the effects of deleting small medical debt collections from credit reports. Contrary to the stated policy goals of recent policy proposals, we find that removing this information has no meaningful impact on credit access or financial health for affected consumers. Our analysis focuses on credit outcomes; we do not study how medical debt or its removal may affect other domains, such as mental or physical health, or other dimensions of hardship.

Leveraging the nationwide removal of medical debts under $500 from credit reports in

2023, we use a regression discontinuity design to estimate the causal effects of information deletion. We find no evidence that consumers benefit from the removal of this information in terms of credit access, repayment behavior, or payday borrowing, and we rule out even small effects. To interpret these null findings, we show that medical debt collections—both small *and* large—carry little incremental predictive value for default by comparing credit scoring models with and without medical debt variables. The negligible performance difference between these models underscores that medical debt information provides little value to lenders' risk assessment.

These results are directly relevant to recent proposals to remove all medical debt collections from credit reports, including the CFPB's rule finalized in 2025 and later struck down. While the policy debate often assumes that deleting medical debt would expand credit access, our evidence suggests that eliminating medical collections—small or large—is unlikely to materially affect credit outcomes.

References

Adams, A., R. Kluender, N. Mahoney, J. Wang, F. Wong, and W. Yin (2022). The impact of financial assistance programs on health care utilization: evidence from Kaiser Permanente. *American Economic Review: Insights 4*(3), 389–407.

Agan, A. and S. Starr (2017, 08). Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment. *The Quarterly Journal of Economics 133*(1), 191–235.

Agan, A. Y., A. Garin, D. K. Koustas, A. Mas, and C. Yang (forthcoming, 2026). Can you erase the mark of a criminal record? Labor market impacts of criminal record remediation. *American Economic Journal: Economic Policy*.

Agarwal, S., S. Alok, P. Ghosh, and S. Gupta (2023). Financial inclusion and alternate credit scoring: Role of big data and machine learning in fintech. Working paper.

Agarwal, S., G. Amromin, I. Ben-David, S. Chomsisengphet, T. Piskorski, and A. Seru (2017). Policy intervention in debt renegotiation: Evidence from the home affordable modification program. *Journal of Political Economy 125*(3), 654–712.

Barreca, A. I., J. M. Lindo, and G. R. Waddell (2015). Heaping-induced bias in regression-discontinuity designs. *Economic Inquiry 53*(2), 1202–1223.

Bartik, A. W. and S. T. Nelson (2025). Deleting a signal: Evidence from pre-employment credit checks. *Review of Economics and Statistics 107*(1), 152–171.

Batty, M., C. Gibbs, and B. Ippolito (2022). Health insurance, medical debt, and financial well-being. *Health Economics 31*(5), 689–728.

Bergstra, J., R. Bardenet, Y. Bengio, and B. Kégl (2011). Algorithms for hyper-parameter optimization. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. C. N. Pereira, and K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011. Proceedings of a meeting held 12–14 December 2011, Granada, Spain*, pp. 2546–2554.

Blattner, L. and S. Nelson (2022). How costly is noise? Data and disparities in consumer credit. Working paper.

Blattner, L., S. T. Nelson, and J. Spiess (2024). Unpacking the black box: Regulating algorithmic decisions. Working paper.

Blavin, F., B. Braga, and M. Karpman (2023). Medical debt was erased from credit records for most consumers, potentially improving many Americans' lives. Urban Wire Blog, Urban Institute.

Bos, M., E. Breza, and A. Liberman (2018, 01). The Labor Market Effects of Credit Market Information. *The Review of Financial Studies 31*(6), 2005–2037.

Braunstein, S. F. (2010, May). Credit scoring. Testimony before the Committee on Financial Services, Subcommittee on Financial Institutions and Consumer Credit, U.S. House of Representatives. Accessed May 2025.

Brevoort, K. and M. Kambara (2014). Data point: Medical debt and credit scores. Technical report, Consumer Financial Protection Bureau.

Brevoort, K. P. and M. Kambara (2015). Are all collections equal? the case of medical debt. *The Journal of Credit Risk 11*(4).

Business Wire (2022, March). Equifax, Experian, and TransUnion support U.S. consumers with changes to medical collection debt reporting. https://www.businesswire.com/news/home/20220318005244/en/Equifax-Experian-and-TransUnion-Support-U.S.-Consumers-With-Changes-to-Medical-Collection-Debt-Reporting. Accessed: 2025-05-30.

Butaru, F., Q. Chen, B. Clark, S. Das, A. W. Lo, and A. Siddique (2016). Risk and risk management in the credit card industry. *Journal of Banking & Finance 72*, 218–239.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica 82*(6), 2295–2326.

Cespedes, J., C. Parra, and C. Sialm (2025). The effect of principal reduction on household distress: Evidence from mortgage cramdown. *Review of Financial Studies*. Forthcoming.

CFPB (2022, March). Medical debt burden in the united states. Research report, Consumer Financial Protection Bureau. Updated April 2022.

Chen, T. and C. Guestrin (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794. ACM.

Cheng, I.-H., F. Severino, and R. R. Townsend (2020, 07). How Do Consumers Fare When Dealing with Debt Collectors? Evidence from Out-of-Court Settlements. *The Review of Financial Studies 34*(4), 1617–1660.

Cherry, S., E. Jiang, G. Matvos, T. Piskorski, and A. Seru (2021). Government and private household debt relief during covid-19. *Brookings Papers on Economic Activity Fall*, 141–199.

Chioda, L., P. Gertler, S. Higgins, and P. C. Medina (2024, November). Fintech lending to borrowers with no credit history. Working Paper 33208, National Bureau of Economic Research.

Correia, F., P. Han, and J. Wang (2023). The online payday loan premium. Working paper.

Davis, J. and M. Goadrich (2006). The relationship between precision-recall and roc curves. In *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, New York, NY, USA, pp. 233–240. Association for Computing Machinery.

Di Maggio, M., A. Kalda, and V. Yao (2020). Second chance: Life without student debt. Working paper.

Dinerstein, M., S. Earnest, D. Koustas, and C. Yannelis (2025). Student loan forgiveness. Working Paper BFI WP 2025-23.

Dobbie, W., P. Goldsmith-Pinkham, N. Mahoney, and J. Song (2020). Bad credit, no problem? Credit and labor market consequences of bad credit reports. *The Journal of Finance 75*(5), 2377–2419.

Dobbie, W. and J. Song (2015, March). Debt relief and debtor outcomes: Measuring the effects of consumer bankruptcy protection. *American Economic Review 105*(3), 1272–1311.

Dobbie, W. and J. Song (2020, April). Targeted debt relief and the origins of financial distress: Experimental evidence from distressed credit card borrowers. *American Economic Review 110*(4), 984–1018.

Dobkin, C., A. Finkelstein, R. Kluender, and M. J. Notowidigdo (2018). Myth and measurement: the case of medical bankruptcies. *The New England Journal of Medicine 378*(12), 1076.

Fedaseyeu, V. (2020). Debt collection agencies and the supply of consumer credit. *Journal of Financial Economics 138*(1), 193–221.

Fedaseyeu, V. and R. Hunt (2018). The economics of debt collection: Enforcement of consumer credit contracts. Working Paper.

Federal Reserve Board (2007). Report to the congress on credit scoring and its effects on the availability and affordability of credit. Technical report, Board of Governors of the Federal Reserve System.

Federal Reserve Board (2013, August). Capital planning at large bank holding companies: Supervisory expectations and range of current practice. Supervisory guidance, Board of Governors of the Federal Reserve System, Washington, D.C.

FICO (2023). Impact of medical collection data removal by the cras on fico scores. https://www.fico.com/en/latest-thinking/executive-brief/impact-medical-collection-data-removal-cras-fico-scores. Executive brief on the effects of medical collection data removal on FICO scores.

FICO (2025). What's not included in your fico® score. https://www.myfico.com/credit-education/credit-scores/whats-not-in-your-credit-score. Accessed: 2025-12-11.

Fonseca, J. (2023). Less Mainstream Credit, More Payday Borrowing? Evidence from Debt Collection Restrictions. *The Journal of Finance 78*(1), 63–103.

Fonseca, J. and L. Liu (2024). Mortgage lock-in, mobility, and labor reallocation. *The Journal of Finance 79*(6), 3729–3772.

Fonseca, J., L. Liu, and P. Mabille (2024). Unlocking mortgage lock-in: Evidence from a spatial housing ladder model. *Working paper*.

Fonseca, J. and J. Wang (2023). How much do small businesses rely on personal credit? Working paper.

Frost, J., L. Gambacorta, Y. Huang, H. S. Shin, and P. Zbinden (2020, 01). Bigtech and the changing structure of financial intermediation. *Economic Policy 34*(100), 761–799.

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther (2022). Predictably unequal? the effects of machine learning on credit markets. *The Journal of Finance 77*(1), 5–47.

Ganong, P. and P. Noel (2020). Liquidity versus wealth in household debt obligations: Evidence from housing policy in the great recession. *American Economic Review 110*(10), 3100–3138.

Gibbs, C., B. Guttman-Kenney, D. Lee, S. Nelson, W. Van der Klaauw, and J. Wang (2025). Consumer credit reporting data. *Journal of Economic Literature 63*(2), 598–636.

Goldsmith-Pinkham, P., M. Pinkovskiy, and J. Wallace (2023). The great equalizer: Medicare and the geography of consumer financial strain. Technical report, National Bureau of Economic Research.

Grembi, V., T. Nannicini, and U. Troiano (2016). Do fiscal rules matter? *American Economic Journal: Applied Economics 8*(3), 1–30.

Gross, T., M. J. Notowidigdo, and J. Wang (2020, April). The marginal propensity to consume over the business cycle. *American Economic Journal: Macroeconomics 12*(2), 351–84.

Guttman-Kenney, B., R. Kluender, N. Mahoney, F. Wong, X. Xia, and W. Yin (2022, 05). Trends in medical debt during the covid-19 pandemic. *JAMA Health Forum 3*(5), e221031–e221031.

Herkenhoff, K., G. Phillips, and E. Cohen-Cole (2021). The impact of consumer credit access on self-employment and entrepreneurship. *Journal of Financial Economics 141*(1), 345–371.

Howard, G. and H. Shao (2022). Internal migration and the microfoundations of gravity. Working paper.

Huh, J. and J. Reif (2021). Teenage driving, mortality, and risky behaviors. *American Economic Review: Insights 3*(4), 523–539.

Jansen, M., F. Nagel, C. Yannelis, and A. Lee Zhang (2024). Data and welfare in credit markets. Working paper.

Keys, B. J., N. Mahoney, and H. Yang (2022, 05). What determines consumer financial distress? place- and person-based factors. *The Review of Financial Studies 36*(1), 42–69.

Khandani, A. E., A. J. Kim, and A. W. Lo (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance 34*(11), 2767–2787.

Kluender, R., N. Mahoney, F. Wong, and W. Yin (2021). Medical debt in the US, 2009-2020. *JAMA 326*(3), 250–256.

Kluender, R., N. Mahoney, F. Wong, and W. Yin (2024, 12). The effects of medical debt relief: Evidence from two randomized experiments. *The Quarterly Journal of Economics 140*(2), qjae045.

Lee, D. S. (2008). Randomized experiments from non-random selection in us house elections. *Journal of Econometrics 142*(2), 675–697.

Liberman, A., C. Neilson, L. Opazo, and S. Zimmerman (2019). The equilibrium effects of asymmetric information: Evidence from consumer credit markets. Working paper.

Lin, J. (2024). Creditor rights, household consumption, and entrepreneurial activity. Working Paper.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

Meursault, V., D. Moulton, L. Santucci, and N. Schor (2022). The time is now: Advancing fairness in lending through machine learning. Working paper.

Miller, S. (2015). Information and default in consumer credit markets: Evidence from a natural experiment. *Journal of Financial Intermediation 24*(1), 45–70.

Miller, S. and C. Soo (2020). Does increasing access to formal credit reduce payday borrowing? Working paper.

Musto, D. K. (2004). What happens when information leaves a market? Evidence from postbankruptcy consumers. *The Journal of Business 77*(4), 725–748.

Quinn, T. (2023). Medical collections less than $500 no longer impact fico scores. https://www.myfico.com/credit-education/blog/medical-collections-removal. Accessed: 2025-05-30.

Robertson, C. T., M. Rukavina, and E. C. F. Brown (2022). New state consumer protections against medical debt. *JAMA 327*(2), 121–122.

Sadhwani, A., K. Giesecke, and J. Sirignano (2020, 07). Deep learning for mortgage risk*. *Journal of Financial Econometrics 19*(2), 313–368.

Sandler, R. and L. Nathe (2022). Paid and low-balance medical collections on consumer credit reports. Technical Report 22-5, Consumer Financial Protection Bureau, Office of Research Reports Series.

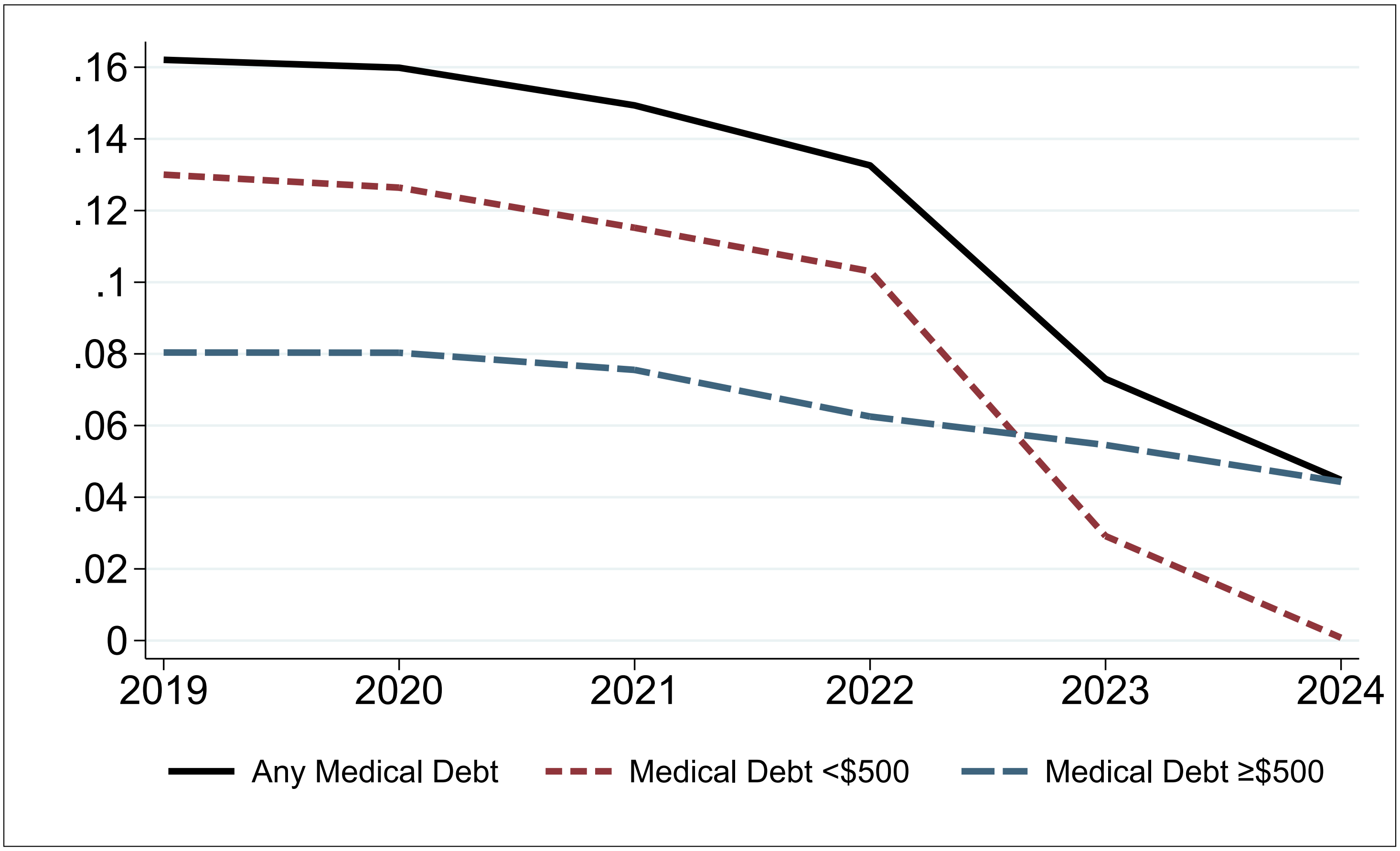
The White House (2025, Jan). Fact sheet: Vice president harris announces final rule removing medical debt from all credit reports. https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2025/01/07/fact-sheet-vice-president-harris-announces-final-rule-removing-medical-debt-from-all-

U.S. Census Bureau (2022). Debt for households, by type of debt and selected characteristics: 2022. Accessed: 2025-01-19.

VantageScore (2022, aug). VantageScore Removes Medical Debt Collection Records From Latest Scoring Models. Published: August 10, 2022; Accessed: 24 November 2025.

Warnock, R., E. Warren, C. E. Schumer, J. A. Merkley, M. Heinrich, S. Whitehouse, K. Gillibrand, R. Gallego, J. Shaheen, A. B. Schiff, J. Fetterman, R. Blumenthal, A. Klobuchar, A. S. J. King, C. Van Hollen, R. Wyden, P. Welch, B. Sanders, C. A. Booker, E. J. Markey, M. K. Hirono, T. Smith, B. R. Luján, A. Alsobrooks, J. Rosen, A. Kim, L. Blunt Rochester, J. Hickenlooper, T. Duckworth, and J. Reed (2025, July 14). Letter to Acting Director Russell Vought regarding CFPB's medical debt rule. https://www.warnock.senate.gov/wp-content/uploads/2025/07/Letter-to-CFPB-Regarding-the-Medical-Debt-Rule.pdf. United States Senate letter to the Consumer Financial Protection Bureau regarding the agency's decision to vacate the medical debt removal rule.

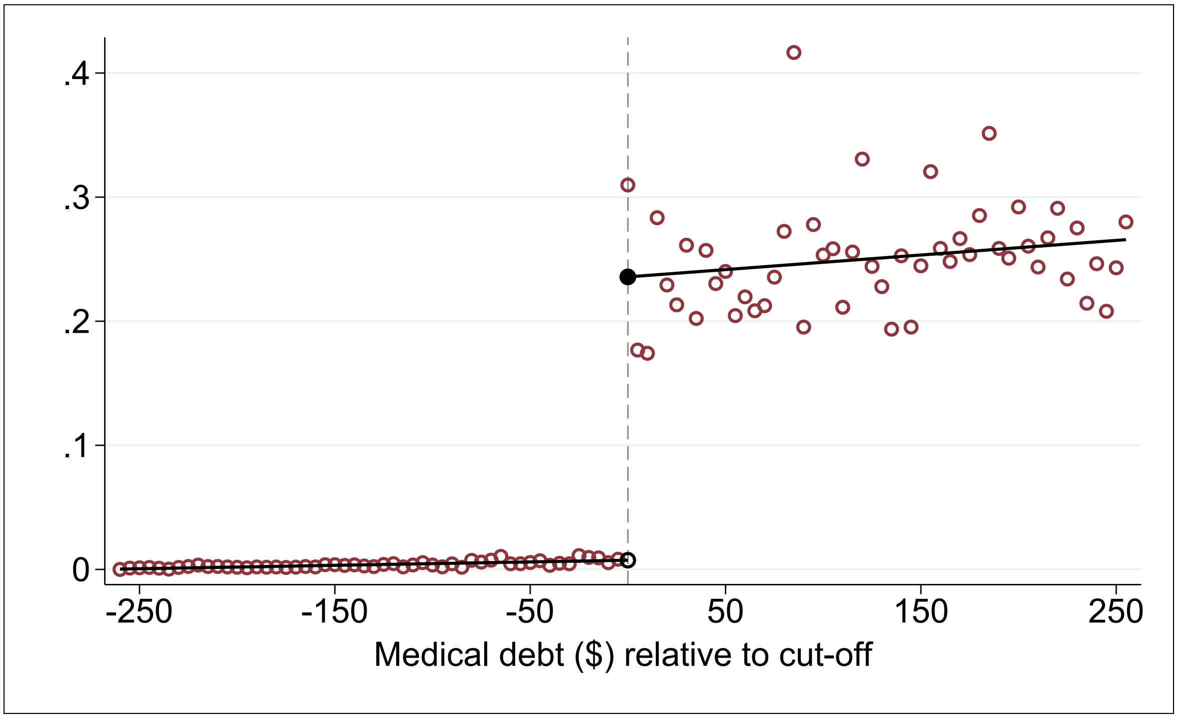
Figure 1: Share of Consumers with Medical Debt Collections, 2019–2024



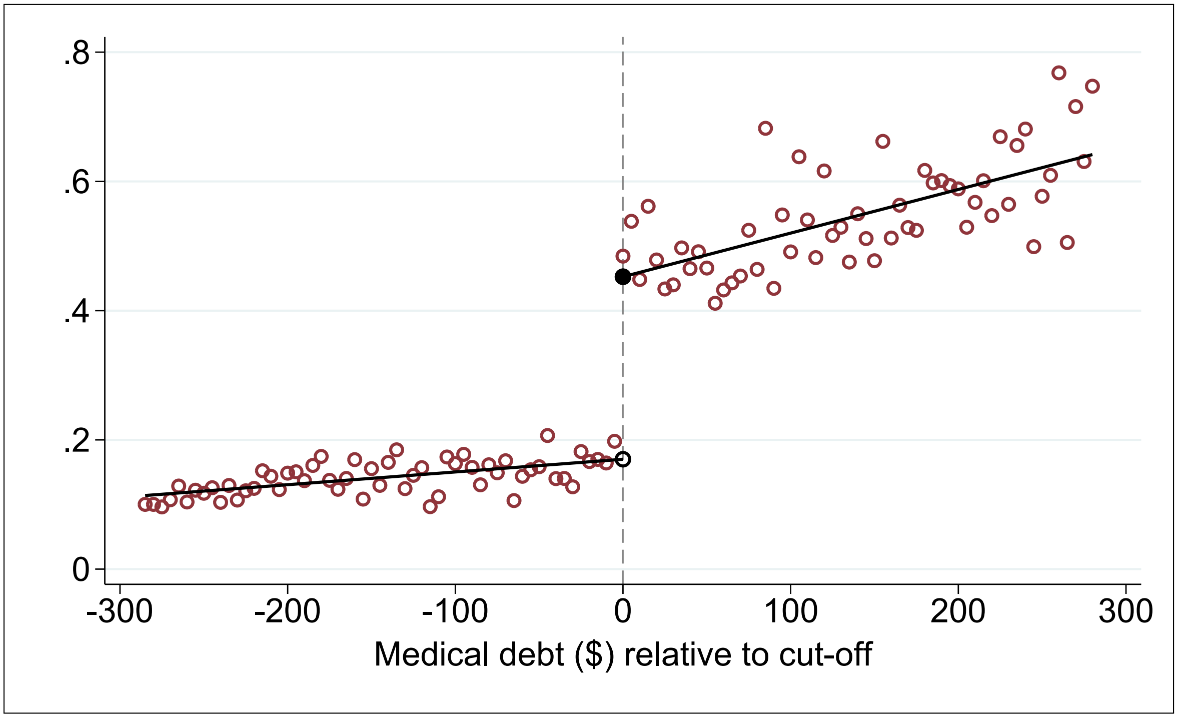
Notes: This figure shows the share of consumers with medical debt collections appearing on credit reports. The black solid line plots the share with any medical debt in collections. The red dashed line plots the share with debt below $500, while the blue long-dashed line plots the share with debt above $500.

Figure 2: Two-Year Evolution of 2022 Medical Debt Collection Accounts

(a) Share of Surviving Accounts



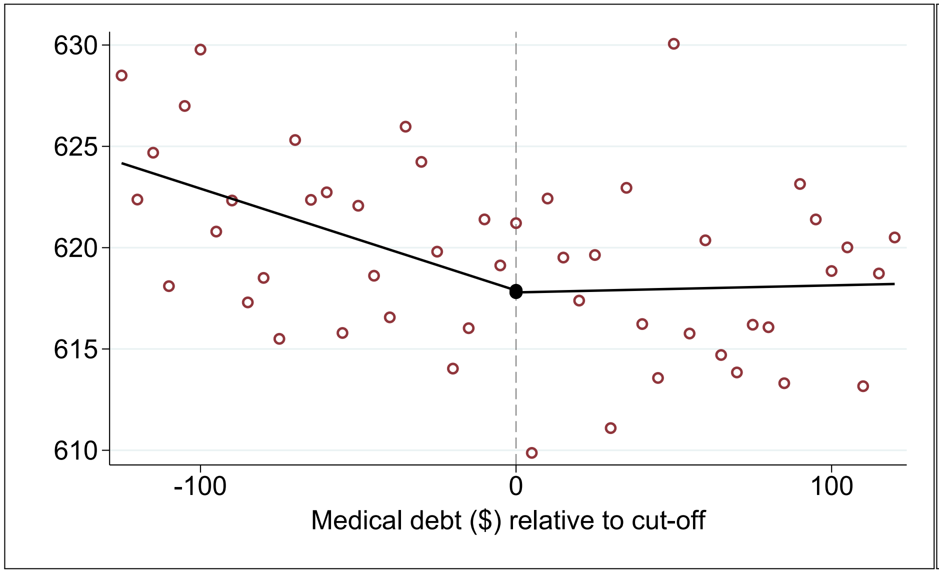
(b) Average Number of Accounts per Person



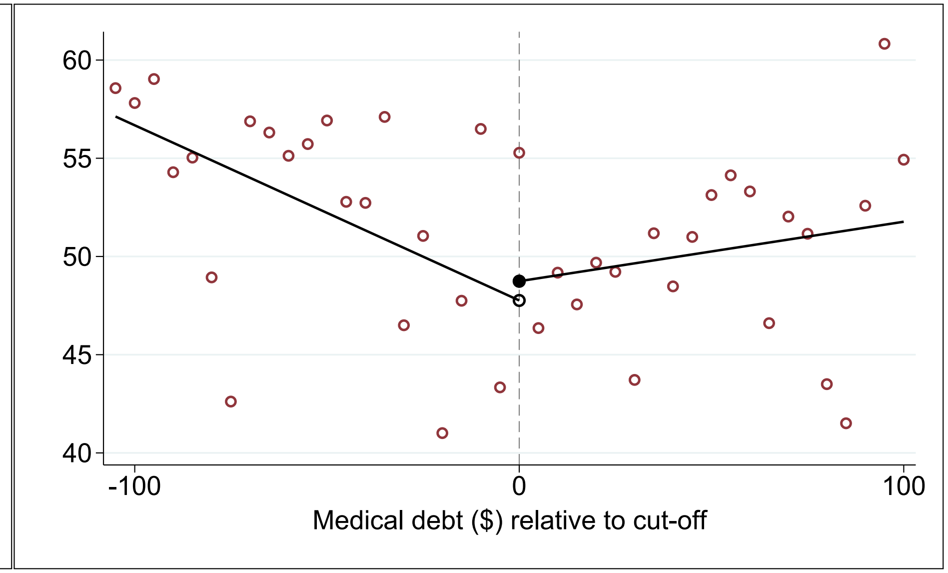
Notes: Panel (a) shows the proportion of 2022 medical debt collection accounts that remain on credit reports in 2024 by account amount, where the amount is measured as distance from the $500 threshold. Panel (b) shows the average number of medical debt collections accounts per person in 2024, where the running variable is the maximum value of the consumer's 2022 medical collections accounts. The fitted lines are estimated using Equation (1) for Panel (a) and Equation (2) for Panel (b). The RD estimate for Panel (a) is reported in Column (11) of Table A.3, and the estimate for Panel (b) appears in Column (2) of Table 3.

Figure 3: Access to Credit, 2024

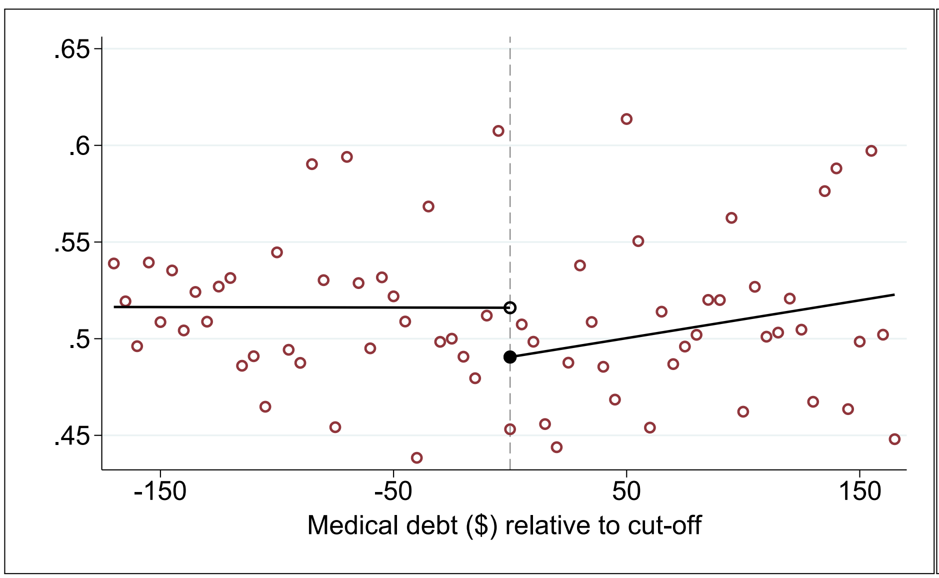
(a) Credit Score



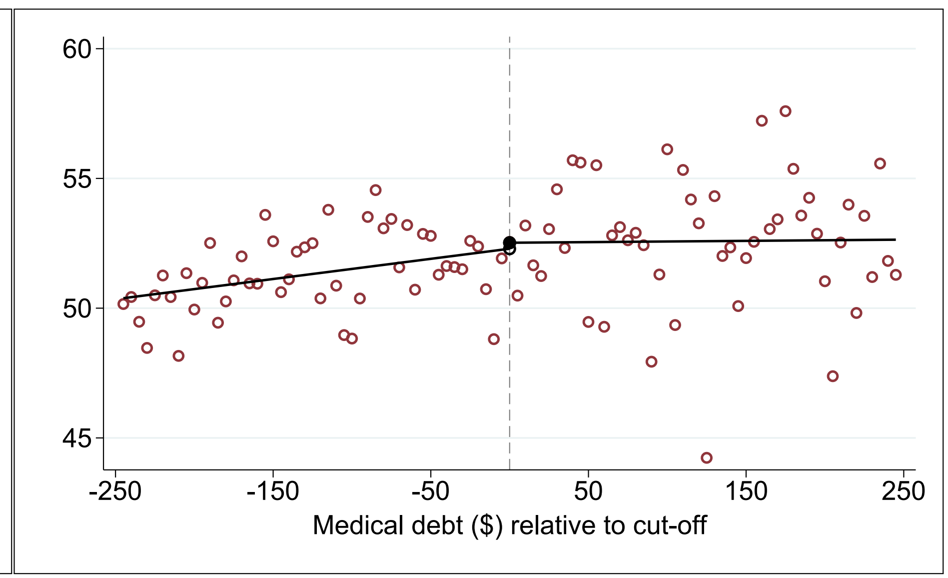
(b) Total Credit Balance ($1,000)



(c) Number of Accounts Opened in Last 6 Months



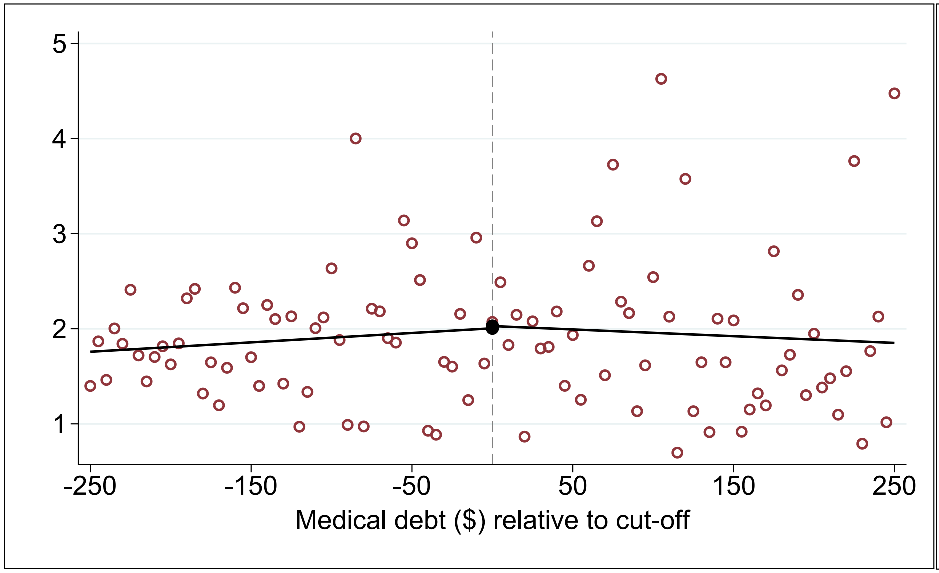
(d) Revolving Utilization (%)



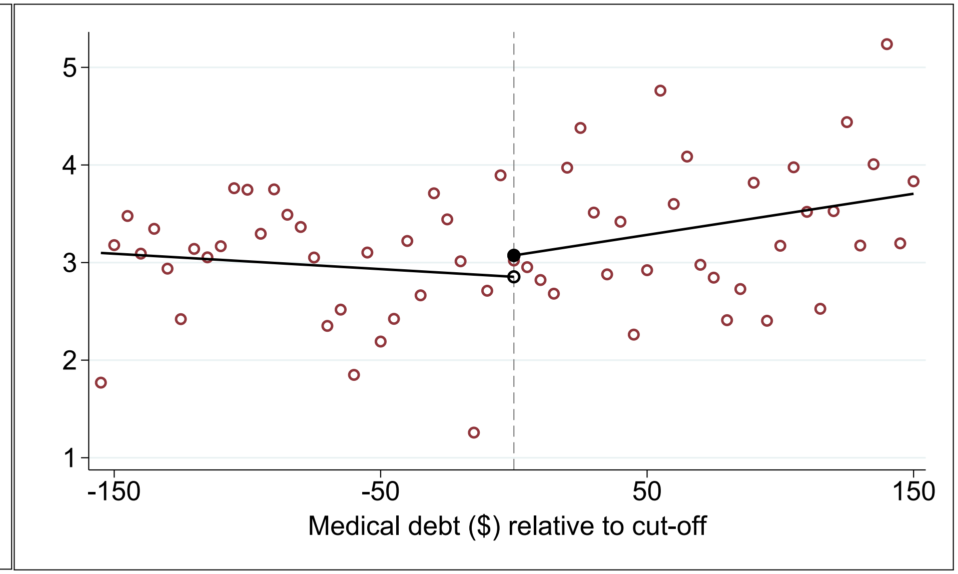
Notes: This figure shows the relationship between medical debt in 2022 and four different measures of credit access in 2024. Medical debt is defined as the maximum value of the consumer's 2022 medical debt collections accounts, measured relative to the $500 threshold. The corresponding RD estimates from Equation (2) are reported in Table 3.

Figure 4: Financial Distress and Access to Alternative Credit, 2024

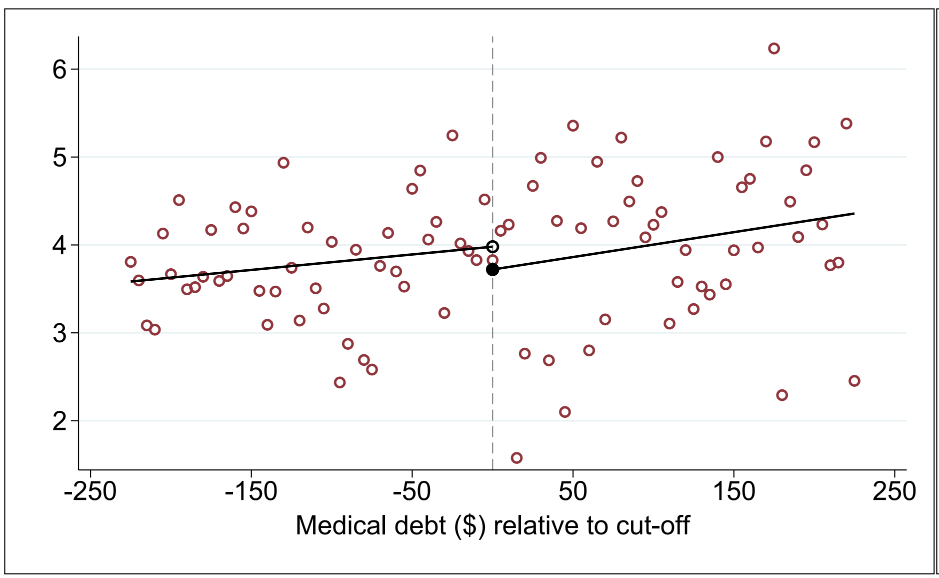
(a) Total Balance 90+ Days Past Due ($1,000)



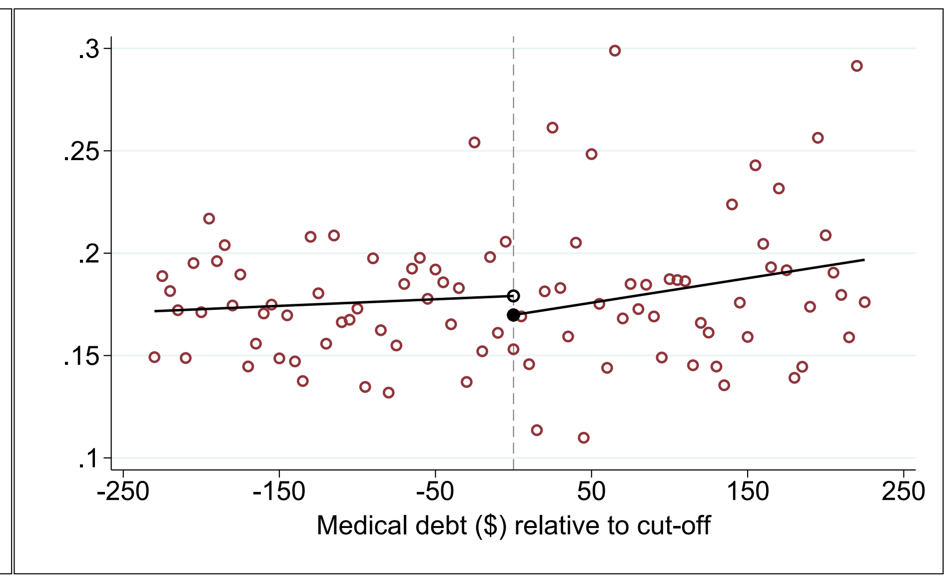
(b) Bankruptcy in Last 7 Years (%)



(c) Has Alternative Credit Balance (%)

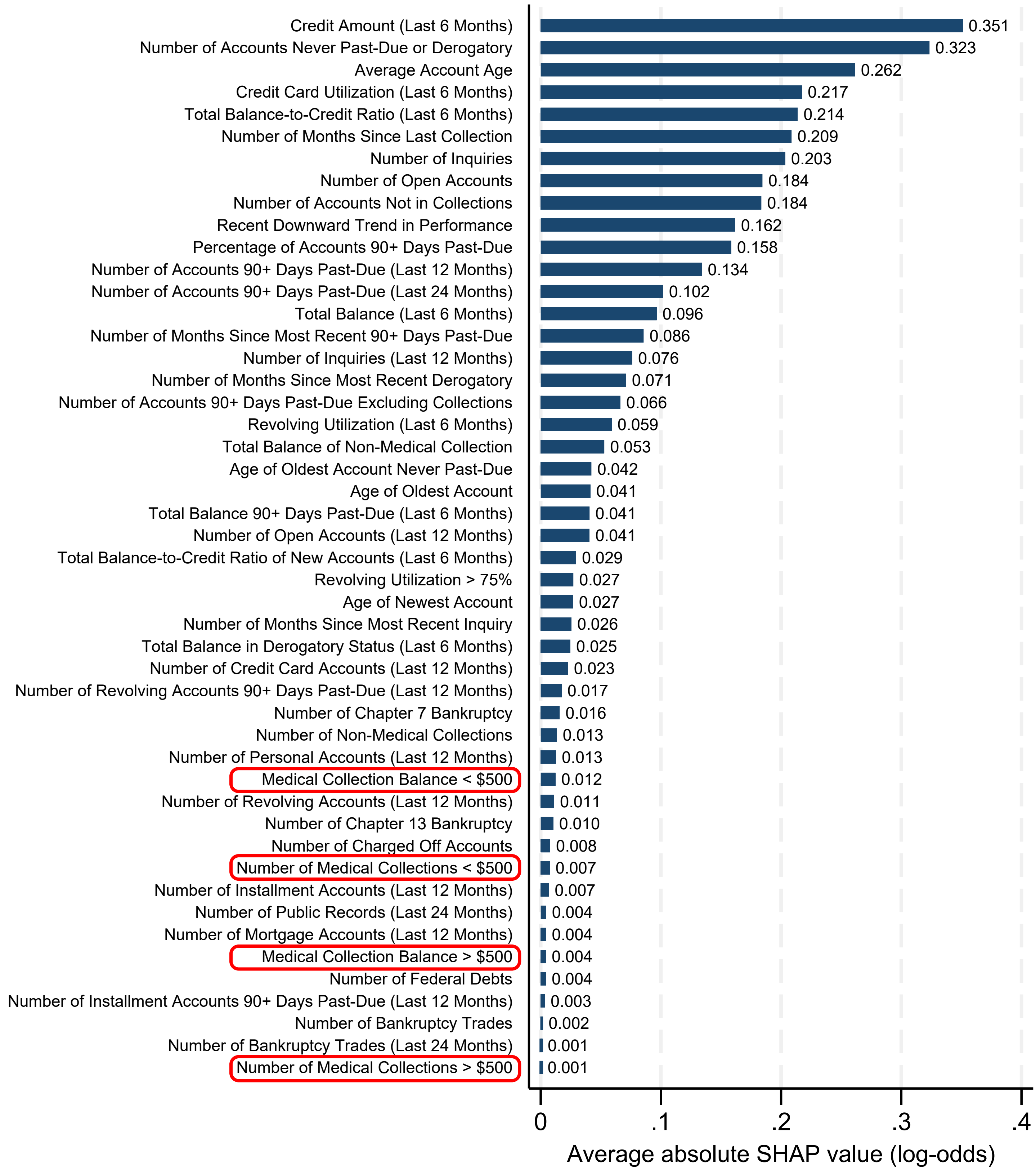


(d) Number of Alternative Credit Accounts



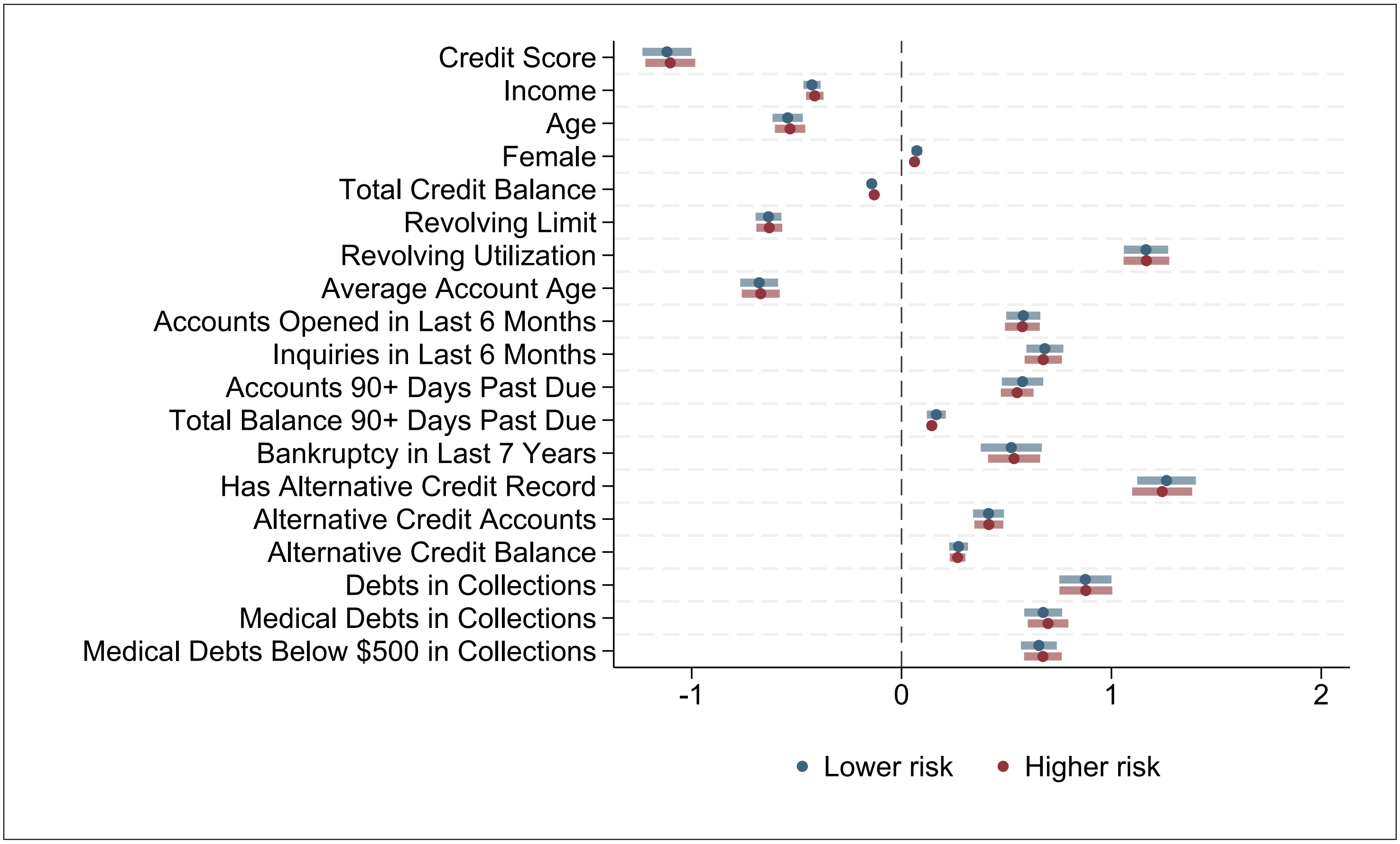
Notes: This figure shows the relationship between medical debt in 2022 and measures of financial distress and access to alternative credit in 2024. Medical debt is defined as the maximum value of the consumer's 2022 medical debt collections accounts, measured relative to the $500 threshold. The corresponding RD estimates from Equation (2) are reported in Table 3.

Figure 5: Importance of Credit-Report Variables for Predicting Default



Notes: This figure reports the relative importance of credit-report variables in the credit scoring framework described in Section 5, measured by average absolute SHAP values (in log-odds units). The XGBoost specification is estimated using 2019 data to predict defaults in 2020–2021. Variables are ordered by their average contribution to predicted default risk.

Figure 6: Covariate Differences Across Reclassified Consumers



Notes: This figure compares observable characteristics across consumers whose predicted default probabilities are most affected by the removal of small medical collections from the credit scoring model in Section 5, relative to unaffected consumers. Higher-risk consumers are those in the top 5%, whose predicted probability of default increases by approximately 2 percentage points or more. Lower-risk consumers are those in the bottom 5%, whose predicted default probability decreases by at least 2 percentage points. Unaffected consumers are those between the 25th and 75th percentiles, whose predicted probabilities change by less than about 0.2 percentage points. Each dot represents the difference in a standardized characteristic between the indicated group and the unaffected group. Blue dots correspond to lower-risk consumers and red dots correspond to higher-risk consumers. Standard errors are clustered at the percentile-bin level after dividing the full sample into 100 equal-sized bins based on the change in predicted default probability. Horizontal bars represent 95% confidence intervals. If small medical collections contain meaningful predictive information, higher- and lower-risk consumers should differ systematically across these characteristics. Instead, the two groups appear similar across all dimensions.

Table 1: Summary Statistics, 2019–2024

	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample			Medical Debt Subsample		
	Mean	St. Dev.	Median	Mean	St. Dev.	Median
A. Demographics						
Income ($1,000)	51.97	32.81	41.00	38.69	19.51	34.00
Age (years)	50.56	19.41	49.00	44.90	15.16	43.00
Female (%)	50.06	50.00	100.00	54.10	49.83	100.00
B. Access to Credit						
Credit Score	702.28	100.86	715.00	611.20	86.69	601.00
Total Credit Balance ($1,000)	76.65	140.96	10.15	41.94	87.15	6.74
Revolving Limit ($1,000)	21.50	33.85	6.43	4.92	14.50	0.00
Revolving Utilization (%)	28.09	32.84	13.00	49.86	39.50	47.00
Average Account Age (months)	105.16	77.03	94.00	73.39	51.59	66.00
Number of Accounts Opened in Last 6 Months	0.41	0.83	0.00	0.47	0.95	0.00
Number of Inquiries in Last 6 Months	0.41	0.74	0.00	0.62	0.90	0.00
Number of New Mortgages in Last 6 Months	0.02	0.16	0.00	0.01	0.11	0.00
C. Access to Alternative Credit						
Has Alternative Credit Record (%)	19.99	39.99	0.00	47.94	49.96	0.00
Has Alternative Credit Balance (%)	1.02	10.05	0.00	2.60	15.92	0.00
Number of Alternative Credit Accounts	0.05	0.43	0.00	0.13	0.68	0.00
Alternative Credit Balance ($1,000)	0.05	0.79	0.00	0.13	1.19	0.00
D. Financial Distress						
Number of Accounts 90+ Days Past Due	0.19	0.90	0.00	0.46	1.33	0.00
Total Balance 90+ Days Past Due ($1,000)	0.81	12.17	0.00	1.96	15.91	0.00
Bankruptcy in Last 7 Years (%)	2.84	16.60	0.00	4.79	21.35	0.00
E. Debt in Collections						
Total Debts ($1,000)	0.56	3.65	0.00	3.14	9.89	1.39
Total Medical Debts ($1,000)	0.16	2.46	0.00	1.55	7.50	0.76
Total Medical Debts Below $500 ($1,000)	0.03	0.13	0.00	0.26	0.32	0.14
Number of Debts	0.60	1.87	0.00	3.81	4.11	3.00
Number of Medical Debts	0.25	1.02	0.00	2.44	2.17	2.00
Number of Medical Debts Below $500	0.15	0.66	0.00	1.45	1.53	1.00
Observations	15,366,574			1,591,560		

Notes: This table presents summary statistics from the 2019–2024 Gies Consumer and Small Business Credit Panel. The first three columns show statistics for the full sample, while the last three focus on consumers with at least one medical collection during the reported year.

Table 2: Summary Statistics for Consumers with Medical Debt Collections, 2022 (RD Sample)

	(1)	(2)	(3)
	Mean	St. Dev.	Median
A. Demographics			
Income ($1,000)	40.69	20.53	35.00
Age (years)	45.11	15.19	43.00
Female (%)	55.30	49.72	100.00
B. Access to Credit			
Credit Score	625.37	87.94	618.00
Total Credit Balance ($1,000)	48.91	94.64	10.66
Revolving Limit ($1,000)	6.03	15.83	0.23
Revolving Utilization (%)	47.93	39.18	44.00
Average Account Age (months)	73.71	49.09	66.00
Number of Accounts Opened in Last 6 Months	0.60	1.09	0.00
Number of Inquiries in Last 6 Months	0.67	0.93	0.00
Number of New Mortgages in Last 6 Months	0.02	0.13	0.00
C. Access to Alternative Credit			
Has Alternative Credit Record (%)	51.74	49.97	100.00
Has Alternative Credit Balance (%)	3.20	17.59	0.00
Number of Alternative Credit Accounts	0.14	0.66	0.00
Alternative Credit Balance ($1,000)	0.18	1.45	0.00
D. Financial Distress			
Number of Accounts 90+ Days Past Due	0.31	0.91	0.00
Total Balance 90+ Days Past Due ($1,000)	1.31	13.16	0.00
Bankruptcy in Last 7 Years (%)	4.11	19.86	0.00
E. Debt in Collections			
Total Debts ($1,000)	2.69	6.50	1.09
Total Medical Debts ($1,000)	1.22	1.55	0.57
Total Medical Debts Below $500 ($1,000)	0.27	0.30	0.17
Number of Debts	3.53	3.77	2.00
Number of Medical Debts	2.37	2.00	1.00
Number of Medical Debts Below $500	1.56	1.43	1.00
Observations	272,490		

Notes: This table presents summary statistics from the Gies Consumer and Small Business Credit Panel. The statistics are based on data from 2022, the year preceding the removal of information on medical collections below $500. The unit of observation is the consumer. The sample is limited to consumers with a non-missing credit score from 2022–2024 who had at least one medical collection account in 2022.

Table 3: RD Estimates of the Effect of Medical Debt Collections Deletion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of Debts	Number of Medical Debts	Credit Score	Total Credit Balance ($1,000)	Number of Accounts Opened in Last 6 Months	Revolving Utilization (%)	Total Balance 90+ Days Past Due ($1,000)	Bankruptcy in Last 7 Years (%)	Has Alternative Credit Balance (%)	Number of Alternative Credit Accounts
ABOVE^{2022}	–0.22**	–0.30**	0.18	–2.3	0.038	–0.47	–0.067	–0.37	0.36	0.011
	[–0.32, –0.14]	[–0.32, –0.27]	[–4.2, 5.5]	[–8.7, 2.5]	[–0.0019, 0.088]	[–2.7, 1.4]	[–0.76, 0.54]	[–1.2, 0.54]	[–0.31, 1.3]	[–0.015, 0.045]
Control Mean	1.5	0.49	618	50	0.50	53	2.0	3.2	3.9	0.18
% of Mean	–15	–61	0.029	–4.7	7.5	–0.89	–3.4	–11	9.3	6.2
Optimal Bandwidth	± 146.08	± 282.39	± 122.15	± 103.82	± 167.87	± 246.00	± 250.43	± 151.71	± 225.05	± 228.93
Observations										
Total	272,490	272,490	272,490	272,490	272,490	169,547	272,490	272,490	272,490	272,490
In-Bandwidth	38,500	84,940	31,760	26,730	44,630	43,334	73,163	40,146	63,354	64,348

Notes: This table presents the estimated coefficient ($-\beta$) and 95% confidence intervals from Equation (2) using outcomes measured in 2024. Outcome variables are defined in Table A.4. The running variable is defined as the maximum 2022 medical debt collection amount across the consumer's accounts, measured relative to the $500 reporting cutoff. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Control Mean reports the mean of the dependent variable for consumers whose maximum medical debt in collections in 2022 falls between $500 and $600. The reported optimal bandwidth corresponds to the symmetric half-width (in dollars) around the $500 cutoff used to estimate each specification. "In-Bandwidth" observations report the number of observations falling within this window, while "Total" observations report the full estimation sample. A */** denotes statistical significance at the 5% and 1% levels respectively, using robust inference.

Table 4: Performance Metrics for Credit Scoring Models With and Without Medical Debt Collections

	(1)	(2)	(3)
	All Predictors	Exclude Medical Debts < \$500	Exclude All Medical Debts
Accuracy	0.907	0.907	0.907
AUC	0.905	0.905	0.905
F1 Score	0.561	0.561	0.562
Precision	0.735	0.737	0.735
Recall (1 − False Negative Rate)	0.454	0.453	0.454
False Positive Rate	0.0247	0.0243	0.0247

Notes: This table reports out-of-sample performance metrics for a credit scoring model predicting defaults occurring between 2020 and 2021, using borrower characteristics from 2019. Column (1) presents metrics for the baseline model, which includes 48 predictors and is estimated using XGBoost. Column (2) reports metrics when small (under \$500) medical collections are excluded from the predictors. Column (3) shows metrics when all medical collections are excluded. The sample consists of 2,482,116 observations, with 90% used for model training and the remaining 10% reserved for out-of-sample performance evaluation. For reference, a naive model that predicts no defaults achieves an accuracy of 0.869, equal to one minus the average default rate (0.131). Definitions of the performance metrics are provided in Appendix C.4.

Table 5: Summary Statistics by Risk-Reclassification Groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Unaffected			Lower Risk			Higher Risk		
	Mean	St. Dev.	Median	Mean	St. Dev.	Median	Mean	St. Dev.	Median
A. Demographics									
Income ($1,000)	60.28	36.20	50.00	44.23	22.80	38.00	44.70	23.17	38.00
Age (years)	57.19	19.41	58.00	45.56	14.95	44.00	45.77	14.97	44.00
Female (%)	50.08	50.00	100.00	54.10	49.83	100.00	53.60	49.87	100.00
B. Access to Credit									
Credit Score	753.37	83.63	789.00	610.93	80.57	609.00	612.57	80.47	611.00
Total Credit Balance ($1,000)	88.53	155.51	7.79	65.41	114.21	18.08	67.19	116.26	19.14
Revolving Limit ($1,000)	32.61	39.38	20.00	6.16	15.62	0.46	6.31	15.88	0.50
Revolving Utilization (%)	15.51	22.47	6.00	54.85	38.57	56.00	54.91	38.47	56.00
Average Account Age (months)	137.11	83.36	123.00	76.88	45.13	69.00	77.56	46.34	69.00
Number of Accounts Opened in Last 6 Months	0.28	0.63	0.00	0.70	1.20	0.00	0.71	1.20	0.00
Number of Inquiries in Last 6 Months	0.26	0.56	0.00	0.77	0.99	0.00	0.77	0.99	0.00
Number of New Mortgages in Last 6 Months	0.03	0.17	0.00	0.02	0.13	0.00	0.02	0.13	0.00
C. Access to Alternative Credit									
Has Alternative Credit Record (%)	6.38	24.43	0.00	53.84	49.85	100.00	53.22	49.90	100.00
Has Alternative Credit Balance (%)	0.19	4.40	0.00	3.85	19.23	0.00	3.88	19.31	0.00
Number of Alternative Credit Accounts	0.01	0.18	0.00	0.19	0.86	0.00	0.19	0.86	0.00
Alternative Credit Balance ($1,000)	0.01	0.33	0.00	0.21	1.56	0.00	0.22	1.60	0.00
D. Financial Distress									
Number of Accounts 90+ Days Past Due	0.04	0.41	0.00	0.65	1.70	0.00	0.64	1.69	0.00
Total Balance 90+ Days Past Due ($1,000)	0.20	5.75	0.00	2.84	24.34	0.00	2.69	21.52	0.00
Bankruptcy in Last 7 Years (%)	0.88	9.33	0.00	9.35	29.11	0.00	9.52	29.35	0.00
E. Debt in Collections									
Total Debts ($1,000)	0.15	1.53	0.00	1.73	5.01	0.14	1.74	5.05	0.08
Total Medical Debts ($1,000)	0.05	0.64	0.00	0.44	1.90	0.00	0.45	1.90	0.00
Total Medical Debts Below $500 ($1,000)	0.01	0.07	0.00	0.08	0.21	0.00	0.08	0.22	0.00
Number of Debts	0.17	0.95	0.00	1.81	3.07	1.00	1.81	3.19	1.00
Number of Medical Debts	0.07	0.53	0.00	0.72	1.65	0.00	0.73	1.73	0.00
Number of Medical Debts Below $500	0.04	0.35	0.00	0.43	1.09	0.00	0.44	1.13	0.00
Observations	6,971,607			691,832			691,691		

Notes: This table reports descriptive statistics from 2019 to 2024 for unaffected consumers (columns 1–3), consumers reclassified as lower risk (columns 4–6), and consumers reclassified as higher risk (columns 7–9). Groups are defined based on changes in predicted default probability when medical collections below $500 are removed from the baseline credit scoring model described in Section 5. Higher-risk consumers are those above the 95th percentile of probability changes, whose predicted probability of default increases by approximately 2 percentage points or more. Lower-risk consumers are those below the 5th percentile, whose predicted probability of default decreases by approximately 2 percentage points or more. Unaffected consumers are those between the 25th and 75th percentiles of the distribution. All variables are drawn from the Gies Consumer and Small Business Credit Panel. If medical debt collections primarily introduce noise rather than predictive signal, these groups should exhibit similar observable characteristics.

Online Appendix

"The Effects of Deleting Medical Debt from Consumer Credit Reports"

Victor Duarte, Julia Fonseca, Divij Kohli, Julian Reif

A GCCP Data Appendix

A.1 Comparing Medical Debt Collection Data to External Sources

To identify consumers with medical debt collections, we use credit account (tradeline) data from the GCCP. We classify a collection as medical if the creditor is categorized as Medical/Health Care or if the furnisher is identified as a business in the medical or health-related sector.[1]

To assess whether the GCCP accurately captures the prevalence and magnitude of medical debt collections, we conduct a benchmarking exercise comparing our data to estimates from external sources. Table A.1 compares the share of consumers with at least one medical debt collection in the GCCP to estimates from other sources. Column (1) reports the annual share of consumers with at least one medical debt collection in the GCCP, which declines from 16.5% in 2018 to 7.0% in 2023. This decline coincides with a series of policy changes affecting medical debt reporting, including the removal of paid medical collections, the extension of the reporting delay for medical collections from six months to one year, and the removal of medical collections below $500.

Similar trends appear in columns (2) and (3), which report estimates from the Urban Institute (Blavin et al., 2023) and the CFPB (Sandler and Nathe, 2022), respectively. The Urban Institute data show a slightly lower share of consumers with medical collections than the GCCP, while the CFPB data report a slightly higher share. These small differences likely reflect differences in measurement timing: the GCCP data are measured in March, the Urban Institute data in August, and the CFPB data in January. Overall, the GCCP closely tracks the levels and trends in these external benchmarks.

Table A.2 presents a similar benchmarking exercise for medical debt balances. Column (1) reports total medical debt collections balances by year in the GCCP, while column (2) reports balances for accounts with amounts below $500. Column (3) reports corresponding

[1] Furnisher categories include Dentists, Chiropractors, Doctors, Medical group, Hospitals and clinics, Osteopaths, Pharmacies and drugstore, Optometrists and optical outlets, and Medical and related health-nonspecific.

estimates of total medical debt balances from the CFPB (CFPB, 2022). The total balances in the GCCP closely match those reported by the CFPB in years where both are available, further supporting the representativeness of the GCCP medical debt data.

B Alternative RD Specifications

B.1 Alternative Definitions of the Running Variable

The RD specification in the main text (Equation (2)) defines the running variable as the maximum medical debt collection balance across all of a consumer's accounts. Under this person-level (max) specification, the underlying policy treatment is the deletion of all medical collection accounts, which occurs only if all of an individual's medical collection accounts fall below the $500 reporting threshold. The regression is parameterized using an indicator for the complementary event—having at least one account at or above $500—so that the coefficient β captures the intent-to-treat effect of having at least one account remain on the credit report. Equivalently, $-\beta$ can be interpreted as the effect of having all accounts deleted.

Alternative constructions of the running variable correspond to different treatment definitions and estimands. To assess the robustness of our results, we consider three such definitions:

1. **Person level (max)**: Treatment occurs only if all of an individual's accounts are deleted (baseline specification).

2. **Person level (min)**: Treatment occurs if any (i.e., at least one) account is deleted.

3. **Account level**: Treatment intensity scales with the proportion of deleted accounts.

The person-level (max) specification is estimated using Equation (2) and is reported in the main text. The person-level (min) specification is estimated analogously, replacing the maximum medical debt collection balance with the minimum balance. Specifically, we define $\text{MINDEBT}_i^{2022} = \min_j \text{DEBT}_{ij}^{2022}$ as the consumer's smallest medical collection balance relative to the $500 cutoff and estimate:

$$Y_i^{2024} = \alpha \text{MINDEBT}_i^{2022} + \beta \text{ABOVE}_i^{2022} + \gamma(\text{ABOVE}_i^{2022} \times \text{MINDEBT}_i^{2022}) + \epsilon_i \qquad (4)$$

The account-level specification allows treatment intensity to vary with the share of accounts deleted and can be estimated using Equation (1). Panel A of Figure 2 illustrates this specification using account survival as the outcome. In this appendix, we extend this account-level framework to examine additional outcomes that are naturally defined at the consumer level. Because outcomes such as credit scores do not vary across accounts, standard errors are clustered at the individual level.

Table A.3 reports estimates for all three treatment definitions. Panel A replicates the main text estimates from Table 3—the person-level (max) specification. Panel B reports estimates for the person-level (min) specification. The first-stage estimates in Columns (1) and (2) are slightly larger in magnitude, likely due to sample composition differences near the threshold. However, as in Panel A, all second-stage estimates in Columns (3)–(10) remain statistically insignificant. Panel C presents results for the account-level specification, showing a similar pattern. Together, these findings show that our results are robust to alternative definitions of the running variable and treatment.

B.2 Robustness to the Credit Score Restriction

Our main sample includes only consumers with non-missing credit scores. This restriction could bias our results if the deletion of small medical debts altered the probability that a consumer is scorable. To assess this possibility, we drop the credit-score restriction and test for discontinuities in the likelihood of having a credit score around the $500 cutoff.

Because there is a pre-existing discontinuity at the threshold in the share of scorable consumers, we estimate a pooled difference-in-discontinuities model using data from 2022 and 2024 to test whether this discontinuity changed following the April 2023 deletion. Let $D_t^{2024} = \mathbf{1}\{t = 2024\}$. Following Grembi et al. (2016), our difference-in-discontinuities specification is:

$$
\begin{aligned}
Y_i^t = {} & \alpha \text{MAXDEBT}_i^{2022} + \beta \text{ABOVE}_i^{2022} + \gamma(\text{ABOVE}_i^{2022} \times \text{MAXDEBT}_i^{2022}) \\
& + \delta D_t^{2024} + \alpha_1(\text{MAXDEBT}_i^{2022} \times D_t^{2024}) + \theta(\text{ABOVE}_i^{2022} \times D_t^{2024}) \\
& + \gamma_1(\text{ABOVE}_i^{2022} \times \text{MAXDEBT}_i^{2022} \times D_t^{2024}) + \epsilon_{it}
\end{aligned}
\tag{5}
$$

Our parameter of interest is θ, which captures the change in the discontinuity at the cutoff after the deletion and thus serves as the difference-in-discontinuities estimate of the deletion's effect on scorable status.

Table A.13 reports the results. In Column (3), the coefficient for *Has Credit Score* is small and statistically insignificant, indicating that the deletion of small medical debts did not meaningfully affect the probability of having a credit score. This result indicates that restricting the main analysis to consumers with non-missing credit scores is unlikely to have introduced bias.

For completeness, we also report difference-in-discontinuities estimates for our other main outcomes. The results closely resemble the RD estimates obtained in the main sample (Table 3), providing evidence that our null findings are not driven by pre-existing discontinuities at the cutoff.

C Gradient-Boosted Tree Model

C.1 Model

In Section 5, we estimate a supervised machine-learning model based on gradient-boosted decision trees (XGBoost). Let $y_i \in \{0, 1\}$ denote the binary outcome for observation i and let X_i denote the corresponding feature vector. XGBoost models the latent score as an additive ensemble of regression trees:

$$f(X_i) = \sum_{m=1}^{M} g_m(X_i), \tag{6}$$

where each $g_m(\cdot)$ is a decision tree and M is the number of boosting rounds. For binary classification, we obtain predicted probabilities via the logistic link:

$$\hat{p}_i \equiv \Pr(y_i = 1 \mid X_i) = \frac{1}{1 + \exp(-f(X_i))}. \tag{7}$$

XGBoost fits trees sequentially to reduce the logistic (cross-entropy) loss, using regularization to control model complexity. In our implementation, we tune L1 and L2 regularization on leaf weights (`alpha` and `lambda`) as well as standard depth, shrinkage, and subsampling parameters.

C.2 Train/validation/test Split

We use a three-way split to separate hyperparameter tuning, final model training, and out-of-sample evaluation. We first reserve 10% of observations as a holdout test set. The remaining 90% of the data are split into training and validation samples, where the validation sample is 10% of the remaining 90% (i.e., 9% of the full sample). Both splits are stratified on the target variable to preserve the same proportions of the positive (default) and negative (non-default) classes as in the full sample.

C.3 Hyperparameter Tuning

We tune hyperparameters to maximize predictive performance on the validation set. We use the F1 score as the tuning criterion to balance precision and recall in a setting with class imbalance, while reporting AUC as a threshold-free measure of predictive performance. Using the `hyperopt` package in Python, we search over a pre-specified hyperparameter space with a Tree-structured Parzen Estimator (TPE) algorithm (`tpe.suggest`), running 200 evaluations. Let $\mathcal{L}(\theta)$ denote the validation loss associated with hyperparameter vector θ. TPE is a form of Bayesian optimization that models the distribution of hyperparameters conditional on performance: after each batch of trials, it partitions observed configurations into "good" and "bad" sets based on a loss quantile and fits separate density estimators. Let $p(\theta)$ denote a density over the hyperparameter space, and let $\ell(\theta)$ and $g(\theta)$ denote the densities conditional on low- and high-loss realizations, respectively:

$$\ell(\theta) = p(\theta \mid \mathcal{L}(\theta) \leq \gamma)$$
$$g(\theta) = p(\theta \mid \mathcal{L}(\theta) > \gamma)$$

New candidates are then proposed by sampling values that are relatively more probable under $\ell(\theta)$ than under $g(\theta)$, allowing the search to concentrate on promising regions of the hyperparameter space while maintaining exploration. For each candidate parameter vector, we fit an `XGBClassifier` on the training sample and compute the F1 score on the validation sample. The tuning objective is the negative F1 score (so that `hyperopt` minimizes loss):

$$\mathcal{L}(\theta) = -\mathrm{F1}(\theta)$$

The search space is:

- `max_depth` $\in \{2, 3, \ldots, 9\}$,
- `learning_rate` drawn from a log-uniform distribution on $[\exp(-5), \exp(-2)]$,
- `subsample` drawn uniformly on $[0.5, 1]$,
- `max_delta_step` $\in \{0, 1, 10\}$,
- `lambda` (L2 regularization; `reg_lambda` in XGBoost) drawn from a log-uniform distribution on $[\exp(-10), \exp(0)]$,
- `alpha` (L1 regularization; `reg_alpha` in XGBoost) drawn from a log-uniform distribution on $[\exp(-10), \exp(0)]$.

We train candidate models using `tree_method=hist` and `objective=binary:logistic`. The following table reports the optimal hyperparameters:

Optimal XGBoost hyperparameters

Hyperparameter	Value
`max_depth`	9
`learning_rate`	0.125499
`subsample`	0.881735
`max_delta_step`	1
`lambda`	0.008075
`alpha`	0.069882

C.4 Final Model Fitting and Evaluation

After selecting hyperparameters, we refit the model using the combined training and validation data (the full 90% non-test sample) and then evaluate predictive performance on the untouched 10% holdout test sample.

The following performance metrics are reported in our tables:

Accuracy measures the share of observations that are correctly classified:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}}$$

Precision measures the share of predicted defaulters who actually defaulted:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

Recall (1 − False Negative Rate) measures the share of actual defaulters who were correctly identified:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

F1 score is the harmonic mean of Precision and Recall:

$$\text{F1} = 2 \cdot \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Area Under the Receiver Operating Characteristic Curve (AUC) measures the probability that the model assigns a higher predicted default probability to a randomly chosen defaulter than to a randomly chosen non-defaulter:

$$\text{AUC} = \Pr\left(\hat{p}_i^D > \hat{p}_j^{ND}\right)$$

where $\hat{p}_i^D$ denotes the predicted default probability for a randomly selected defaulter and $\hat{p}_j^{ND}$ denotes the predicted default probability for a randomly selected non-defaulter.

False positive rate is the share of non-defaulters incorrectly classified as defaulters:

$$\text{False Positive Rate} = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}}$$

Equivalently, using Precision, Recall, and the average default rate π, the false positive rate can be written as:

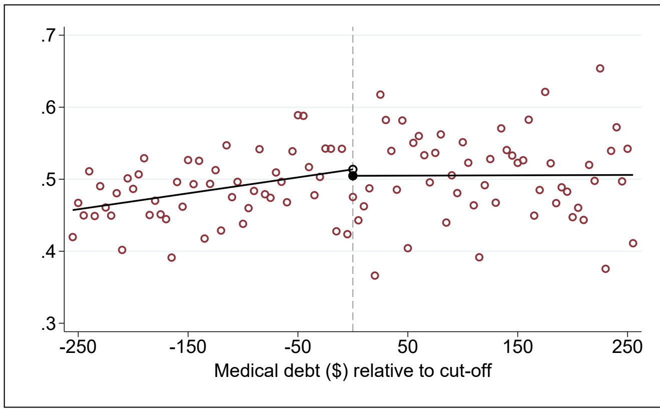
$$\text{False Positive Rate} = \frac{\pi\,\text{Recall}\,(1 - \text{Precision})}{(1 - \pi)\,\text{Precision}}$$

where π denotes the average default rate in the evaluation sample.

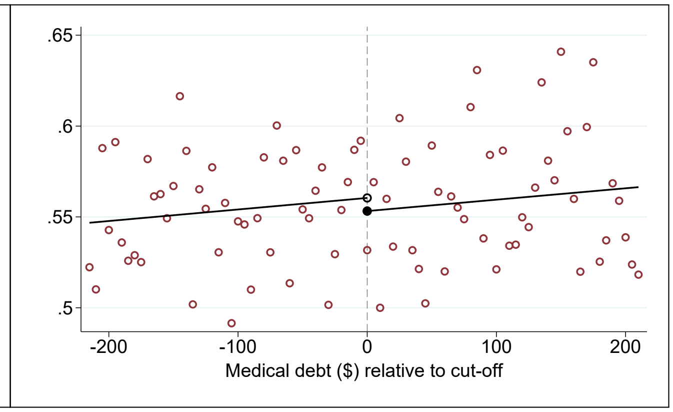
D Additional Figures and Tables

Figure A.1: Additional Credit Outcomes for RD Analysis, 2024

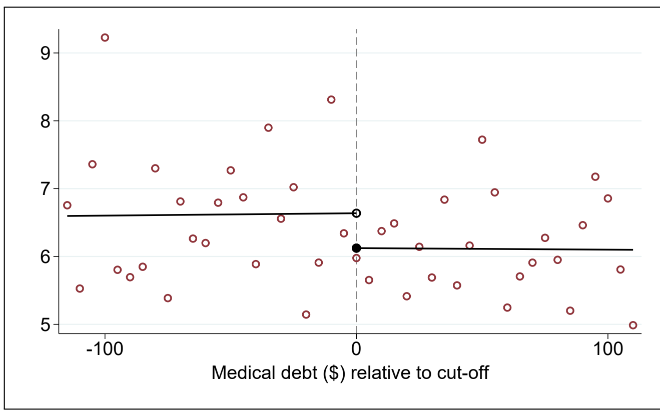
(a) Number of Accounts 90+ Days Past Due



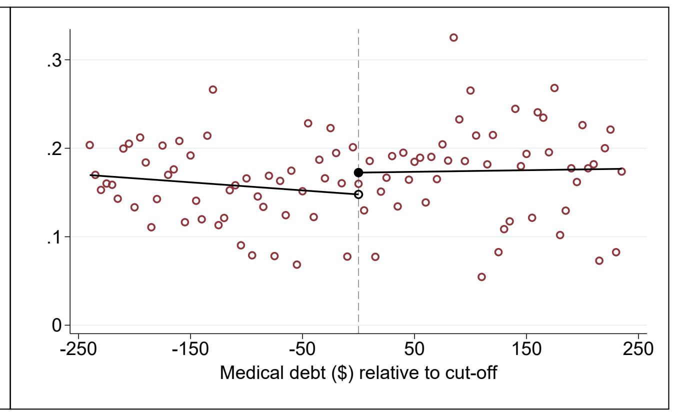
(b) Number of Inquiries in Last 6 Months



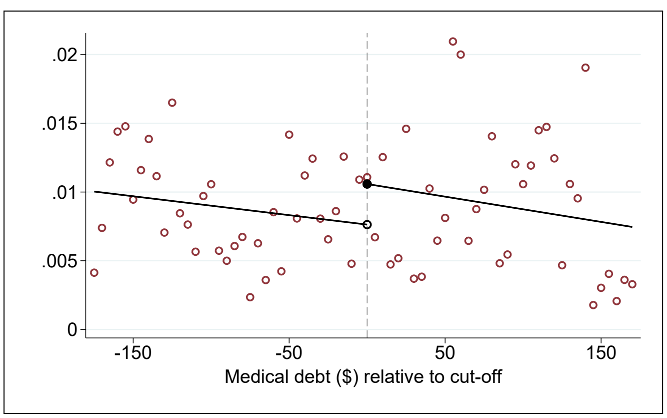
(c) Revolving Limit ($1,000)



(d) Alternative Credit Balance ($1,000)



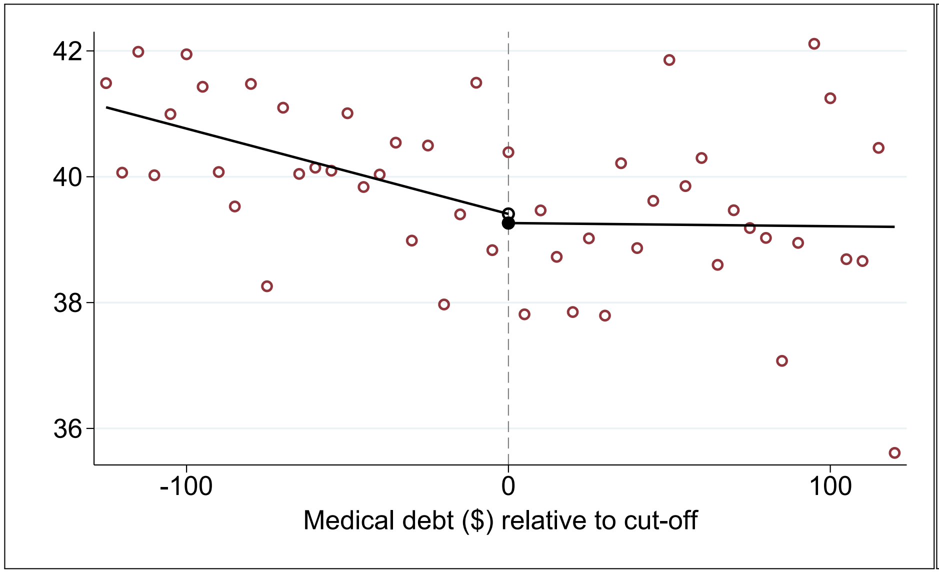
(e) Number of New Mortgages in Last 6 Months



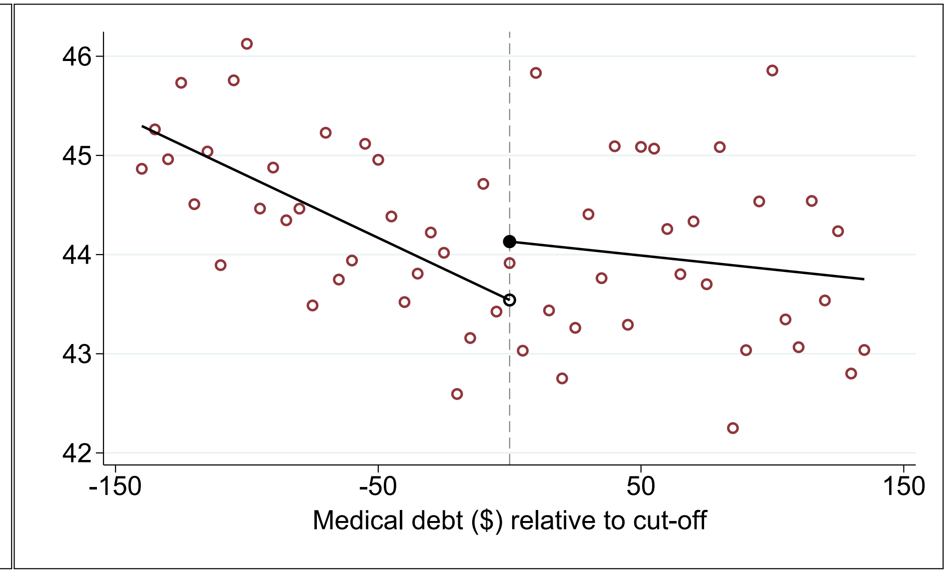
Notes: This figure shows the relationship between 2022 medical debt and five supplementary credit outcomes in 2024. Medical debt is defined as the maximum value of the consumer's 2022 medical debt collections accounts, measured relative to the $500 threshold. RD estimates from Equation (2) are reported in Table A.6.

Figure A.2: Covariate Smoothness Test: Demographics

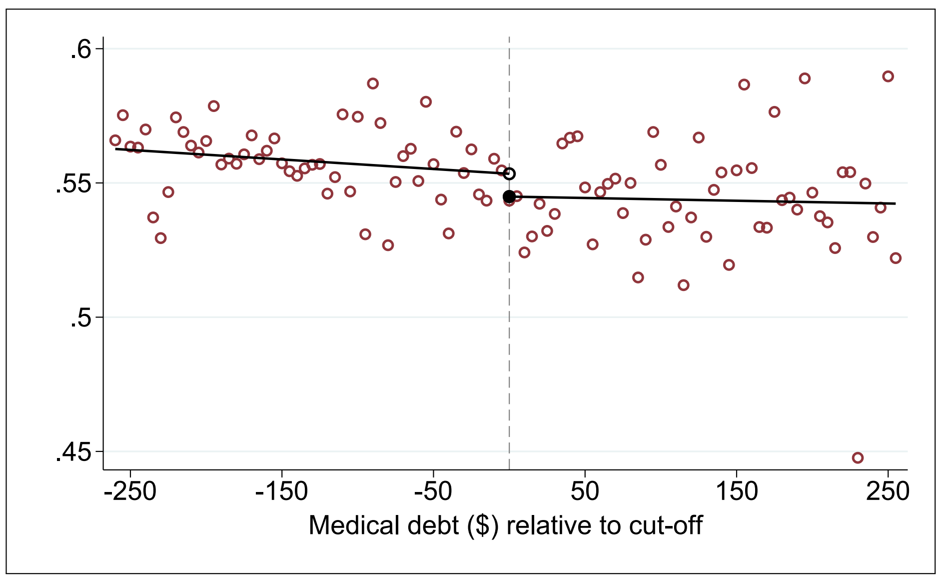
(a) Income ($1000)



(b) Age (years)

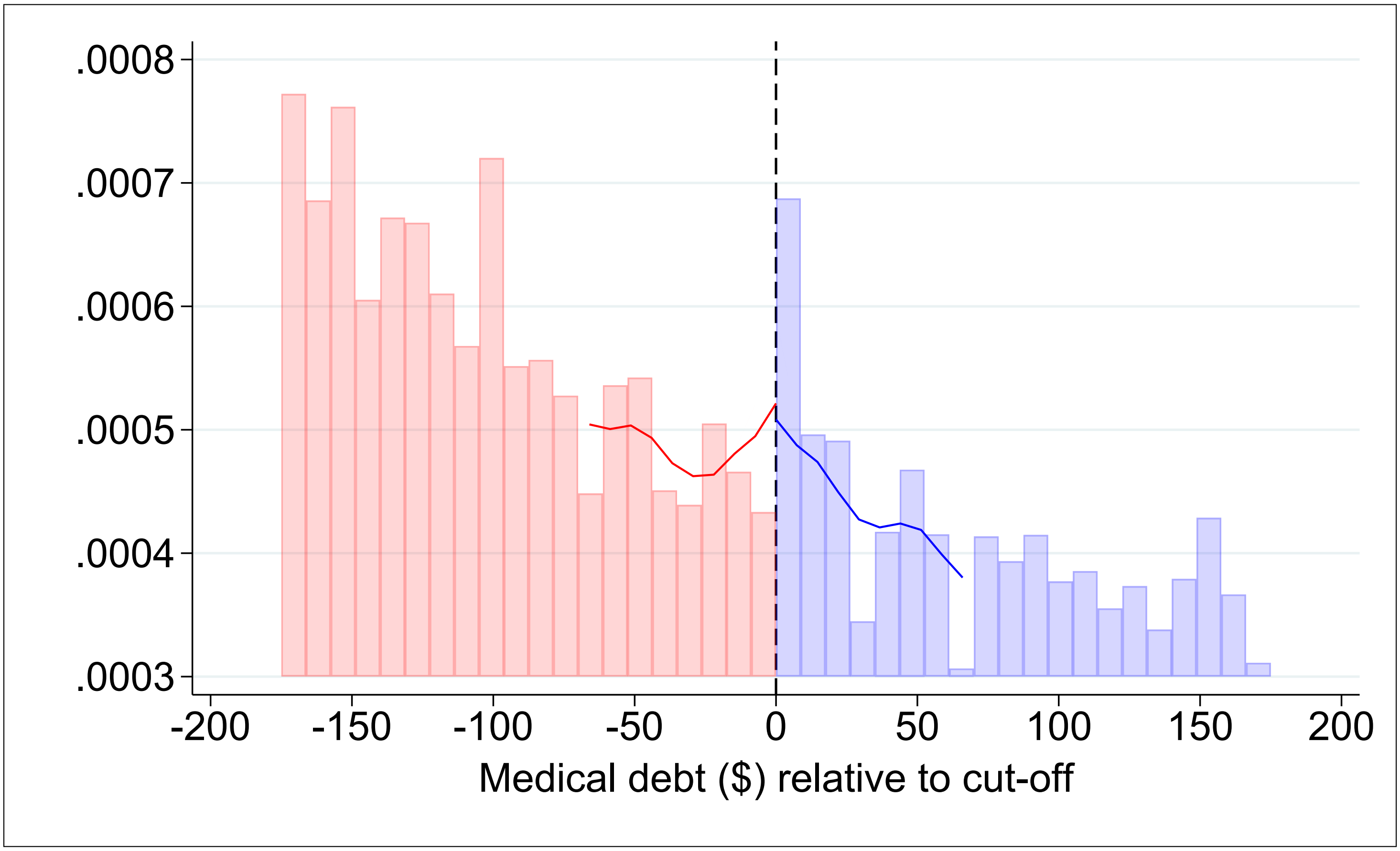


(c) Female (%)



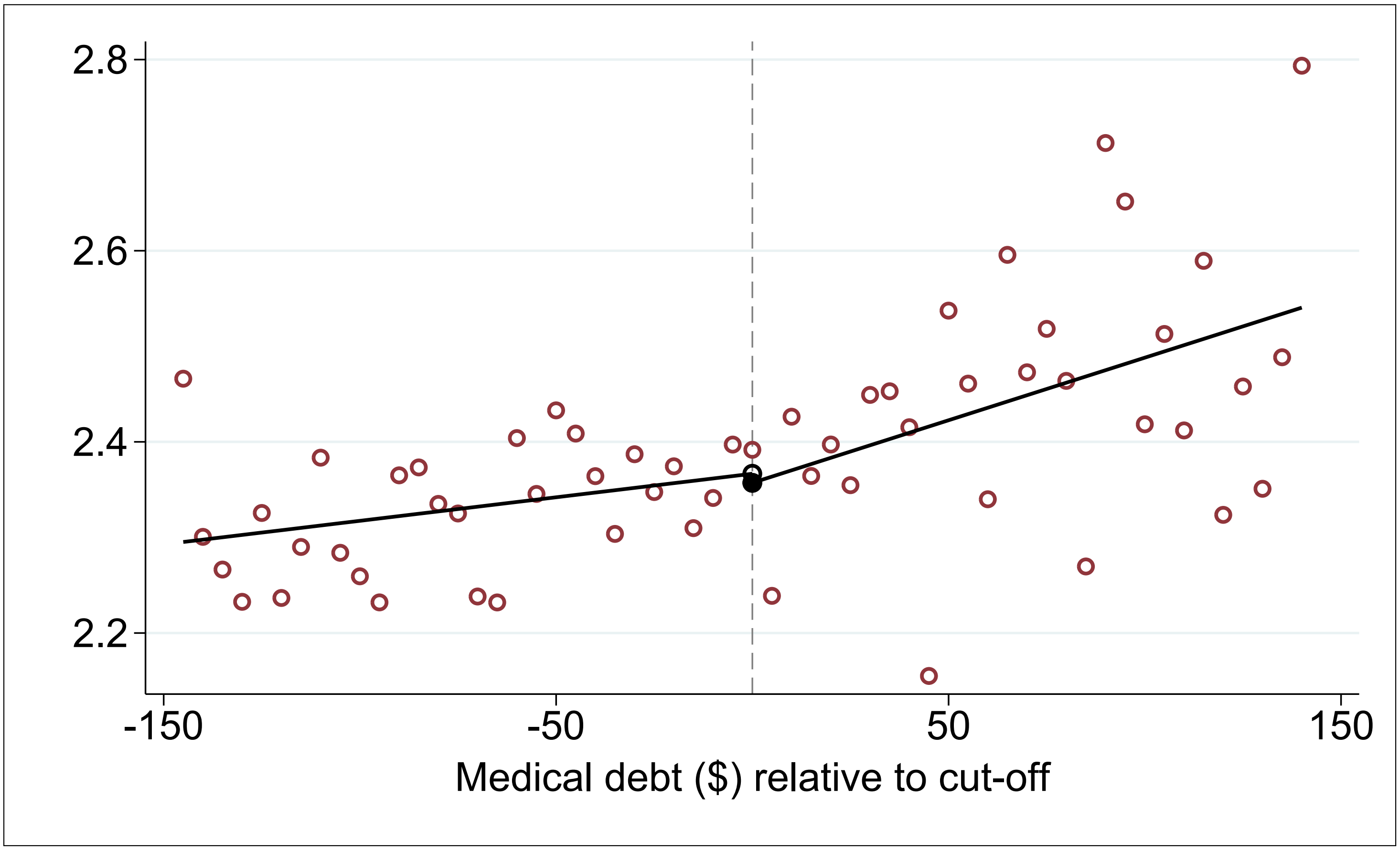
Notes: This figure shows the relationship between 2022 medical debt and three demographic variables in 2022. Medical debt is defined as the maximum value of the consumer's 2022 medical debt collections accounts, measured relative to the $500 threshold. RD estimates from Equation (2) are reported in Table A.5.

Figure A.3: Density Test of the Running Variable (Medical Debt), 2022



Notes: This figure shows the results of the McCrary density test for a discontinuity in the distribution of medical debt at the $500 cutoff (normalized to zero on the x-axis). The running variable is defined as the maximum value of the consumer's 2022 medical debt collections accounts, measured relative to the $500 threshold. The red and blue lines are the local polynomial density estimates on each side of the cutoff. The distribution shows no evidence of bunching below the cutoff, which would be consistent with strategic manipulation to qualify for deletion. Instead, the pattern is consistent with rounding behavior or reporting practices if some providers systematically refrain from reporting debts under $500.

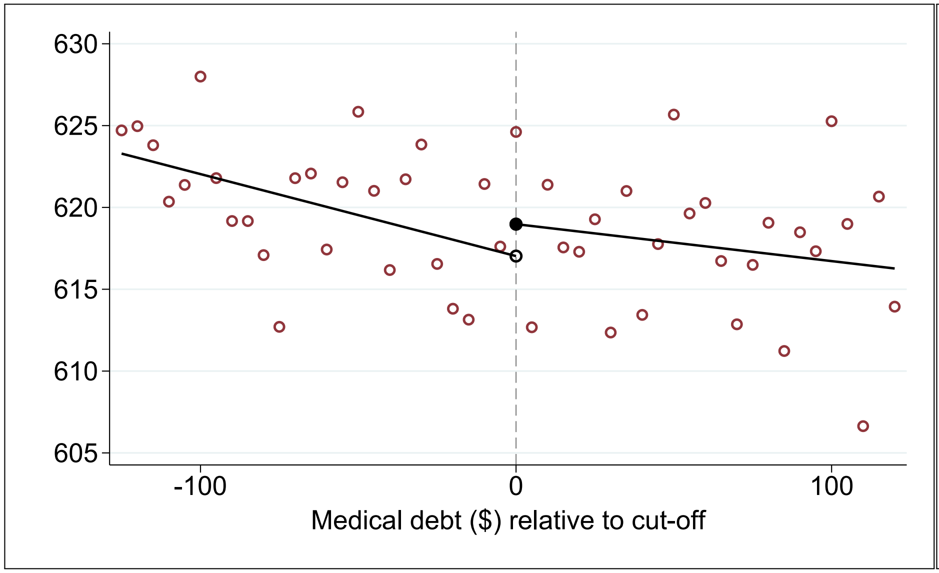
Figure A.4: Falsification test: Average Number of Accounts per Person, 2022



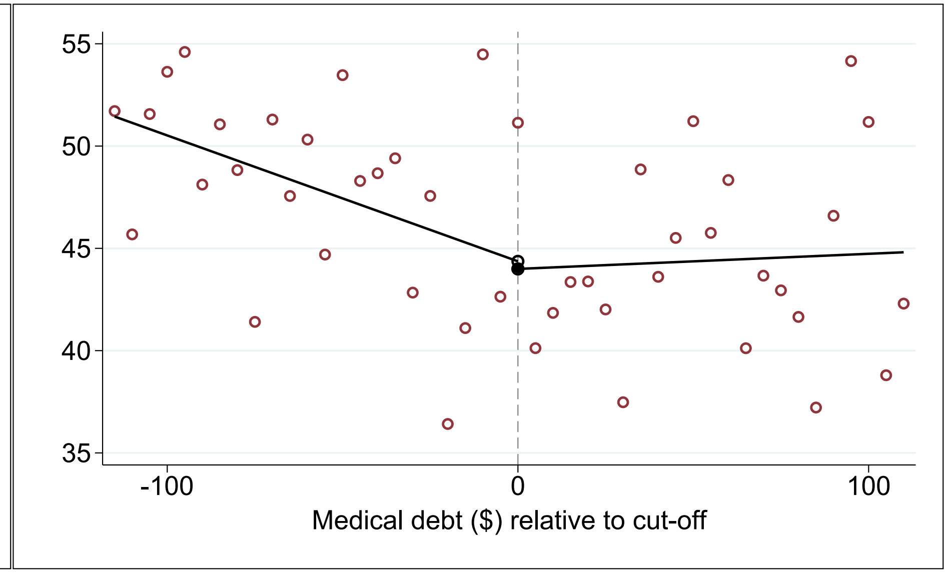
Notes: This figure shows the relationship between the medical debt running variable and the average number of medical debt collections accounts per person in 2022. The medical debt running variable is defined as the maximum value of the consumer's 2022 medical debt collections accounts, measured relative to the $500 threshold. RD estimates from Equation (2) are reported in Table A.7.

Figure A.5: Falsification Test: Access to Credit, 2022

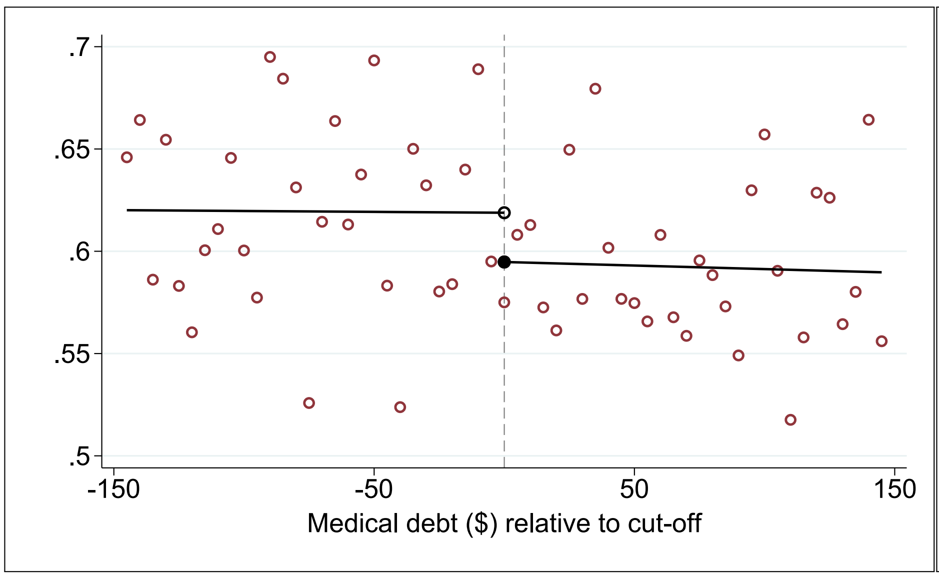
(a) Credit Score



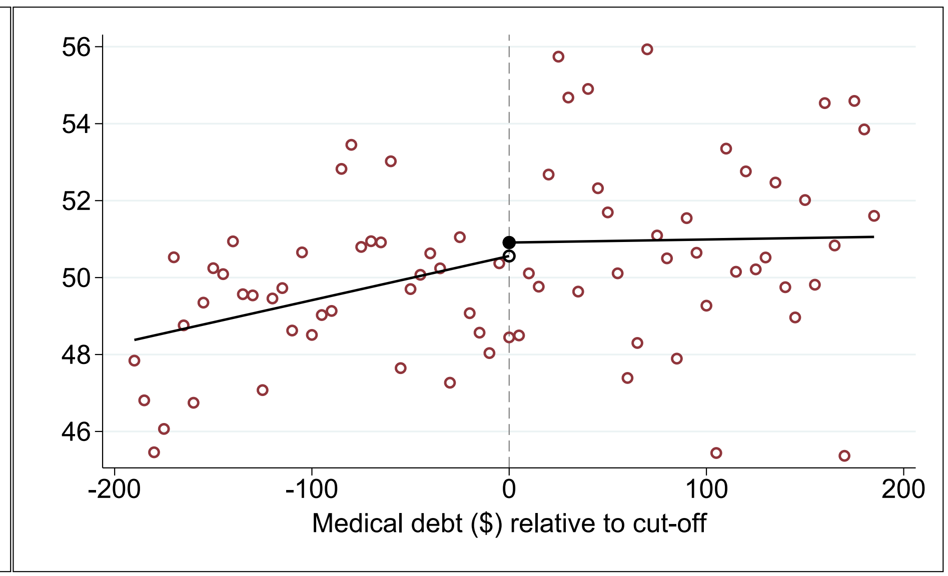
(b) Total Credit Balance ($1,000)



(c) Number of Accounts Opened in Last 6 Months

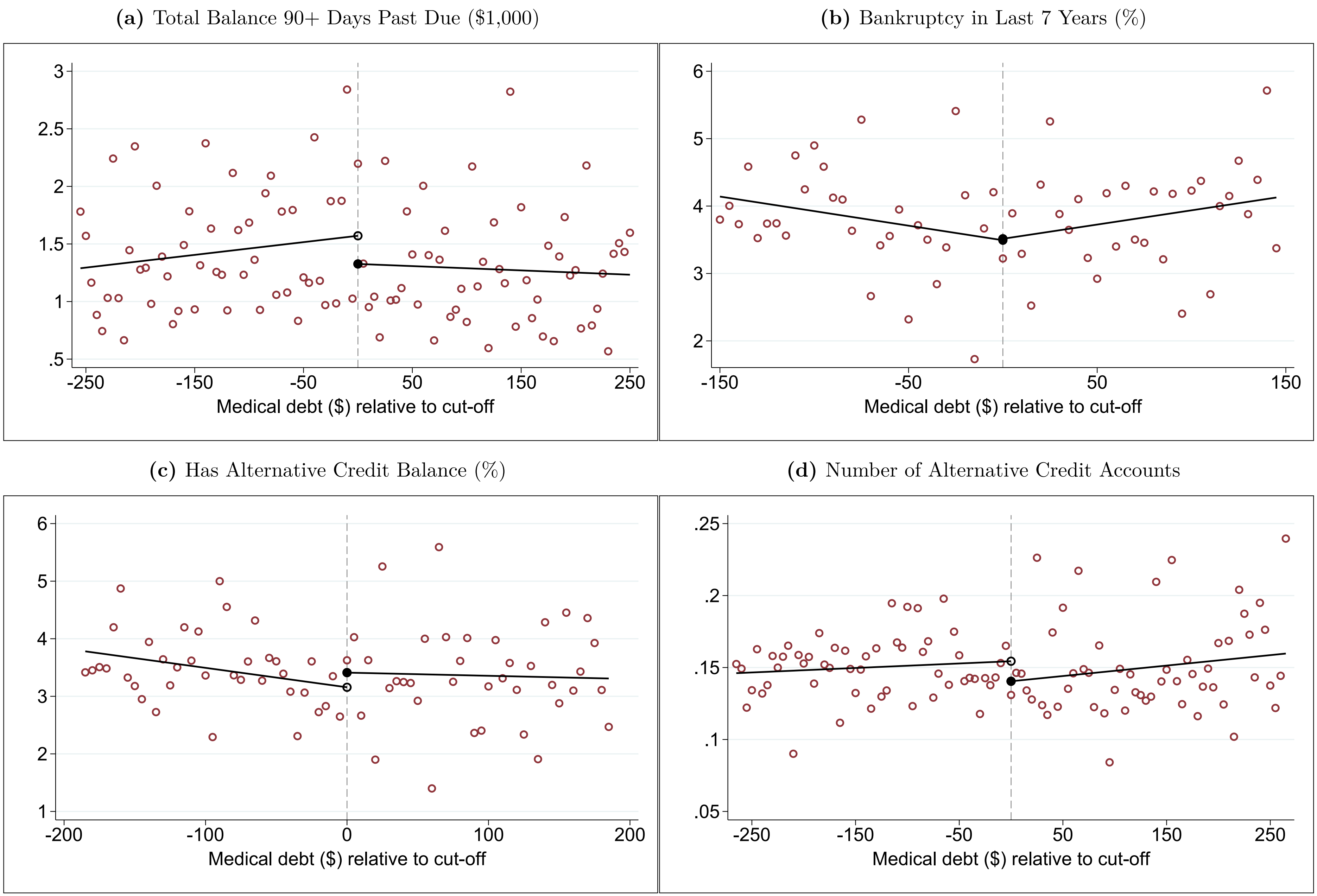


(d) Revolving Utilization (%)



Notes: This figure shows the relationship between 2022 medical debt and four credit measures in 2022. Medical debt is defined as the maximum value of the consumer's 2022 medical debt collections accounts, measured relative to the $500 threshold. RD estimates from Equation (2) are reported in Table A.7.

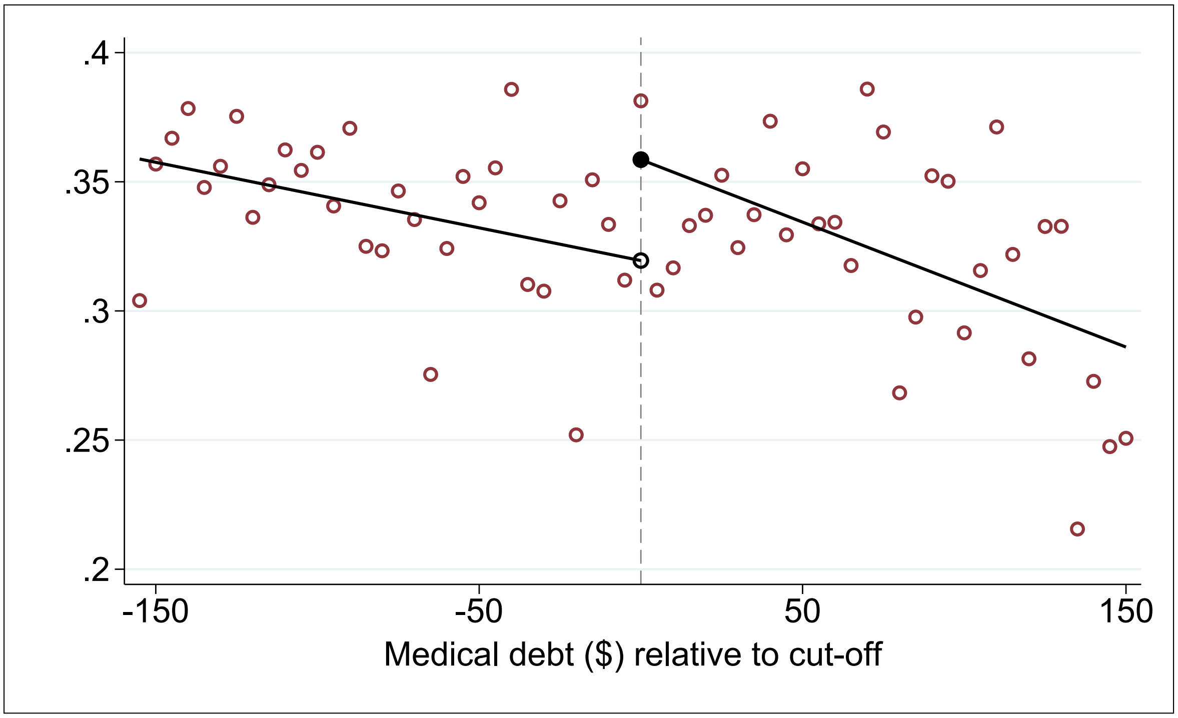
Figure A.6: Falsification Test: Financial Distress and Access to Alternative Credit, 2022



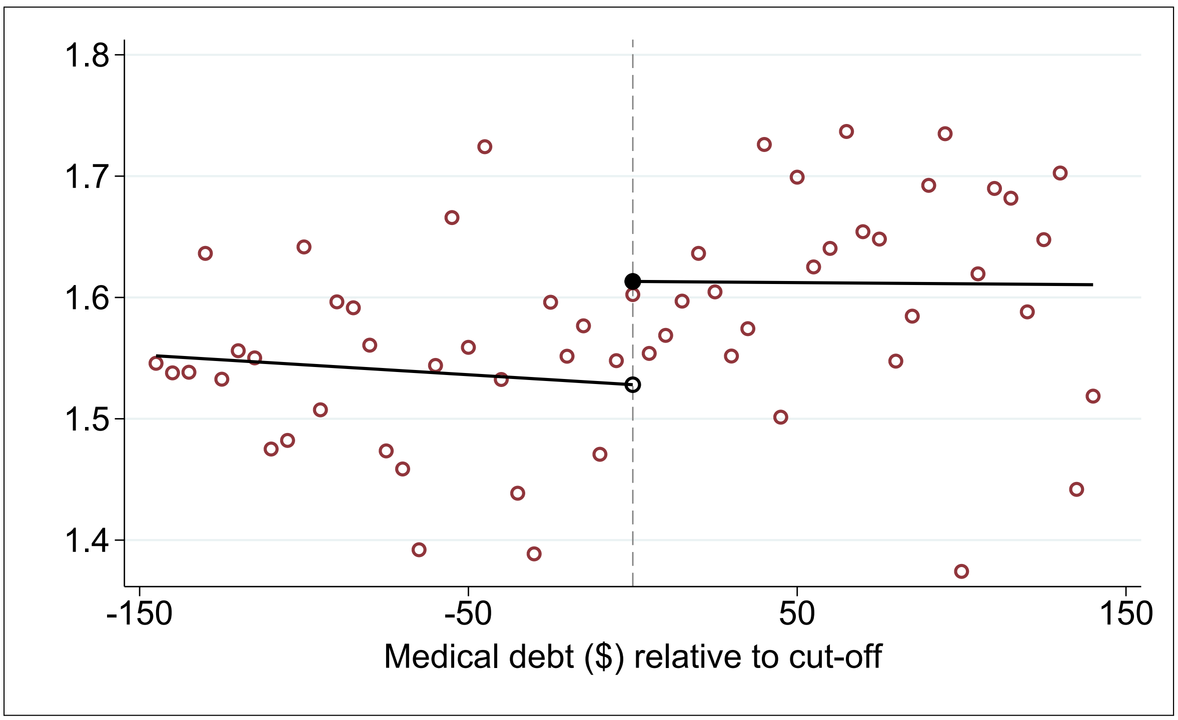
Notes: This figure shows the relationship between medical debt in 2022 and measures of financial distress and access to alternative credit in 2022. Medical debt is defined as the maximum value of the consumer's 2022 medical debt collections accounts, measured relative to the $500 threshold. RD estimates from Equation (2) are reported in Table A.7.

Figure A.7: Placebo Test: Two-Year Evolution of 2020 Medical Debt Collections Accounts

(a) Share of Surviving Accounts

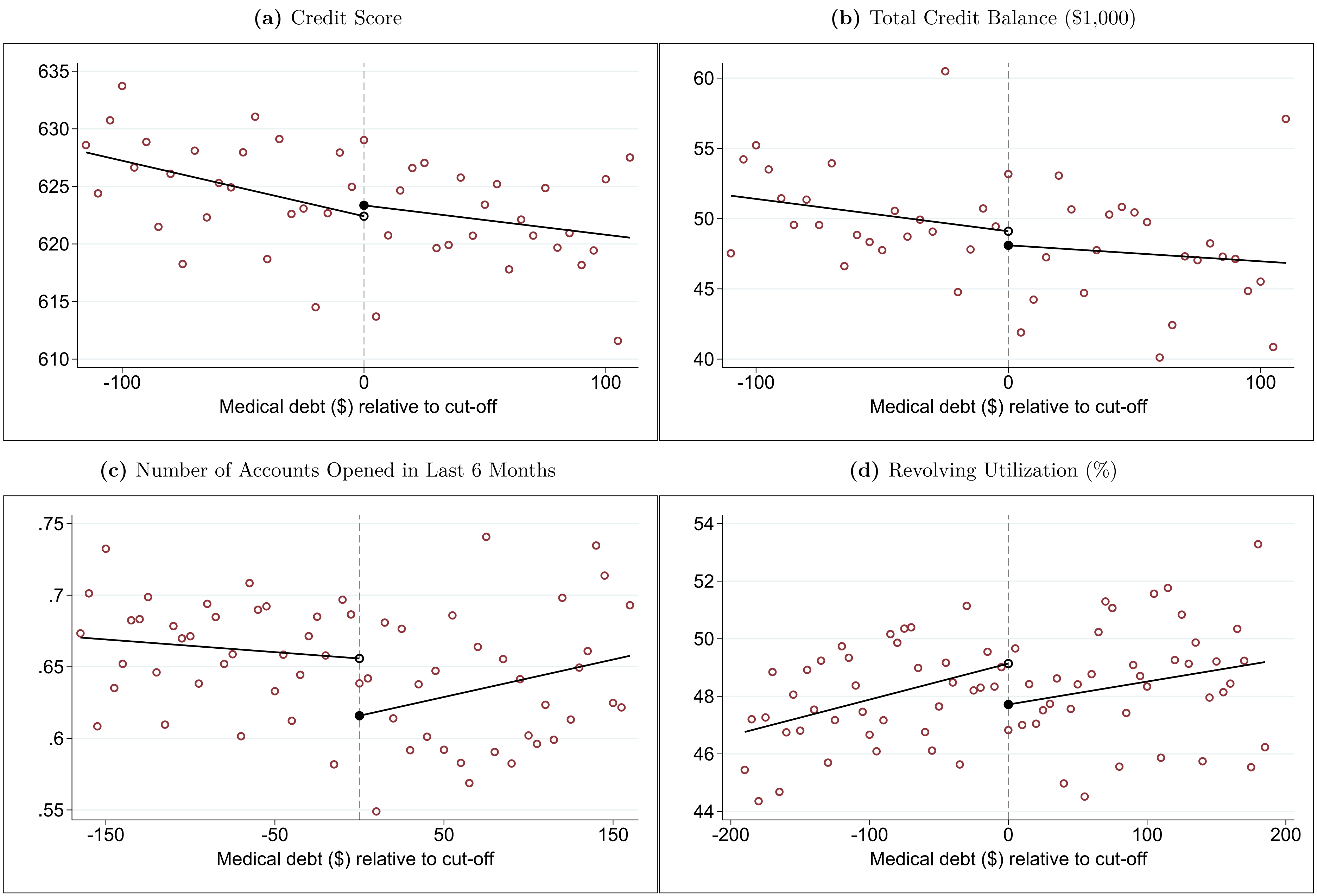


(b) Average Number of Accounts per Person



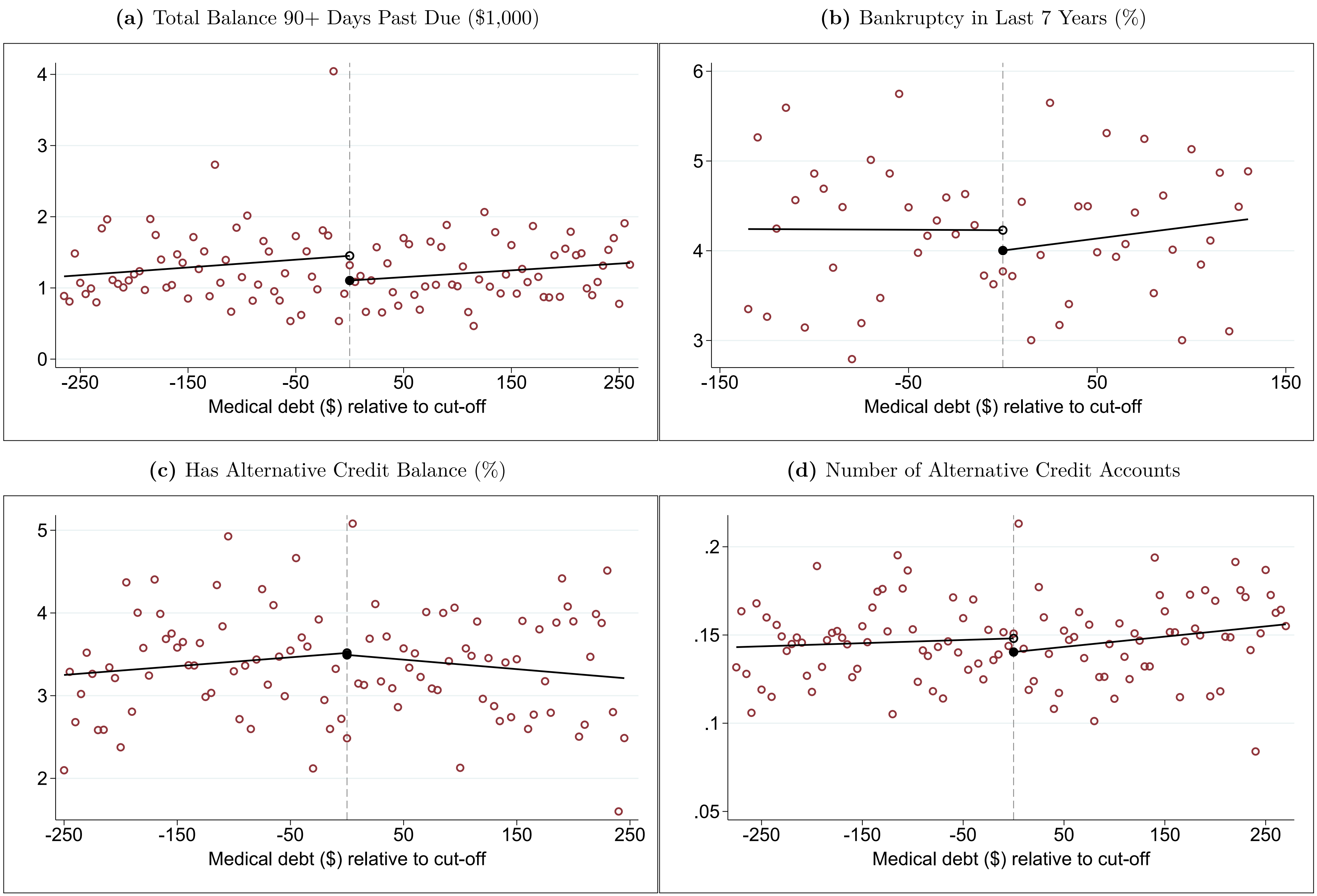
Notes: Panel (a) shows the proportion of 2020 medical debt collection accounts which remain present on 2022 credit reports by account amount, where the amount is measured as distance from the $500 threshold. Panel (b) shows the average number of accounts per person. In panel (b), the running variable is equal to the maximum value of the consumer's medical debt collections accounts. RD estimates from Equation (2) are reported in Table A.8.

Figure A.8: Placebo Test: Access to Credit



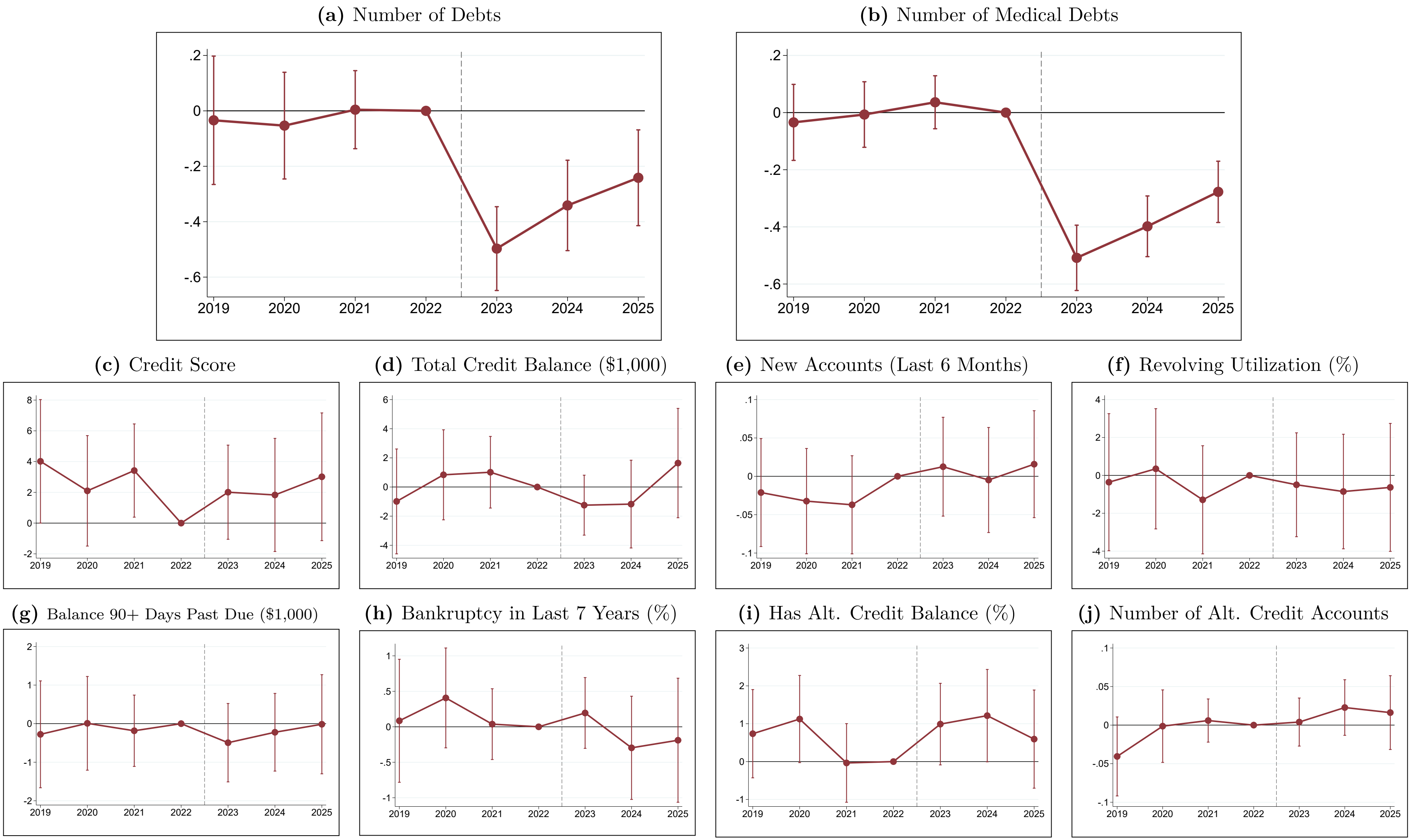
Notes: This figure shows the relationship between 2020 medical debt and four supplementary credit outcomes in 2022. Medical debt is defined as the maximum value of the consumer's 2020 medical debt collections accounts, measured relative to the $500 threshold. RD estimates from Equation (2) are reported in Table A.8.

Figure A.9: Placebo Test: Financial Distress and Access to Alternative Credit



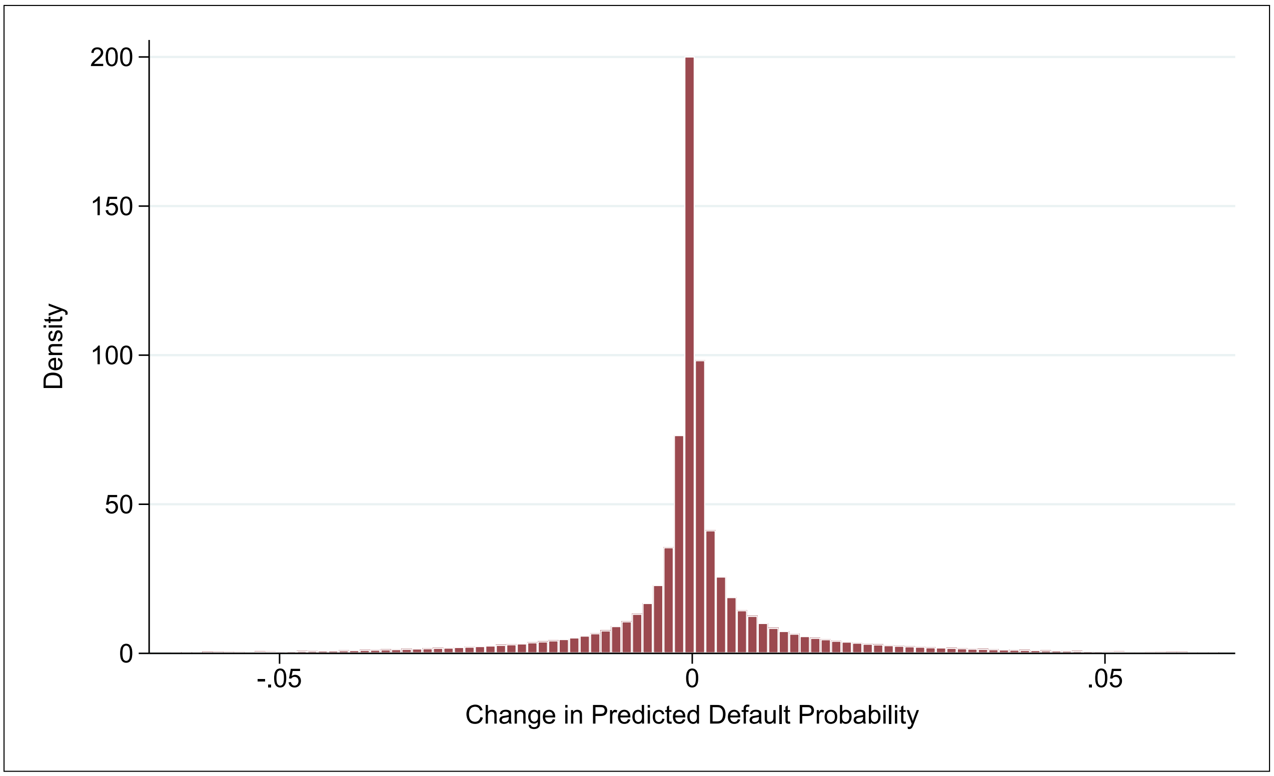
Notes: This figure shows the relationship between medical debt in 2020 and measures of financial distress and access to alternative credit in 2022. Medical debt is defined as the maximum value of the consumer's 2020 medical debt collections accounts, measured relative to the $500 threshold. RD estimates from Equation (2) are reported in Table A.8.

Figure A.10: Event Study: Effects of Medical Debt Removal on Credit Outcomes



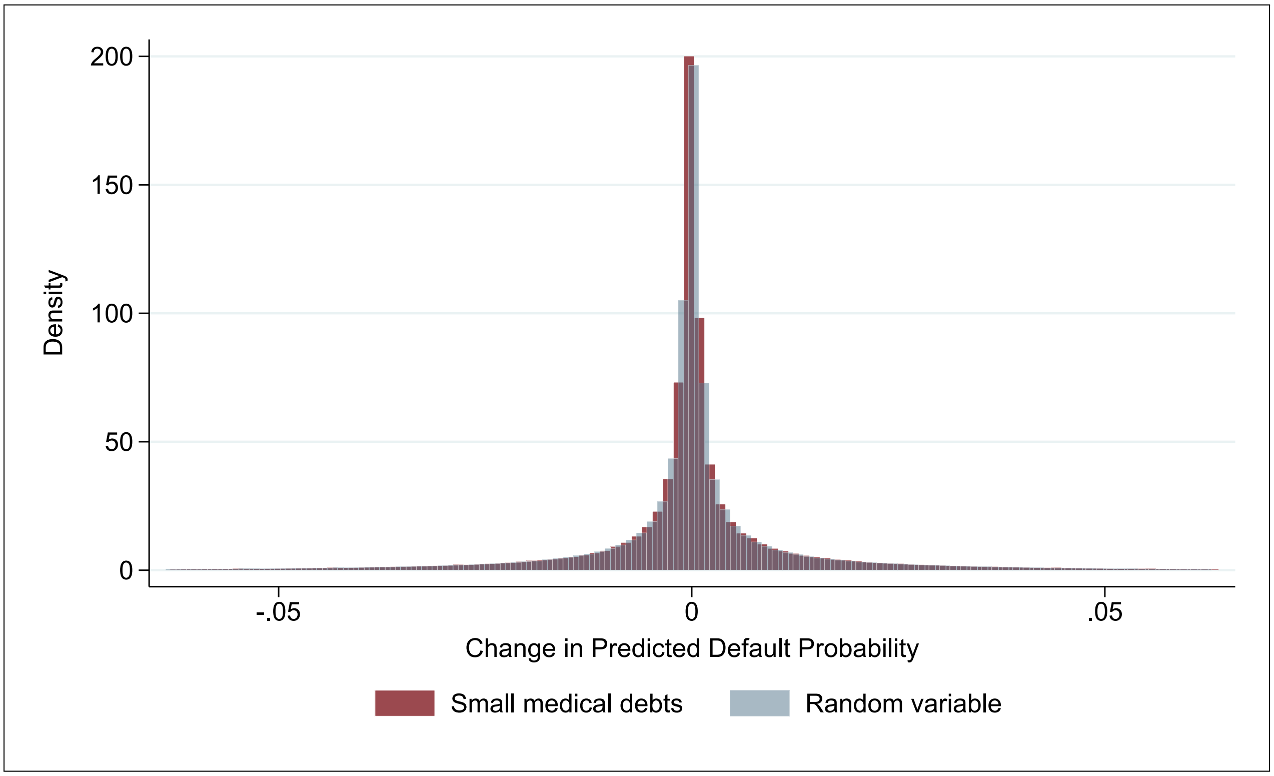
Notes: This figure plots $\hat{\beta}_k$ from the event study regression $y_{it} = \alpha_i + \alpha_t + \sum_{k \neq 2022} \beta_k \big(D_i \cdot \mathbf{1}[t = k]\big) + \varepsilon_{it}$, where $D_i = 1$ if consumer i's largest outstanding medical collection balance in 2022 was below the $500 deletion threshold and $D_i = 0$ otherwise. α_i and α_t denote individual and year fixed effects. The sample includes consumers within $20 of the $500 threshold who held at least one medical debt collection in 2022. Bars show 95% confidence intervals with standard errors clustered by consumer.

Figure A.11: Effect of Removing Small Medical Debt Collections on Predicted Default Probabilities



Notes: This figure shows the change in the predicted probability of default over 24 months following the removal of small (< $500) medical debt collections for 2.4 million consumers in the GCCP. Predictions, generated using the credit scoring model described in Section 5, are based on 2019 data.

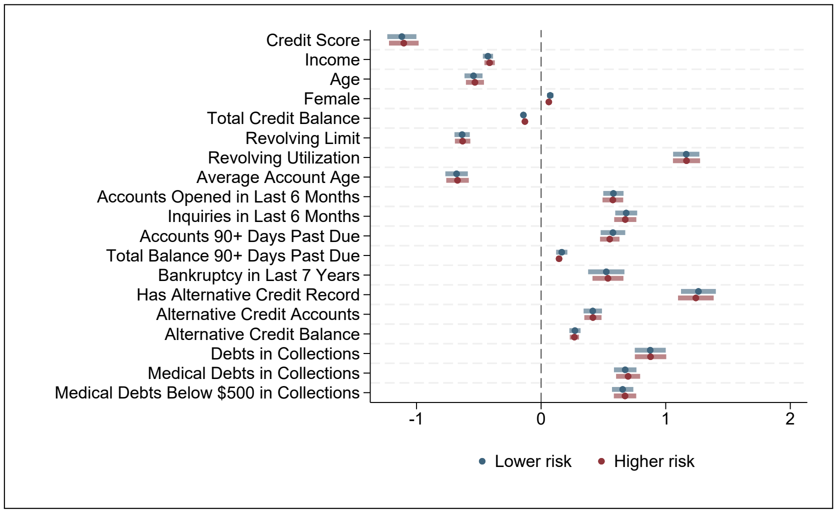
Figure A.12: Effect of Removing Small Medical Debt Collections vs. Removing Noise



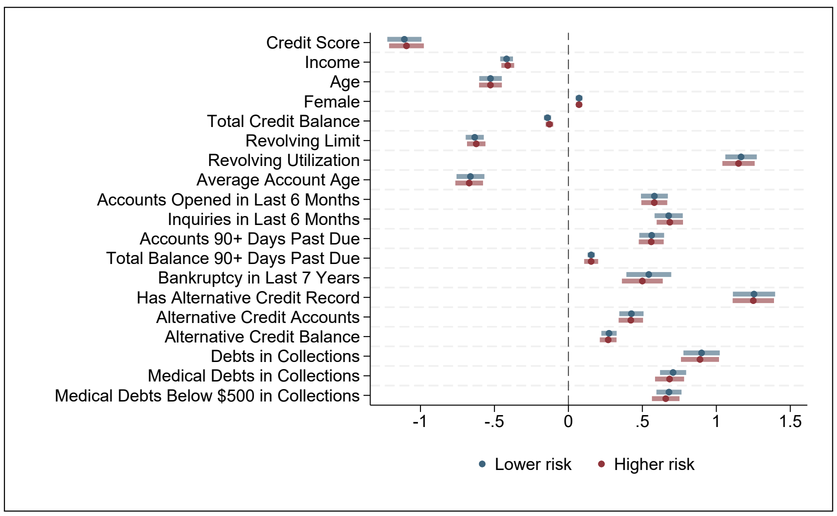
Notes: This figure shows the change in the predicted probability of default over 24 months following the removal of a variable from the credit scoring model described in Section 5. The red histogram reproduces the plot from Figure A.11, showing the effect of removing small (< $500) medical debt collections. The blue histogram shows the effect of removing a random noise predictor that was drawn randomly from a uniform distribution.

Figure A.13: Covariate Differences Across Reclassified Consumers: Small Medical Debt Collections vs. Noise

(a) Small (< \$500) Medical Collections

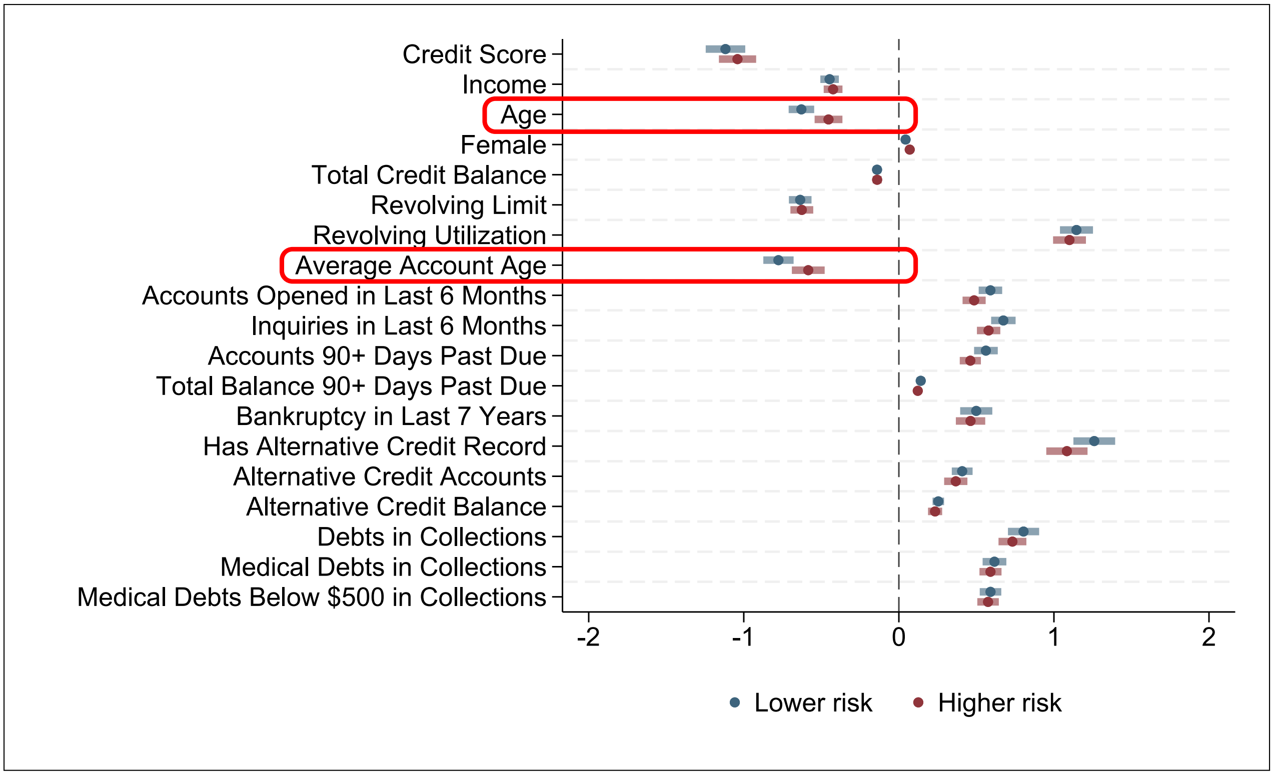


(b) Random Variable



Notes: This figure compares observable characteristics across consumers whose predicted default probabilities are most affected by the removal of small medical collections (Panel a) or a randomly generated predictor (Panel b) from the credit scoring model described in Section 5, relative to unaffected consumers. Higher-risk consumers are those in the top 5%, whose predicted probability of default increases by approximately 2 percentage points or more when small medical debt collections (Panel a) or a randomly generated predictor (Panel b) are removed from the credit scoring model. Lower-risk consumers are those in the bottom 5%, whose predicted default probability decreases by at least two percentage points. Unaffected consumers are those between the 25th and 75th percentiles, whose predicted probabilities change by less than about 0.2 percentage points. Each dot represents the difference in a standardized characteristic between the indicated group and the unaffected group. Horizontal bars represent 95% confidence intervals. We divide consumers in the full sample into 100 equal-sized bins based on changes in predicted default probability and cluster standard errors at the bin level.

Figure A.14: Covariate Differences Across Reclassified Consumers: Length of Credit History



Notes: This figure compares observable characteristics across consumers whose predicted default probabilities are most affected by the removal of credit history length variables from the credit scoring model described in Section 5, relative to unaffected consumers. Higher-risk consumers are those in the top 5% of changes in predicted default probability when information on credit history length is removed from the model. Lower-risk consumers are those in the bottom 5% of this distribution. Unaffected consumers are those between the 25th and 75th percentiles, whose predicted probabilities change by less than about 0.2 percentage points. Each dot represents the difference in a standardized characteristic between the indicated group and the unaffected group. Blue dots correspond to lower-risk consumers and red dots correspond to higher-risk consumers. Horizontal bars represent 95% confidence intervals. We divide consumers in the full sample into 100 equal-sized bins based on changes in predicted default probability and cluster standard errors at the bin level.

Table A.1: Share of Consumers with Medical Debt Collections: GCCP vs. External Sources

	(1)	(2)	(3)
Year	GCCP	Urban Institute	CFPB
2018	16.5%	16%	17.6%
2019	15.7%	16%	17.5%
2020	15.4%	15%	16%
2021	14.3%	14%	15.5%
2022	12.7%	12%	14%
2023	7.0%	5%	

Notes: This table compares the share of individuals with medical debt collections in the GCCP to estimates from other sources. Column (1) reports the share of consumers with at least one account in medical debt collections. Columns (2) and (3) present estimates from Blavin et al. (2023) and Sandler and Nathe (2022), respectively. The GCCP data are measured in March, the Urban Institute data in August, and the CFPB data in January.

Table A.2: Medical Debt Collection Balances ($bn): GCCP vs. CFPB

	(1)	(2)	(3)
Year	All Medical Debt Balance	Balance for Accounts <$500	CFPB
2018	95.92	17.29	96.78
2019	95.79	16.18	96.50
2020	94.91	15.78	95.79
2021	89.35	13.99	89.29
2022	68.93	11.63	-
2023	52.12	2.91	-
2024	39.07	0.06	-

Notes: This table compares aggregate medical debt collection balances in the GCCP to external estimates from the CFPB. Column (1) reports the total medical debt collections balance. Column (2) reports the total medical debt collections balance for accounts with balances below $500. Column (3) presents corresponding aggregate balance estimates reported by the CFPB (CFPB, 2022). Both the GCCP and CFPB balances are measured in March of the indicated year.

Table A.3: Alternative RD Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Number of Debts	Number of Medical Debts	Credit Score	Total Credit Balance ($1,000)	Number of Accounts Opened in Last 6 Months	Revolving Utilization (%)	Total Balance 90+ Days Past Due ($1,000)	Bankruptcy in Last 7 Years (%)	Has Alternative Credit Balance (%)	Number of Alternative Credit Accounts	Share of Surviving Accounts
A. Person-Level (Max Debt)											
ABOVE^{2022}	–0.22**	–0.30**	0.18	–2.3	0.038	–0.47	–0.067	–0.37	0.36	0.011	N/A
	[–0.32, –0.14]	[–0.32, –0.27]	[–4.2, 5.5]	[–8.7, 2.5]	[–0.0019, 0.088]	[–2.7, 1.4]	[–0.76, 0.54]	[–1.2, 0.54]	[–0.31, 1.3]	[–0.015, 0.045]	
Control Mean	1.5	0.49	618	50	0.50	53	2.0	3.2	3.9	0.18	
% of Mean	–15	–61	0.029	–4.7	7.5	–0.89	–3.4	–11	9.3	6.2	
Optimal Bandwidth	± 146.08	± 282.39	± 122.15	± 103.82	± 167.87	± 246.00	± 250.43	± 151.71	± 225.05	± 228.93	
Observations											
Total	272,490	272,490	272,490	272,490	272,490	169,547	272,490	272,490	272,490	272,490	
In-Bandwidth	38,500	84,940	31,760	26,730	44,630	43,334	73,163	40,146	63,354	64,348	
B. Person-Level (Min Debt)											
ABOVE^{2022}	–0.41**	–0.44**	0.17	–2.4	0.0083	–0.34	0.19	0.14	–0.026	–0.0082	N/A
	[–0.56, –0.30]	[–0.53, –0.37]	[–5.5, 5.9]	[–9.4, 2.9]	[–0.037, 0.060]	[–3.0, 2.5]	[–0.49, 0.78]	[–0.67, 1.1]	[–0.75, 0.93]	[–0.041, 0.030]	
Control Mean	1.7	0.72	614	46	0.47	52	2.0	2.5	3.8	0.18	
% of Mean	–24	–61	0.028	–5.3	1.7	–0.66	9.5	5.6	–0.68	–4.7	
Optimal Bandwidth	± 119.65	± 119.27	± 122.98	± 105.74	± 189.36	± 201.75	± 252.06	± 161.98	± 260.51	± 282.43	
Observations											
Total	272,490	272,490	272,490	272,490	272,490	169,547	272,490	272,490	272,490	272,490	
In-Bandwidth	22,710	22,710	23,345	20,007	38,246	24,278	56,875	31,621	59,524	66,834	
C. Account-Level											
ABOVE^{2022}	–0.44**	–0.44**	2.7	0.17	0.013	–0.83	0.086	0.026	0.29	0.0088	0.23**
	[–0.64, –0.29]	[–0.63, –0.30]	[–0.26, 6.7]	[–3.6, 3.3]	[–0.015, 0.050]	[–3.0, 0.91]	[–0.33, 0.44]	[–0.55, 0.62]	[–0.28, 0.92]	[–0.011, 0.032]	[0.22, 0.23]
Control Mean	1.8	0.74	614	46	0.49	53	1.9	2.9	3.9	0.18	0.24
% of Mean	–25	–60	0.45	0.37	2.7	–1.6	4.5	0.90	7.4	4.9	93
Optimal Bandwidth	± 93.14	± 89.32	± 103.04	± 100.75	± 173.42	± 176.12	± 234.72	± 150.30	± 248.15	± 240.21	± 256.32
Observations											
Total	726,501	726,501	726,501	726,501	726,501	409,292	726,501	726,501	726,501	726,501	726,501
In-Bandwidth	59,468	56,920	66,538	64,690	116,565	63,581	171,319	99,987	185,660	177,707	196,782

Notes: This table presents the estimated coefficient ($-\beta$) and 95% confidence intervals from Equation (2) using outcomes measured in 2024. Outcome variables are defined in Table A.4. The running variable in Panel A corresponds to the maximum 2022 medical debt collection amount across the consumer's accounts, measured relative to the $500 reporting cutoff. The running variable in Panel B corresponds to the minimum 2022 medical debt collection amount across the consumer's accounts, measured relative to the $500 reporting cutoff. The running variable in Panel C, an account-level specification, corresponds to the debt amount in the consumer's 2022 medical debt collection accounts, measured relative to the $500 reporting cutoff. "Share of Surviving Accounts" varies across accounts and is not reported for the person-level specifications presented in Panels A and B. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Control Mean in Panel A reports the mean of the dependent variable for consumers whose maximum medical debt in collections in 2022 falls between $500 and $600. Control Mean in Panel B reports the mean of the dependent variable for consumers whose minimum medical debt in collections in 2022 falls between $500 and $600. Control Mean in Panel C reports the mean of the dependent variable for consumers whose medical debt in collections in 2022 falls between $500 and $600. The reported optimal bandwidth corresponds to the symmetric half-width (in dollars) around the $500 cutoff used to estimate each specification. "In-Bandwidth" observations report the number of observations falling within this window, while "Total" observations report the full estimation sample. A */** denotes statistical significance at the 5% and 1% levels respectively, using robust inference.

Table A.4: Definitions of Main RD Outcomes

Outcome	Definition	Notes
Number of Debts	Total number of debt collection accounts	
Number of Medical Debts	Total number of medical debt collection accounts	
Credit Score	Consumer credit score (VantageScore)	
Total Credit Balance ($1,000)	Total balance across all open credit accounts	
Number of Accounts Opened in Last 6 Months	Total number of newly opened credit accounts	
Revolving Utilization (%)	Balance-to-limit ratio on open revolving credit accounts	Equal to missing if limit equals 0
Total Balance 90+ Days Past Due ($1,000)	Total balance across all accounts 90 or more days past due	
Bankruptcy in Last 7 Years (%)	Indicator for any bankruptcy filing (Chapter 7 or Chapter 13) in past 7 years	
Has Alternative Credit Balance (%)	Indicator for positive alternative (non-traditional) credit balance, including payday loans and title loans	
Number of Alternative Credit Accounts	Total number of alternative credit accounts	

Notes: Outcome names match those reported in the main text tables and figures. All outcomes are measured at the consumer level using GCCP data.

Table A.5: Covariate Smoothness: RD Estimates of the Effect of Medical Debt Collections Deletion

	(1)	(2)	(3)
	Income ($1,000)	Age (years)	Female (%)
ABOVE^{2022}	0.39	–0.44	0.011
	[–0.74, 1.4]	[–1.1, 0.40]	[–0.0085, 0.031]
Control Mean	39	44	0.54
% of Mean	1.0	–0.99	2.0
Optimal Bandwidth	± 121.13	± 139.18	± 258.38
Observations			
Total	272,490	272,490	265,995
In-Bandwidth	31,547	36,401	73,984

Notes: This table presents the estimated coefficient ($-\beta$) and 95% confidence intervals from Equation (2) using outcomes measured in 2022. The running variable is defined as the maximum 2022 medical debt collection amount across the consumer's accounts, measured relative to the $500 reporting cutoff. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Control Mean reports the mean of the dependent variable for consumers whose maximum medical debt in collections in 2022 falls between $500 and $600. The reported optimal bandwidth corresponds to the symmetric half-width (in dollars) around the $500 cutoff used to estimate each specification. "In-Bandwidth" observations report the number of observations falling within this window, while "Total" observations report the full estimation sample. A */** denotes statistical significance at the 5% and 1% levels respectively, using robust inference.

Table A.6: Additional Credit Outcomes: RD Estimates of the Effect of Medical Debt Collections Deletion

	(1)	(2)	(3)	(4)	(5)
	Number of Accounts 90+ Days Past Due	Number of Inquiries in Last 6 Months	Revolving Limit ($1,000)	Alternative Credit Balance ($1,000)	Number of Mortgage Accounts Opened in Last 6 Months
ABOVE^{2022}	0.018	0.016	0.60	–0.0073	–0.00074
	[–0.030, 0.073]	[–0.020, 0.051]	[–0.24, 1.7]	[–0.050, 0.049]	[–0.0051, 0.0046]
Control Mean	0.50	0.55	6.2	0.18	0.0094
% of Mean	3.5	3.0	9.8	–4.1	–7.9
Optimal Bandwidth	± 255.43	± 212.56	± 114.92	± 237.21	± 173.93
Observations					
Total	272,490	272,490	272,490	272,490	272,490
In-Bandwidth	74,886	58,910	29,691	67,500	46,293

Notes: This table presents the estimated coefficient ($-\beta$) and 95% confidence intervals from Equation (2) using outcomes measured in 2024. The running variable is defined as the maximum 2022 medical debt collection amount across the consumer's accounts, measured relative to the $500 reporting cutoff. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Control Mean reports the mean of the dependent variable for consumers whose maximum medical debt in collections in 2022 falls between $500 and $600. The reported optimal bandwidth corresponds to the symmetric half-width (in dollars) around the $500 cutoff used to estimate each specification. "In-Bandwidth" observations report the number of observations falling within this window, while "Total" observations report the full estimation sample. A */** denotes statistical significance at the 5% and 1% levels respectively, using robust inference.

Table A.7: Falsification Test: RD Estimates of the Effect of Medical Debt Collections Deletion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of Debts	Number of Medical Debts	Credit Score	Total Credit Balance ($1,000)	Number of Accounts Opened in Last 6 Months	Revolving Utilization (%)	Total Balance 90+ Days Past Due ($1,000)	Bankruptcy in Last 7 Years (%)	Has Alternative Credit Balance (%)	Number of Alternative Credit Accounts
ABOVE^{2022}	0.080 [–0.072, 0.19]	0.028 [–0.070, 0.12]	–2.1 [–6.8, 2.7]	–0.45 [–6.2, 4.0]	0.024 [–0.031, 0.081]	–0.97 [–3.5, 1.2]	0.15 [–0.47, 0.70]	–0.12 [–1.1, 0.78]	–0.45 [–1.3, 0.18]	0.0063 [–0.022, 0.031]
Control Mean	3.6	2.4	618	45	0.59	51	1.3	3.7	3.4	0.15
% of Mean	2.3	1.1	–0.34	–1.0	4.1	–1.9	12	–3.2	–13	4.2
Optimal Bandwidth	± 191.76	± 142.95	± 123.89	± 111.18	± 145.26	± 187.80	± 251.51	± 149.71	± 185.96	± 265.11
Observations										
Total	272,490	272,490	272,490	272,490	272,490	167,789	272,490	272,490	272,490	272,490
In-Bandwidth	52,018	37,300	31,971	28,940	38,251	30,755	73,456	39,240	50,090	78,540

Notes: This table presents the estimated coefficient ($-\beta$) and 95% confidence intervals from Equation (2) using outcomes measured in 2022. Outcome variables are defined in Table A.4. The running variable is defined as the maximum 2022 medical debt collection amount across the consumer's accounts, measured relative to the $500 reporting cutoff. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Control Mean reports the mean of the dependent variable for consumers whose maximum medical debt in collections in 2022 falls between $500 and $600. The reported optimal bandwidth corresponds to the symmetric half-width (in dollars) around the $500 cutoff used to estimate each specification. "In-Bandwidth" observations report the number of observations falling within this window, while "Total" observations report the full estimation sample. A */** denotes statistical significance at the 5% and 1% levels respectively, using robust inference.

Table A.8: Placebo Test: RD Estimates of the Effect of Medical Debt Collections Deletion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of Debts	Number of Medical Debts	Credit Score	Total Credit Balance ($1,000)	Number of Accounts Opened in Last 6 Months	Revolving Utilization (%)	Total Balance 90+ Days Past Due ($1,000)	Bankruptcy in Last 7 Years (%)	Has Alternative Credit Balance (%)	Number of Alternative Credit Accounts
ABOVE^{2020}	–0.034	–0.065	–1.2	0.56	0.033	1.4	0.33	0.25	–0.13	0.00067
	[–0.19, 0.072]	[–0.17, 0.0017]	[–5.2, 3.7]	[–3.9, 6.1]	[–0.023, 0.077]	[–0.45, 3.6]	[–0.16, 0.86]	[–0.58, 1.3]	[–0.85, 0.39]	[–0.027, 0.024]
Control Mean	2.7	1.6	622	48	0.63	48	1.2	4.2	3.5	0.14
% of Mean	–1.3	–4.0	–0.20	1.2	5.2	3.0	28	6.1	–3.9	0.46
Optimal Bandwidth	± 154.04	± 144.05	± 112.34	± 110.47	± 163.58	± 189.60	± 264.02	± 134.30	± 246.53	± 273.15
Observations										
Total	324,012	324,012	324,012	324,012	324,012	198,456	324,012	324,012	324,012	324,012
In-Bandwidth	50,892	46,832	36,090	35,484	54,034	38,989	94,851	43,331	86,240	98,672

Notes: This table presents the estimated coefficient ($-\beta$) and 95% confidence intervals from Equation (2) using outcomes measured in 2022. Outcome variables are defined in Table A.4. The running variable is defined as the maximum 2020 medical debt collection amount across the consumer's accounts, measured relative to the $500 reporting cutoff. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Control Mean reports the mean of the dependent variable for consumers whose maximum medical debt in collections in 2020 falls between $500 and $600. The reported optimal bandwidth corresponds to the symmetric half-width (in dollars) around the $500 cutoff used to estimate each specification. "In-Bandwidth" observations report the number of observations falling within this window, while "Total" observations report the full estimation sample. A */** denotes statistical significance at the 5% and 1% levels respectively, using robust inference.

Table A.9: Medical Collections Sub-Sample: RD Estimates of the Effect of Medical Debt Collections Deletion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of Debts	Number of Medical Debts	Credit Score	Total Credit Balance ($1,000)	Number of Accounts Opened in Last 6 Months	Revolving Utilization (%)	Total Balance 90+ Days Past Due ($1,000)	Bankruptcy in Last 7 Years (%)	Has Alternative Credit Balance (%)	Number of Alternative Credit Accounts
ABOVE^{2022}	–0.26**	–0.28**	2.4	–3.5	0.0050	–2.8*	–0.14	0.58	–0.0059	–0.010
	[–0.32, –0.20]	[–0.31, –0.25]	[–2.7, 9.3]	[–13, 3.9]	[–0.047, 0.065]	[–5.9, –0.69]	[–1.1, 0.52]	[–0.38, 1.9]	[–0.85, 0.86]	[–0.042, 0.022]
Control Mean	0.76	0.44	650	65	0.49	48	1.7	2.7	2.8	0.13
% of Mean	–34	–65	0.37	–5.4	1.0	–5.9	–8.4	21	–0.21	–7.8
Optimal Bandwidth	± 240.00	± 305.14	± 153.25	± 126.92	± 189.39	± 194.96	± 226.53	± 142.14	± 284.91	± 278.22
Observations										
Total	150,244	150,244	150,244	150,244	150,244	108,467	150,244	150,244	150,244	150,244
In-Bandwidth	36,420	51,558	21,393	17,194	27,159	19,849	33,966	19,601	46,206	44,905

Notes: This table presents the estimated coefficient ($-\beta$) and 95% confidence intervals from Equation (2) using outcomes measured in 2024. Outcome variables are defined in Table A.4. The medical collections sub-sample restricts the RD sample to consumers for whom the total number of accounts in collections is the same as the total number of medical debt collection accounts in 2022. The running variable is defined as the maximum 2022 medical debt collection amount across the consumer's accounts, measured relative to the $500 reporting cutoff. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Control Mean reports the mean of the dependent variable for consumers whose maximum medical debt in collections in 2022 falls between $500 and $600. The reported optimal bandwidth corresponds to the symmetric half-width (in dollars) around the $500 cutoff used to estimate each specification. "In-Bandwidth" observations report the number of observations falling within this window, while "Total" observations report the full estimation sample. A */** denotes statistical significance at the 5% and 1% levels respectively, using robust inference.

Table A.10: Same-Year RD Estimates of the Effect of Medical Debt Collections Deletion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of Debts	Number of Medical Debts	Credit Score	Total Credit Balance ($1,000)	Number of Accounts Opened in Last 6 Months	Revolving Utilization (%)	Total Balance 90+ Days Past Due ($1,000)	Bankruptcy in Last 7 Years (%)	Has Alternative Credit Balance (%)	Number of Alternative Credit Accounts
ABOVE^{2022}	–0.34**	–0.38**	–0.22	–2.3	0.031	–0.66	–0.087	0.048	0.22	0.0025
	[–0.46, –0.23]	[–0.46, –0.32]	[–4.6, 5.0]	[–8.7, 2.4]	[–0.014, 0.082]	[–3.1, 1.3]	[–0.88, 0.53]	[–0.79, 1.0]	[–0.43, 0.93]	[–0.028, 0.029]
Control Mean	2.0	1.0	617	48	0.52	52	1.7	3.3	3.4	0.16
% of Mean	–17	–37	–0.035	–4.8	6.0	–1.3	–5.2	1.4	6.5	1.6
Optimal Bandwidth	± 138.56	± 183.35	± 123.42	± 99.84	± 154.52	± 212.03	± 302.99	± 151.27	± 263.63	± 249.39
Observations										
Total	272,490	272,490	272,490	272,490	272,490	169,517	272,490	272,490	272,490	272,490
In-Bandwidth	36,117	49,494	31,971	25,316	40,982	36,339	93,904	40,146	77,793	71,597

Notes: This table presents the estimated coefficient ($-\beta$) and 95% confidence intervals from Equation (2) using outcomes measured in 2023. Outcome variables are defined in Table A.4. The running variable is defined as the maximum 2022 medical debt collection amount across the consumer's accounts, measured relative to the $500 reporting cutoff. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Control Mean reports the mean of the dependent variable for consumers whose maximum medical debt in collections in 2022 falls between $500 and $600. The reported optimal bandwidth corresponds to the symmetric half-width (in dollars) around the $500 cutoff used to estimate each specification. "In-Bandwidth" observations report the number of observations falling within this window, while "Total" observations report the full estimation sample. A */** denotes statistical significance at the 5% and 1% levels respectively, using robust inference.

Table A.11: Longer-run (Two-Year) RD Estimates of the Effect of Medical Debt Collections Deletion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of Debts	Number of Medical Debts	Credit Score	Total Credit Balance ($1,000)	Number of Accounts Opened in Last 6 Months	Revolving Utilization (%)	Total Balance 90+ Days Past Due ($1,000)	Bankruptcy in Last 7 Years (%)	Has Alternative Credit Balance (%)	Number of Alternative Credit Accounts
ABOVE^{2022}	–0.11*	–0.16**	0.63	–0.60	0.023	0.82	0.26	–0.51	–0.033	0.0087
	[–0.22, –0.025]	[–0.19, –0.14]	[–3.9, 5.8]	[–6.8, 4.7]	[–0.018, 0.075]	[–1.3, 2.6]	[–0.52, 1.0]	[–1.3, 0.34]	[–0.75, 0.75]	[–0.023, 0.046]
Control Mean	1.4	0.31	619	50	0.53	51	3.6	3.1	4.2	0.21
% of Mean	–8.1	–50	0.10	–1.2	4.4	1.6	7.4	–16	–0.79	4.2
Optimal Bandwidth	± 155.27	± 282.15	± 130.23	± 112.89	± 175.95	± 278.96	± 260.76	± 152.40	± 279.86	± 253.92
Observations										
Total	264,753	264,753	264,753	264,753	264,753	166,189	264,753	264,753	264,753	264,753
In-Bandwidth	40,136	82,561	33,088	28,430	45,729	50,477	74,589	39,256	81,390	72,016

Notes: This table presents the estimated coefficient ($-\beta$) and 95% confidence intervals from Equation (2) using outcomes measured in 2025. Outcome variables are defined in Table A.4. The running variable is defined as the maximum 2022 medical debt collection amount across the consumer's accounts, measured relative to the $500 reporting cutoff. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Control Mean reports the mean of the dependent variable for consumers whose maximum medical debt in collections in 2022 falls between $500 and $600. The reported optimal bandwidth corresponds to the symmetric half-width (in dollars) around the $500 cutoff used to estimate each specification. "In-Bandwidth" observations report the number of observations falling within this window, while "Total" observations report the full estimation sample. A */** denotes statistical significance at the 5% and 1% levels respectively, using robust inference.

Table A.12: RD Estimates of the Effect of Medical Debt Collections Deletion Excluding Observations Just Above the Cutoff

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of Debts	Number of Medical Debts	Credit Score	Total Credit Balance ($1,000)	Number of Accounts Opened in Last 6 Months	Revolving Utilization (%)	Total Balance 90+ Days Past Due ($1,000)	Bankruptcy in Last 7 Years (%)	Has Alternative Credit Balance (%)	Number of Alternative Credit Accounts
ABOVE^{2022}	–0.23** [–0.35, –0.14]	–0.29** [–0.32, –0.26]	–1.8 [–7.1, 4.0]	–2.3 [–9.6, 3.8]	0.025 [–0.020, 0.080]	–0.46 [–2.8, 1.6]	–0.062 [–0.80, 0.59]	–0.53 [–1.5, 0.49]	0.46 [–0.26, 1.4]	0.0049 [–0.025, 0.042]
Control Mean	1.5	0.49	618	50	0.50	53	2.0	3.3	3.9	0.18
% of Mean	–16	–60	–0.30	–4.5	4.9	–0.87	–3.2	–16	12	2.7
Optimal Bandwidth	± 139.78	± 264.71	± 124.50	± 105.45	± 168.98	± 241.64	± 243.11	± 149.61	± 213.68	± 233.95
Observations										
Total	271,180	271,180	271,180	271,180	271,180	168,801	271,180	271,180	271,180	271,180
In-Bandwidth	35,091	76,828	30,953	25,974	43,626	41,801	68,220	37,930	57,848	64,703

Notes: This table presents the estimated coefficient ($-\beta$) and 95% confidence intervals from Equation (2) using outcomes measured in 2024. Outcome variables are defined in Table A.4. The running variable is defined as the maximum 2022 medical debt collection amount across the consumer's accounts, measured relative to the $500 reporting cutoff. Consumers whose maximum medical debt collections in 2022 falls between $500 and $505 have been excluded from the estimation sample to address local bunching above the cutoff. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Control Mean reports the mean of the dependent variable for consumers whose maximum medical debt in collections in 2022 falls between $506 and $600. The reported optimal bandwidth corresponds to the symmetric half-width (in dollars) around the $500 cutoff used to estimate each specification. "In-Bandwidth" observations report the number of observations falling within this window, while "Total" observations report the full estimation sample applying the donut exclusion. A */** denotes statistical significance at the 5% and 1% levels respectively, using robust inference.

Table A.13: Difference-in-Discontinuities Estimates of the Effect of Medical Debt Collections Deletion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Number of Debts	Number of Medical Debts	Has Credit Score (%)	Credit Score	Total Credit Balance ($1,000)	Number of Accounts Opened in Last 6 Months	Revolving Utilization (%)	Total Balance 90+ Days Past Due ($1,000)	Bankruptcy in Last 7 Years (%)	Has Alternative Credit Balance (%)	Number of Alternative Credit Accounts
$\text{ABOVE}^{2022} \times \text{D}^{2024}$	–0.36**	–0.47**	0.0036	2.1	–0.82	–0.0018	0.21	–0.15	–0.31	0.030	–0.0028
	[–0.47, –0.25]	[–0.52, –0.42]	[–0.0014, 0.0086]	[–0.99, 5.3]	[–3.4, 1.7]	[–0.043, 0.040]	[–1.7, 2.1]	[–0.68, 0.39]	[–0.83, 0.21]	[–0.55, 0.61]	[–0.021, 0.016]
Control Mean	1.4	0.54	0.93	617	45	0.46	53	1.9	2.9	3.6	0.16
% of Mean	–25	–86	0.38	0.35	–1.8	–0.39	0.41	–7.6	–11	0.83	–1.7
Optimal Bandwidth	± 176.41	± 410.00	± 388.51	± 119.73	± 123.74	± 230.20	± 247.87	± 322.12	± 189.05	± 400.61	± 384.62
Observations	296,311	296,311	296,311	281,530	296,311	296,311	172,455	296,311	296,311	296,311	296,311

Notes: This table presents the estimated coefficient ($-\theta$) and 95% confidence intervals from Equation (5) using outcomes measured in 2024. Outcome variables are defined in Table A.4. The running variable is defined as the maximum 2022 medical debt collection amount across the consumer's accounts, measured relative to the $500 reporting cutoff. We report MSE-optimal estimates with 95% confidence intervals in brackets, standard errors are clustered at the consumer level. Control Mean reports the mean of the dependent variable for consumers whose maximum medical debt in collections in 2022 falls between $500 and $600. The reported optimal bandwidth corresponds to the symmetric half-width (in dollars) around the $500 cutoff used to estimate each specification. A */** denotes statistical significance at the 5% and 1% levels respectively, using conventional inference.

Table A.14: RD Estimates of the Effect of Medical Debt Collections Deletion: Excluding New York and Colorado

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of Debts	Number of Medical Debts	Credit Score	Total Credit Balance ($1,000)	Number of Accounts Opened in Last 6 Months	Revolving Utilization (%)	Total Balance 90+ Days Past Due ($1,000)	Bankruptcy in Last 7 Years (%)	Has Alternative Credit Balance (%)	Number of Alternative Credit Accounts
ABOVE^{2022}	–0.25**	–0.31**	1.2	–2.3	0.033	–0.67	–0.065	–0.44	0.34	0.0079
	[–0.35, –0.17]	[–0.34, –0.28]	[–3.1, 6.7]	[–8.9, 2.5]	[–0.0080, 0.084]	[–3.0, 1.2]	[–0.80, 0.54]	[–1.3, 0.45]	[–0.36, 1.3]	[–0.019, 0.043]
Control Mean	1.5	0.51	618	50	0.50	53	1.9	3.2	4.0	0.18
% of Mean	–17	–62	0.20	–4.7	6.5	–1.3	–3.4	–14	8.5	4.4
Optimal Bandwidth	± 143.40	± 293.04	± 120.16	± 101.80	± 171.62	± 237.91	± 238.64	± 154.64	± 227.75	± 235.60
Observations										
Total	260,653	260,653	260,653	260,653	260,653	161,762	260,653	260,653	260,653	260,653
In-Bandwidth	35,901	85,222	29,857	25,102	43,826	39,747	64,752	39,176	61,145	63,853

Notes: This table presents the estimated coefficient ($-\beta$) and 95% confidence intervals from Equation (2) using outcomes measured in 2024. Outcome variables are defined in Table A.4. This subsample restricts the RD analysis to consumers residing outside Colorado and New York. The running variable is defined as the maximum 2022 medical debt collection amount across the consumer's accounts, measured relative to the $500 reporting cutoff. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Control Mean reports the mean of the dependent variable for consumers whose maximum medical debt in collections in 2022 falls between $500 and $600. The reported optimal bandwidth corresponds to the symmetric half-width (in dollars) around the $500 cutoff used to estimate each specification. "In-Bandwidth" observations report the number of observations falling within this window, while "Total" observations report the full estimation sample. A */** denotes statistical significance at the 5% and 1% levels respectively, using robust inference.

Table A.15: Association Between Number of Medical Debts and Two-Year Default Rates

	(1)	(2)	(3)	(4)	(5)
Number of medical debts < \$500					
1	0.145**	0.109**	0.090**	0.029**	0.010**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
2	0.158**	0.120**	0.099**	0.038**	0.012**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
3+	0.147**	0.110**	0.091**	0.020**	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number of medical debts $\geq$ \$500					
1	0.126**	0.101**	0.076**	0.040**	0.019**
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
2	0.116**	0.096**	0.067**	0.045**	0.024**
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
3+	0.096**	0.084**	0.049**	0.017**	0.022**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Control Variables Included					
Num Accounts 90+ Days Past-Due (24 Months)		X	X	X	X
Percent Accounts 90+ Days Past-Due			X	X	X
All Credit-Report Controls				X	X
Flexible Controls (Polynomial Terms)					X
In-Sample Performance Metrics					
Accuracy	0.869	0.873	0.873	0.890	0.899
AUC	0.606	0.753	0.780	0.879	0.888
F1 Score	-	0.153	0.176	0.379	0.483
Precision	-	0.604	0.577	0.725	0.726
Recall (1 − False Negative Rate)	0	0.088	0.104	0.256	0.362
False Positive Rate	0	0.009	0.011	0.015	0.021
R-squared	0.038	0.130	0.148	0.289	0.337
Mean outcome	0.131	0.131	0.131	0.131	0.131
Observations	2,482,116	2,482,116	2,482,116	2,482,116	2,482,116

Notes: This table reports estimates from a linear probability model relating the number of medical debt collections in 2019 to default occurring between 2020 and 2021. Medical debt collections are separated into balances below \$500 and balances at or above \$500; for each group, the omitted category is zero collections. Column (1) includes no control variables. Column (2) controls for the number of accounts 90+ days past due in the last 24 months. Column (3) additionally controls for the percentage of accounts 90+ days past due. Column (4) includes the complete set of credit-report controls listed in Figure 5. Column (5) augments this specification with a fifth-order polynomial in credit-report variables. For the purpose of computing classification metrics, observations are classified as predicted defaults if the fitted value from the linear probability model is at least 0.5. A */** denotes statistical significance at the 5% and 1% levels, respectively. Standard errors are robust to heteroskedasticity. Definitions of the performance metrics are provided in Appendix C.4.

Table A.16: Association Between Maximum Medical Debt Amount and Two-Year Default Rates

	(1)	(2)	(3)	(4)	(5)
Medical debt $< \$250$	0.182**	0.139**	0.116**	0.045**	0.020**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Medical debt $\$250 - \$500$	0.201**	0.155**	0.125**	0.053**	0.024**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Medical debt $\$500 - \$1,000$	0.213**	0.168**	0.131**	0.061**	0.031**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Medical debt $\geq \$1,000$	0.208**	0.165**	0.124**	0.052**	0.026**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Control Variables Included					
Num Accounts 90+ Days Past-Due (24 Months)		X	X	X	X
Percent Accounts 90+ Days Past-Due			X	X	X
All Credit-Report Controls				X	X
Flexible Controls (Polynomial Terms)					X
In-Sample Performance Metrics					
Accuracy	0.869	0.873	0.873	0.890	0.899
AUC	0.602	0.751	0.779	0.879	0.888
F1 Score	-	0.153	0.174	0.380	0.484
Precision	-	0.605	0.586	0.726	0.726
Recall (1 − False Negative Rate)	0	0.087	0.102	0.257	0.362
False Positive Rate	0	0.009	0.011	0.015	0.021
R-squared	0.041	0.131	0.149	0.290	0.337
Mean outcome	0.131	0.131	0.131	0.131	0.131
Observations	2,482,116	2,482,116	2,482,116	2,482,116	2,482,116

Notes: This table reports linear probability model estimates of the association between consumers' maximum medical debt collection balance and default within two years. Medical debt is categorized into mutually exclusive balance bins. The omitted category is consumers with no medical debt collections. Column (1) includes no control variables. Column (2) controls for the number of accounts 90+ days past due in the last 24 months. Column (3) additionally controls for the percentage of accounts 90+ days past due. Column (4) includes the complete set of credit-report controls listed in Figure 5. Column (5) augments this specification with a fifth-order polynomial in credit-report variables. For the purpose of computing classification metrics, observations are classified as predicted defaults if the fitted value from the linear probability model is at least 0.5. A */** denotes statistical significance at the 5% and 1% levels, respectively. Standard errors are robust to heteroskedasticity. Definitions of the performance metrics are provided in Appendix C.4.

Table A.17: Performance Metrics for Credit Scoring Models With and Without Medical Debt Collections Versus a Random Variable

	(1)	(2)	(3)
	All Predictors + Noise	Exclude Noise (Baseline)	Exclude Medical Debts < \$500
Accuracy	0.907	0.907	0.907
AUC	0.905	0.905	0.905
F1 Score	0.561	0.561	0.561
Precision	0.735	0.735	0.737
Recall	0.454	0.454	0.453
False Positive Rate	0.0247	0.0247	0.0243

Notes: This table reports out-of-sample performance metrics for a credit scoring model predicting defaults occurring between 2020 and 2021, using borrower characteristics from 2019. Column (1) reports metrics for a model that adds a random variable to the baseline set of 48 predictors, estimated using XGBoost. Column (2) presents metrics for the baseline model, with 48 predictors. Column (3) reports metrics when small (under \$500) medical collections are excluded from the baseline set of predictors. The sample consists of 2,482,116 observations, with 90% used for model training and the remaining 10% reserved for out-of-sample performance evaluation. For reference, a naive model that predicts no defaults achieves an accuracy of 0.869, equal to one minus the average default rate (0.131). Definitions of the performance metrics are provided in Appendix C.4.

Table A.18: Performance Metrics for Credit Scoring Models With and Without Medical Debt Collections: Logit Model

	(1)	(2)	(3)
	All Predictors	Exclude Medical Debts < $500	Exclude All Medical Debts
Accuracy	0.899	0.899	0.898
AUC	0.892	0.892	0.892
F1 Score	0.495	0.495	0.494
Precision	0.723	0.723	0.722
Recall	0.376	0.376	0.376
False Positive Rate	0.0217	0.0217	0.0218

Notes: This table reports out-of-sample performance metrics for a credit scoring model predicting defaults occurring between 2020 and 2021, using borrower characteristics from 2019. Column (1) presents metrics for the model including 48 predictors and estimated using logistic regression. Column (2) reports metrics when small (under $500) medical collections are excluded from the predictors. Column (3) shows metrics when all medical collections are excluded. The sample consists of 2,482,116 observations, with 90% used for model training and the remaining 10% reserved for out-of-sample performance evaluation. For reference, a naive model that predicts no defaults achieves an accuracy of 0.869, equal to one minus the average default rate (0.131). Definitions of the performance metrics are provided in Appendix C.4.