The Effects of Deleting Medical Debt from Consumer Credit Reports*

Victor Duarte
University of Illinois

Julia Fonseca
University of Illinois

Divij Kohli
University of Illinois

Julian Reif
University of Illinois and NBER

March 2025

Abstract

One in seven Americans carry medical debt, with $88 billion reported on consumer credit reports. In April 2023, the three major credit bureaus stopped reporting medical debts below $500. We study the effects of this information deletion on consumer credit access and financial health. Using a machine learning model, we show that small medical debts are not meaningfully predictive of defaults, suggesting their deletion should have minimal effect on lending decisions. We test this prediction using two complementary research designs. First, a regression discontinuity analysis comparing individuals just above and below the $500 threshold finds no direct benefits from the information deletion, ruling out small changes in credit access and financial health. Second, to assess potential indirect effects, we classify consumers based on whether their predicted probability of default increases or decreases when medical debts are deleted. A difference-in-differences analysis comparing these groups before and after the 2023 policy change reveals no evidence of negative spillover effects. Finally, we show that larger medical debts ($\geq$ $500) are also not meaningfully predictive of default, suggesting that eliminating medical debts entirely from credit reports, as planned under a January 2025 decision by the Consumer Financial Protection Bureau, is unlikely to affect credit outcomes.

*We thank Ray Kluender, Neale Mahoney, and seminar participants at the University of Illinois and Georgetown for helpful comments and discussions.

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 $88 billion in medical debt appearing on consumer credit reports as of 2021 (Consumer Financial Protection Bureau, 2022). Policymakers have raised growing concerns about these reporting practices, arguing that making medical debt visible to lenders could impair access to credit following unexpected medical shocks. 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 measure, the Consumer Financial Protection Bureau (CFPB) issued a final rule in January 2025 to eliminate all remaining medical debt collections from credit reports, arguing that this change would enhance credit access and improve loan terms for consumers burdened with medical debt.[1] To address concerns that deleting this information could harm consumers without medical debt, the CFPB cited evidence that medical debt collections are poor predictors of default and thus unlikely to negatively affect lending decisions (Consumer Financial Protection Bureau, 2024). However, research in other credit contexts finds that information deletion can sometimes reduce borrowing opportunities for certain groups (Liberman et al., 2019). Moreover, if medical debt truly lacks predictive power, its removal from credit reports may not improve credit access, even for those burdened by it.

Against this backdrop, this paper studies the direct and indirect effects of deleting information about medical debt collections from credit reports, using 2019–2024 data from the Gies Consumer and Small Business Credit Panel (GCCP). Specifically, we investigate whether consumers whose medical debt information was deleted experienced any direct benefits and whether this deletion produced negative spillovers for other consumers reclassified as higher risk. Using machine learning techniques, we build credit scoring models to evaluate the predictive value of medical debt collections for default risk prior to 2023. We compare two models: one trained on borrower credit histories including medical debts below $500,

[1] This rule was set to go into effect 60 days after publication in the Federal Register but implementation was delayed until June 15, 2025 by the U.S. District Court for the Eastern District of Texas. For the announcement of the rule, see https://www.consumerfinance.gov/about-us/newsroom/cfpb-finalizes-rule-to-remove-medical-bills-from-credit-reports/.

and another trained on histories excluding these small medical debts. We find that excluding small medical debts has no meaningful effect on default prediction, highlighting their minimal value for lending decisions. Furthermore, we show that larger medical debts ($\geq$ \$500) also have little predictive value, suggesting that eliminating all medical debts from credit reports is similarly unlikely to influence credit access or financial health.

We test this prediction using two distinct research designs. First, we employ a regression discontinuity (RD) approach to estimate the direct effects of the 2023 deletion of small medical debts from credit reports. Comparing individuals just above and below the \$500 threshold, we find no evidence that deleting this information improved credit access or financial health. Our null estimates are precise: the 95% confidence intervals rule out increases in credit scores greater than 6.03 points (0.97%) and decreases in the balance-to-limit ratio of revolving credit exceeding 1.54 percentage points (4.80%).

Next, we use a difference-in-differences approach to estimate the indirect effects of removing small medical debt collections from credit reports on consumers who, as a result, are re-classified as higher default risk. We use our two credit default prediction models—one incorporating medical debts below \$500 and the other excluding them—to identify two groups: consumers whose predicted probability of default increases by at least 2 percentage points (the 95th percentile of the distribution) when medical debts are removed from the model, and those whose predicted probability falls by at least 2 percentage points. We show that these two groups are observationally similar across key characteristics and exhibited parallel trends prior to the 2023 information deletion.[2] Consistent with our RD results, we find no evidence of negative spillover effects from deleting small medical debts, with precise estimates that again rule out small effects.

Overall, we conclude that the 2023 decision to delete small medical debt collections from credit reports offered no measurable benefits to affected consumers and had no negative indirect effects on those reclassified as higher risk. These findings suggest that policies promoting information deletion may fail to achieve their intended goals, particularly for

[2]Both groups consist primarily of low-income consumers with thin credit files. As shown in the main text, when reliable information is scarce, even random noise can influence default predictions. Thus, an uninformative predictor like medical debt serves as a noisy partitioning mechanism, effectively creating two randomly assigned groups.

consumers burdened by medical debt. Thus, our findings underscore the importance of exploring alternative strategies to support financially vulnerable consumers more effectively.

We make several contributions to the literature. First, we investigate the effects of information deletion in a new context: medical debt. Prior studies in other settings have found mixed effects. Liberman et al. (2019) study the deletion of credit default information in Chile and, like us, use credit scoring models to assess how changes in predicted probabilities affect borrowing behavior. They find that deletion increases borrowing for consumers whose predicted default risk declines, but reduces borrowing for those reclassified as higher-risk. Similarly, Jansen et al. (2024) find that removing bankruptcy flags lowers interest rates for affected consumers while raising them for those with no history of bankruptcy, resulting in a small decline in social surplus.[3] Beyond credit markets, Agan and Starr (2017) find that removing criminal history information from job applications reduces callbacks for Black applicants, and Bartik and Nelson (2024) show that bans on employers' use of credit reports lower job-finding rates and increase involuntary separations for Black workers. Unlike these studies, we show that the deleted information in our setting—small medical debts—has minimal predictive value, resulting in neither direct nor indirect effects. A unique feature of our setting is the presence of a cutoff value for information deletion, which allows us to estimate direct effects using a rigorous RD design.

Second, we contribute to the literature on medical debt forgiveness, a policy often discussed alongside information deletion. Kluender et al. (2024) conduct two large-scale randomized experiments on medical debt forgiveness and find that while forgiveness modestly improves credit access for consumers whose medical debt collections were reported to credit bureaus, it has no effect on other consumers. Their results suggest that reporting medical debt to credit bureaus could play a key role in determining the impact of relief. However, our analysis—which is focused specifically on medical debt reported on credit reports—fails to detect any meaningful effects on credit access. Taken together, these findings suggest that neither deleting medical debt information from credit reports nor forgiving the debt itself alleviates financial distress for people burdened by it. These results contrast with evidence

[3]For studies examining the direct (but not indirect) effects of removing bankruptcy flags and other unpaid debts from credit reports, see Musto (2004), Bos et al. (2018), Dobbie et al. (2020), Gross et al. (2020), and Herkenhoff et al. (2021).

from other debt relief contexts, which generally find positive results (Dobbie and Song, 2015; Di Maggio et al., 2020; Cespedes et al., 2025).

Third, we advance the literature on using machine learning to analyze credit markets. We are the first to demonstrate that medical debt collections are poor predictors of default risk. However, we also show that even unreliable information like medical debt can still influence credit scores, particularly for consumers with thin credit files, a phenomenon previously highlighted by Blattner and Nelson (2022). Prior studies 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 substantially better performance than prior studies across multiple 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; Batty et al., 2022; Guttman-Kenney et al., 2022; Keys et al., 2022; Fonseca, 2023; Lin, 2024). The study most closely related to ours is Batty et al. (2022), who show that expanding health insurance coverage reduces medical debts in collection, but does not improve other financial outcomes. 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 describes the data used in our analysis. Section 3 investigates whether medical debt is predictive of default. Section 4 presents RD estimates of the direct effect of deleting medical debt. Section 5 estimates the indirect effect using a differences-in-differences analysis. Section 6 concludes.

2 Data

Our study uses the Gies Consumer and Small Business Credit Panel (GCCP), a panel dataset of anonymized credit record data 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.[4] Covering the years 2004–2024, the dataset provides annual snapshots of credit records taken at the end of each year's first quarter. Consumers are randomly sampled based on the last two digits of their 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 dataset accurately reflects the evolving dynamics of the underlying population.

The GCCP provides detailed debt information at the credit account ("tradeline") level, including outstanding balances and payment histories for mortgages, student loans, and credit cards. It also includes individuals' VantageScore credit scores, along with records of bankruptcies, judgments, and other public records. Additionally, the dataset include information on the individuals' 5-digit zip codes of residence, age, and gender. For a subset of observations, further demographic information is available, including education level, marital status, homeownership status, and occupation. We identify medical debt collections as those where the creditor is categorized as Medical/Health Care or the furnisher is identified as a business within the medical or health-related sector.[5] In Table A.2, we benchmark the share of people with medical collections according to our classification against other sources.

We restrict the sample to the years 2019–2024 and to consumers aged 18 or older. We exclude people with missing data on age, credit score, or income, as well as those whose reported age increases by 10 years or more within a 12-month period. This filtering results

[4] 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.

[5] These category labels are Dentists, Chiropractors, Doctors, Medical group, Hospitals and clinics, Osteopaths, Pharmacies and drugstore, Optometrists and optical outlets, and Medical and related health-nonspecific.

in a final sample of 15,313,700 observations, summarized in Table 1. The first three columns present statistics for the full sample. Approximately half of the sample is female. On average, consumers have a credit score of 702, an annual income of $51,960, and a total balance of $76,460 across all credit products. Approximately 20% of consumers have an alternative credit record, and the average number of medical collections is 0.25.

The next three columns of Table 1 present statistics for individuals with medical collections. Consumers with medical collections have, on average, lower credit scores, lower income, and lower balances compared to the full sample. They are also more likely to have subprime credit records. They have an average of 2.44 medical collections, of which 1.45 are for amounts below $500.

2.1 Regression Discontinuity Sample

For our RD analysis, we further restrict the sample to consumers with at least one medical collection in 2022 and a non-missing credit score between 2022 and 2024. The resulting sample includes 271,305 consumers, totaling 813,915 observations across the three years. Table 2 presents summary statistics for this sample as of 2022, the year prior to the deletion of medical collections under $500. Consumers in this sample have an average of 3.53 accounts in collections in 2022, of which 2.37 are medical collections and 1.56 are medical collections below $500.

3 Do Medical Collections Predict Default?

3.1 Background on medical collections

Medical debt arises when patients are unable to pay the out-of-pocket portions of their medical bills. Typically, healthcare providers first attempt to recover unpaid amounts directly from patients. If these efforts are unsuccessful, the provider may hire a third-party collection agency, which employs various strategies to secure payment. These strategies include initiating lawsuits to obtain court judgments for repayment and reporting unpaid debts to credit bureaus.[6] In some cases, medical debts are sold to debt buyers who continue recovery efforts. To protect consumers, the Fair Debt Collection Practices Act prohibits abusive or deceptive practices by third-party debt collectors. Importantly, as of April 2023, debt collectors can no longer report medical debts under $500 to credit bureaus, thereby reducing their leverage in such cases.

The consequences of medical debt are complex and challenging to quantify, partly because payment rates are exceedingly low—medical debt can be purchased for pennies on the dollar (Kluender et al., 2024). This stands in contrast to other forms of unsecured debt, such as student loans and credit card debt. Student debt is not automatically dischargeable in bankruptcy; eliminating it requires proving undue hardship, a stringent legal standard. Credit card debt also has much higher repayment rates, as issuers can threaten to impair or revoke a delinquent borrower's access to credit. Additionally, some states prohibit or restrict wage garnishments related to medical debt.

Media discussions frequently highlight the relationship between medical debt and personal bankruptcies. While many bankruptcy filers do carry medical debt, this correlation does not necessarily imply causation. To address this issue, Dobkin et al. (2018) examine the impact of hospitalizations in California on the likelihood of filing for bankruptcy within four years of admission. Their analysis reveals 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 suggest that medical debt could potentially serve as a helpful predictor of future defaults. However, further research is needed to determine whether its predictive value diminishes when other credit-related variables are considered, as these may capture similar information.

3.2 Should deleting medical collections affect underwriting?

The effect of deleting medical collections hinges on their predictive power for future default. If medical collections accurately predict default risk, lenders who incorporate these data into their proprietary credit scoring models might adjust their lending decisions following

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

its removal, potentially limiting access to credit for some consumers. Conversely, if medical collections offer limited predictive value, their removal should not affect credit underwriting, even for consumers whose information is deleted.

Not all lenders use proprietary credit scoring models. Some rely exclusively on scores provided by major credit bureaus, such as VantageScore and FICO. For these lenders, the impact of deleting medical collections on their origination decisions depends on whether this deletion affects these widely used credit scores. However, VantageScore stopped including medical collections below $500 in its model in January 2023, and FICO followed suit a few months later.[7] As a result, the deletion of small medical collections is unlikely to affect the origination decisions of lenders relying solely on these bureau-provided credit scores.

For lenders that develop proprietary credit scoring models, the decision to include medical collections depends on whether these data help predict defaults. To examine this, we investigate the predictive power of medical collections by simulating the effects of the April 2023 deletion of medical collections below $500. We train two credit scoring models: one that includes data on medical collections below $500 and another that excludes it. While our models do not exactly replicate those used by any specific lender, they rely on similar data and algorithms, and they perform favorably compared to existing models in the literature. Our approach assumes well-designed credit scoring models such as ours can accurately detect the predictive power of medical collections, to the extent that they have any.

As we demonstrate below, we find that medical collections below $500 provide no meaningful predictive value beyond what other standard credit variables already offer. This finding implies that the April 2023 intervention should have no direct or indirect effects on credit access or financial health, predictions which we test in Sections 4 and 5, respectively. Finally, we show that medical collections above $500 also fail to predict default, suggesting that the CFPB's 2025 final rule to delete all remaining medical collections is unlikely to substantially affect credit access or financial health.

[7]See announcements at https://www.vantagescore.com/major-credit-score-news-vantagescore-removes-medical-debt-collection-records-from-latest-scoring-models/ and https://www.myfico.com/credit-education/blog/medical-collections-removal.

3.3 Credit scoring with and without medical collections

Credit scoring models predict whether a borrower will default based on his/her financial and credit history. Formally, these models estimate:

$$Y = f(X_1, X_2, ...X_n) + e \tag{1}$$

where Y is a credit outcome, X_i are borrower characteristics, and e captures irreducible noise. The function $f(\cdot)$ represents the mapping from borrower attributes to a predicted outcome, which may be specified parametrically or estimated flexibly using machine learning techniques.

Traditional credit scoring models, such as the FICO and the VantageScore, typically use logit models estimated using person-level data (Federal Reserve Board, 2007). These models generally aim to predict "default," defined as any account that becomes 90 days or more past due within the next 18–24 months, using observed variables relating to payment history, amounts owed, length of credit history, new credit, and credit mix (Federal Reserve Board, 2007).

Following this approach, we select $n = 46$ predictors capturing information on current accounts and balances past due, the number of medical and non-medical collections, bankruptcies and other public records, balances and accounts of different credit types, average account age, age of oldest account, and new inquiries and accounts.[8] Consistent with prior work, our model excludes variables prohibited by the Equal Credit Opportunity Act—such as sex, marital status, and age—as well as variables that may serve as proxies for these characteristics, such as geographic identifiers and income (Federal Reserve Board, 2007; Blattner and Nelson, 2022). Unlike traditional models, our baseline algorithm is XGBoost, a state-of-the-art machine learning algorithm for classification problems.

We train two person-level credit scoring models: one including the number of medical collections below \$500 and one excluding this information. Both models include information on the number of medical collections above \$500.[9] Using data from 2019 to 2021—prior to

[8]For more information on the predictors included in traditional credit scoring models, see https://www.myfico.com/credit-education/whats-in-your-credit-score.

[9]We also investigated the effect of including information on the balance amounts of medical collections.

the removal of information on medical collections below \$500—we predict the probability of a default occurring between 2020 and 2021 using borrower characteristics from 2019. The dataset contains records for over 2.8 million consumers. We allocate 90% of these observations for training and reserve 10% for out-of-sample performance evaluation. Predicted default probabilities are converted into binary predictions using a threshold of 50%.

We report model performance across various metrics in Table 3. The first column shows results for the model that includes medical collections below \$500, while the second column presents results for the model that excludes this information. In our context, the accuracy score—defined as the share of correct predictions—provides limited insight because default is a relatively rare event. For example, a model that predicts no consumers will default achieves an accuracy score of 86.69%, reflecting the share of consumers in the 2019 sample who did not default in 2020–2021. Similarly, 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, is less informative in this setting. Instead, metrics such as precision or recall, which better capture a model's ability to predict rare outcomes, are more relevant for evaluating performance (Davis and Goadrich, 2006).

Precision and recall are defined as:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{2}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{3}$$

Precision measures the proportion of predicted defaults that were correctly classified, while recall measures the proportion of actual defaults that were correctly classified. Both metrics are important in our setting: high precision minimizes the risk of incorrectly classifying creditworthy borrowers as defaulters, helping lenders avoid missed profitable opportunities, while high recall ensures that the model successfully identifies most defaulters, reducing the likelihood of extending credit to high-risk borrowers. To balance these goals, we also compute

Surprisingly, incorporating these data worsened the predictive accuracy of the model, even for balances over \$500. We therefore don't include balance information in this analysis.

the F1 score, which is the harmonic mean of precision and recall:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

Table 3 shows that our algorithm performs well in predicting default, achieving an F1 Score of 0.557. To assess this result, we focus on its components, precision and recall, which are more commonly reported in other papers. Our recall of 0.448 ranks among the highest, with prior studies typically reporting values 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 achieve recalls of 0.654 and 0.749, respectively, but over shorter prediction horizons of 3 and 6 months.[10] Shorter prediction windows generally yield higher precision and recall, which likely contributes the stronger performance of the models in Khandani et al. (2010) and Chioda et al. (2024). Moreover, Chioda et al. (2024) use a 20% threshold to classify predicted default probabilities, increasing recall and decreasing precision relative to the 50% threshold we use.

Our precision score of 0.736 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 horizon. Although precision and recall are more informative than AUC in settings with rare outcomes, our model's AUC of 0.712 further demonstrates its effectiveness, as it falls within the typical range of 0.66 to 0.88 reported in the literature.

Comparing the first two columns in Table 3 shows that removing information on medical collections below \$500 has no measurable impact on model performance. All metrics remain unchanged up to the third decimal, except for accuracy, which slightly *increases* by 0.001 when small medical collections are removed. This result provides strong evidence that medical collections below \$500 are not meaningfully predictive of default. The third column further shows that deleting information on all medical collections, including those exceeding \$500, also has no measurable impact on model performance. These results suggest that the

[10] 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.

CFPB's 2025 final rule to eliminate all remaining medical collections from credit reports is unlikely to have any meaningful effect on the predictive accuracy of credit scoring models.

To further support the conclusion that medical collections are not an important predictor of default risk, Figure 1 reports variable importance measures based on average SHAP values, ranking variables in order of predictive importance.[11] Small medical collections rank near the bottom, with an average SHAP value of 0.0011, compared to an average value of 0.278 for the 10 most important features. Additionally, Figure A.1 shows that removing information on medical collections below $500 has minimal effect on predicted default probabilities, with only approximately 10% of consumers experiencing a change greater than 2 p.p. in absolute value.[12]

Although Figure A.1 suggests that small medical collections might improve predictive performance for a subset of individuals represented in the tails of that distribution, these effects are more likely due to noise than meaningful differences in default risk. To investigate this further, we categorize consumers into three groups based on the changes in their predicted default probabilities across the two models:

Negatively treated: Consumers above the 95th percentile in the distribution of probability difference, whose predicted probability of default increases by approximately 2 p.p. or more when medical collections below $500 are removed from the model.

Positively treated: Consumers below the 5th percentile in the distribution of probability difference, whose predicted probability of default decreases by approximately 2 p.p. or more when medical collections below $500 are removed from the model.

Unaffected: Consumers between the 25th and 75th percentiles, who experience a change in predicted default probabilities (in absolute value) of approximately 0.002 p.p. or less.

[11] A feature's SHAP value quantifies its contribution to a specific model prediction, indicating how much the feature shifts the prediction from the mean. The average SHAP value reflects the feature's mean contribution across all predictions.

[12] For context, 2 p.p. 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/).

If small medical collections were truly predictive of default, we would expect clear differences between these groups. For instance, we would expect consumers with medical debt to see lower predicted default probabilities once that information was eliminated from the credit scoring model. In that case, consumers in the positively treated group should have a significantly higher number of medical collections than those in the negatively treated group.

However, comparing these groups does not support this hypothesis that small medical collections meaningfully contribute to default predictions. Table 4 shows summary statistics for all three groups. While both the positively and negatively treated groups differ significantly from the unaffected group, they are remarkably similar to each other. Figure 2 illustrates these results using balancing regressions. All variables are standardized and each dot represents the regression coefficient of the variable labeled on the y-axis, regressed on either the positive (blue) or negatively treated (red) group indicator. If small medical collections were strongly predictive of default, we would expect a pronounced sorting effect along that variable. However, Figure 2 shows no evidence of sorting; positively and negatively treated consumers are similar across a wide range of characteristics, including the presence of medical collections.

While the positively and negatively treated consumers are quite similar to each other, they differ significantly from the unaffected group, indicating that comparing predicted probabilities with and without medical collections under $500 results does introduce some non-random sorting. This sorting arises because default probability estimates for low-income consumers with thin credit files are noisy (Blattner and Nelson, 2022). Consequently, even features with little or no predictive power, such as small medical collections, may receive positive loadings and influence predictions for these consumers.[13] Excluding the feature will then distinguish between the less noisy (Unaffected) and more noisy (Positively Treated and Negatively Treated) groups, but fail to further differentiate between positively and negatively treated consumers, effectively assigning consumers to these two groups randomly.

To confirm this explanation, Section 3.4 introduces a randomly generated variable into the model and shows that excluding it produces the same pattern observed when excluding

[13] In theory, machine-learning algorithms such as XGBoost should learn to disregard variables with no predictive value during training. However, in practice, these variables can still exhibit non-zero importance due to finite sample limitations.

medical collections under \$500. This finding underscores the limited predictive value of small medical collections.

3.4 Placebo test: credit scoring with and without random noise

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

Table A.3 presents the performance metrics. Columns (1)–(2) replicate the estimates from the first two columns of Table 3, showing that removing small medical collections has no discernible impact on model performance. Column (3) shows a similar pattern when including a random variable.[14] Figure A.2 reinforces this result by overlaying the histogram of probability differences from Figure A.1 with the corresponding histogram obtained after removing the noise variable. The two distributions are very similar.

Furthermore, Figure A.3 shows that removing the random variable sorts consumers into our three groups—unaffected, positively treated, and negatively treated—in the same way as removing medical collections under \$500. Panel A reproduces the balance plot from Figure 2, where each dot represents the coefficient from a regression of the standardized variable labeled (y-axis) on either the positively treated (blue) or negatively treated (red) group indicator. Panel B presents an analogous plot, comparing models one with and without the random variable and re-sorting consumers into three groups as described in Section 3.3. We again find that "treated" consumers have, on average, lower credit scores, lower income, and lower balances, reinforcing the conclusion that default probability estimates are substantially noisier for these consumers (Blattner and Nelson, 2022).

One potential concern with this placebo test is that removing any variable—regardless of its predictive power—could fail to generate systematic differences between the positively and negatively treated groups. To address this concern, Figure A.4, examines the impact of

[14]Comparing Column (3) to Column (1), we find that removing the random variable increases the F1 Score by 0.002, suggesting a slight improvement in prediction due to reduced overfitting. In contrast, there is no change in the F1 Score when comparing Column (2) to Column (1), indicating that small medical collections may have minimal but nonzero predictive power in the credit scoring model.

removing credit history length, as measured by average account age, the age of the oldest account, and the age of the oldest account never delinquent or derogatory. Credit history length is widely used in credit scoring models as a predictor of default (Federal Reserve Board, 2007) and average account age is one of the most important features in our model as measured by SHAP values (Figure 1). We train a version of our baseline model that omits average account age and the length of the oldest account, and then apply the same methodology to sort consumers into the three groups—unaffected, positively treated, and negatively treated. Unlike the removal of small medical collections or the random variable, excluding credit history length produces significant differences between the positively and negatively treated groups: the positively treated group is younger and has a shorter credit history, as measured by average account age. This contrast underscores that our placebo test is meaningful, as removing truly predictive information leads to distinct sorting patterns, while removing noise does not.

3.5 Do medical collections predict default in the absence of better information?

The previous results suggest that medical collections are not meaningfully predictive of default. This finding may seem surprising given that hospitalizations account for 4–6% of personal bankruptcies (Dobkin et al., 2018), which suggests that medical debt should have some predictive power for future defaults. However, it is possible that other credit report variables offer a more precise signal of financial distress than medical debt alone.

To test whether medical debt predicts default in the absence of more useful information, we train a restricted version of our baseline XGBoost model using only four predictors. These predictors include the number of medical debts below and above $500, along with the two variables with even less predictive power than medical debt, as measured by the SHAP feature importance measure (Figure 1): bankruptcy trades and bankruptcy trades in the past 24 months. Thus, this restricted model relies solely on medical debt and variables with even weaker predictive power for default.

Table A.4 presents the performance metrics for the restricted model. As expected, column

(1) shows that this model performs very poorly. The accuracy of 0.8669 accuracy matches that achieved by a naive model that predicts no defaults. The model's recall score is just 0.0008, meaning that it correctly identifies only 0.08% of true defaulters. Among borrowers classified as defaulters, only 34.62% are correctly classified. The model achieves an F1 Score of 0.0016, far below our baseline of 0.557.

Column (2) reports metrics for the restricted model when we exclude medical debts below \$500. The model's performance drops by 50% according to our preferred metric—the F1 Score. In column (3), we remove all medical debts from the restricted model, leading to a further 75% drop in the F1 Score. These results suggest that medical debts have some predictive power in the absence of better predictors, but are not meaningfully predictive when more informative credit report variables are available.

4 Direct Effect: Regression Discontinuity

4.1 Empirical strategy

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

$$Y_j^{2024} = \alpha_1 DEBT_j^{2022} + \beta_1 ABOVE_j^{2022} + \gamma_1 (ABOVE_j^{2022} \times DEBT_j^{2022}) + \epsilon_j \tag{5}$$

The dependent variable, Y_j^{2024}, represents an outcome for account j measured in 2024. The running variable, $DEBT_j^{2022}$, is defined as the account's balance relative to the \$500 cutoff in 2022, the year prior to the deletion of small medical collections. The indicator variable $ABOVE_j^{2022}$ is equal to one if $DEBT_j^{2022} \geq 0$. Our model employs a local linear approximation to the unknown regression functions underlying the average causal effect at the threshold. We allow the slope of our linear approximation to vary on either side of the cutoff.

Our second-stage outcomes are measured at the consumer level. We therefore aggregate the running variable by taking the maximum debt amount across all accounts. We then estimate the following model at the consumer level:

$$Y_i^{2024} = \alpha MAXDEBT_i^{2022} + \beta ABOVE_i^{2022} + \gamma(ABOVE_i^{2022} \times MAXDEBT_i^{2022}) + \epsilon_i \quad (6)$$

The running variable, $MAXDEBT_i^{2022}$, is the balance of consumer i's largest medical collection relative to the $500 cutoff. Our focal parameter of interest is β, which we interpret as the intent-to-treat effect of having at least one account not deleted. Equivalently, we interpret $-\beta$ as the effect of having all accounts deleted. In the appendix, we estimate a model using the minimum rather than the maximum debt value across all accounts, as well as a model estimated at the account level instead of the individual level. While these models estimate different treatment effects, the results remain qualitatively similar.

Our main identifying assumption is that assignment to either side of the medical debt cutoff is effectively random. This assumption is plausible given that individuals have limited ability to precisely manipulate their medical debt, which is determined by fixed prices that are often unknown in advance. Additionally, our data come from administrative records, which minimizes concerns about sample selection bias.

The main threat to this research design is the possibility of other interventions occurring at the $500 threshold. For instance, if hospitals implement policies that restrict services to individuals once their unpaid medical bills exceed $500, then any observed effects in our RD design could reflect hospital policies rather than credit bureau reporting rules. If this concern is valid, we would expect to detect similar effects data from before 2023, when the credit bureaus revised their reporting policies. To test this possibility, we estimate additional RD specifications using pre-2023 data.

A related concern is that debt collectors may have systematically treated medical collections differently at the $500 threshold even before the 2023 policy change. For example, if debt collectors routinely refrained from reporting medical debts under $500 to credit bureaus, then any observed effects could stem from debt collector behavior rather than changes in credit bureau policies. We assess this possibility using the same approach as above: estimating RD specifications using pre-2023 data to determine whether similar discontinuities existed before the policy change.

We use a triangular kernel in all RD regressions. Our preferred specification uses a mean-squared error optimal bandwidth that remains constant on either side of the cutoff but can vary across outcomes. We report robust bias-corrected confidence intervals that accommodate 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 3 shows that by 2024, nearly all accounts with balances below \$500 in 2022 had been removed from credit reports, whereas more than 10 percent of accounts with balances above \$500 remained. Panel B demonstrates that this effect also appears at the consumer level. After aggregating the running variable by taking the maximum debt amount across all accounts, the intervention is shown to have reduced the number of medical collections per person in 2024 by 0.29 (107%).[15]

Figure 4 shows the direct effect of the 2023 deletion on credit access and utilization. We find no significant changes in the outcomes around the debt threshold. Table 5 presents formal estimates, with 95% confidence intervals ruling out improvements in credit scores greater than 6.03 points (0.97%), increases in balances greater than \$2,602 (5.05%), new credit accounts by 0.09 (17.65%), and decreases in revolving utilization by 1.54 (4.80%).[16] Figure 5 shows the direct effect of the 2023 deletion on delinquency, bankruptcy, and alternative credit use. Again, the 95% confidence intervals can rule out that deletion improved delinquency balance by \$802 (42.36%), the probability of bankruptcy by 1.19 percentage points (37.42%), the probability of having alternative credit balance by -0.29 percentage points (7.51%), and an increase in the number of alternative credit accounts by 0.04 (22.60%).

Finally, Figure A.6 presents estimates for additional credit outcomes. We fail to detect effects on number of accounts 90+ days past due, the number of new inquiries, revolving limits, total balance in alternative credit accounts, or the number of new mortgage accounts. Our

[15] A consumer whose largest medical collection account was under \$500 in 2022 but subsequently acquires new medical collection accounts exceeding \$500 in 2023 or 2024 will still be recorded as having medical collections in 2024. These newly acquired accounts account for the positive values plotted to the left of the cutoff in Panel (b).

[16] Medical collections have not been used in the VantageScore model since January 2023; there is therefore no mechanical relationship between these two variables in this analysis.

null estimates remain precise, with 95% confidence intervals ruling out meaningful changes in the outcome.

Table A.9 presents results for the subsample of consumers whose debts in collections consist solely of medical collections—a group for which Kluender et al. (2024) found modest positive effects of debt relief, including slight improvements in credit scores. However, our analysis finds no such benefits, ruling out a credit score increase of 3.07 points (0.47%) at the 95% confidence level. One possible reason for this discrepancy is that the two major credit scoring models—VantageScore and FICO—stopped using medical debt as a predictor in 2023, after the study period in Kluender et al. (2024) but before our post-period outcome measurements. Thus, while removing medical debt from credit reports may have affected credit scores in the earlier period, it no longer has a direct effect on credit scores in our study period.[17]

Under the assumptions of our RD design, outcomes unrelated to medical debt should remain unchanged at the threshold. To test this, we conduct a series of covariate smoothness tests, presented in Figure A.5. We find no significant differences at the threshold in average age, average income, or the share female.

We also conduct a series of falsification and placebo tests. Our falsification tests, shown in Figures A.7 and A.8, confirm that no effects are present at the threshold when measuring our main outcomes in 2022—the year before deletion—instead of 2024, the year after deletion. Similarly, our placebo tests, presented in Figures A.10 and A.11, replicate our RD analysis using the 2020–2022 period instead of 2022–2024. Once again, we find no significant effects, reinforcing the validity of our design.

[17]While we find that medical debt is not a significant predictor in our credit scoring model, older models may have treated it differently for two reasons. First, medical debt may have historically been a stronger predictor of default risk. Second, if earlier models used fewer variables or less sophisticated algorithms than ours, they may have assigned greater importance to medical debt.

5 Indirect Effect: Difference-in-differences

5.1 Empirical strategy

We use a differences-in-differences research design to study the indirect effects (negative spillovers) of deleting small medical collections from credit reports. To identify consumers most at risk for negative spillovers, we define the treatment group as consumers whose predicted probability of default increases when medical collections under $500 are excluded from our credit scoring model, i.e., the "negatively treated" group described in Section 3. Because this group is defined based on 2019 characteristics, we restrict the sample period for our analysis to 2020–2024.

A natural choice for a control group would be consumers classified as "unaffected."[18] However, as shown in Table 4, these consumers differ significantly from the "negatively treated" group across most observables. Instead, we use the "positively treated" group as a control group. These consumers—whose predicted probability of default decreases when small medical collections are removed—closely resembles the "negatively treated" group in terms of observables. As shown in Section 3, because small medical collections have minimal predictive power, the assignment to the "negatively" and "positively" treated groups is effectively random. This randomness supports the validity of our key identifying assumption: in the absence of the information deletion, outcomes for the two groups would have followed similar trends.[19]

We estimate the following regression model at the individual level:

$$Y_{ict} = \alpha + \beta TREATED_i \times POST_t + \lambda_i + \delta_{ct} + \epsilon_{it}, \tag{7}$$

where Y_{ict} is an outcome for consumer i, residing in county c, in year t; $TREATED_i$ is an indicator equal to one if the consumer belongs to the "negatively treated" group and zero if they belong to the "positively treated" group; $POST_t$ is an indicator equal to one beginning

[18] This is the control group used in Liberman et al. (2019), who implement a similar difference-in-differences analysis.

[19] If small medical collections have any predictive power, their deletion would be expected to positively affect the treatment group and negatively affect the control group, providing an upper bound for the estimated effects.

in 2023, the year of information deletion; and λ_i and δ_{ct} denote consumer and county-year fixed effects, respectively.

Treatment is assigned based on differences in the predicted probability of default with and without medical collections. To account for this, we divide the full sample into 1,000 equal-sized bins based on these differences and cluster our standard errors at the bin level. The negative and positively treated groups correspond to approximately 10% of the full sample, resulting in 100 clusters.

Our coefficient of interest is β, which we interpret as the average effect of deleting medical collections for the treatment group—consumers whose predicted probability of default increases when medical collections are removed—relative to the control group. To assess the validity of the parallel trends assumption, we also estimate an event-study version of Equation (7):

$$Y_{ict} = \alpha + \sum_{\substack{\tau=2020, \\ \tau \neq 2022}}^{2024} \beta_\tau TREATED_i \times \mathbb{I}_{t=\tau} + \lambda_i + \delta_{ct} + \epsilon_{it}, \tag{8}$$

where $\mathbb{I}_{t=\tau}$ is an indicator equal to one in year t and zero otherwise. We use 2022, the year prior to information deletion, as the reference period so that β_τ captures how the change in the treatment group's outcome between 2022 and year t differs from changes in the control group's outcome over the same period.

5.2 Results

Figure 6 presents estimates of the indirect effect of information deletion (Equation (8)) on outcomes relating to credit availability, including credit scores (panel A), total balance across all credit products (panel B), accounts opened in the past 6 months (panel C), and revolving utilization (panel D). Outcome trends are similar between the treatment and control groups prior to 2023, with no statistically significant differences. We observe no significant trend changes after 2023, providing evidence consistent with the absence of any indirect effects.

Figure 7 shows similar results for measures of payment history and subprime borrowing, including the number of trades 90 days or more past due (panel A), a binary variable equal to one if the consumer has a subprime balance (panel B), total subprime credit conditional on being a subprime borrower (panel C), and the number of subprime trades conditional on being a subprime borrower (panel D). We again find no evidence of pre-trends or of an effect of the intervention.[20] Overall, these results suggest that deleting medical collections from credit reports does not have negative spillover effects on other consumers.

Table 6 presents estimates from Equation (7), which assumes a constant treatment effect over time and aggregates years into longer periods to increase statistical power. Across the 8 outcomes analyzed, the estimated effects are small and almost all statistically insignificant. The sole significant estimate is a 0.721-point increase in credit scores, equivalent to just 0.12% of the sample mean (617.13). Overall, our estimates are highly precise, allowing us to rule out even small negative spillovers from the deletion of small medical collections.

6 Conclusion

This paper studies the effects of deleting medical debts from credit reports using a combination of machine-learning credit scoring models, regression discontinuity, and differences-in-differences analysis. Contrary to stated policy goals, we find that deleting medical collections from credit reports has no meaningful impact on credit access or financial health. Our analysis focuses on medical collections under $500, which were removed from credit reports by all three major credit bureaus in 2023. First, we show that this information has little predictive value for default by comparing the output of two credit scoring models: one that includes small medical collections and one that excludes them.

We then test the implication of this finding: if small medical collections are not relevant for risk pricing, their removal should not affect credit access or loan terms. Using a regression discontinuity design, we find no evidence that consumers benefit from the removal of this information in terms of credit access, repayment behavior, or subprime borrowing, ruling out even small effects. Next, we analyze potential spillover effects using a differences-in-differences framework. We compare consumers whose predicted default probability rises when small medical collections are removed with observationally similar consumers whose predicted default probability declines. We find no evidence of negative spillover effects from

[20] We show additional outcomes in Appendix Figure A.16.

the information deletion, again ruling out small effects.

Our findings contribute to the policy debate on alleviating the burden of medical debt. While standard economic models emphasize ex-ante solutions such as increasing access to health insurance, this approach has proven challenging: about 30 million Americans remain uninsured, and many insured individuals face substantial out-of-pocket costs (Einav and Finkelstein, 2023). Recent policy efforts have shifted toward ex-post solutions, such as debt forgiveness and the removal of medical debt from credit reports. However, these alternatives may not deliver their intended benefits. For example, Kluender et al. (2024) find that purchasing and forgiving medical debt in the secondary market has no impact on credit access or financial distress in two large-scale randomized experiments. Our findings reinforce this evidence, suggesting that deleting records of these medical debts is similarly ineffective.

References

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.

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.

Bartik, A. W. and S. T. Nelson (2024). Deleting a signal: Evidence from pre-employment credit checks. Working paper.

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

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.

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

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.

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.

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.

Consumer Financial Protection Bureau (2022, March). Medical debt burden in the united states.

Consumer Financial Protection Bureau (2024, June). Prohibition on creditors and consumer reporting agencies concerning medical information (Regulation V). Federal Register. Docket No. CFPB-2024-0023, RIN 3170-AA54.

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.

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.

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.

Einav, L. and A. Finkelstein (2023). *We've Got You Covered: Rebooting American Health Care*. New York: Portfolio.

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.

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.

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 C.-C. Ethan (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.

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*, qjae045.

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.

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

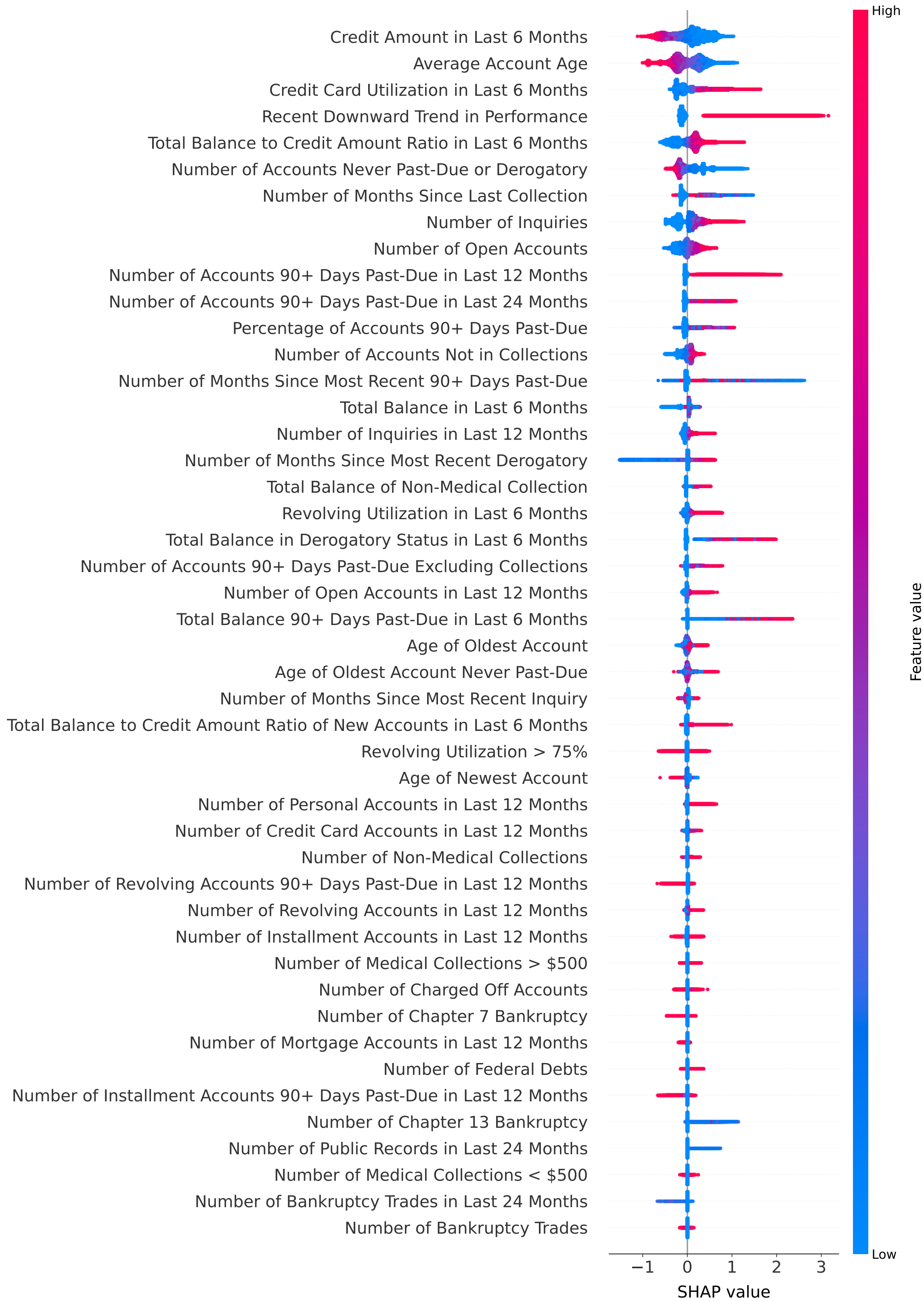
Musto, D. K. (2004). What happens when information leaves a market? evidence from postbankruptcy consumers. *The Journal of Business 77*(4), 725–748.

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.

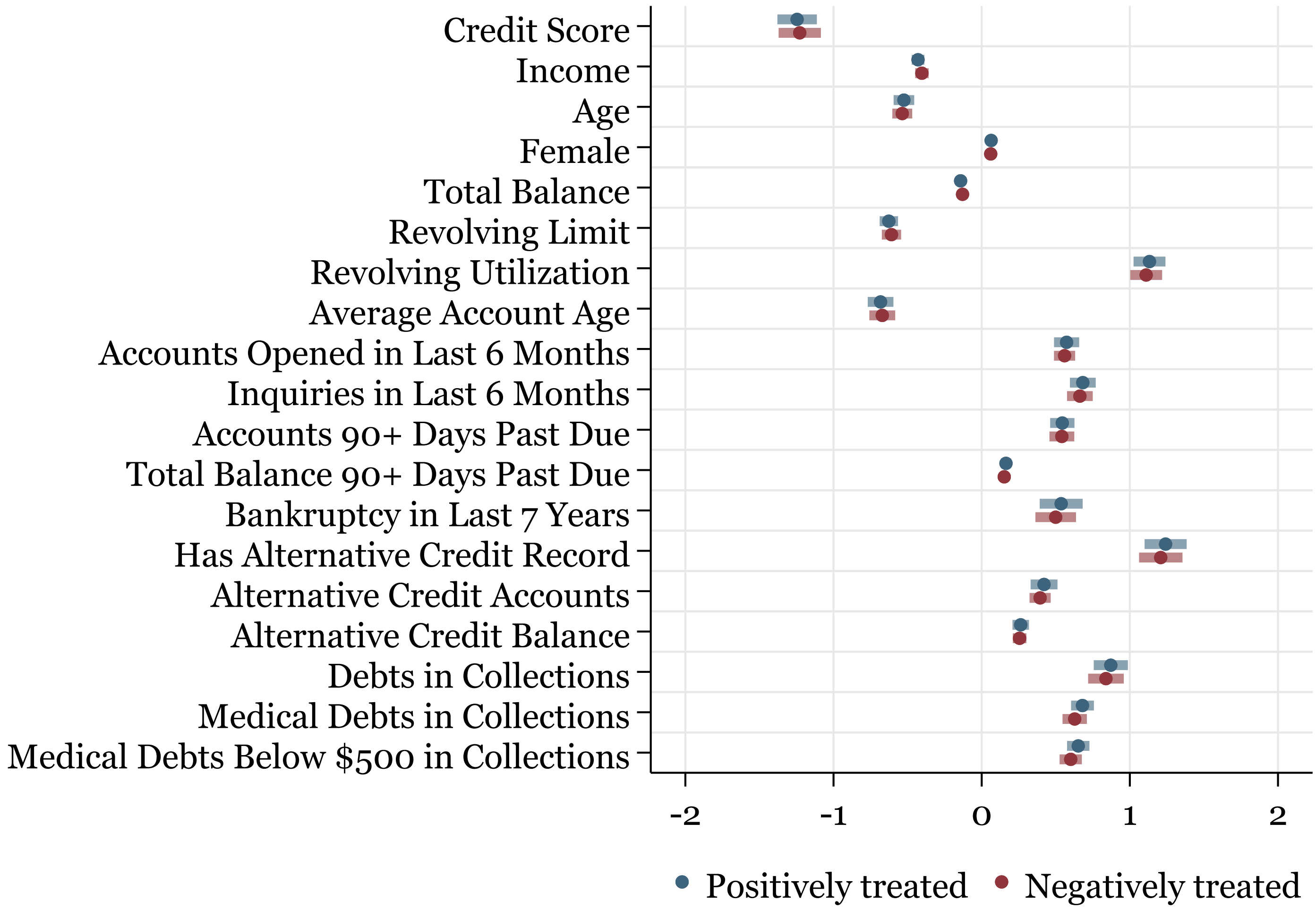
U.S. Census Bureau (2022). Debt for households, by type of debt and selected characteristics: 2022. Accessed: 2025-01-19.

Figure 1: Variable Importance in the Credit Scoring Model



Notes: This figure displays variable importance measures, expressed as average SHAP values, for the credit scoring model presented in Section 3. The model, trained on 2019 data, predicts defaults occurring in 2020–2021 and is estimated using XGBoost. It incorporates 46 predictors, including medical collections under $500. Predictors ("features") are listed from top to bottom based on their average contribution to the model's predictions (average absolute SHAP value). Each row shows the distribution of SHAP values for individual observations, with the predictor's value (X_i) color-coded according to the heat map on the right. Narrow horizontal lines centered around 0 indicate that the predictor has little effect on predictions.

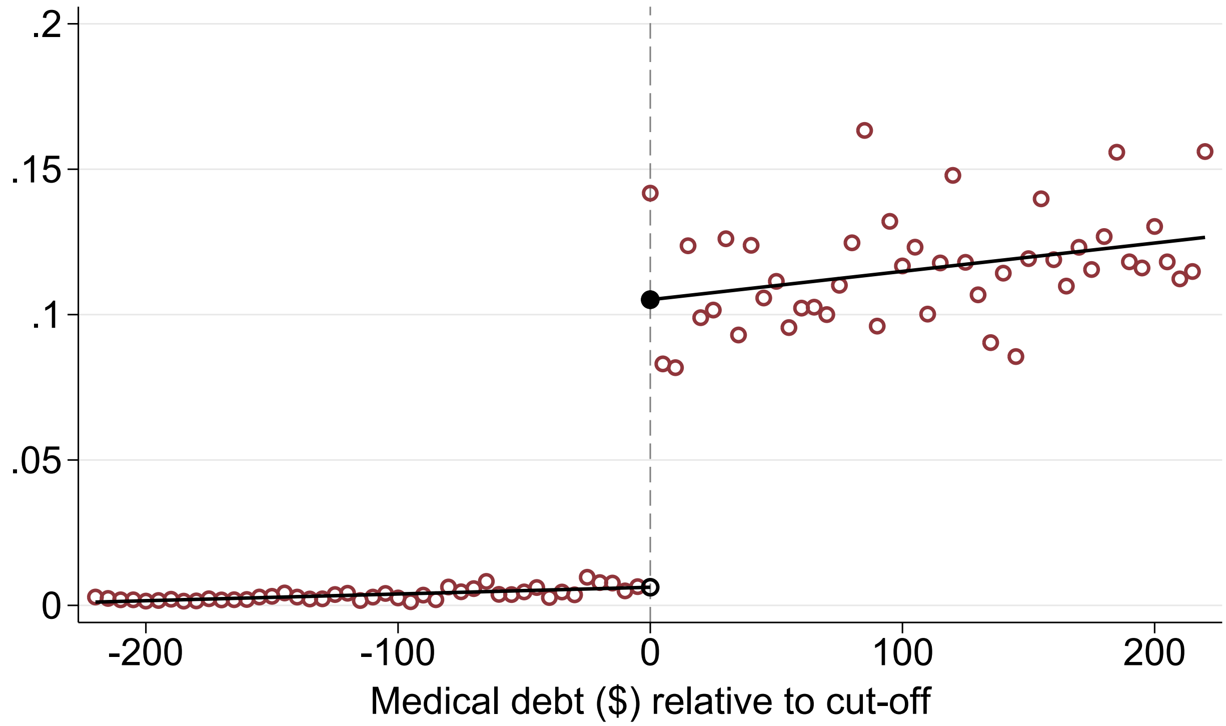
Figure 2: Covariate Balance by Changes in Default Probabilities



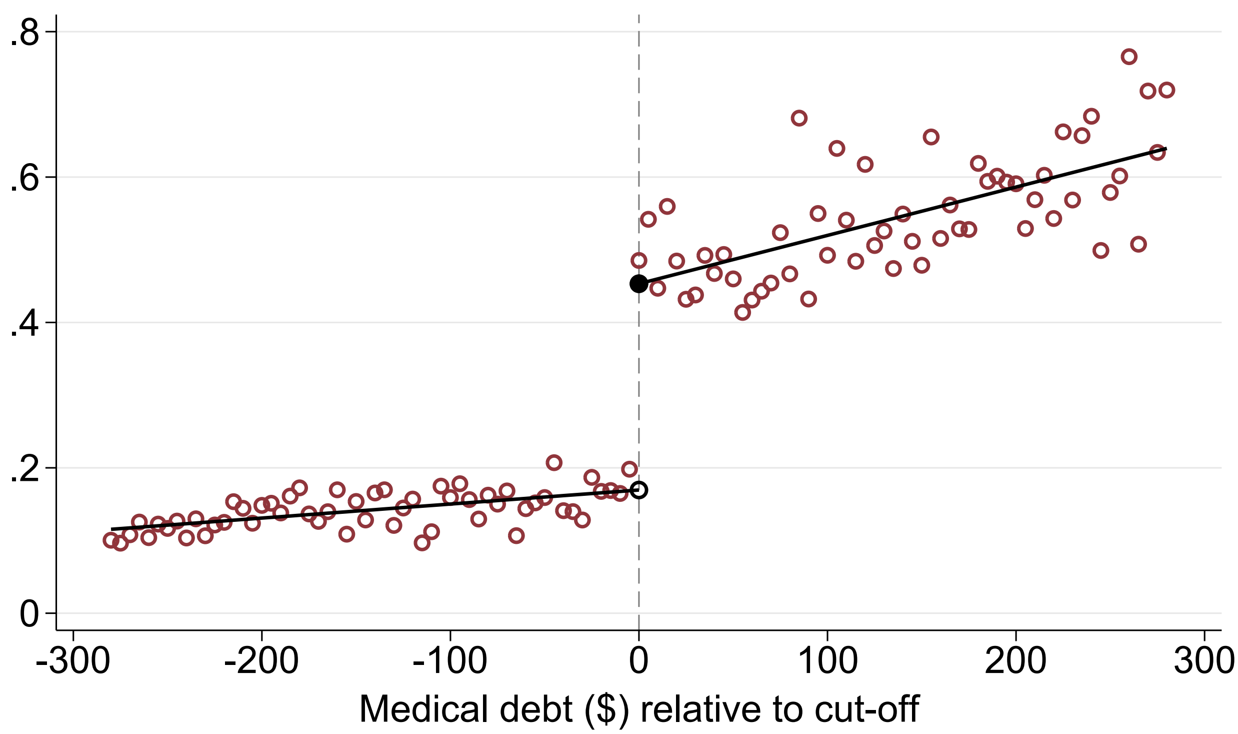
Notes: This figure shows estimates from balancing regressions for selected outcomes. Each balancing regression compares positively or negatively treated consumers to unaffected consumers. Negatively treated consumers are those whose predicted probability of default increases by 2 percentage points or more when small medical collections are removed from the credit scoring model described in Section 3. Positively treated consumers are those whose predicted default probability decreases by at least two percentage points. Unaffected consumers experience changes of less than 0.002 percentage points. All variables are standardized, and each dot represents the regression coefficient of the variable labeled on the y-axis, regressed on either the positive (blue) or negatively treated (red) group indicator. We divide consumers into 100 equal-sized bins based on changes in predicted default probability and cluster standard errors at the bin level.

Figure 3: Two-Year Evolution of 2022 Medical Collections Accounts

(a) Share of Surviving Accounts

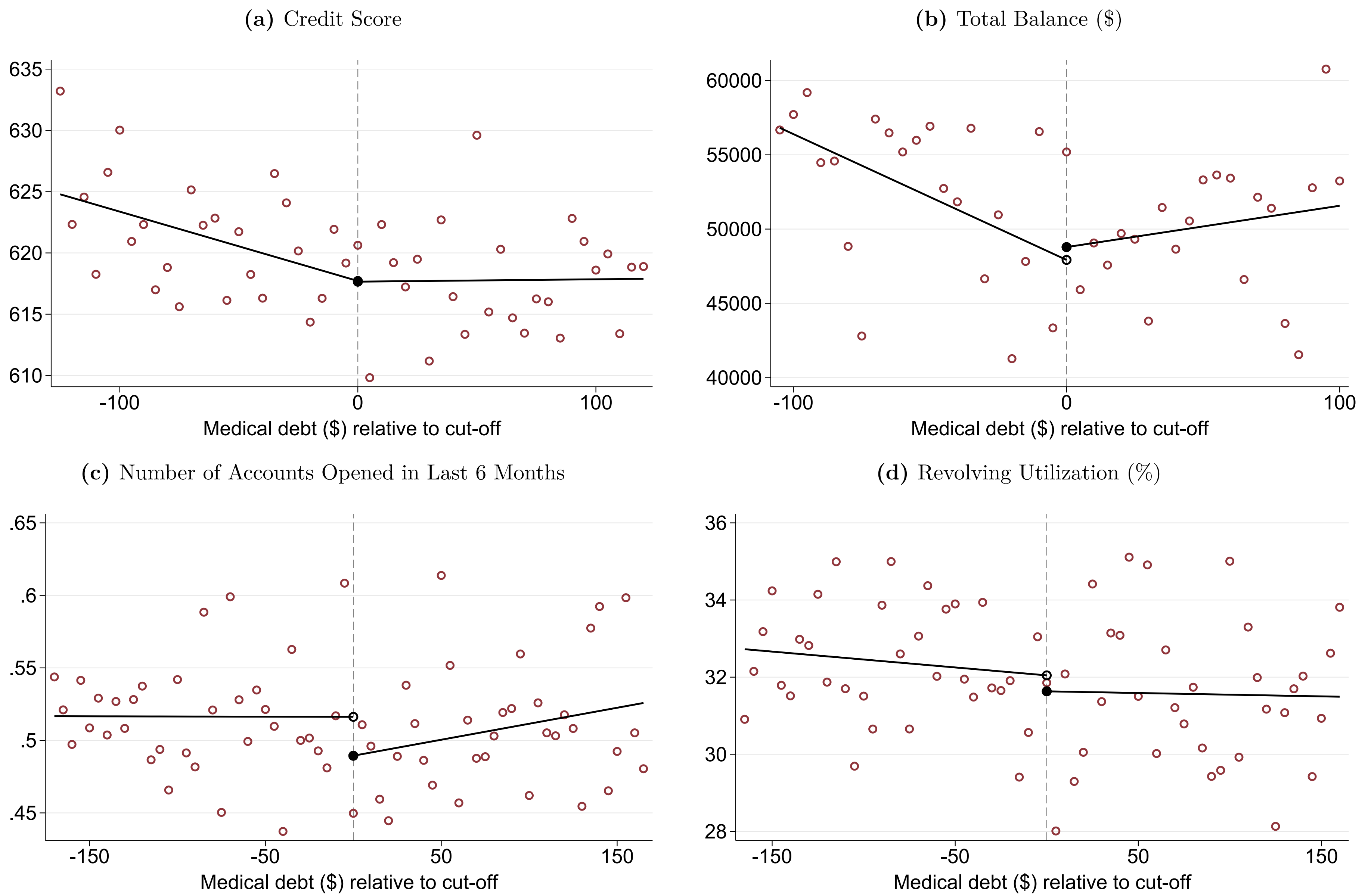


(b) Average Number of Accounts per Person



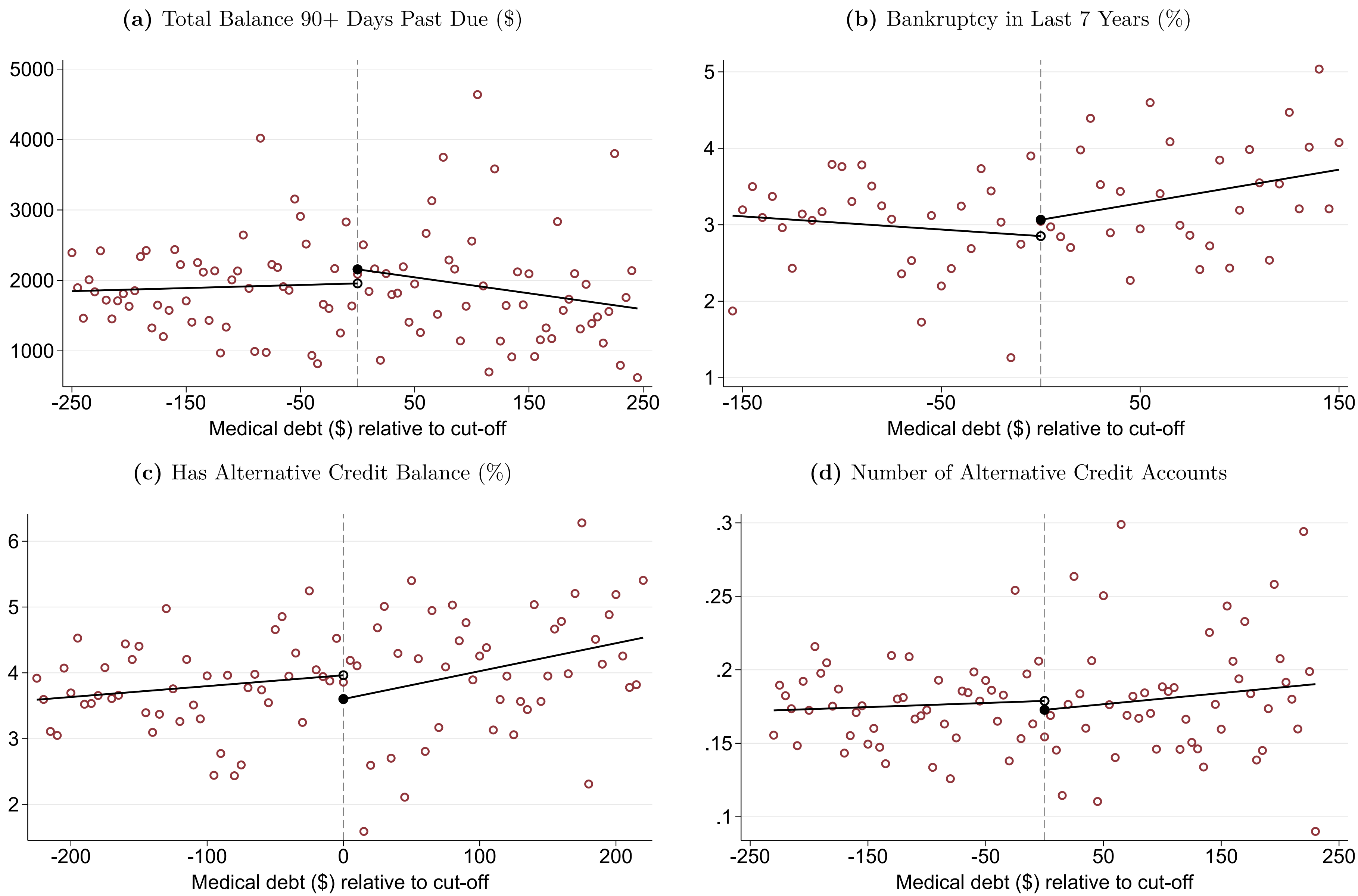
Notes: Panel (a) shows the proportion of 2022 medical collection accounts that remain on credit reports in 2024 by account amount, where the amount is measured as distance from the $500 threshold. Panel (b) shows the average number of medical collections accounts per person in 2024, where the running variable is the maximum value of the consumer's 2022 medical collections accounts. The fitted lines are estimated using Equation (5) for Panel (a) and Equation (6) for Panel (b). RD estimates for Panel (b) are reported in Table 5.

Figure 4: Access to Credit, 2024



Notes: This figure shows the relationship between 2022 medical debt and four credit measures in 2024: consumer credit score, total balance, number of accounts opened in last 6 months, and revolving utilization. Medical debt is defined as the maximum value of the consumer's 2022 medical collections accounts, measured relative to the $500 threshold. The corresponding RD estimates from Equation (6) are reported in Table 5.

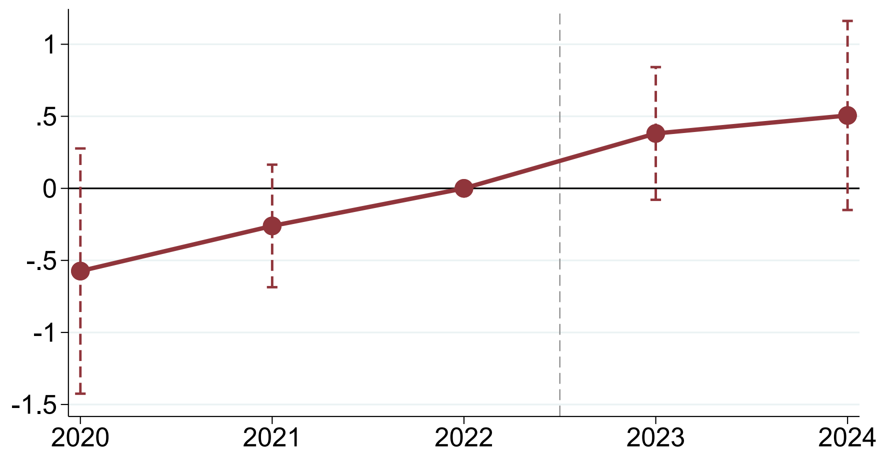
Figure 5: Financial Distress and Access to Alternative Credit, 2024



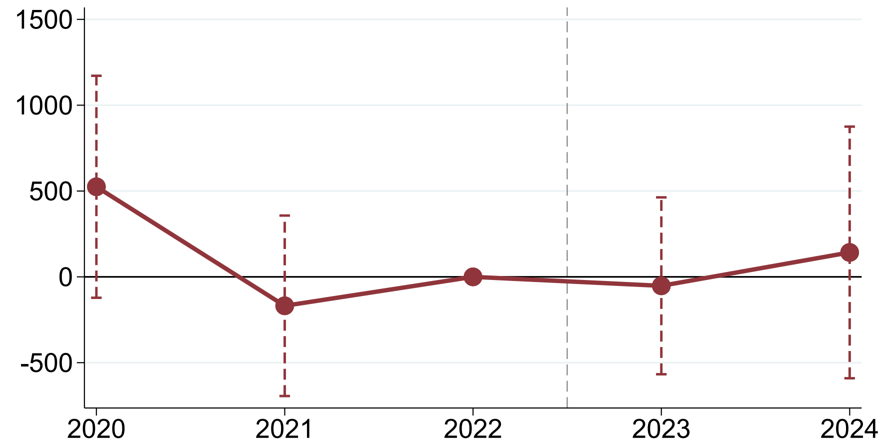
Notes: This figure shows the relationship between 2022 medical debt and measures of delinquency, bankruptcy, and alternative credit use. Medical debt is defined as the maximum value of the consumer's 2022 medical collections accounts, measured relative to the $500 threshold. The corresponding RD estimates from Equation (6) are reported in Table 5.

Figure 6: Effect of Removing Medical Debts on Access to Credit for Consumers Reclassified as Higher Risk

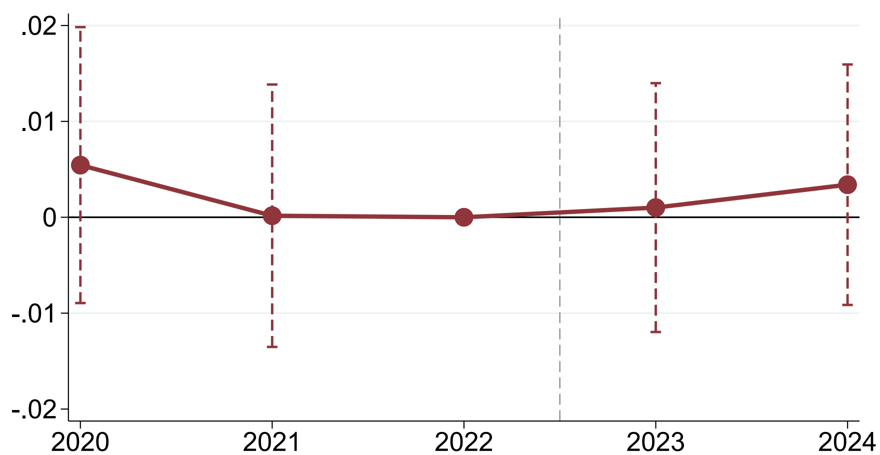
(a) Credit Score



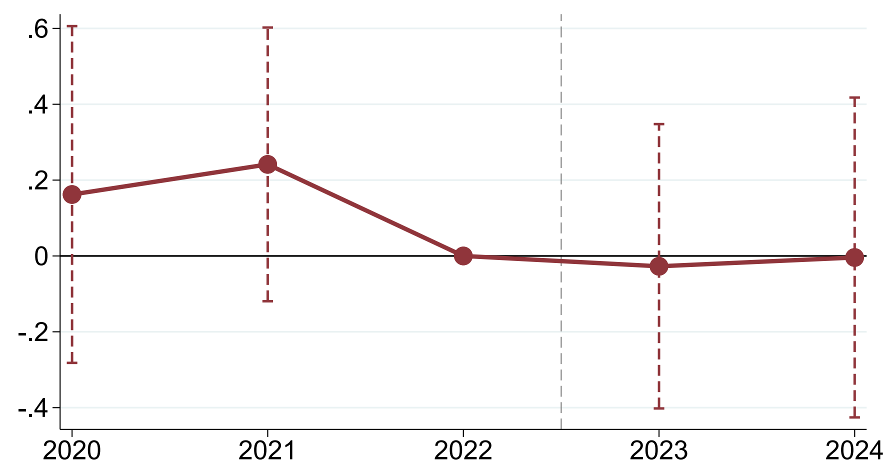
(b) Total Balance ($)



(c) Number of Accounts Opened in Last 6 Months

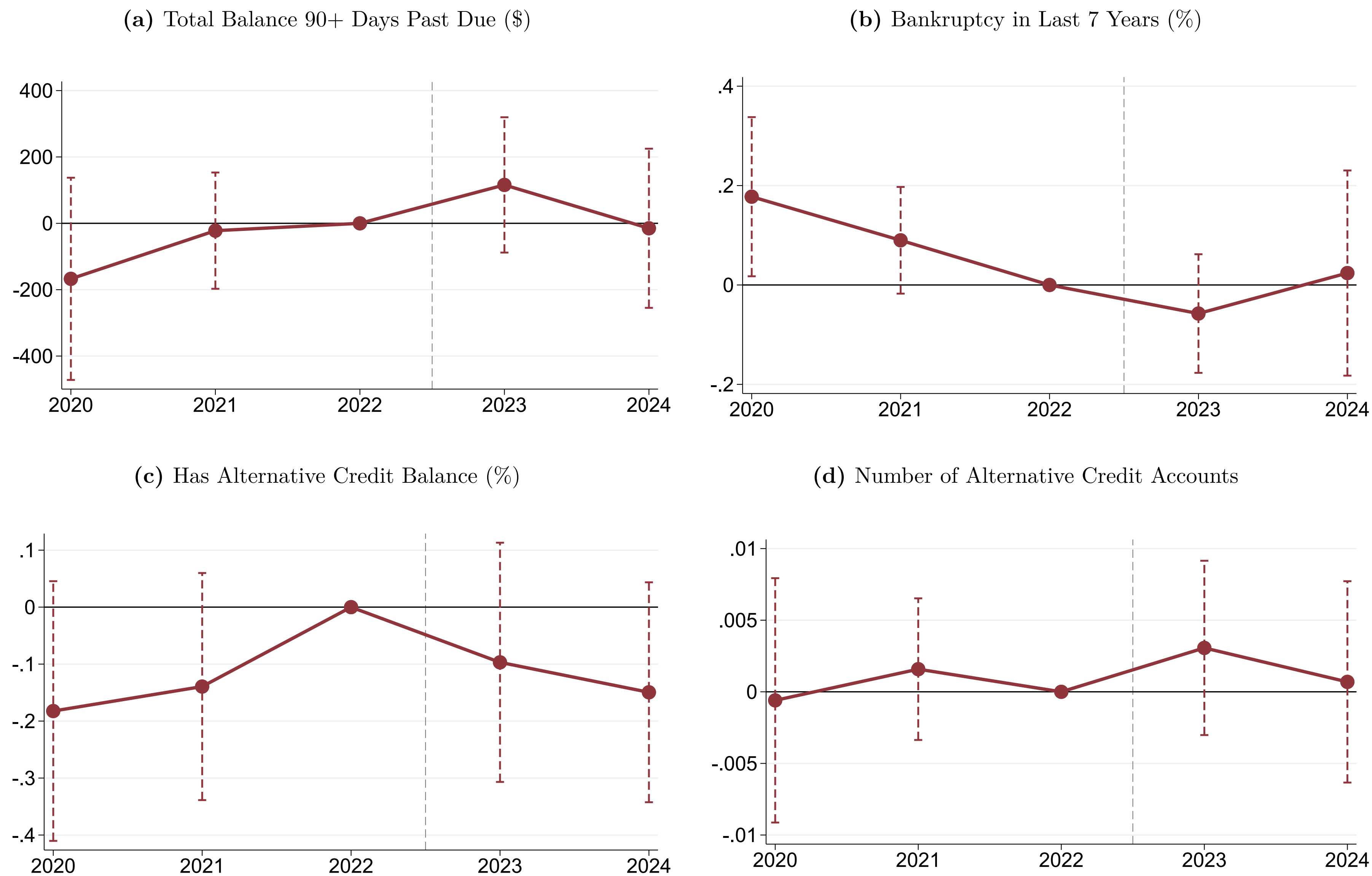


(d) Revolving Utilization (%)



Notes: This figure presents difference-in-differences estimates based on Equation (8). The treatment group consists of consumers whose predicted probability of default increases by 2 percentage points or more when small medical collections are removed from the credit scoring model described in Section 3. The control group includes consumers whose predicted probability falls by at least 2 percentage points. The dashed vertical line indicates when small medical debts were removed from credit reports. Standard errors are clustered based on 100 bins of predicted default probabilities.

Figure 7: Effect of Removing Medical Debts on Financial Distress and Alternative Credit Access for Consumers Reclassified as Higher Risk



(a) Total Balance 90+ Days Past Due ($)

(b) Bankruptcy in Last 7 Years (%)

(c) Has Alternative Credit Balance (%)

(d) Number of Alternative Credit Accounts

Notes: This figure presents difference-in-differences estimates based on Equation (8). The treatment group consists of consumers whose predicted probability of default increases by 2 percentage points or more when small medical collections are removed from the credit scoring model described in Section 3. The control group includes consumers whose predicted probability falls by at least 2 percentage points. The dashed vertical line indicates when small medical debts were removed from credit reports. Standard errors are clustered based on 100 bins of predicted default probabilities.

Table 1: Summary Statistics, 2019–2024

	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample			Medical Debt Subsample		
	Mean	Median	St. Dev.	Mean	Median	St. Dev.
A. Demographics						
Income ($1,000)	51.96	41.00	32.81	38.69	34.00	19.51
Age (years)	50.56	49.00	19.41	44.90	43.00	15.16
Female (%)	50.02	100.00	50.00	54.06	100.00	49.83
B. Access to Credit						
Credit Score	702.26	715.00	100.85	611.21	601.00	86.69
Total Balance ($1,000)	76.46	10.14	140.15	41.91	6.75	86.90
Revolving Limit ($1,000)	21.47	6.42	33.72	4.92	0.00	14.46
Revolving Utilization (%)	28.09	13.00	32.84	49.86	47.00	39.50
Average Account Age (months)	105.15	94.00	77.05	73.39	66.00	51.58
Number of Accounts Opened in Last 6 Months	0.41	0.00	0.83	0.47	0.00	0.95
Number of Inquiries in Last 6 Months	0.41	0.00	0.74	0.62	0.00	0.90
Number of New Mortgages in Last 6 Months	0.02	0.00	0.16	0.01	0.00	0.11
C. Access to Alternative Credit						
Has Alternative Credit Record (%)	19.99	0.00	39.99	47.93	0.00	49.96
Has Alternative Credit Balance (%)	1.02	0.00	10.05	2.60	0.00	15.92
Number of Alternative Credit Accounts	0.05	0.00	0.43	0.13	0.00	0.68
Alternative Credit Balance ($1,000)	0.05	0.00	0.79	0.13	0.00	1.19
D. Financial Distress						
Number of Accounts 90+ Days Past Due	0.19	0.00	0.90	0.46	0.00	1.33
Total Balance 90+ Days Past Due ($1,000)	0.81	0.00	12.16	1.95	0.00	15.90
Bankruptcy in Last 7 Years (%)	2.84	0.00	16.60	4.79	0.00	21.35
E. Debt in Collections						
Number of Debts	0.60	0.00	1.87	3.81	3.00	4.11
Total Debts ($1,000)	0.56	0.00	3.65	3.14	1.39	9.90
Number of Medical Debts	0.25	0.00	1.02	2.44	2.00	2.17
Number of Medical Debts Below $500	0.15	0.00	0.66	1.45	1.00	1.53
Observations	15,313,700			1,585,485		

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 Collections, 2022 (RD sample)

	(1)	(2)	(3)
	Mean	St. Dev.	Median
A. Demographics			
Income ($1,000)	40.70	20.53	35.00
Age (years)	45.11	15.19	43.00
Female (%)	55.28	49.72	100.00
B. Access to Credit			
Credit Score	625.38	87.95	618.00
Total Balance ($1,000)	48.89	94.40	10.66
Revolving Limit ($1,000)	6.03	15.79	0.23
Revolving Utilization (%)	29.51	38.58	6.00
Average Account Age (months)	73.74	49.08	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 (%)	51.72	49.97	100.00
Has Alternative Credit Balance (%)	3.20	17.59	0.00
Number of Alternative Credit Accounts	0.14	0.67	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	13.17	0.00
Bankruptcy in Last 7 Years (%)	4.11	19.86	0.00
E. Debt in Collections			
Number of Debts	3.53	3.77	2.00
Total Debts ($1,000)	2.69	6.51	1.09
Number of Medical Debts	2.37	2.00	1.00
Number of Medical Debts Below $500	1.56	1.43	1.00
Observations	271,305		

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

	(1)	(2)	(3)
	All variables	Exclude Medical Debts < $500	Exclude All Medical Debts
Accuracy	0.905	0.906	0.906
Recall	0.448	0.448	0.448
Precision	0.736	0.736	0.737
F1 Score	0.557	0.557	0.557
AUC	0.712	0.712	0.712

Notes: This table reports performance metrics for a credit scoring model predicting defaults occurring between 2020 and 2021, based on 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 accuracy score represents the share of correct predictions. For comparison, a naive model predicting no defaults achieves an accuracy of 0.867. Precision is the proportion of predicted defaults that were correctly classified. Recall is the proportion of actual defaults correctly classified. F1 score is the harmonic mean of Precision and Recall. The AUC (Area Under the Receiver Operating Characteristic Curve) indicates the probability that the model assigns a higher default probability to a true defaulter than to a non-defaulter.

Table 4: Summary Statistics by Treatment Groups Based on Changes in Default Probabilities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Unaffected			Positively Treated			Negatively Treated		
	Mean	Median	St. Dev.	Mean	Median	St. Dev.	Mean	Median	St. Dev.
A. Demographics									
Income ($1,000)	60.00	50.00	36.31	44.00	38.00	22.68	44.84	38.00	24.00
Age (years)	57.06	58.00	19.42	45.77	44.00	14.95	45.59	44.00	15.08
Female (%)	50.03	100.00	50.00	53.52	100.00	49.88	53.35	100.00	49.89
B. Access to Credit									
Credit Score	750.49	787.00	85.44	611.01	609.00	80.44	612.67	610.00	82.00
Total Balance ($1,000)	88.24	7.82	154.59	65.33	17.97	115.15	67.43	18.33	118.42
Revolving Limit ($1,000)	31.79	19.11	38.96	6.06	0.43	15.52	6.72	0.49	17.30
Revolving Utilization (%)	16.09	6.00	23.16	54.68	55.00	38.61	54.06	55.00	38.76
Average Account Age (months)	136.56	123.00	83.72	77.00	68.00	46.49	77.43	69.00	46.31
Number of Accounts Opened in Last 6 Months	0.28	0.00	0.64	0.70	0.00	1.19	0.70	0.00	1.19
Number of Inquiries in Last 6 Months	0.26	0.00	0.57	0.78	0.00	1.00	0.76	0.00	0.99
Number of New Mortgages in Last 6 Months	0.03	0.00	0.17	0.02	0.00	0.13	0.02	0.00	0.14
C. Access to Alternative Credit									
Has Alternative Credit Record (%)	6.90	0.00	25.35	54.01	100.00	49.84	52.86	100.00	49.92
Has Alternative Credit Balance (%)	0.22	0.00	4.72	3.88	0.00	19.32	3.76	0.00	19.03
Number of Alternative Credit Accounts	0.01	0.00	0.19	0.19	0.00	0.86	0.18	0.00	0.85
Alternative Credit Balance ($1,000)	0.01	0.00	0.35	0.22	0.00	1.59	0.21	0.00	1.55
D. Financial Distress									
Number of Accounts 90+ Days Past Due	0.05	0.00	0.44	0.64	0.00	1.68	0.64	0.00	1.67
Total Balance 90+ Days Past Due ($1,000)	0.22	0.00	6.42	2.85	0.00	22.81	2.64	0.00	21.64
Bankruptcy in Last 7 Years (%)	0.95	0.00	9.68	9.61	0.00	29.48	9.17	0.00	28.85
E. Debt in Collections									
Number of Debts	0.20	0.00	1.07	1.78	1.00	3.01	1.73	1.00	2.96
Total Debts ($1,000)	0.17	0.00	1.71	1.69	0.11	4.88	1.67	0.08	4.53
Number of Medical Debts	0.09	0.00	0.60	0.71	0.00	1.66	0.66	0.00	1.60
Number of Medical Debts Below $500	0.05	0.00	0.39	0.42	0.00	1.09	0.39	0.00	1.05
Observations	6,914,163			691,413			691,415		

Notes: This table shows descriptive statistics between 2019 and 2024 for Unaffected consumers in the first three columns, Positively Treated consumers in the next three columns, and Negatively Treated consumers in the last three columns. Negatively Treated refers to consumers above the 95th percentile in the distribution of probability difference, whose predicted probability of default increases by approximately 2 p.p. or more when medical collections below $500 are removed from our baseline credit scoring model. Positively Treated refers to consumers below the 5th percentile in the distribution of probability difference, whose predicted probability of default decreases by approximately 2 p.p. or more when medical collections below $500 are removed from our baseline credit scoring model. Unaffected consists of consumers between the 25th and 75th percentiles. All variables come from the Gies Consumer and Small Business Credit Panel.

Table 5: RD Estimates of the Direct Effect of Medical Debt Deletion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of Debts	Number of Medical Debts	Credit Score	Total 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^{2024}	0.226***	0.299***	–0.683	2,235	–0.0386*	–0.393	92.6	0.367	–0.381	–0.0109
	[0.145, 0.324]	[0.267, 0.322]	[–6.03, 3.63]	[–2,602, 8,590]	[–0.0895, 0.000785]	[–2.38, 1.54]	[–510, 802]	[–0.542, 1.19]	[–1.28, 0.293]	[–0.0453, 0.0153]
Sample mean	1.31	0.279	620	51,469	0.510	32.1	1,893	3.18	3.86	0.177
% of Mean	17.3	107	–0.110	4.34	–7.55	–1.22	4.89	11.6	–9.87	–6.15
Bandwidth	146	281	122	104	167	161	246	152	223	230
Observations	271,305	271,305	271,305	271,305	271,305	271,305	271,305	271,305	271,305	271,305

Notes: This table shows the coefficient (β) estimates and 95% confidence intervals of Equation (6). The running variable for medical debt corresponds to the highest debt amount across the consumer's medical collections accounts. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Sample Mean reports the mean of the dependent variable in 2024. A ***, **, and * indicate significance at the 1%, 5%, and 10% level respectively, using conventional inference.

Table 6: Difference-in-Differences estimates of the Indirect Effect of Medical Debt Deletion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Credit Score	Total 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
$TREATED \times POST$	0.721***	–74.5	0.000329	–0.144	114	–0.106	–0.0159	0.00156
	(0.271)	(302)	(0.00455)	(0.166)	(102)	(0.0957)	(0.0634)	(0.00309)
Sample mean	617	68,860	0.688	53.0	2,468	9.05	3.84	0.190
% of mean	0.117	–0.108	0.0478	–0.272	4.62	–1.17	–0.414	0.820
Observations	1,143,272	1,143,272	1,143,272	794,118	1,143,272	1,143,272	1,143,272	1,143,272

Notes: This table presents difference-in-differences estimates based on Equation (7). Standard errors are clustered based on 100 bins of predicted default probabilities. A ***, **, and * indicate significance at the 1%, 5%, and 10% level respectively.

Online Appendix

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

Victor Duarte, Julia Fonseca, Divij Kohli, Julian Reif

A Alternative RD specifications

The effect of deleting small medical collections on individual-level outcomes depends on the underlying treatment mechanism. We consider three possibilities:

1. **Account level**: Treatment scales with the proportion of deleted accounts.
2. **Person level (max)**: Treatment occurs only if all of an individual's medical collections are deleted.
3. **Person level (min)** Treatment occurs if any (i.e., at least one) medical collection is deleted.

The account-level specification can be estimated using Equation (5). First-stage estimates for this case are reported in Panel A of Figure 3. In this appendix, we extend this specification to estimate second-stage outcomes. Since outcomes such as an individual's credit score do not vary across accounts, we cluster standard errors at the individual level.

The person-level (max) specification, where treatment occurs only if all medical collections are deleted, is estimated using Equation (6), as reported in the main text. The person-level (min) specification, where treatment occurs if any medical collection is deleted, can be estimated using a variation of Equation (6), with the running variable, $MINDEBT_i^{2022}$, defined as the balance of the consumer's smallest medical collection relative to the \$500 cutoff:

$$Y_i^{2024} = \alpha MINDEBT_i^{2022} + \beta ABOVE_i^{2022} + \gamma(ABOVE_i^{2022} \times MINDEBT_i^{2022}) + \epsilon_i \tag{9}$$

We present estimates for these three treatment definitions in Table A.1. Panel A replicates the main text estimates from Table 5, which correspond to 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. Overall, these findings suggest that our main estimates are robust to alternative RD specifications.

Table A.1: Direct Effect: Alternative RD Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of Debts	Number of Medical Debts	Credit Score	Total Balance ($1,000s)	Number of Accounts Opened in Last 6 Months	Revolving Utilization (%)	Total Balance 90+ Days Past Due ($1000s)	Bankruptcy in Last 7 Years (%)	Has Alternative Credit Balance (%)	Number of Alternative Credit Accounts
					A: Running Variable is Maximum Debt					
ABOVE^{2022}	0.23***	0.29***	-0.68	2,234	-0.03*	-0.39	92.64	0.36	-0.38	-0.01
	[0.14, 0.32]	[0.26, 0.32]	[-6.03, 3.62]	[-32,601, 8,590]	[-0.08, 0.00]	[-2.37, 1.54]	[-509.60, 801.60]	[-0.54, 1.19]	[-1.28, .29]	[-0.04, 0.01]
	1.30	0.27	619	51,469	0.51	32.05	1893.30	3.17	3.85	0.17
% of Mean	17.2	107	-0.11	4.34	-7.55	-1.22	4.89	11.55	-9.87	-6.15
Observations	38,303	84,177	32,250	26,884	44,126	42,766	70,196	40,178	62,110	64,776
Optimal Bandwidth Size	146.17	281.38	121.63	104.09	166.94	161.48	246.05	152.34	223.17	230.44
					B: Running Variable is Minimum Debt					
ABOVE^{2022}	0.41***	0.43***	-0.41	2,586	-0.01	-0.94	-159	-0.13	-0.01	0.01
	[0.29, 0.56]	[0.36, 0.52]	[-6.13, 5.27]	[-2692, 9,641]	[-0.05, 0.03]	[-3.46, 1.08]	[-745.06, 526.37]	[-1.08, 0.68]	[-0.97, 0.71]	[-0.03, 0.04]
Sample Mean	1.54	0.52	615.77	46,677	0.47	29.71	1929.21	2.69	3.63	0.16
% of Mean	26.62	82.69	-0.06	5.54	-2.12	-3.16	-8.24	-4.83	-0.27	6.25
Observations	22,814	22,814	23,385	19,495	38,630	29,242	54,144	31,698	59,233	66,524
Optimal Bandwidth Size	120.95	120.63	123.2	103.75	191.65	150.74	247.90	162.46	260.92	282.69
					C: Running Variable is Amount of Debt in Medical Account					
ABOVE^{2022}	0.44***	0.38***	-2.85*	-160	-0.01	-0.98	-78.6	-0.04	-0.28	-0.01
	[0.28, 0.64]	[0.25, 0.55]	[-6.78, 0.15]	[-3,235, 3,619]	[-0.05, 0.01]	[-2.69, 0.44]	[-434, 342.4]	[-0.64, 0.53]	[-0.91, 0.29]	[-0.03, 0.01]
Sample Mean	1.58	0.54	615.41	47,365	0.49	30.51	1904.86	2.93	3.78	0.17
% of Mean	27.47	71.16	-0.46	-0.33	-2.62	-3.21	-0.41	-1.25	-7.49	5.20
Observations	59,101	70,962	66,119	64,284	118,082	91,462	168,314	99,417	185,582	178,873
Optimal Bandwidth Size	93.25	110.88	103.12	100.08	175.48	140.46	232.22	150.61	249.29	242.78

Notes: This table shows the coefficient (β) estimates and 95% confidence intervals of Equation (6). The running variable in Panel A corresponds to the highest debt amount across the consumer's medical collections accounts. The running variable in Panel B corresponds to the smallest debt amount across the consumer's medical collections accounts. Whereas, the running variable in Panel C corresponds to the debt amount in the consumer's medical collections account. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Sample Mean reports the mean of the dependent variable in 2024. A ***, **, and * indicate significance at the 1%, 5%, and 10% level respectively, using conventional inference.

B Comparing GCCP Medical Collection Data to External Sources

To identify consumers with medical 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 the furnisher is identified as a business in the medical or health-related sector.[1] To assess whether our sample accurately captures the proportion of consumers with medical collections, we conduct a benchmarking exercise.

Table A.2 compares the share of consumers with medical collections in the GCCP to estimates from other sources. Column (1) reports the annual share of consumers with at least one medical collection in the GCCP from 2018 to 2023, showing a decline from 16.8% in 2018 to 7.1% in 2023. This decline reflects policy changes made during this period, 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).

A similar trend appears in columns (2) and (3), which report estimates from Blavin et al. (2023) (Urban Institute) and Sandler and Nathe (2022) (CFPB), 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 might reflect differences in reporting timelines: the GCCP data are measured in March, the Urban Institute data in August, and the CFPB data in January. Overall, the GCCP data aligns well with these external benchmarks.

[1]These category labels are Dentists, Chiropractors, Doctors, Medical group, Hospitals and clinics, Osteopaths, Pharmacies and drugstore, Optometrists and optical outlets, and Medical and related health-nonspecific.

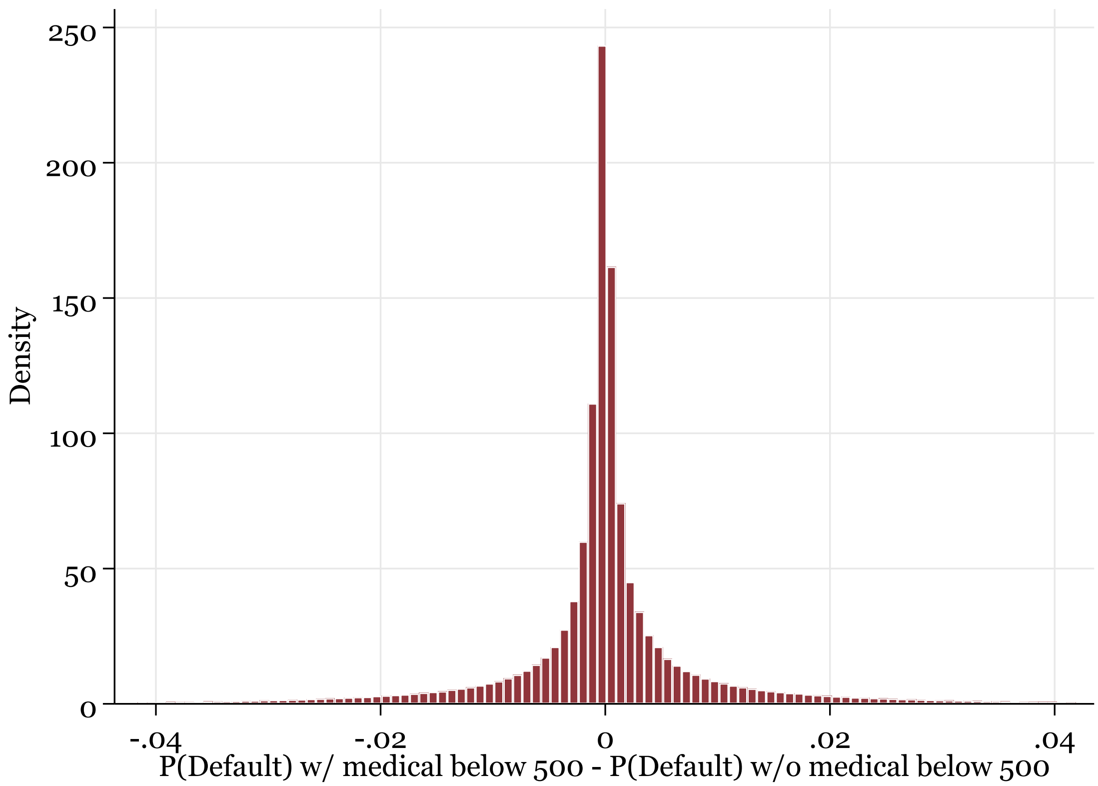
Table A.2: Comparing GCCP Medical Collection Data to External Sources

	(1)	(2)	(3)
Year	GCCP	Urban Institute	CFPB
2018	16.8%	16%	17.6%
2019	15.9%	16%	17.5%
2020	15.6%	15%	16%
2021	14.6%	14%	15.5%
2022	12.9%	12%	14%
2023	7.1%	5%	

Notes: This table compares the share of individuals with medical collections in the GCCP to estimates from other sources. Column (1) reports the share of consumers with at least one account in medical 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.

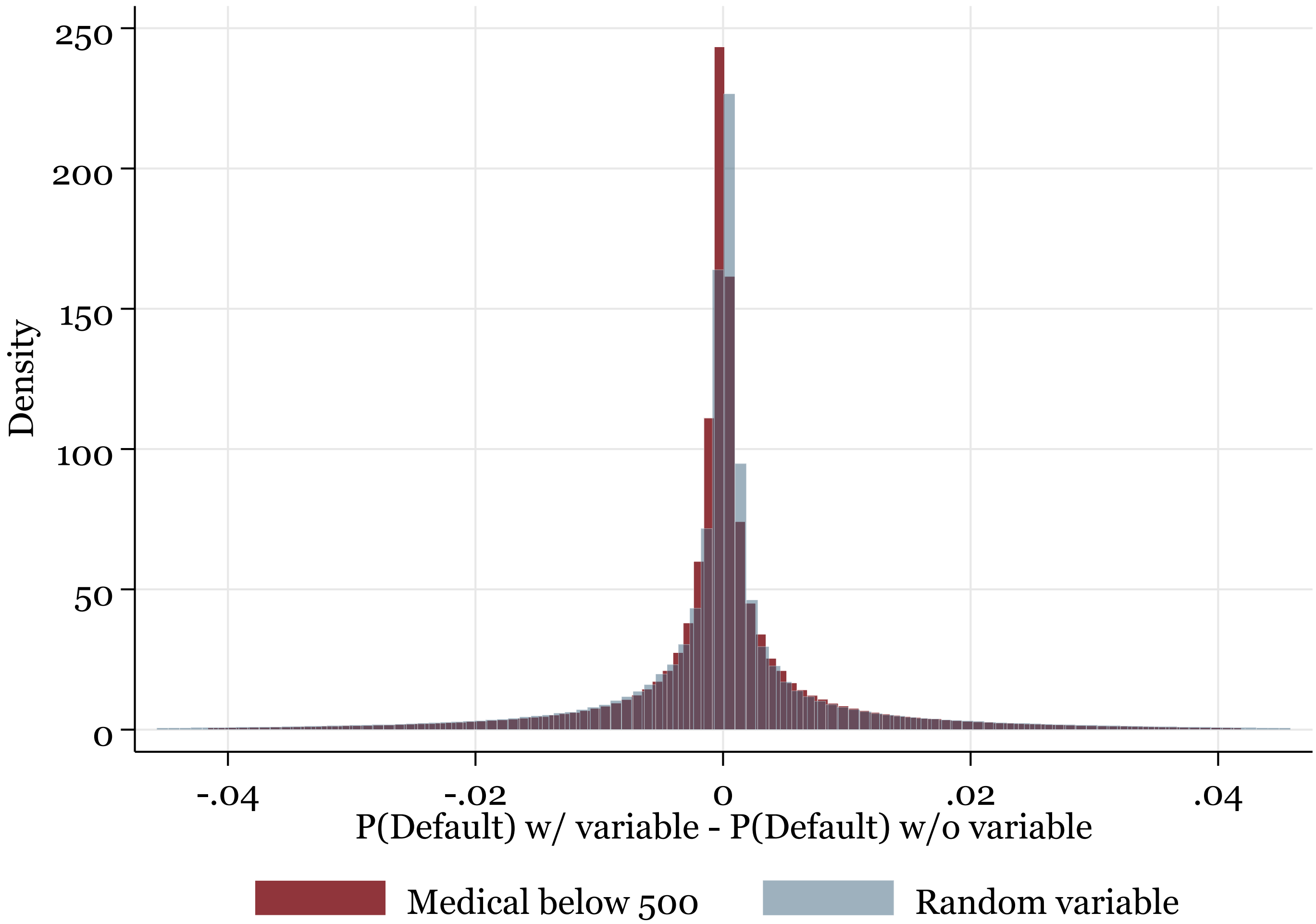
C Additional Figures and Tables

Figure A.1: Effect of Removing Small Medical 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 collections for 2.8 million consumers in the GCCP. Predictions, generated using the credit scoring model described in Section 3, are based on 2019 data.

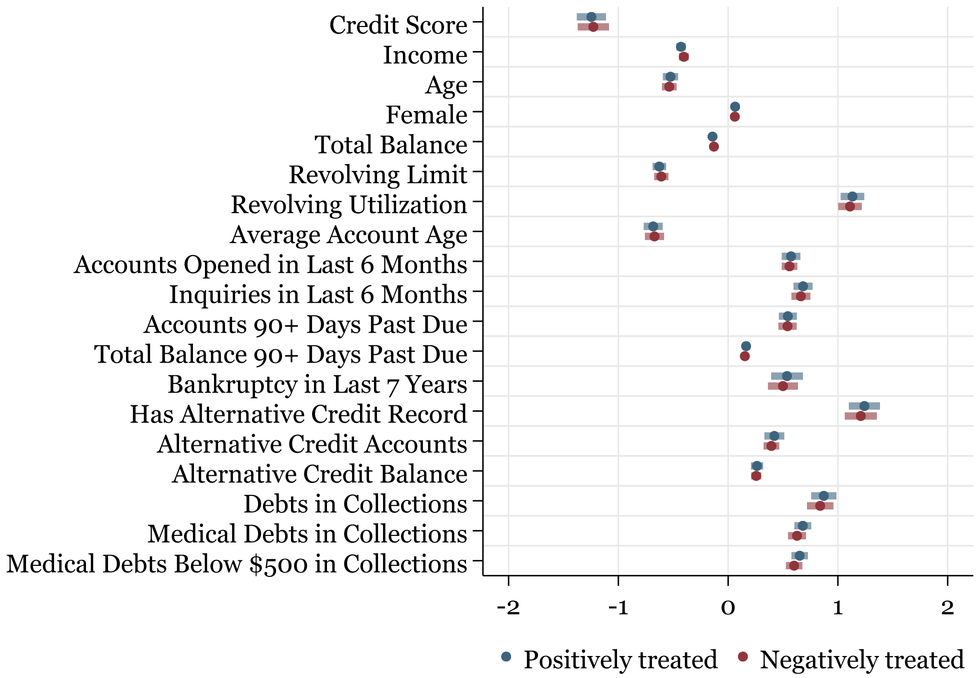
Figure A.2: Effect of Removing Small Medical Collections vs. Removing Noise



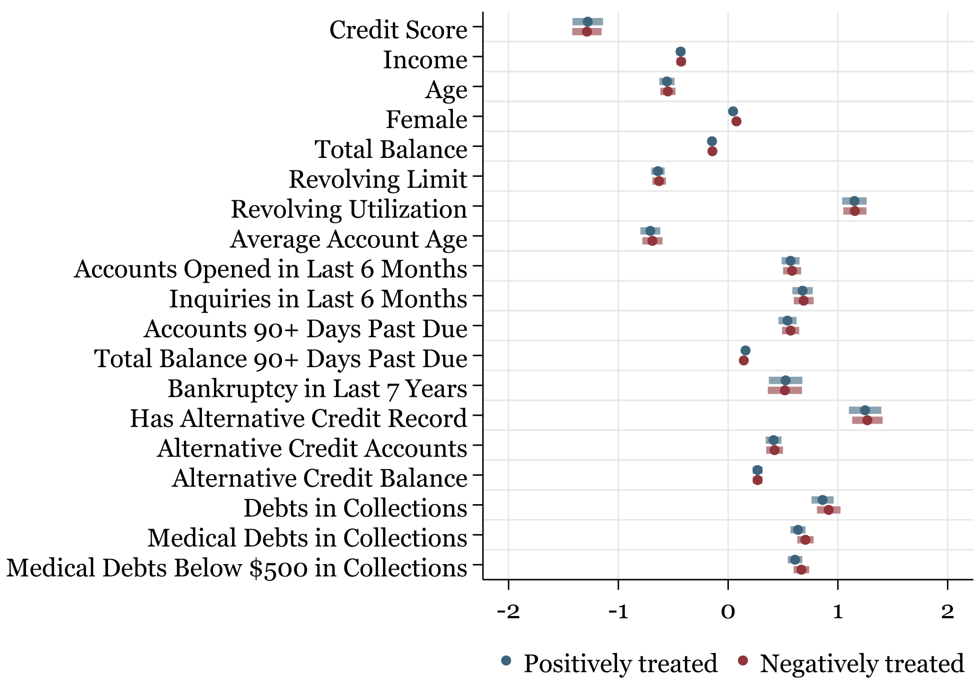
Notes: This figure shows the change in the predicted probability of default over 24 months following the removal of a variable from the credit scoring model described in Section 3. The red histogram reproduces the plot from Figure A.1, showing the effect of removing small (< $500) medical collections. The blue histogram shows the effect of removing a random noise predictor that was drawn randomly from a uniform distribution.

Figure A.3: Covariate Balance by Changes in Default Probabilities: Small Medical Collections vs. Noise

(a) Small (< $500) Medical Collections

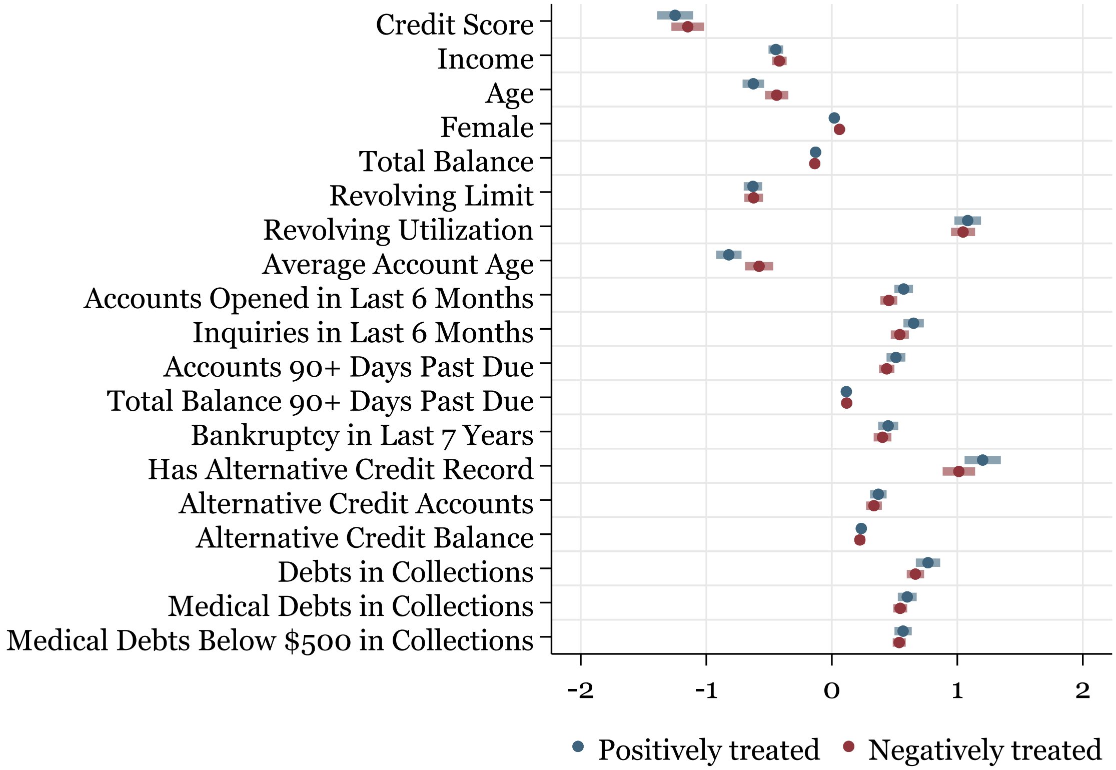


(b) Random Variable



Notes: This figure shows estimates from balancing regressions for selected outcomes. Each balancing regression compares positively or negatively treated consumers to unaffected consumers. Negatively treated consumers are those whose predicted probability of default increases by 2 percentage points or more when small medical collections (Panel a) or a randomly generated predictor (Panel b) are removed from the credit scoring model described in Section 3. Positively treated consumers are those whose predicted default probability decreases by at least two percentage points. Unaffected consumers experience changes of less than 0.002. All variables are standardized, and each dot represents the regression coefficient of the variable labeled on the y-axis, regressed on either the positive (blue) or negatively treated (red) group indicator. We divide consumers into 100 equal-sized bins based on changes in predicted default probability and cluster standard errors at the bin level.

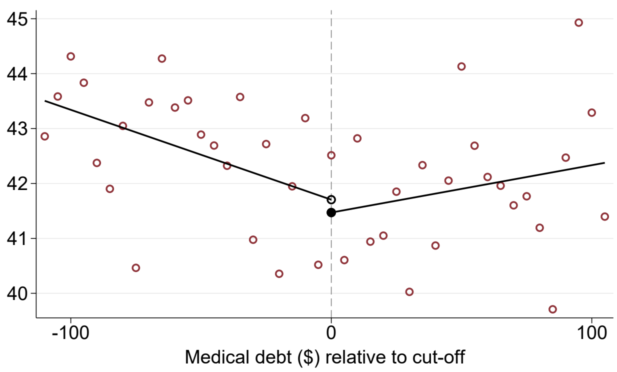
Figure A.4: Covariate Balance by Changes in Default Probabilities: Length of Credit History



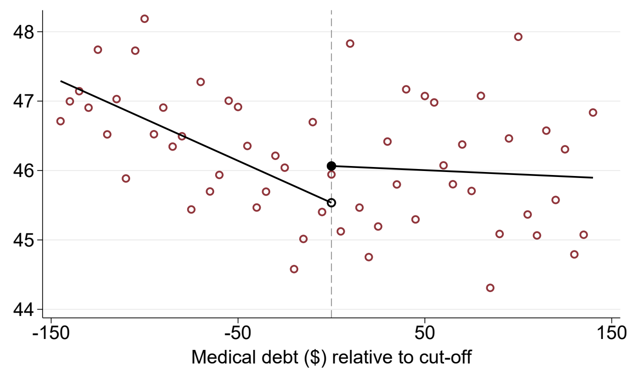
Notes: This figure shows estimates from balancing regressions for selected outcomes. Each balancing regression compares positively or negatively treated consumers to unaffected consumers. Negatively treated refers to consumers above the 95th percentile in the distribution of probability differences according to credit scoring models with and without the variable. Positively treated refers to consumers below the 5th percentile in the distribution of probability differences. Unaffected refers to consumers between the 25th and 75th percentiles. We construct 100 equal-sized bins of the difference in predicted probability of default All variables are standardized, and each dot represents the regression coefficient of the variable labeled on the y-axis, regressed on either the positive (blue) or negatively treated (red) group indicator. We divide consumers into 100 equal-sized bins based on changes in predicted default probability and cluster standard errors at the bin level.

Figure A.5: Covariate Smoothness Test: Demographics

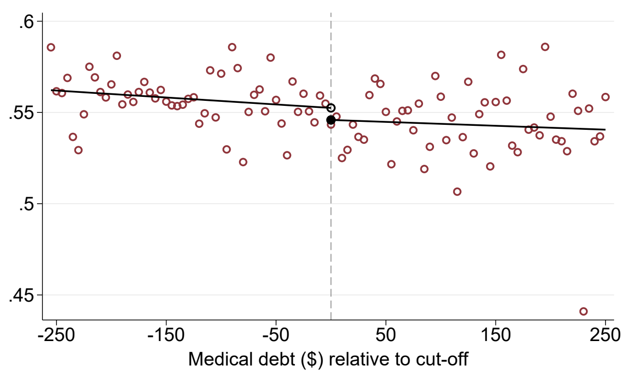
(a) Income ($1000s)



(b) Age (years)



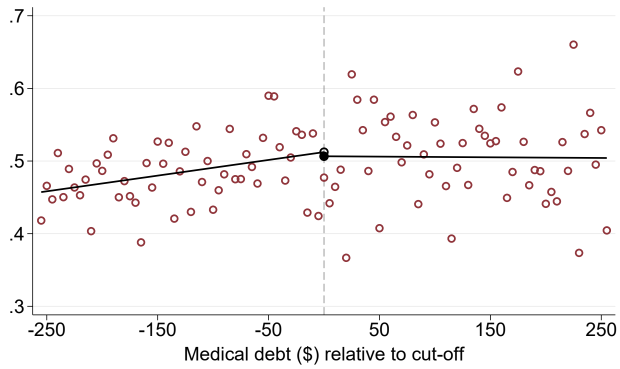
(c) Female (%)



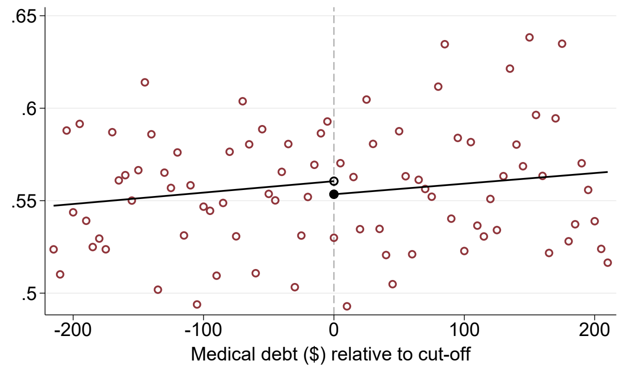
Notes: This figure shows average income, age, and share female by medical debt amount, where the amount is measured as distance from the $500 threshold. The running variable is equal to the maximum value of the consumer's medical collections accounts in 2022. The covariates are measured in 2022. RD estimates from Equation (6) are reported in Table A.5.

Figure A.6: Additional Credit Outcomes for RD Analysis, 2024

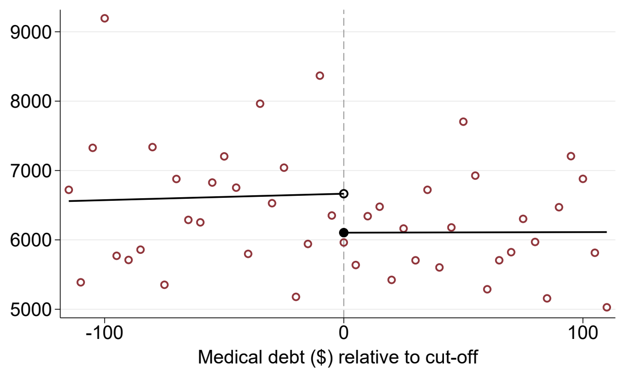
(a) Number of Accounts 90+ Days Past Due



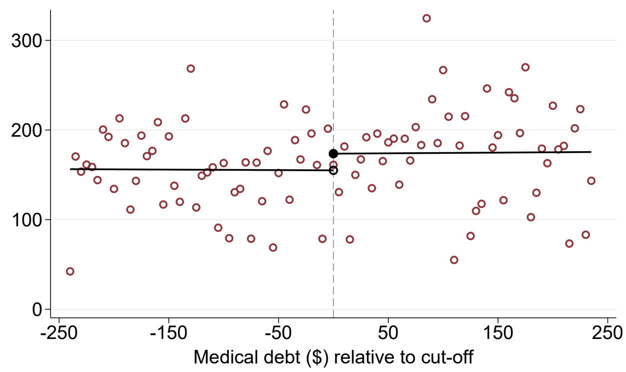
(b) Number of Inquiries in Last 6 Months



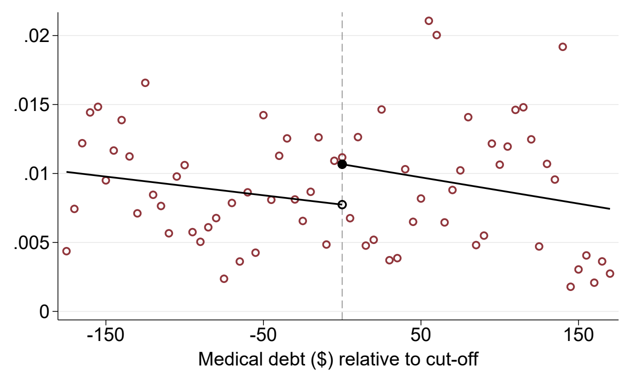
(c) Revolving Limit ($)



(d) Alternative Credit Balance ($)

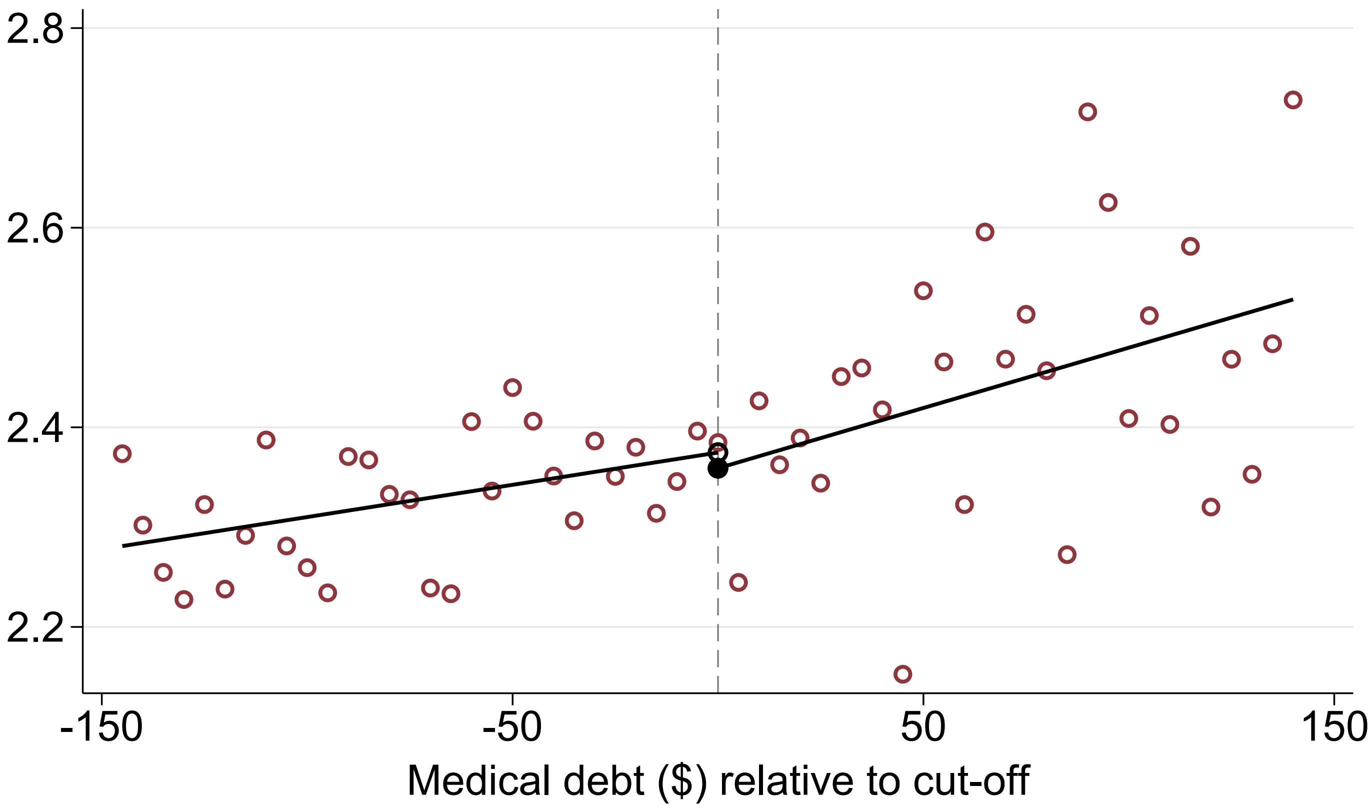


(e) Number of New Mortgages in Last 6 Months



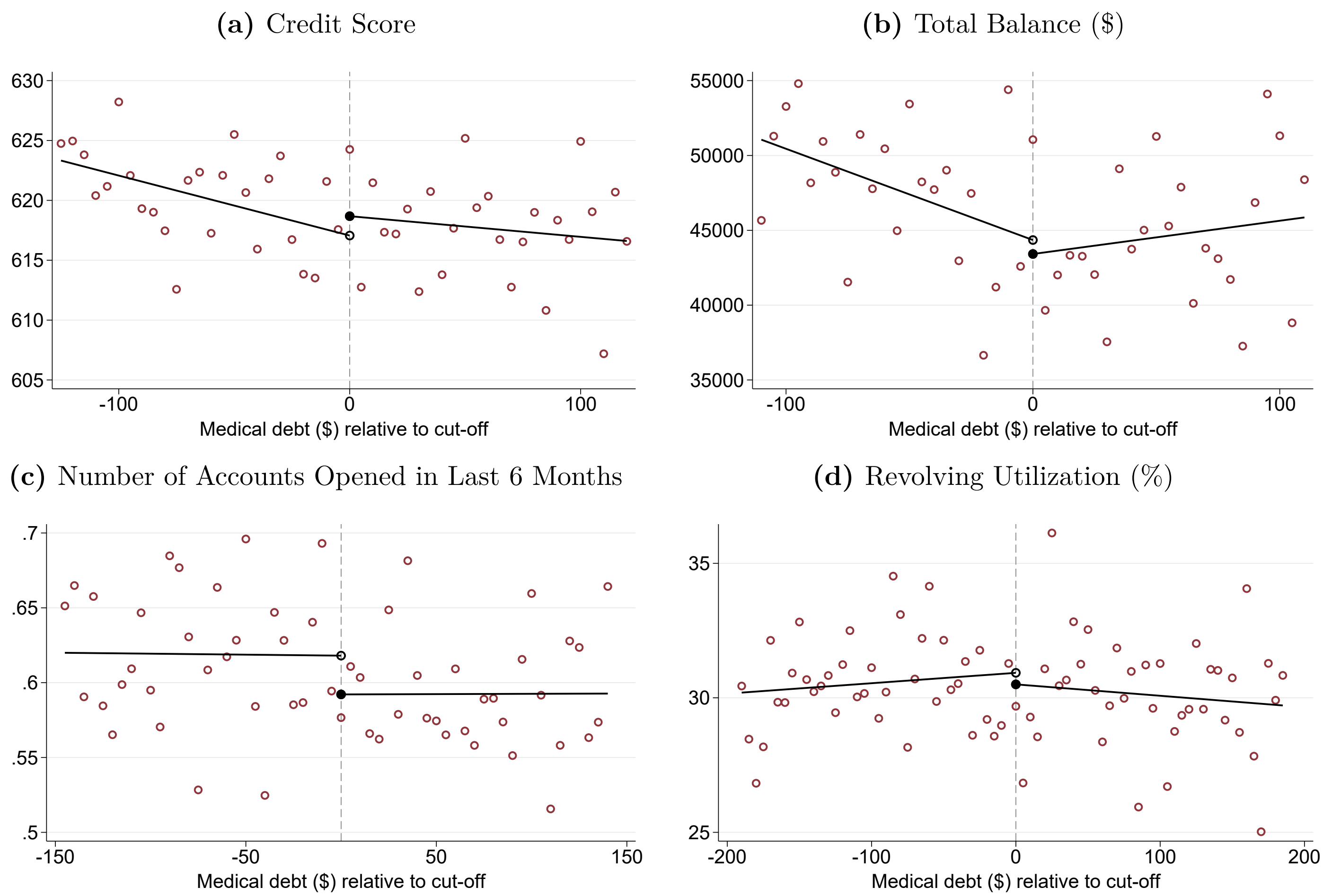
Notes: This figure shows the relationship between 2022 medical debt and five supplementary credit outcomes. Medical debt is defined as the maximum value of the consumer's 2022 medical collections accounts, measured relative to the $500 threshold. Number of Accounts 90+ days Past Due is the number of credit accounts (trades) that are at least 90 days delinquent. Number of New Inquiries is the number of credit inquiries in the past 6 months. Revolving Limit is the total credit limit across all revolving accounts. New Mortgage Accounts is the number of mortgage trades opened in the prior 6 months. RD estimates from Equation (6) are reported in Table A.6.

Figure A.7: 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 collections accounts per person in 2022. The medical debt running variable is defined as the maximum value of the consumer's 2022 medical collections accounts, measured relative to the $500 threshold. RD estimates from Equation (6) are reported in Table 5.

Figure A.8: Falsification Test: Access to Credit, 2022



(a) Credit Score

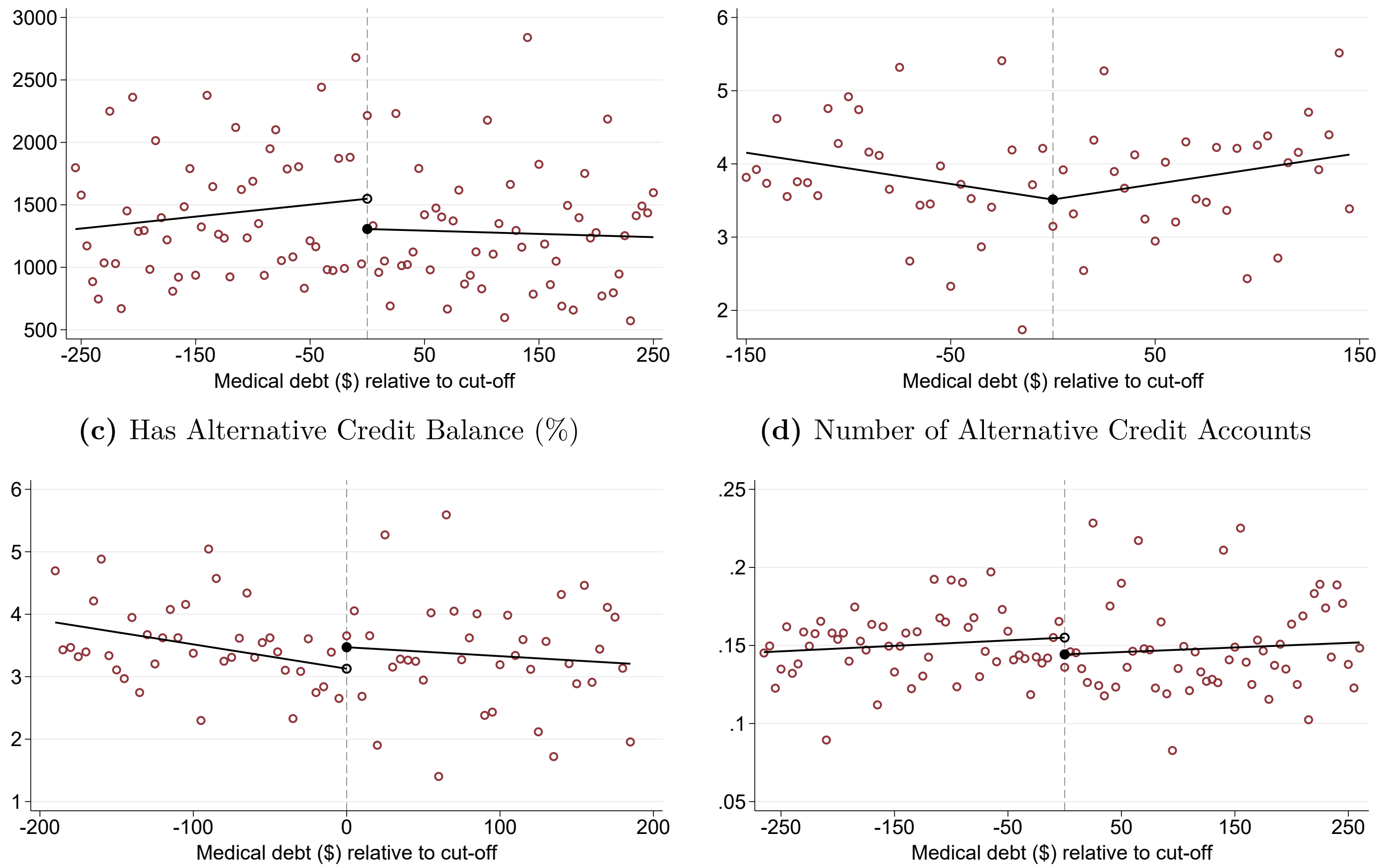
(b) Total Balance ($)

(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: consumer credit score, total balance, number of new credit accounts, and revolving utilization. Medical debt is defined as the maximum value of the consumer's 2022 medical collections accounts, measured relative to the $500 threshold. Credit Scores are based on the VantageScore model. Total Balance is the combined balance across all credit accounts. Number of New Credit Accounts is the number of new trades opened in the past 6 month. Revolving Utilization is the ratio of outstanding balances to credit limits across all revolving accounts. RD estimates from Equation (6) are reported in Table A.7.

Figure A.9: Falsification Test: Financial Distress and Access to Alternative Credit, 2022



(a) Total Balance 90+ Days Past Due ($)

(b) Bankruptcy in Last 7 Years (%)

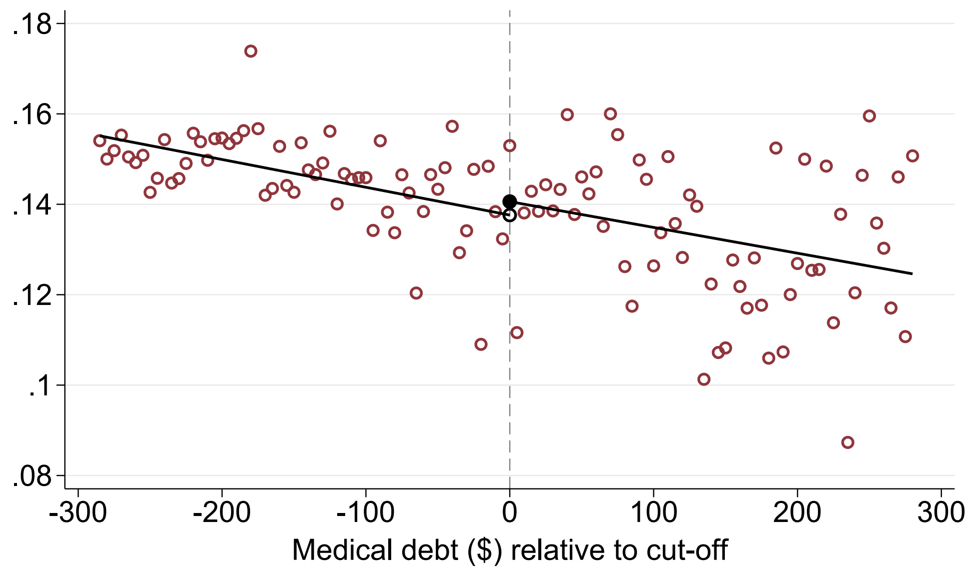
(c) Has Alternative Credit Balance (%)

(d) Number of Alternative Credit Accounts

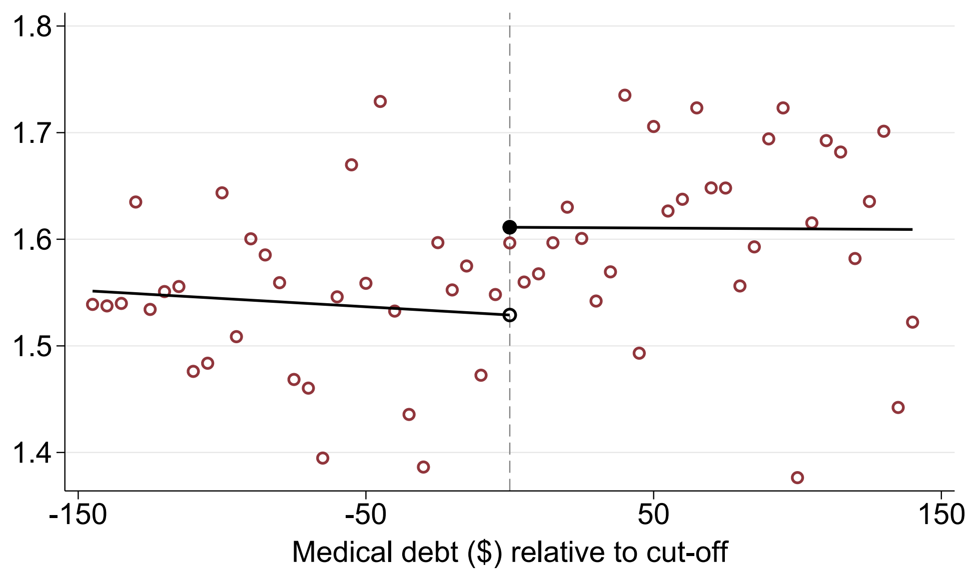
Notes: This figure shows the relationship between 2022 medical debt and measures of delinquency, bankruptcy, and alternative credit use. Medical debt is defined as the maximum value of the consumer's 2022 medical collections accounts, measured relative to the $500 threshold. Total Balance 90+ Days Past Due is the total balance of credit accounts that are at least 90 days delinquent. Bankruptcy in Last 7 Years indicates the percentage of consumers with an active bankruptcy flag. Has Alternative Credit Balance is the percentage of consumers holding a positive alternative credit balance. Number of Alternative Credit Accounts is the total count of alternative credit accounts in the consumer's Clarity report. RD estimates from Equation (6) are reported in Table A.7.

Figure A.10: Placebo Test: Two-Year Evolution of 2020 Medical Collections Accounts

(a) Share of Surviving Accounts

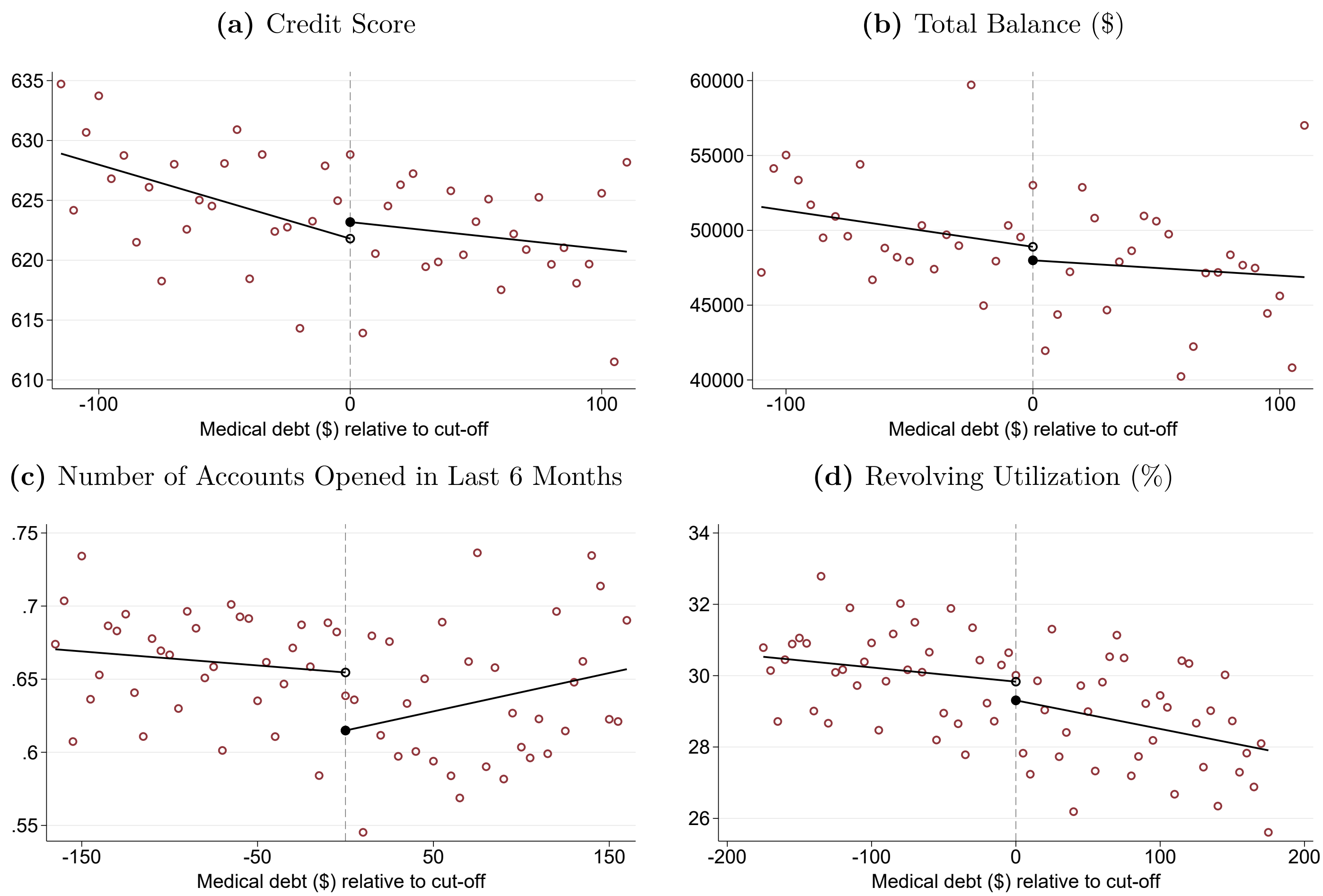


(b) Average Number of Accounts per Person



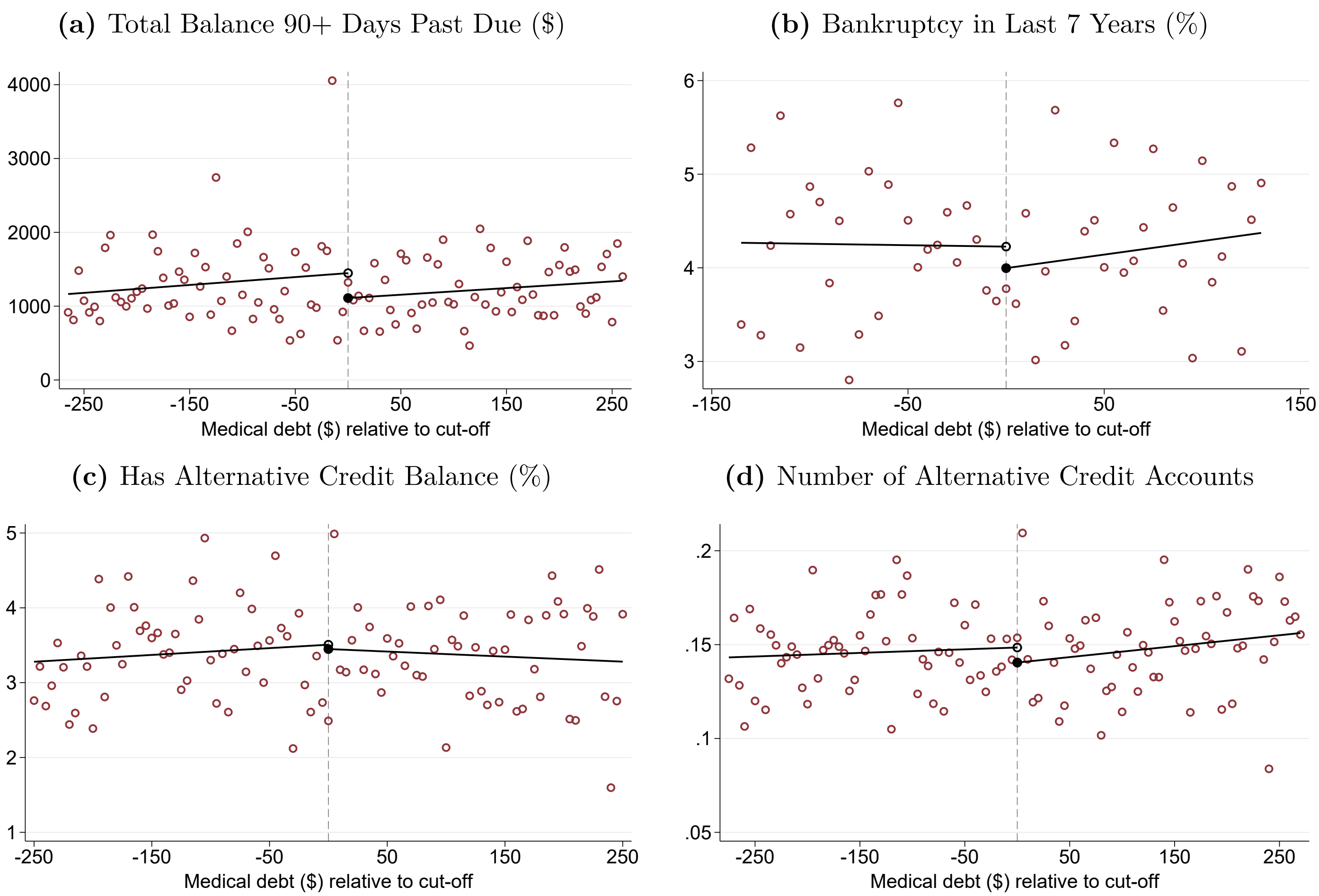
Notes: Panel (a) shows the proportion of 2020 medical collection accounts which remain present on 2022 credit reports by account amount, where the amount is measured as distance from the $500 threshold. Panel (b) shows the average number accounts per person. In panel (b), the running variable is equal to the maximum value of the consumer's medical collections accounts. RD estimates from Equation (6) are reported in Table A.8.

Figure A.11: Placebo Test: Access to Credit



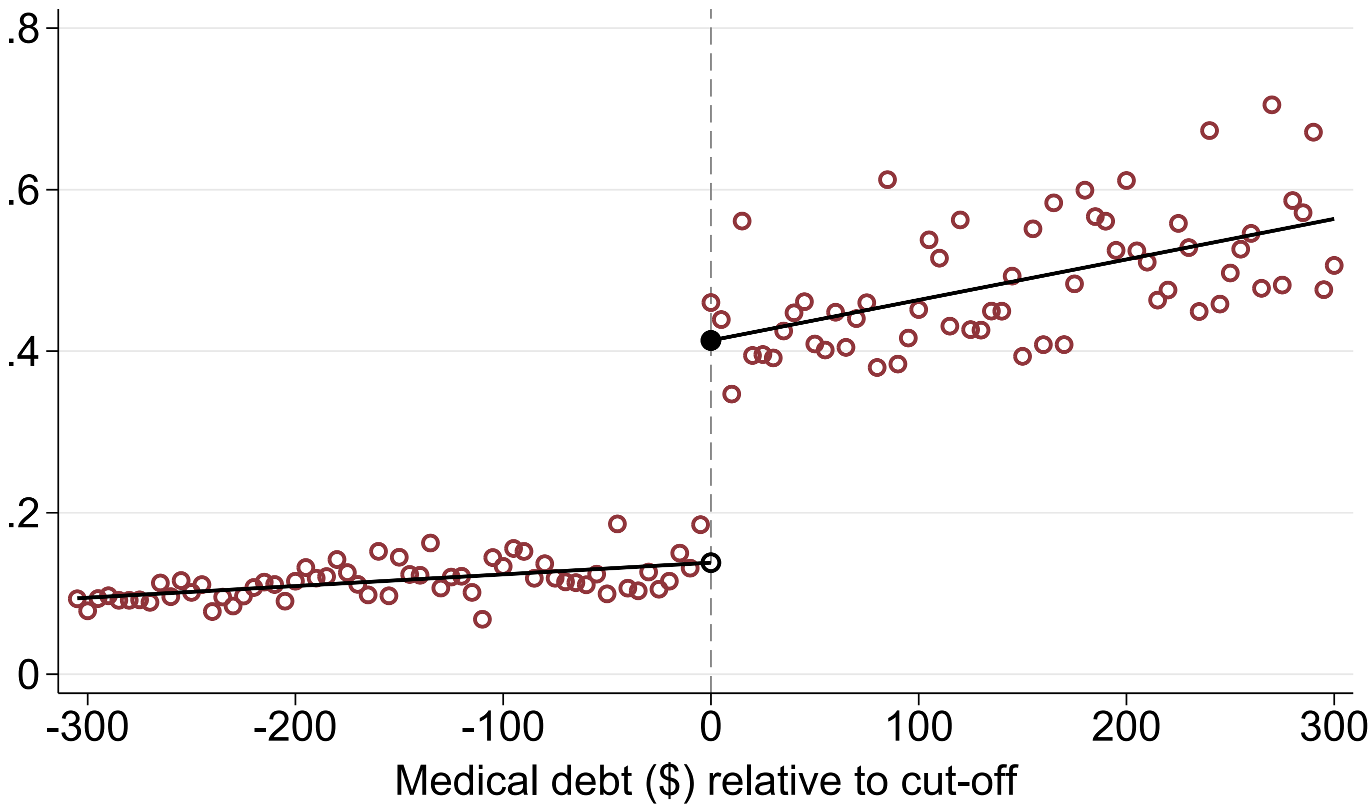
Notes: This figure shows the average consumer credit score, total balance, revolving limit, and revolving utilization in 2022 by medical debt, relative to the $500 threshold. The running variable for medical debt corresponds to the highest debt amount across the consumer's medical collections accounts in 2020. Credit Score are credit scores according to the VantageScore model. Total Balance is the total balance across all credit accounts. New Trades is the number of trades opened in the prior 6 month. Revolving Utilization is the balance-to-limit ratio across all revolving credit accounts. RD estimates from Equation (6) are reported in Table A.8.

Figure A.12: Placebo Test: Financial Distress and Access to Alternative Credit



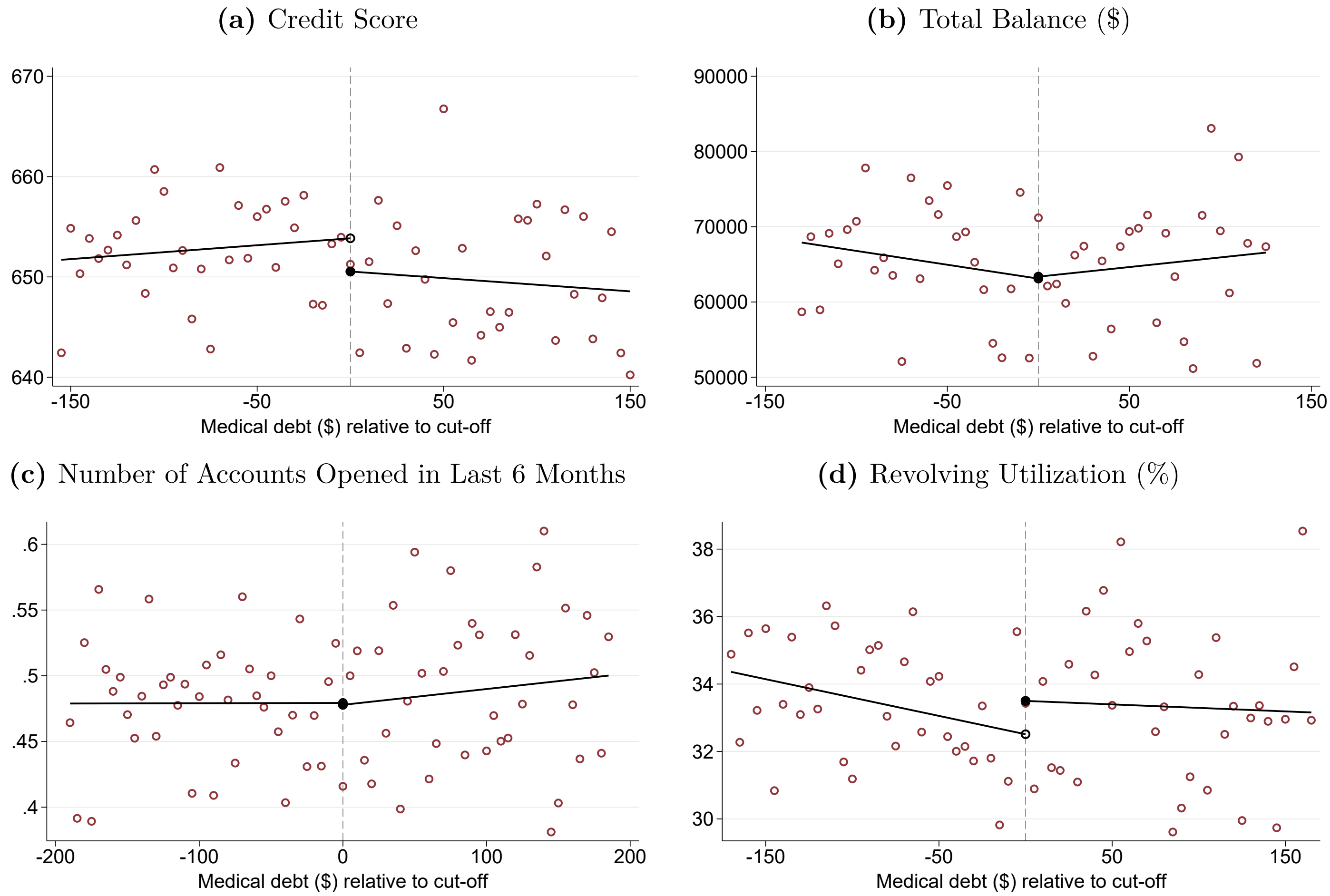
Notes: This figure shows the relationship between 2022 medical debt and measures of delinquency, bankruptcy, and alternative credit use. Medical debt is defined as the maximum value of the consumer's 2022 medical collections accounts, measured relative to the $500 threshold. Total Balance 90+ Days Past Due is the total balance of credit accounts that are at least 90 days delinquent. Bankruptcy in Last 7 Years indicates the percentage of consumers with an active bankruptcy flag. Has Alternative Credit Balance is the percentage of consumers holding a positive alternative credit balance. Number of Alternative Credit Accounts is the total count of alternative credit accounts in the consumer's Clarity report. RD estimates from Equation (6) are reported in Table A.8.

Figure A.13: Medical Collections Sub-Sample: Average Number of Accounts per Person, 2022



Notes: This figure shows the average number of medical collections accounts per person in 2024, where the running variable is the maximum value of the consumer's 2022 medical collections account, measured relative to the $500 threshold. The medical collections sub-sample restricts the RD sample to consumers for whom the total number of accounts in collection is same as the total number of medical collections accounts. RD estimates from Equation (6) are reported in Table A.9.

Figure A.14: Medical Collections Sub-Sample: Access to Credit



(a) Credit Score

(b) Total Balance ($)

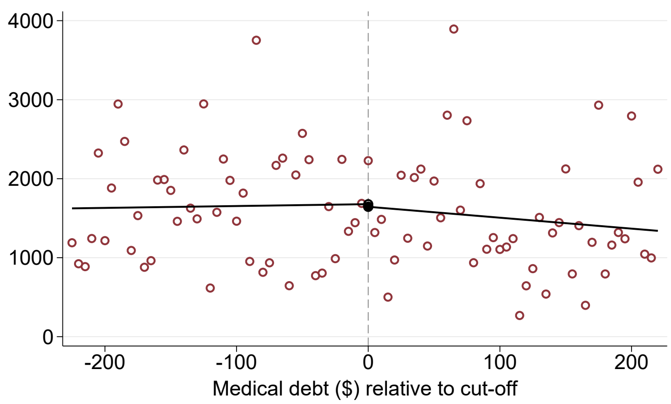
(c) Number of Accounts Opened in Last 6 Months

(d) Revolving Utilization (%)

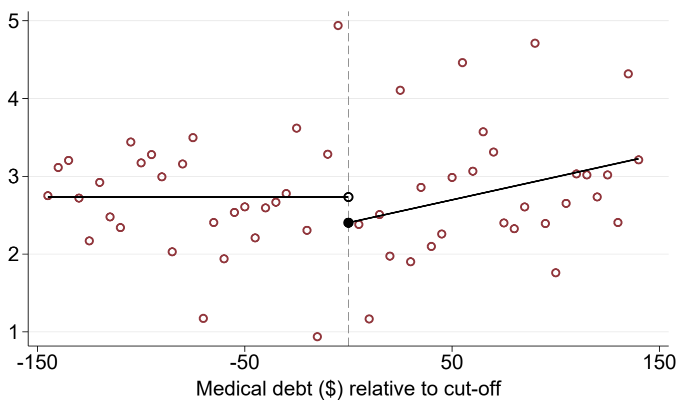
Notes: This figure shows the average consumer credit score, total balance, revolving limit, and revolving utilization in 2022 by medical debt, relative to the $500 threshold. The medical collections sub-sample restricts the RD sample to consumers for whom the total number of accounts in collections is same as the total number of medical collection accounts. The running variable for medical debt corresponds to the highest debt amount across the consumer's medical collections accounts in 2020. Credit Score are credit scores according to the VantageScore model. Total Balance is the total balance across all credit accounts. New Trades is the number of trades opened in the prior 6 month. Revolving Utilization is the balance-to-limit ratio across all revolving credit accounts. RD estimates from Equation (6) are reported in Table A.9.

Figure A.15: Medical Collections Sub-Sample: Financial Distress and Access to Alternative Credit

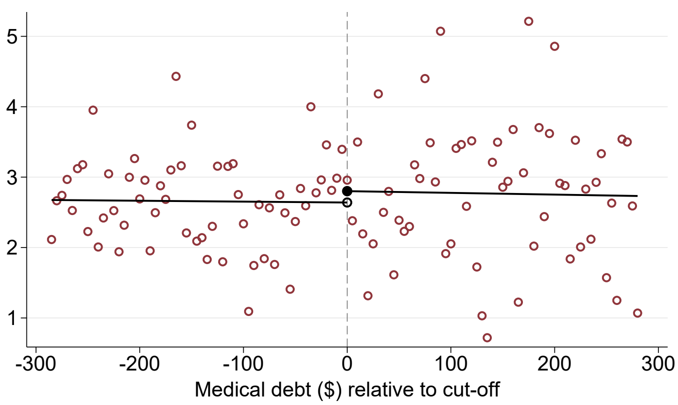
(a) Total Balance 90+ Days Past Due ($)



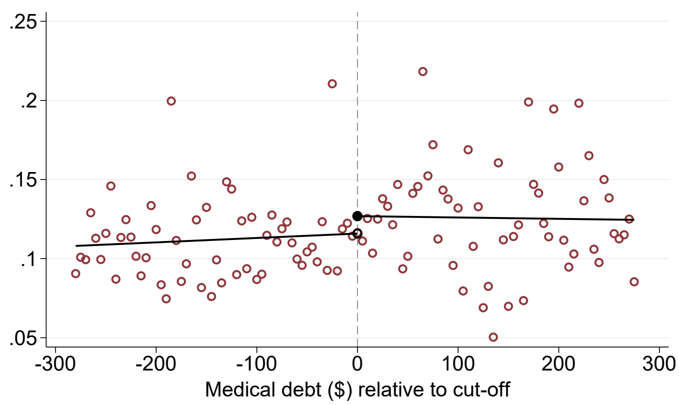
(b) Bankruptcy in Last 7 years (%)



(c) Has Alternative Credit Balance (%)



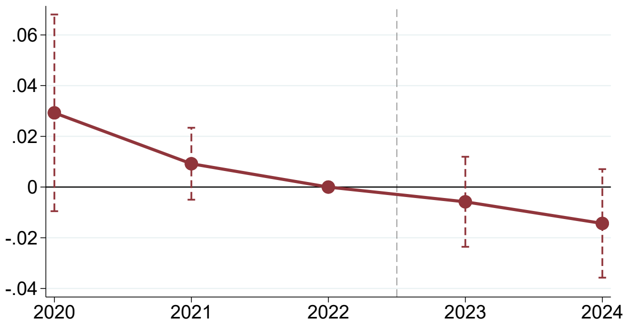
(d) Number of Alternative Credit Accounts



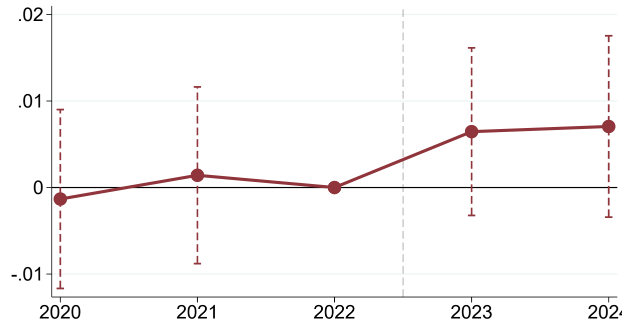
Notes: This figure shows the relationship between 2022 medical debt and measures of delinquency, bankruptcy, and alternative credit use. Medical debt is defined as the maximum value of the consumer's 2022 medical collections accounts, measured relative to the $500 threshold. The medical collections sub-sample restricts the RD sample to consumers for whom the total number of accounts in collections is same as the total number of medical collection accounts. Balance 90+ Days Past Due is the total balance of credit accounts that are at least 90 days delinquent. Share Bankrupt indicates the percentage of consumers with an active bankruptcy flag. Share with Positive Alternative Credit Balance is the percentage of consumers holding a positive alternative credit balance. Number of Alternative Credit Accounts is the total count of alternative credit accounts in the consumer's Clarity report. RD estimates from Equation (6) are reported in Table A.9.

Figure A.16: Effect of Removing Small Medical Collections on Additional Outcomes

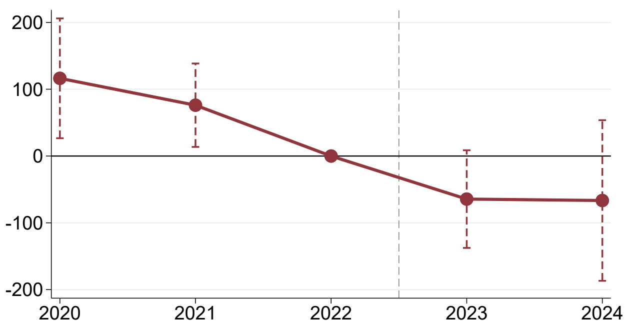
(a) Number of Accounts 90+ Days Past Due



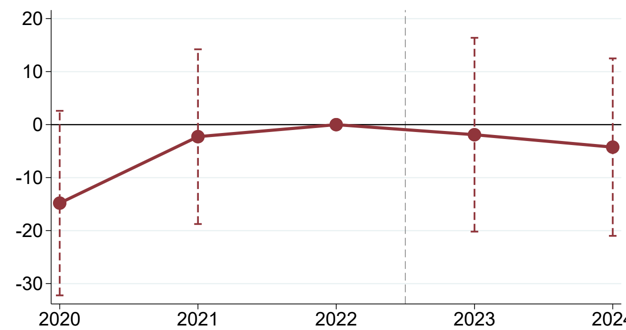
(b) Number of Inquiries in Last 6 Months



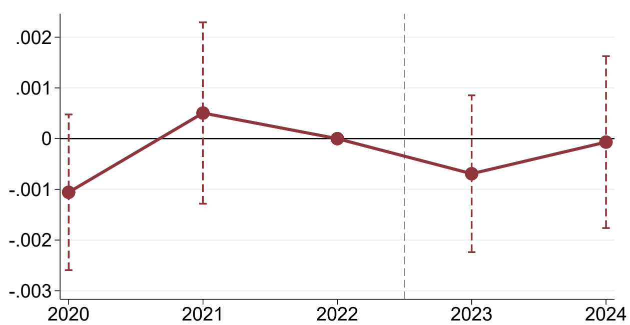
(c) Revolving Limit ($)



(d) Alternative Credit Balance ($)



(e) Number of New Mortgages in Last 6 Months



This figure shows coefficient estimates and 95% confidence intervals of Equation (8) for the outcomes denoted in panel captions. Trades 90 days+ Past Due is the number of trades 90 days or more past due. New Inquiries is the number of inquiries in the past 6 months. Revolving Limit is the total limit across all revolving accounts. New Mortgage Trades is the number of mortgage trades opened in the past 6 months. We create 1,000 equal-sized bins of the difference in predicted probability of default in our full sample and cluster our standard errors at the bin level. Our differences-in-differences sample corresponds to 10% of the full sample, and we are thus left with 100 clusters.

Table A.3: Performance Metrics for Credit Scoring Models With and Without Medical Collections Versus a Random Variable

	(1)	(2)	(3)
	All variables	Exclude Medical Debts < \$500	Include Random Variable
Accuracy	0.905	0.906	0.905
Recall	0.448	0.448	0.445
Precision	0.736	0.736	0.737
F1 Score	0.557	0.557	0.555
AUC	0.712	0.712	0.710

Notes: This table reports performance metrics for a credit scoring model predicting defaults occurring between 2020 and 2021, based on 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 a random variable is included to the set of predictors. The accuracy score represents the share of correct predictions. For comparison, a naive model predicting no defaults achieves an accuracy of 0.867. Precision is the proportion of predicted defaults that were correctly classified. Recall is the proportion of actual defaults correctly classified. F1 score is the harmonic mean of Precision and Recall. The AUC (Area Under the Receiver Operating Characteristic Curve) indicates the probability that the model assigns a higher default probability to a true defaulter than to a non-defaulter.

Table A.4: Performance Metrics for Credit Scoring Models With and Without Medical Collections: Restricted Model

	(1)	(2)	(3)
	Restricted Model	Exclude Medical Debts < \$500	Exclude All Medical Debts
Accuracy	0.8669	0.8670	0.8670
Recall	0.0008	0.0004	0.0001
Precision	0.3462	0.3250	0.5000
F1 Score	0.0016	0.0008	0.0002
AUC	0.5003	0.5001	0.5001

Notes: This table reports performance metrics for a credit scoring model predicting defaults occurring between 2020 and 2021, based on borrower characteristics from 2019. Column (1) presents metrics for the a model with six predictors—medical collections below and above \$500, bankruptcy trades, bankruptcy trades in the past 24 months, tax liens in the past 24 months, and judgments trades in the past 24 months.—estimated using XGBoost. Column (2) reports metrics when small (under \$500) medical collections are excluded from the set of six predictors. Column (3) shows metrics when all medical collections are excluded. The accuracy score represents the share of correct predictions. For comparison, a naive model predicting no defaults achieves an accuracy of 0.867. Precision is the proportion of predicted defaults that were correctly classified. Recall is the proportion of actual defaults correctly classified. F1 score is the harmonic mean of Precision and Recall. The AUC (Area Under the Receiver Operating Characteristic Curve) indicates the probability that the model assigns a higher default probability to a true defaulter than to a non-defaulter.

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

	(1)	(2)	(3)
	Income ($1,000)	Age (years)	Female (%)
ABOVE^{2024}	0.0663	0.512	–0.00970
	[–1.04, 1.44]	[–0.314, 1.17]	[–0.0301, 0.00990]
Sample mean	42.3	46.3	0.552
% of Mean	0.157	1.11	–1.76
Bandwidth	109	141	254
Observations	271,305	271,305	263,895

Notes: This table shows the coefficient (β) estimates and 95% confidence intervals of Equation (6). The running variable for medical debt corresponds to the highest debt amount across the consumer's medical collections accounts. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Sample Mean reports the mean of the dependent variable in 2024. A ***, **, and * indicate significance at the 1%, 5%, and 10% level respectively, using conventional inference.

Table A.6: Additional Credit Outcomes: RD Estimates of the Direct Effect of Medical Debt 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^{2024}	–0.0146	–0.0167	–635	6.72	0.000727
	[–0.0698, 0.0334]	[–0.0520, 0.0198]	[–1,710, 204]	[–49.8, 49.0]	[–0.00459, 0.00509]
Sample mean	0.490	0.556	6,391	164	0.00909
% of Mean	–2.97	–3.01	–9.94	4.09	8.01
Bandwidth	255	212	114	236	174
Observations	271,305	271,305	271,305	271,305	271,305

Notes: This table shows the coefficient (β) estimates and 95% confidence intervals of Equation (6). The running variable for medical debt corresponds to the highest debt amount across the consumer's medical collections accounts. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Sample Mean reports the mean of the dependent variable in 2024. A ***, **, and * indicate significance at the 1%, 5%, and 10% level respectively, using conventional inference.

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of Debts	Number of Medical Debts	Credit Score	Total 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^{2024}	–0.0875 [–0.195, 0.0644]	–0.0310 [–0.128, 0.0672]	1.99 [–2.85, 6.70]	427 [–4,038, 6,226]	–0.0255 [–0.0840, 0.0285]	–0.0617 [–1.95, 1.67]	–141 [–677, 492]	0.0906 [–0.809, 1.08]	0.457 [–0.174, 1.36]	–0.00521 [–0.0302, 0.0240]
Sample mean	3.45	2.37	619	46,421	0.607	30.3	1,360	3.85	3.45	0.149
% of Mean	–2.53	–1.31	0.321	0.921	–4.20	–0.203	–10.4	2.35	13.2	–3.50
Bandwidth	193	143	124	111	144	187	251	150	188	264
Observations	271,305	271,305	271,305	271,305	271,305	271,305	271,305	271,305	271,305	271,305

Notes: This table shows the coefficient (β) estimates and 95% confidence intervals of Equation (6). The running variable for medical debt corresponds to the highest debt amount across the consumer's medical collections accounts. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Sample Mean reports the mean of the dependent variable in 2024. A ***, **, and * indicate significance at the 1%, 5%, and 10% level respectively, using conventional inference.

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of Debts	Number of Medical Debts	Credit Score	Total 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^{2024}	0.0287 [–0.0775, 0.184]	0.0631 [–0.00398, 0.166]	1.15 [–3.78, 5.09]	–433 [–5,948, 4,064]	–0.0333 [–0.0779, 0.0230]	–0.660 [–2.17, 1.04]	–322 [–853, 168]	–0.260 [–1.30, 0.575]	0.0933 [–0.437, 0.803]	–0.000989 [–0.0242, 0.0268]
Sample mean	2.59	1.57	624	48,913	0.651	29.6	1,262	4.22	3.35	0.146
% of Mean	1.11	4.02	0.185	–0.886	–5.11	–2.23	–25.5	–6.17	2.79	–0.676
Bandwidth	154	145	111	110	164	176	264	134	250	274
Observations	322,756	322,756	322,756	322,756	322,756	322,756	322,756	322,756	322,756	322,756

Notes: This table shows the coefficient (β) estimates and 95% confidence intervals of Equation (6). The running variable for medical debt corresponds to the highest debt amount across the consumer's medical collections accounts. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Sample Mean reports the mean of the dependent variable in 2024. A ***, **, and * indicate significance at the 1%, 5%, and 10% level respectively, using conventional inference.

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Number of Debts	Number of Medical Debts	Credit Score	Total 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^{2024}	0.259***	0.283***	–3.07	3,409	–0.00518	1.04	173	–0.597	–0.0284	0.00930
	[0.201, 0.327]	[0.247, 0.312]	[–10.1, 1.98]	[–4,032, 12,609]	[–0.0654, 0.0464]	[–1.20, 3.77]	[–493, 1,096]	[–1.85, 0.333]	[–0.901, 0.804]	[–0.0231, 0.0419]
Sample mean	0.568	0.228	652	65,542	0.481	33.4	1,593	2.58	2.68	0.116
% of Mean	45.6	124	–0.471	5.20	–1.08	3.12	10.8	–23.2	–1.06	8.04
Bandwidth	237	306	150	128	190	169	223	145	285	279
Observations	149,185	149,185	149,185	149,185	149,185	149,185	149,185	149,185	149,185	149,185

Notes: This table shows the coefficient (β) estimates and 95% confidence intervals of Equation (6). The medical collections sub-sample restricts the RD sample to consumers for whom the total number of accounts in collections is same as the total number of medical collection accounts. The running variable for medical debt corresponds to the highest debt amount across the consumer's medical collections accounts. We report MSE-optimal estimates with robust, bias-corrected 95% confidence intervals in brackets. Sample Mean reports the mean of the dependent variable in 2024. A ***, **, and * indicate significance at the 1%, 5%, and 10% level respectively, using conventional inference.