The Direct and Indirect Effects of Information Deletion: Evidence from Medical Collections*

Victor Duarte
University of Illinois

Julia Fonseca
University of Illinois

Divij Kohli
University of Illinois

Julian Reif
University of Illinois and NBER

February 2025

PRELIMINARY AND INCOMPLETE

Abstract

We study the effect of deleting medical collections from credit reports on consumer credit access and financial health. In April 2023, the three major credit bureaus announced that they had ceased reporting information about medical collections below $500. Using a machine learning model, we show that these small medical collections have minimal predictive value for defaults, suggesting their deletion should have little to no effect on credit origination. We confirm this prediction using two separate research designs. First, using a regression discontinuity design that compares individuals just above and below the $500 threshold, we find no direct benefits from the information deletion, ruling out even small improvements in access to credit, repayment behavior, or subprime borrowing. Second, to assess potential indirect effects, we use our prediction model to classify consumers based on whether their predicted probability of default increases or decreases when medical collections are removed from the model. A difference-in-differences analysis comparing these groups before and after the April 2023 policy change reveals no evidence of negative spillover effects. Our findings indicate that the 2023 decision to delete small medical collections had no adverse indirect effects, but also did not benefit consumers whose information was removed. Finally, we show that large ($\geq$ $500) medical collections are also not meaningfully predictive of default, suggesting that the CFPB's proposal to delete remaining medical collections from credit reports would likely have no significant impacts.

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

1 Introduction

One in seven Americans carry medical debt (U.S. Census Bureau, 2022). When medical bills go unpaid, they may be sent to collection agencies and reported to credit bureaus. Policymakers have raised growing concerns about these reporting practices, noting 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 reporting medical collections below $500.

Building on this measure, the Consumer Financial Protection Bureau (CFPB) proposed a rule in June 2024 to eliminate all remaining medical collections from credit reports, arguing that this change would enhance credit access and improve terms for consumers burdened with medical debt. To address concerns that deleting this information could harm consumers without medical debt, the CFPB cited evidence suggesting that medical collections are poor predictors of future default and thus provide limited value for credit origination decisions (Consumer Financial Protection Bureau, 2024). However, research in other credit contexts has shown that information deletion can reduce borrowing opportunities for certain groups (Liberman et al., 2019). Moreover, if medical collections truly lack predictive power, their removal may not improve credit access, even for those with medical debt.

Against this backdrop, this paper studies the direct and indirect effects of deleting information about medical collections from credit reports, using 2019–2024 data from the Gies Consumer and Small Business Credit Panel (GCCP). Specifically, we investigate the direct benefits to consumers whose medical collection information is deleted, as well as the potential negative spillovers to other consumers. Using machine learning techniques, we build credit scoring models to evaluate the predictive value of medical collections for default risk prior to 2023. We compare two models: one trained on borrower credit history that includes medical collections below $500, and another trained on histories excluding these small medical collections. The results reveal that excluding small medical collections has no meaningful effect on default prediction, highlighting their minimal value for credit origination decisions. Thus, the April 2023 intervention that deleted this information is unlikely to affect credit access, either directly or indirectly. Furthermore, we show that large medical collections also

have minimal predictive value, suggesting that eliminating all medical collections from credit reports is similarly unlikely to influence credit access or financial health.

We confirm this insight using two distinct research designs. First, we employ a Regression Discontinuity (RD) approach to estimate the direct effects of removing small medical collections. Comparing individuals just above and below the $500 information deletion threshold, we find no evidence that information deletion improved credit access or financial health. Our null estimates are precise: the 95% confidence intervals rule out increases in credit scores greater than 3.62 points (0.58%) and increases in the balance-to-limit ratio of revolving credit exceeding 1.54 percentage points (4.8%).

Next, we use a difference-in-differences approach to estimate the indirect effects of removing small medical collections from credit reports on consumers without these collections. Similar to Liberman et al. (2019), we use our two default prediction models—one incorporating medical collections below $500 and the other excluding them—to identify two groups: consumers whose predicted probability of default increases by approximately 2 percentage points (the 95th percentile of the distribution) or more when medical collections below $500 are removed from the model, and those whose predicted probability falls by 2 percentage points or more. We show that these two groups are observationally similar across a wide range of characteristics and exhibited parallel trends prior to the 2023 information deletion.[1] Consistent with our RD results, we find no evidence of negative spillover effects from the April 2023 deletion, with precise estimates that again rule out small effects.

We conclude that the 2023 decision to delete small medical collections from credit reports had no negative indirect effects and offered no measurable benefits to those whose information was removed. These findings suggest that policies promoting information deletion may fail to achieve their intended goals, particularly for consumers burdened by medical debt. Our study highlights the importance of exploring alternative strategies to support financially vulnerable consumers more effectively.

We make several contributions to the literature. First, we break new ground by investi-

[1]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 collections serves as a noisy partitioning mechanism, effectively creating two randomly assigned groups. By contrast, Liberman et al. (2019)—who study a setting where informative data was deleted—use unaffected individuals as their control group.

gating both the direct and indirect effects of information deletion in the context of medical debt. Prior studies in other contexts have documented mixed effects. Liberman et al. (2019) study the deletion of default information in Chile, finding that it increases borrowing for consumers whose predicted probability of default decreases after deletion, but reduces borrowing for those reclassified as higher default risks. Similarly, Jansen et al. (2024) find that removing bankruptcy flags directly benefits affected consumers through lower interest rates, while raising rates for those with no history of bankruptcy, resulting in a small decline in social surplus.[2] Agan and Starr (2017) study the effect of removing criminal history information from job applications, finding that it reduces callbacks for Black applicants. Similarly, Bartik and Nelson (2024) show that bans on employers' use of credit report data result in lower job-finding rates and more involuntary separations for Black workers. In contrast to these studies, we show that the deleted information in our setting—small medical collections—has minimal predictive value. As a result, we find 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 alternative often discussed alongside information deletion. Kluender et al. (2024) conduct two large-scale randomized experiments on medical debt forgiveness and find no impact on credit access, utilization, and financial distress. Together, their findings and ours suggest that neither deleting medical debt information nor forgiving the debt itself alleviates financial distress for people burdened by it. These results contrast with evidence 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 the use of machine learning techniques to better understand credit markets. We are the first to demonstrate that medical collections—regardless of size—are poor predictors of default risk. Nonetheless, we also show that noisy information like medical collections can still influence credit scores for consumers with thin credit files, a phenomenon previously highlighted by Blattner and Nelson (2022). Prior studies have

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

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 predictive performance than prior studies across multiple metrics.

Finally, we contribute to a growing literature that studies debts in collections and the 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). Closest to our paper is Batty et al. (2022), who show that expanding health insurance coverage reduces medical debts in collection, but does not improve other financial outcomes. Similar to Fonseca (2023), we study both mainstream and subprime credit outcomes by linking traditional credit reports from one of the major credit bureaus to reports from Clarity, a bureau specializing in alternative financial services. This linkage allows us to capture 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 collections are predictive of default. Section 4 presents our RD results on the direct effect of deleting medical collections. Section 5 reports estimates of the indirect effect from 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 Experian, one of the three major U.S. credit reporting agencies. The GCCP features a one-percent random sample of individuals with a credit report, linked to alternative credit records from Experian's alternative credit bureau, Clarity Services, and business credit records for indi-

viduals who own a business.[3] Spanning 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, allowing for natural flows into the panel as new Social Security numbers are issued and out of the panel through death or inactivity. This design ensures that the sample reflects the dynamics of the underlying population.

The GCCP provides detailed debt information at the credit account ("tradeline") level, such as amounts owed for mortgages, student loans, and credit cards. It also provides individuals' VantageScore credit scores, along with payment histories, bankruptcies, judgments, and other public records. Finally, the data include information on the individual's 5-digit zip code of residence, age, and gender, with a subset of observations containing data on education, marital status, homeownership status, and occupation. We classify a medical collection as one where the creditor is categorized as Medical/Health Care or the furnisher is identified as a business in the medical or health-related sector.[4] In Table A.1, 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 yields 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 (Clarity) 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

[3]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.

[4]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.

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. Initially, healthcare providers attempt to recover unpaid amounts directly from patients. If these efforts are unsuccessful, the debt may be transferred to third-party collection agencies, which employ various strategies to secure payment. These strategies include initiating lawsuits to obtain court judgments for repayment and reporting unpaid debts to credit bureaus.[5] 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 are prohibited from reporting 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

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

(Kluender et al., 2024). 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 predict default, some lenders may incorporate this information into their in-house credit scoring models. Deleting this information could alter their origination decisions and limit access to credit for some consumers. Conversely, if medical collections have limited predictive value, their removal should not affect credit underwriting, even for consumers whose information is deleted.

Not all lenders use in-house credit scoring models; some rely exclusively on scores provided by 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 common credit scores. VantageScore stopped including medical collections below $500 in its model in January 2023, followed by FICO a few months later.[6] As a result, the deletion of small medical collections should not affect the origination decisions of lenders relying solely on these models.

For lenders that develop proprietary credit scoring models, including medical collections depends on whether this information helps predict defaults. To examine this, we begin

[6]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.

by investigating the predictive power of medical collections. Specifically, we simulate the consequences of the April 2023 decision to delete medical collections below $500 by training two credit scoring models: one that includes data on medical collections below $500 and one that excludes it. While our models are not identical to those used by any specific lender, they rely on similar data and algorithms and demonstrate performance comparable to other models in the literature. This approach assumes that the output of our models is correlated with those of other well-designed credit scoring models.

As we demonstrate below, we find that medical collections below $500 are not meaningfully predictive of default. This implies that the April 2023 intervention should have no direct or indirect effects on credit access or terms. We test these predictions in Sections 4 and 5, respectively. Finally, we show that medical collections above $500 are also not predictive of default, suggesting that the CFPB's proposed rule to delete all remaining medical collections is unlikely to produce substantial effects on credit access and terms.

3.3 Credit scoring with and without medical collections

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

$$Y = f(X_1, X_2, ...X_n) + e$$

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 $n = 48$ 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.[7] 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.[8] Using data from 2019 to 2021—prior to the removal of information on medical collections below $500—we predict defaults 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 87%, 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,

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

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

are more relevant for evaluating performance (Davis and Goadrich, 2006).

Precision and recall are defined as:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{1}$$

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

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

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

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., Agarwal et al. (2023), Butaru et al. (2016)). Two exceptions are Khandani et al. (2010) and Chioda et al. (2024), who achieve recalls of 0.654 and 0.478, respectively, when predicting default over shorter time horizons of 3 months and 6–9 months ahead.[9] Shorter prediction windows, such as 3, 6, or 9 months, generally yield higher precision and recall compared to longer horizons such as our 24-month window, which likely explains the superior performance of the models in Khandani et al. (2010) and Chioda et al. (2024).

Our precision score of 0.736 also ranks highly, exceeding many prior studies, which often

[9]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. Chioda et al. (2024) use a 20% threshold, which raises recall and decreases precision relative to a 50% threshold.

report values between 0.06 and 0.50 (e.g., Fuster et al. (2022), Agarwal et al. (2023), Butaru et al. (2016)). Khandani et al. (2010) is again an exception, obtaining a precision of 0.853 when predicting default over a 3-month horizon. Precision and recall are more appropriate performance metrics than AUC in settings with rare outcomes. Nevertheless, our AUC of 0.712 lies within the typical range of 0.66 to 0.88 obtained in the literature, further demonstrating the effectiveness of our model.

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 CFPB's proposed 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.[10] 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 2 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.[11]

Although Figure 2 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 rather than meaningful differences in default risk. To investi-

[10] 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.

[11] 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/).

gate 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.

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 3 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 3 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 prob-

abilities 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. 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 medical collections under $500. This finding underscores the limited predictive value of small medical collections.[12]

3.4 Robustness: credit scoring with and without random noise

Next, we compare the effect of deleting medical collections below $500 with deleting pure noise. To do so, we train a version of our model that includes a feature drawn randomly from a uniform distribution. We then compare the performance of that model to our baseline model, which excludes this random variable. The performance metrics, presented in Table A.2, indicate that while removing small medical collections doesn't impact model performance, removing the random variable slightly improves it, with the F1 Score increasing slightly from 0.557 to 0.555. This potentially suggests that small medical collections are not purely noise and do have non-zero predictive power. However, the changes in predicted default probabilities that we obtain by removing the noise variable are strikingly similar to those from removing small medical collections. Figure A.1 illustrates this similarity by overlaying the histogram of probability differences from Figure 2 with the corresponding histogram obtained by removing the noise variable. The patterns are very similar in both cases.

Moreover, Figure A.2 shows that removing noise sorts consumers across our three groups—

[12] 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.

unaffected, positively treated, and negatively treated—in a manner that mirrors the effects of removing medical collections under $500. Panel A reproduces the balance plot from Figure 3, where each dot represents the coefficient from a regression of the (standardized) variable labeled on the y-axis on either the positive (blue) or negatively treated (red) group indicator. Panel B presents an analogous plot, comparing models one with and without the noise 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).

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:

$$Y_j^{2024} = \alpha_1 DEBT_j^{2022} + \beta_1 ABOVE_j^{2022} + \gamma_1 (ABOVE_j^{2022} \times DEBT_j^{2022}) + \epsilon_j \quad (4)$$

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 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:

$$Y_i^{2024} = \alpha MAXDEBT_i^{2022} + \beta ABOVE_i^{2022} + \gamma (ABOVE_i^{2022} \times MAXDEBT_i^{2022}) + \epsilon_i \quad (5)$$

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 additional, unreported analyses, we estimated 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 Experian's 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 effect of the 2023 deletion on the number of accounts in collections. Panel A of Figure 4 indicates that this intervention reduced the total number of collections in 2024 by 0.24. Focusing specifically on medical collections, Panel B shows that the intervention reduced medical collections by 0.3 (107%). These findings confirm that the credit reporting agencies largely removed small collections accounts from their reports.

Figure 5 shows the effect of the 2023 deletion on measures of credit access and utilization. We do not observe any significant changes in the outcomes around the debt threshold. Table 5 presents formal estimates. Our 95% confidence intervals can rule out that deletion improved credit scores by 0.68 (0.11%), balances by $2,234 (4.34%), new trades by 0.03 (7.55%), and revolving utilization by 0.39 (1.22%).[13] Figure 6 shows the effect of the 2023 deletion on derogatories and alternative credit. Again, the 95% confidence intervals can rule out that deletion improved delinquency balance by $92.64 (4.89%), bankruptcy by 0.36 (11.55%), likelihood of having alternative credit balance by 0.38 (9.87%) and alternative trades by 0.01 (6.15%).

Under the assumptions of our RD design, outcomes unrelated to medical debt should remain unchanged at the threshold. We present a series of covariate smoothness tests in Figure A.3. In line with our expectations, we find no significant differences in average age, average income, or the proportion of individuals married at the threshold, for both 2022 and 2024.

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 onto consumers without pre-existing small medical collections. For this analysis, we define the treatment group as consumers whose predicted probability of default increases when medical collections under

[13] Note that we should not expect a mechanical effect on credit scores from removing small medical collections since medical collections have not been used in the VantageScore model since January 2023.

$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."[14] 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.[15]

We estimate the following regression model at the individual level:

$$Y_{ict} = \alpha + \beta TREATED_i \times POST_t + \lambda_i + \delta_{ct} + \epsilon_{it}, \tag{6}$$

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 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.

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

[15]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.

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 (6):

$$Y_{ict} = \alpha + \sum_{\substack{\tau=2020, \\ \tau \neq 2022}}^{2024} \beta_\tau TREATED_i \times \mathbb{I}_{t=\tau} + \lambda_i + \delta_{ct} + \epsilon_{it}, \tag{7}$$

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 7 presents estimates of the indirect effect of information deletion (Equation (7)) 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 8 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.[16] 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 (6), which assumes a constant treatment effect over time and aggregates years into longer periods to increase statistical power. Across the 8

[16] We show additional outcomes in Appendix Figure A.9.

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 direct and indirect effects of deleting medical collections from credit reports using a combination of machine-learning algorithms for predicting default, regression discontinuity, and differences-in-differences analysis. Contrary to intentions, we demonstrate that deleting medical collections from credit reports likely has no meaningful impact on credit access or terms. Our analysis focuses on medical collections under $500, which were removed from reports by all three major credit bureaus in 2023. First, we show that this information is not meaningfully predictive of default by comparing the output of two machine-learning credit scoring models: one with and one without this information.

We then confirm our hypothesis: if small medical records are not relevant for risk pricing, their deletion should not affect credit access and terms. Using a regression discontinuity design, we find no evidence that consumers benefit from the removal of derogatory 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 from the credit scoring model with observationally similar consumers whose predicted default probability declines. We find no evidence of negative spillover effects from 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 advocate for ex-ante solutions like health insurance, this approach has proven difficult to achieve in practice: about 30 million Americans remain without formal health insurance, and even the insured often face substantial out-of-pocket expenses (Einav and Finkelstein, 2023). Recent policy efforts have centered around alternative, ex-post solutions such as debt forgiveness and deleting medical debt from credit reports.

However, these alternatives may not deliver their intended benefit to consumers. For example, Kluender et al. (2024) find that purchasing medical debt in the secondary market and forgiving it has no impact on credit access and financial distress in two large-scale randomized experiments. Our findings add to 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 (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.

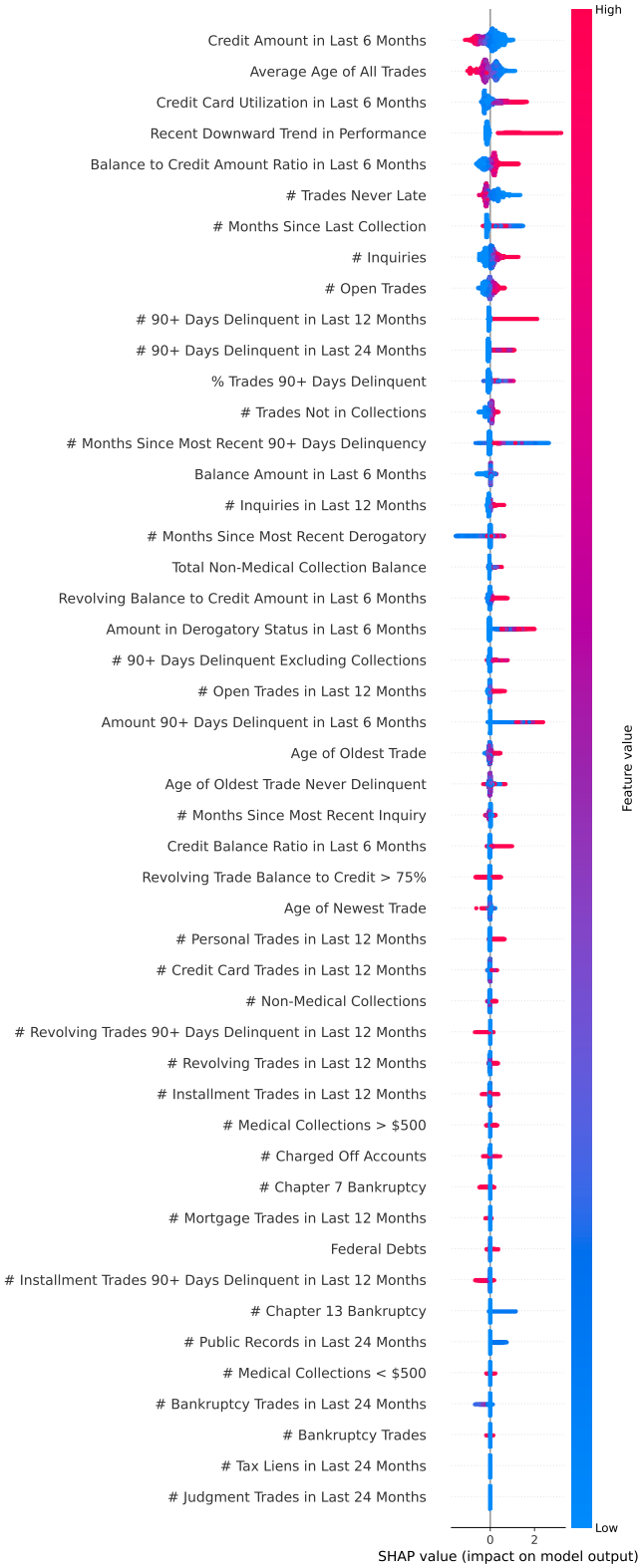
Musto, D. K. (2004). What happens when information leaves a market? evidence from postbankruptcy consumers. *The Journal of Business 77*(4), 725–748.

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.

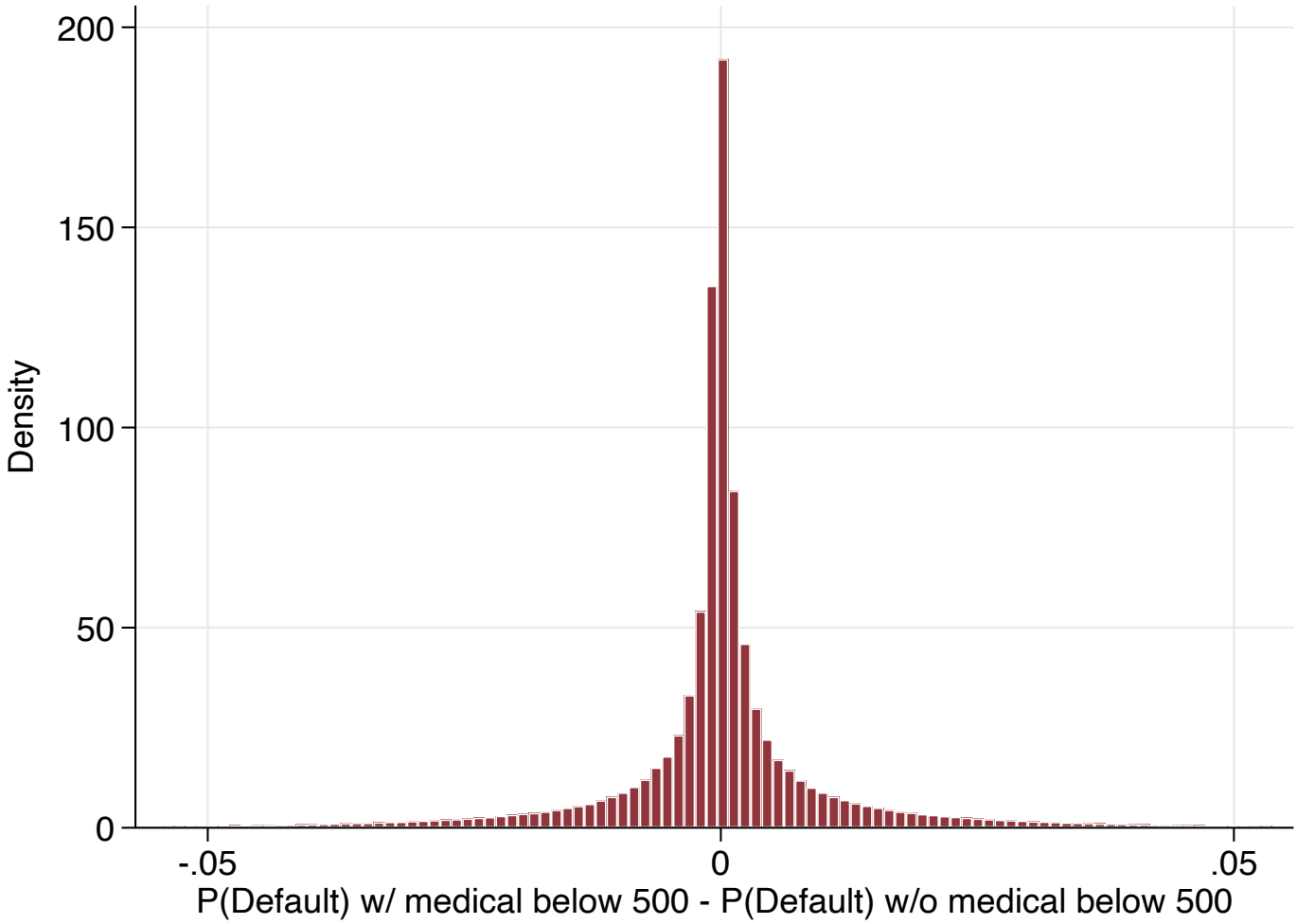
U.S. Census Bureau (2022). Debt for households, by type of debt and selected characteristics: 2022. Accessed: 2025-01-19.

Figure 1: Variable Importance in the Credit Scoring Model



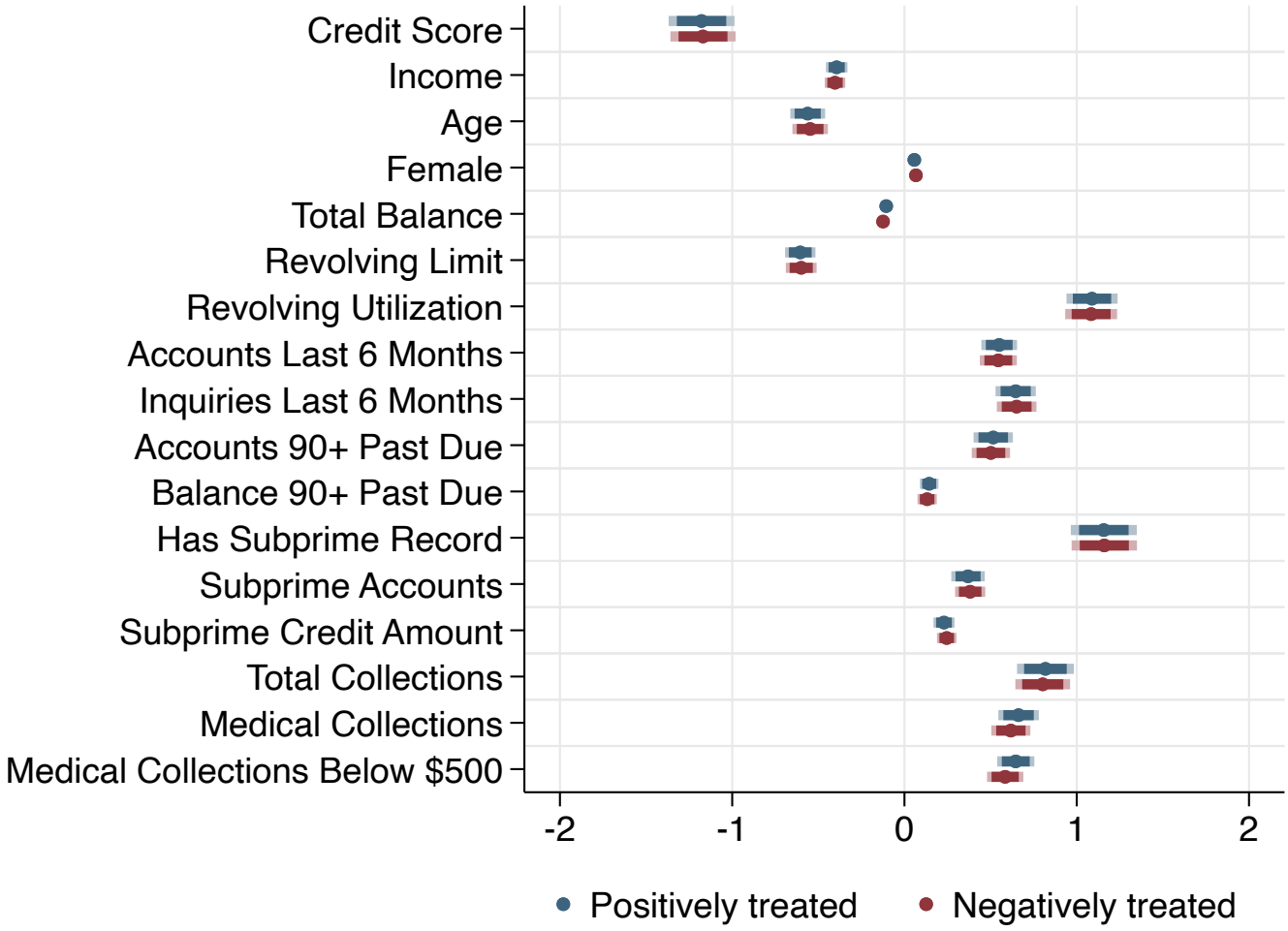
Notes: This figure displays variable importance measures, expressed as average SHAP values, for the credit scoring model presented in Section 3. The model, trained on 2019 data, predicts defaults occurring in 2020–2021 and is estimated using XGBoost. It incorporates 48 predictors, including medical collections 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: 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.

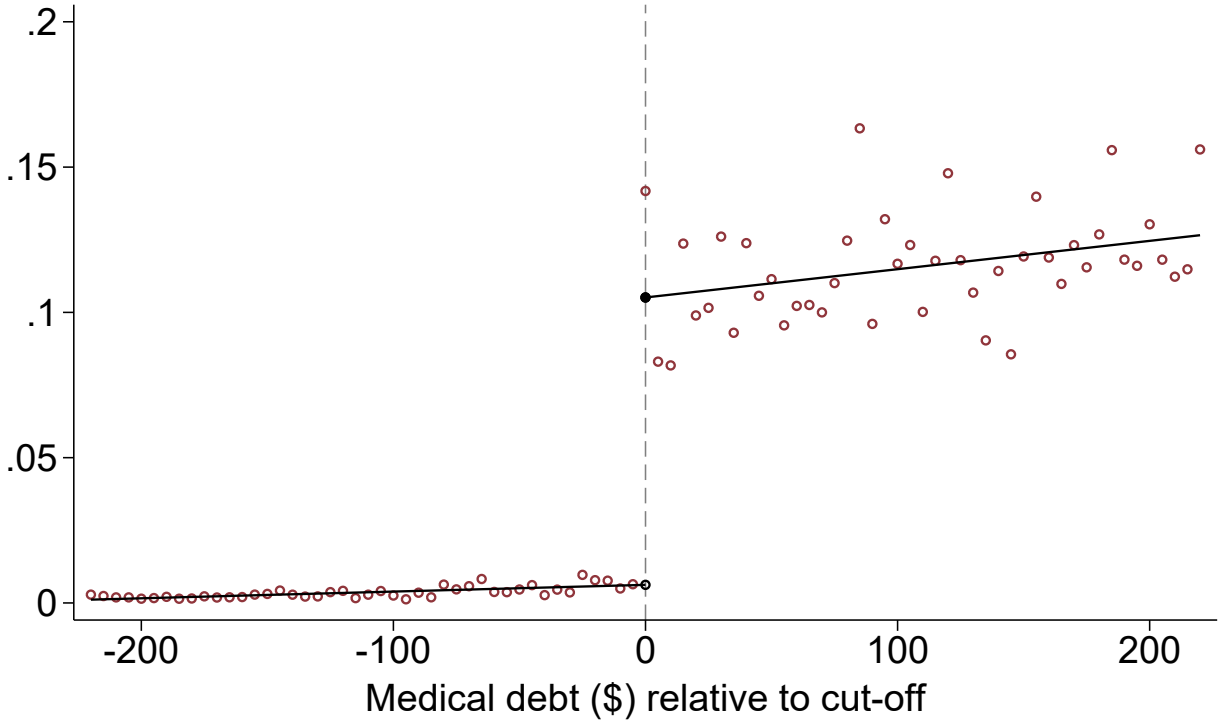
Figure 3: Covariate Balance by Changes in Default Probabilities



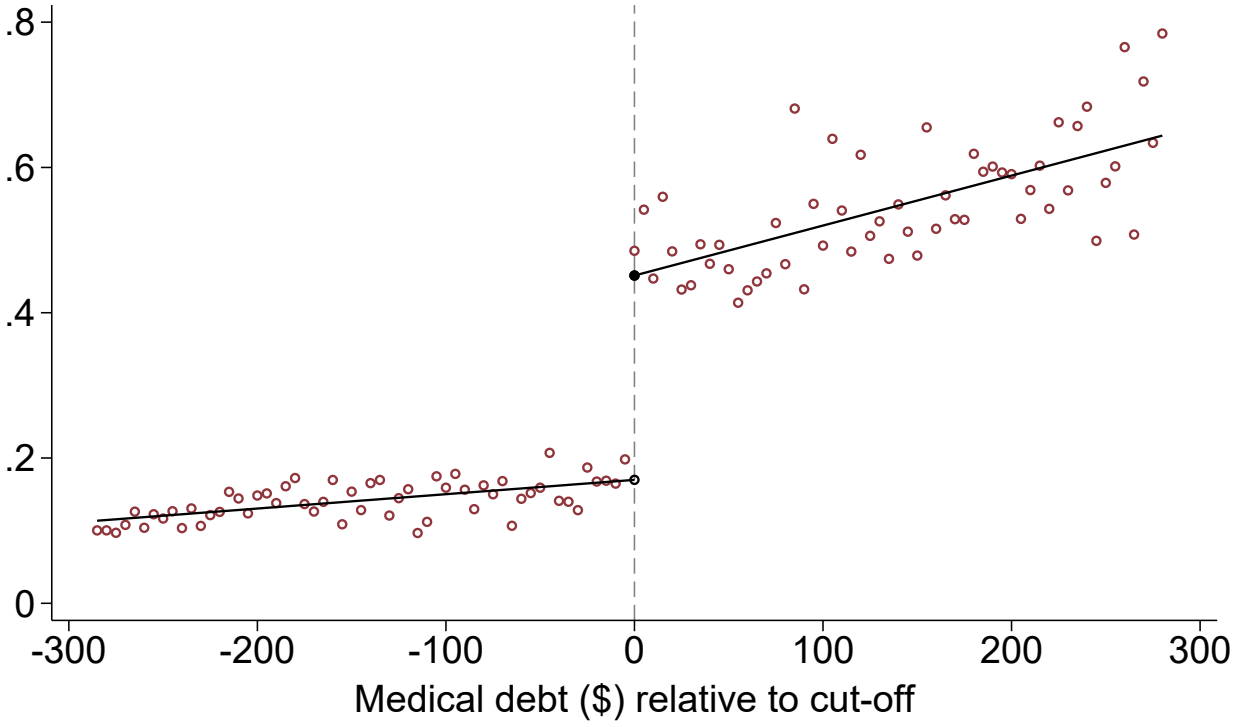
Notes: This figure shows estimates from balancing regressions for selected outcomes. Each balancing regression compares positively or negatively treated consumers to unaffected consumers. Negatively treated consumers are those whose predicted probability of default increases by 2 percentage points or 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 0.0015 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 4: Two-Year Evolution of 2022 Medical Collections Accounts

(a) Share of Surviving Accounts

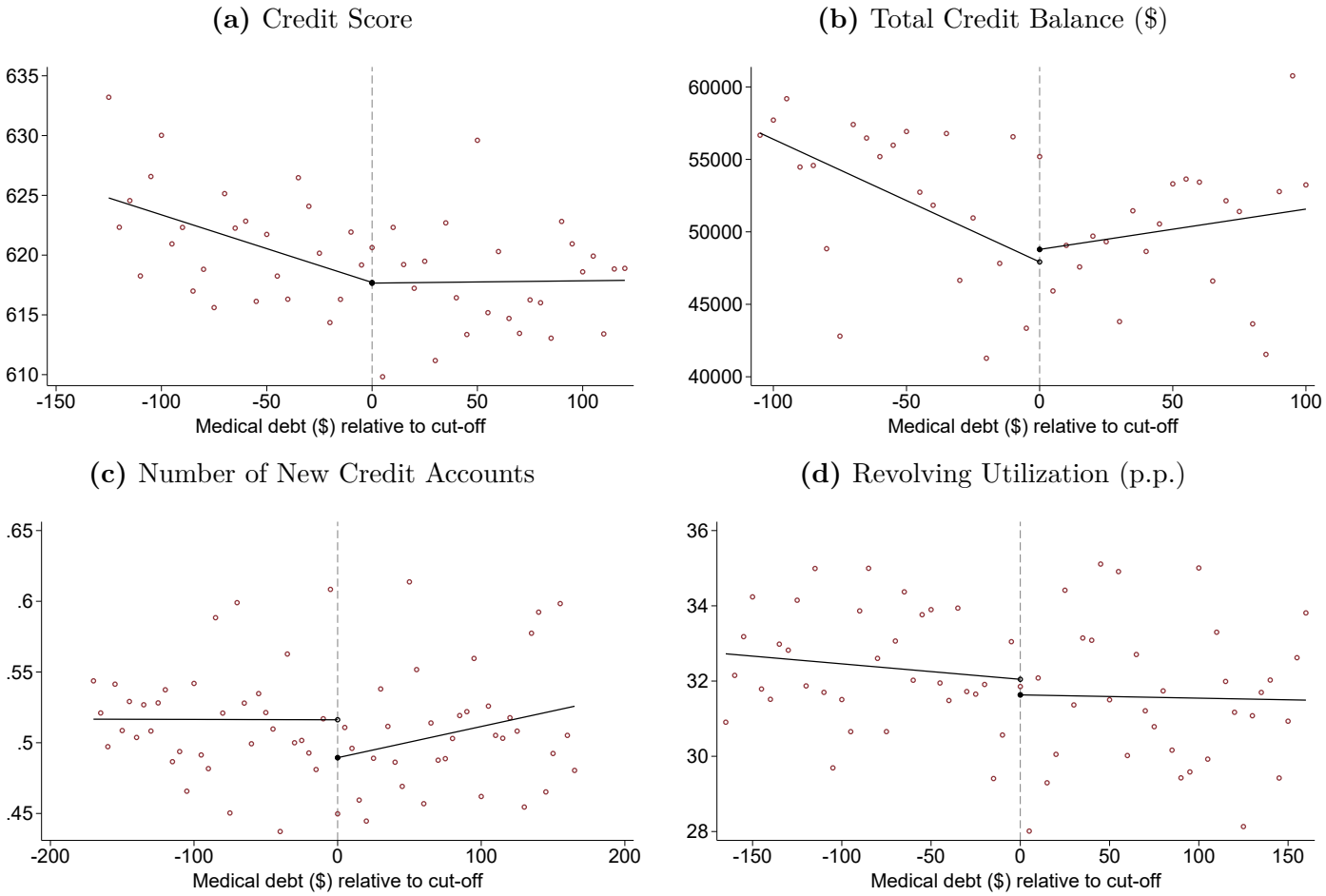


(b) Average Number of Accounts per Person



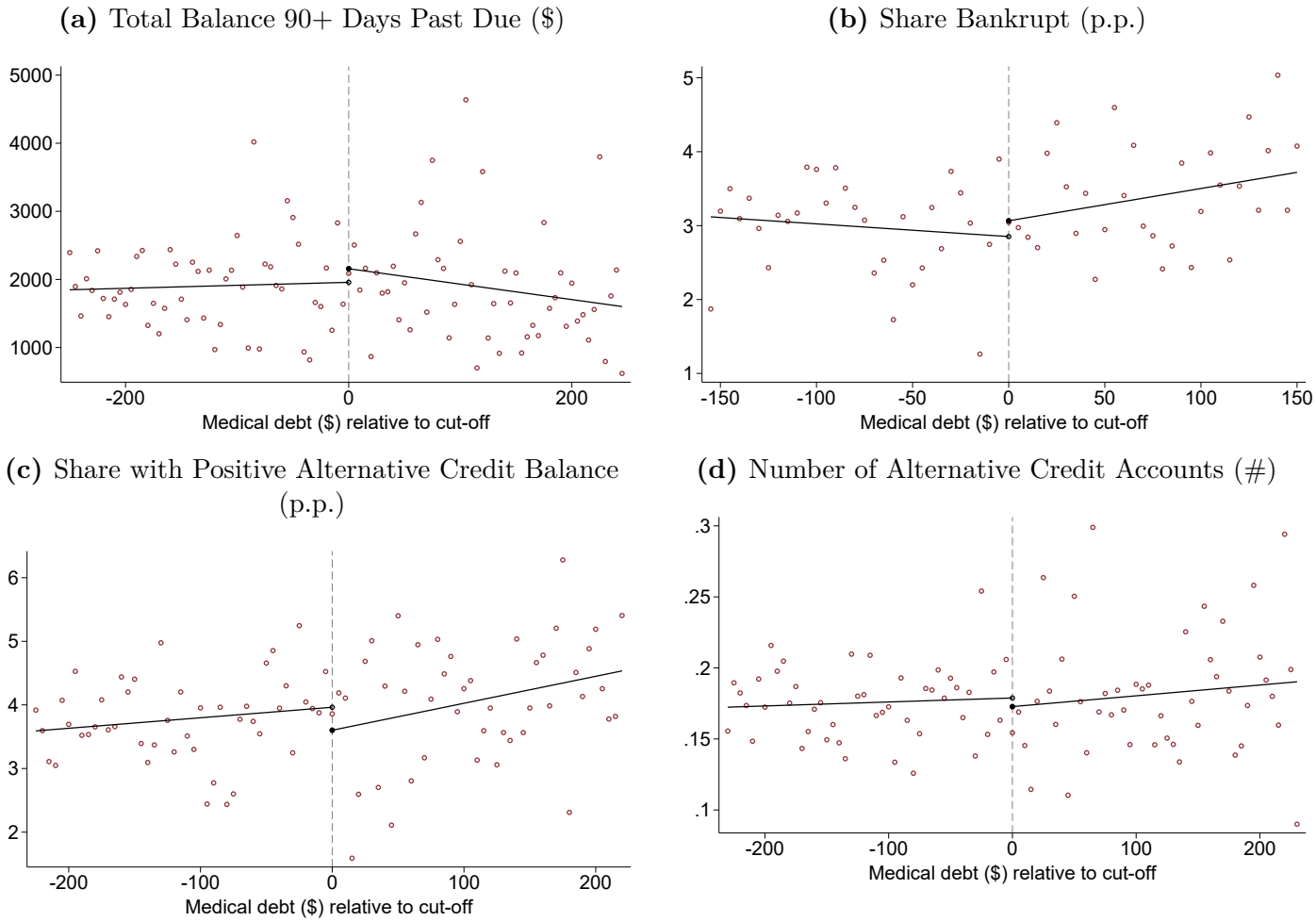
Notes: Panel (a) shows the proportion of 2022 medical collection accounts that remain on credit reports in 2024 by account amount, where the amount is measured as distance from the $500 threshold. Panel (b) shows the average number of medical collections accounts per person in 2024, where the running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

Figure 5: Credit Access and Utilization, 2024



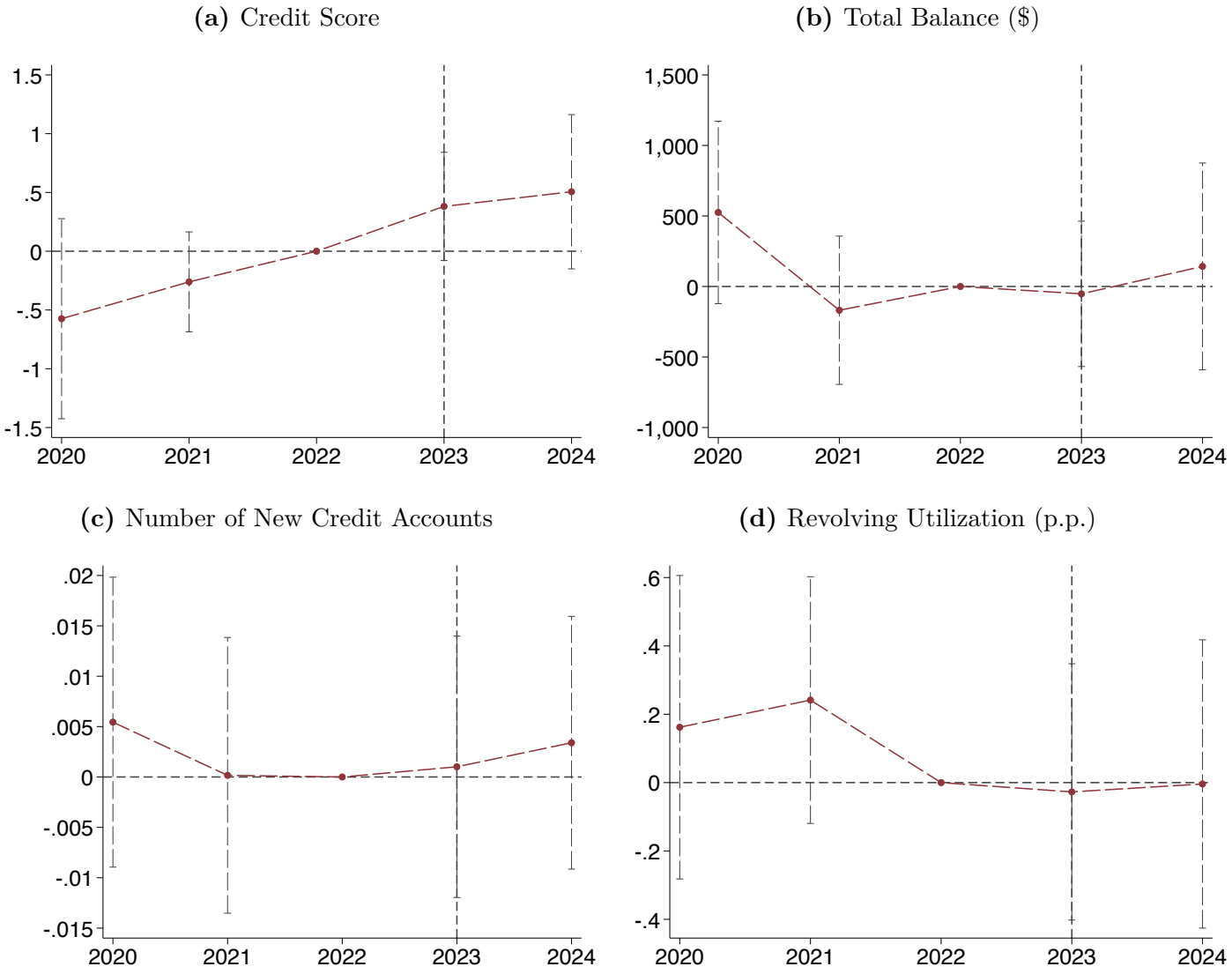
Notes: This figure shows the relationship between 2022 medical debt and four credit measures in 2024: consumer credit score, total balance, new trades, and revolving utilization. Medical debt is defined as the maximum value of the consumer's 2022 medical collections accounts, measured relative to the $500 threshold. Credit Scores are based on the VantageScore model. Total Balance is the combined balance across all credit accounts. New Trades is the number of new credit accounts opened in the past 6 month. Revolving Utilization is the ratio of outstanding balances to credit limits across all revolving accounts. The fitted lines are estimated using Equation (5).

Figure 6: Delinquency, Bankruptcy, and Alternative Credit Use, 2024



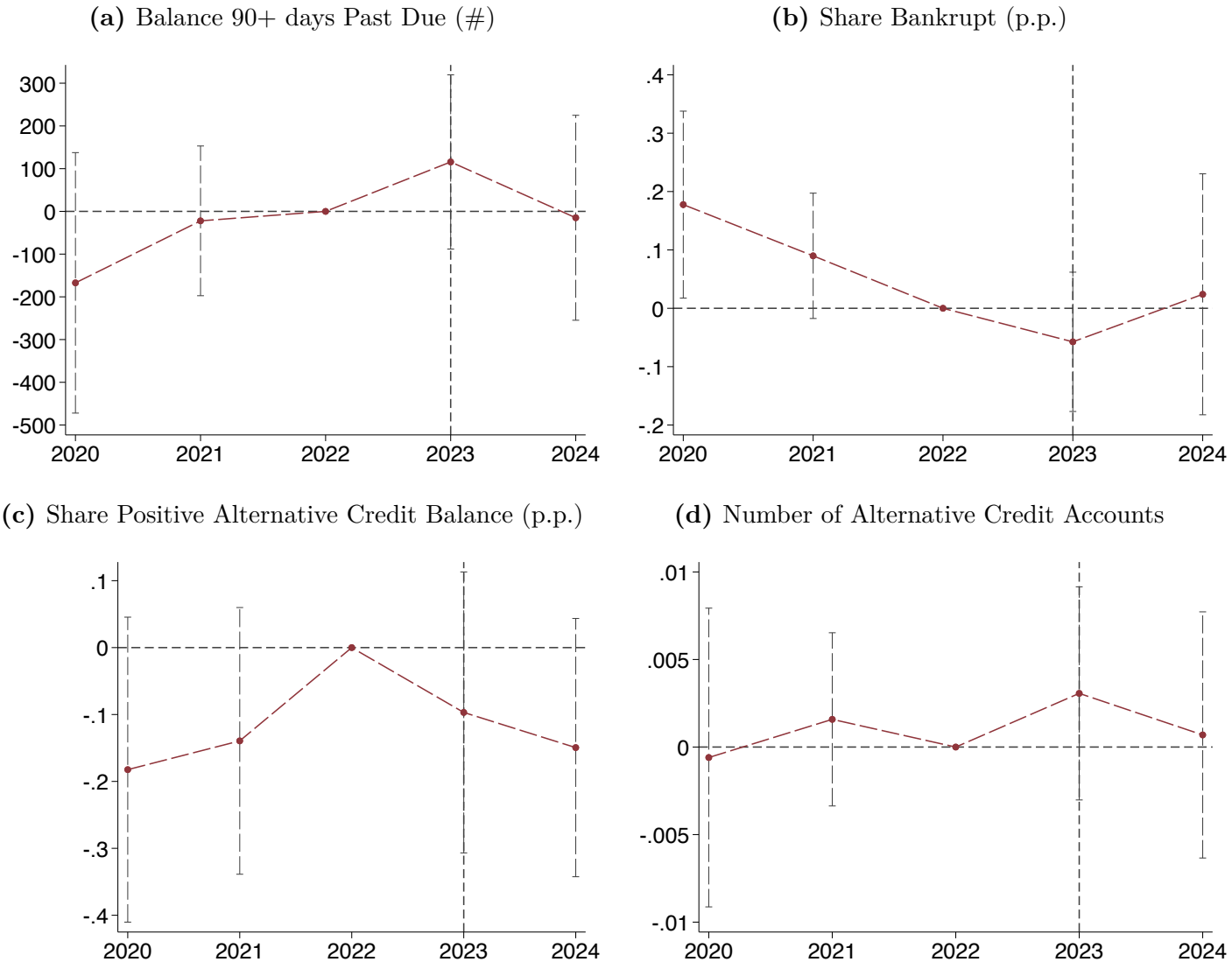
Notes: This figure shows the relationship between 2022 medical debt and measures of delinquency, bankruptcy, and alternative credit use. Medical debt is defined as the maximum value of the consumer's 2022 medical collections accounts, measured relative to the $500 threshold. Balance 90+ Days Past Due 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. The fitted lines are estimated using Equation (5).

Figure 7: Effect of Removing Small Medical Collections on Credit Access and Utilization



Notes: This figure presents difference-in-differences estimates based on Equation (7). 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. Credit Score is based on the VantageScore model. Total Balance is the combined balance across all credit accounts. New Trades is the number of new credit accounts opened in the past 6 month. Revolving Utilization is the ratio of outstanding balances to credit limits across all revolving accounts. The dashed vertical line marks the year 2023, when small medical collections were removed from credit reports. Standard errors are clustered based on 100 bins of predicted default probabilities.

Figure 8: Effect of Removing Small Medical Collections on Delinquency, Bankruptcy and Alternative Credit Use



Notes: This figure presents difference-in-differences estimates based on Equation (7). 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. 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. The dashed vertical line marks the year 2023, when small medical collections were removed from credit reports. Standard errors are clustered based on 100 bins of predicted default probabilities.

Table 1: Summary Statistics, 2019–2024

	All Consumers			Consumers With Medical Collections		
	Mean	Med.	St. Dev.	Mean	Med.	St. Dev.
Credit Score	702.26	715.00	100.85	611.21	601.00	86.69
Income ($1,000)	51.96	41.00	32.81	38.69	34.00	19.51
Age	50.56	49.00	19.41	44.90	43.00	15.16
Female (p.p.)	50.02	100.00	50.00	54.06	100.00	49.83
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 (p.p.)	28.09	13.00	32.84	49.86	47.00	39.50
Accounts Opened Last 6 Months	0.41	0.00	0.83	0.47	0.00	0.95
Inquiries Last 6 Months	0.41	0.00	0.74	0.62	0.00	0.90
Accounts 90+ Past Due	0.19	0.00	0.90	0.46	0.00	1.33
Balance 90+ Past Due ($1,000)	0.81	0.00	12.16	1.95	0.00	15.90
Total Collections	0.60	0.00	1.87	3.81	3.00	4.11
Total Collection Balance ($1,000)	0.56	0.00	3.65	3.14	1.39	9.90
Bankruptcy (p.p.)	2.84	0.00	16.60	4.79	0.00	21.35
Has Alternative Credit Record (p.p.)	19.99	0.00	39.99	47.93	0.00	49.96
Has Alternative Balance (p.p.)	1.02	0.00	10.05	2.60	0.00	15.92
Alternative Credit Accounts	0.05	0.00	0.43	0.13	0.00	0.68
Alternative Credit Amount ($1,000)	0.05	0.00	0.79	0.13	0.00	1.19
Medical Collections	0.25	0.00	1.02	2.44	2.00	2.17
Medical Collections 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)

	Mean	Med.	St. Dev.
Credit Score	625.38	618.00	87.95
Income ($1000)	40.70	35.00	20.53
Age	45.11	43.00	15.19
Female (p.p.)	0.55	1.00	0.50
All Open Accounts	4.35	3.00	4.97
Total Balance ($1,000)	48.89	10.66	94.40
Revolving Limit ($1000)	6.03	0.23	15.79
Revolving Utilization (p.p.)	29.51	6.00	38.58
Accounts Opened Last 6 Months	0.60	0.00	1.09
Inquiries Past 6 Months	0.67	0.00	0.93
Accounts 90+ Past Due	0.31	0.00	0.91
Balance ($1,000) 90+ Past Due	1.31	0.00	13.17
Accounts in Collections	3.53	2.00	3.77
Collection Balance ($1,000)	2.69	1.09	6.51
Mortgage Accounts Opened Past 6 Months	0.02	0.00	0.13
Bankruptcy (p.p.)	4.11	0.00	19.56
Has Alternative Credit (p.p.)	3.19	0.00	17.58
Alternative Credit Accounts	0.14	0.00	0.67
Alternative Credit Amount ($1,000)	0.18	0.00	1.45
Medical Collections	2.37	1.00	2.00
Medical Collections Below $500	1.56	1.00	1.43
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

Metric	Baseline	Medical < $500	All Medical
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. The accuracy score represents the share of correct predictions. 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

	Unaffected			Positively Treated			Negatively Treated		
	Mean	Med.	St. Dev.	Mean	Med.	St. Dev.	Mean	Med.	St. Dev.
Credit Score	750.49	787.00	85.44	611.01	609.00	80.44	612.67	610.00	82.00
Income ($1,000)	60.00	50.00	36.31	44.00	38.00	22.68	44.84	38.00	24.00
Age	57.06	58.00	19.42	45.77	44.00	14.95	45.59	44.00	15.08
Female (p.p.)	50.03	100.00	50.00	53.52	100.00	49.88	53.35	100.00	49.89
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	0.42	15.52	6.72	0.49	17.30
Revolving Utilization (p.p.)	16.09	6.00	23.16	54.68	55.00	38.61	54.06	55.00	38.76
Accounts Opened Last 6 Months	0.28	0.00	0.64	0.70	0.00	1.19	0.70	0.00	1.19
Inquiries Last 6 Months	0.26	0.00	0.57	0.78	0.00	1.00	0.76	0.00	0.99
Accounts 90+ Past Due	0.05	0.00	0.44	0.64	0.00	1.68	0.64	0.00	1.67
Balance 90+ Past Due ($1,000)	0.22	0.00	6.42	2.85	0.00	22.81	2.64	0.00	21.64
Total Collections	0.20	0.00	1.07	1.78	1.00	3.01	1.73	1.00	2.96
Total Collection Balance ($1,000)	0.17	0.00	1.71	1.69	0.11	4.88	1.67	0.08	4.53
Bankruptcy (p.p.)	0.95	0.00	9.68	9.61	0.00	29.48	9.17	0.00	28.85
Has Alternative Credit Record (p.p.)	6.90	0.00	25.35	54.01	100.00	49.84	52.86	100.00	49.92
Has Alternative Balance (p.p.)	0.22	0.00	4.72	3.88	0.00	19.32	3.76	0.00	19.03
Alternative Credit Accounts	0.01	0.00	0.19	0.19	0.00	0.86	0.18	0.00	0.85
Alternative Credit Amount ($1,000)	0.01	0.00	0.35	0.22	0.00	1.59	0.21	0.00	1.55
Medical Collections	0.09	0.00	0.60	0.71	0.00	1.66	0.66	0.00	1.60
Medical Collections 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

	Total Collections (#)	Medical Collections (#)	Credit Score	Balance ($)	New Trades (#)	Utilization (%)	Del. Balance ($)	Bankruptcy (%)	Has Alternative (%)	Alt. Trades (#)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
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]
Sample Mean	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

This table shows the coefficient (β) estimates and 95% confidence intervals of Equation (5). 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. Total collections is the total number of accounts in collections on a consumer's credit report, Medical Collections is the total number of trades in collection in Medical or Healthcare categories. Credit Score are credit scores according to the VantageScore model. Balance is the total balance across all credit accounts, New Trades is the total number of trades opened in prior 6 months. Utilization is the balance-to-limit ratio across all revolving credit accounts (scaled by 100). Del. Balance is the total balance on trades 90 days or more past due. Bankruptcy is a binary variable that equals one if the consumer has any bankruptcies. Has Alternative is a binary variable that equals one if the consumer has a positive alternative credit amount. Alternative Trades is the total number of alternative credit accounts in the consumer's Clarity report. A ***, **, and * indicate significance at the 1%, 5%, and 10% level respectively, using conventional inference.

Table 6: Indirect Effect: Difference-in-Differences

	Credit Score	Balance ($)	New Trades (#)	Utilization (%)	Del. Balance ($)	Bankruptcy (%)	Has Alternative (%)	Alt. Trades (#)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated × Post	0.721***	-74.496	0.000	-0.144	113.886	-0.106	-0.016	0.002
	(0.271)	(301.540)	(0.005)	(0.166)	(102.173)	(0.096)	(0.063)	(0.003)
Sample Mean	617.13	68,860.42	0.69	52.97	2,467.57	9.05	3.84	0.19
% of Mean	0.12	-0.11	0.05	-0.27	4.62	-1.17	-0.41	0.82
Consumer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
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 shows estimates Equation (6). Credit Score are credit scores according to the VantageScore model. Balance is the total balance across all credit accounts, New Trades is the total number of trades opened in prior 6 months. Utilization is the balance-to-limit ratio across all revolving credit accounts (scaled by 100). Del. Balance is the total balance on trades 90 days or more past due. Bankruptcy is a binary variable that equals one if the consumer has any bankruptcies (scaled by 100). Has Alternative is a binary variable that equals one if the consumer has a positive alternative credit amount (scaled by 100). Alternative Trades is the total number of alternative credit accounts in the consumer's Clarity report. We create 1,000 equal-sized bins of the difference in predicted probability of default in our full sample and cluster our standard errors, reported in parentheses, at the bin level. Our differences-in-differences sample corresponds to approximately 10% of the full sample, and we are thus left with approximately 100 clusters. A ***, **, and * indicate significance at the 1%, 5%, and 10% level respectively

Online Appendix

"The Direct and Indirect Effects of Information Deletion: Evidence from Medical Collections"

Victor Duarte, Julia Fonseca, Divij Kohli, Julian Reif

A Alternative RD specifications

The effect of deleting small medical collections—an account-level treatment—on individual-level outcomes depends on the underlying treatment mechanism. We consider three potential scenarios:

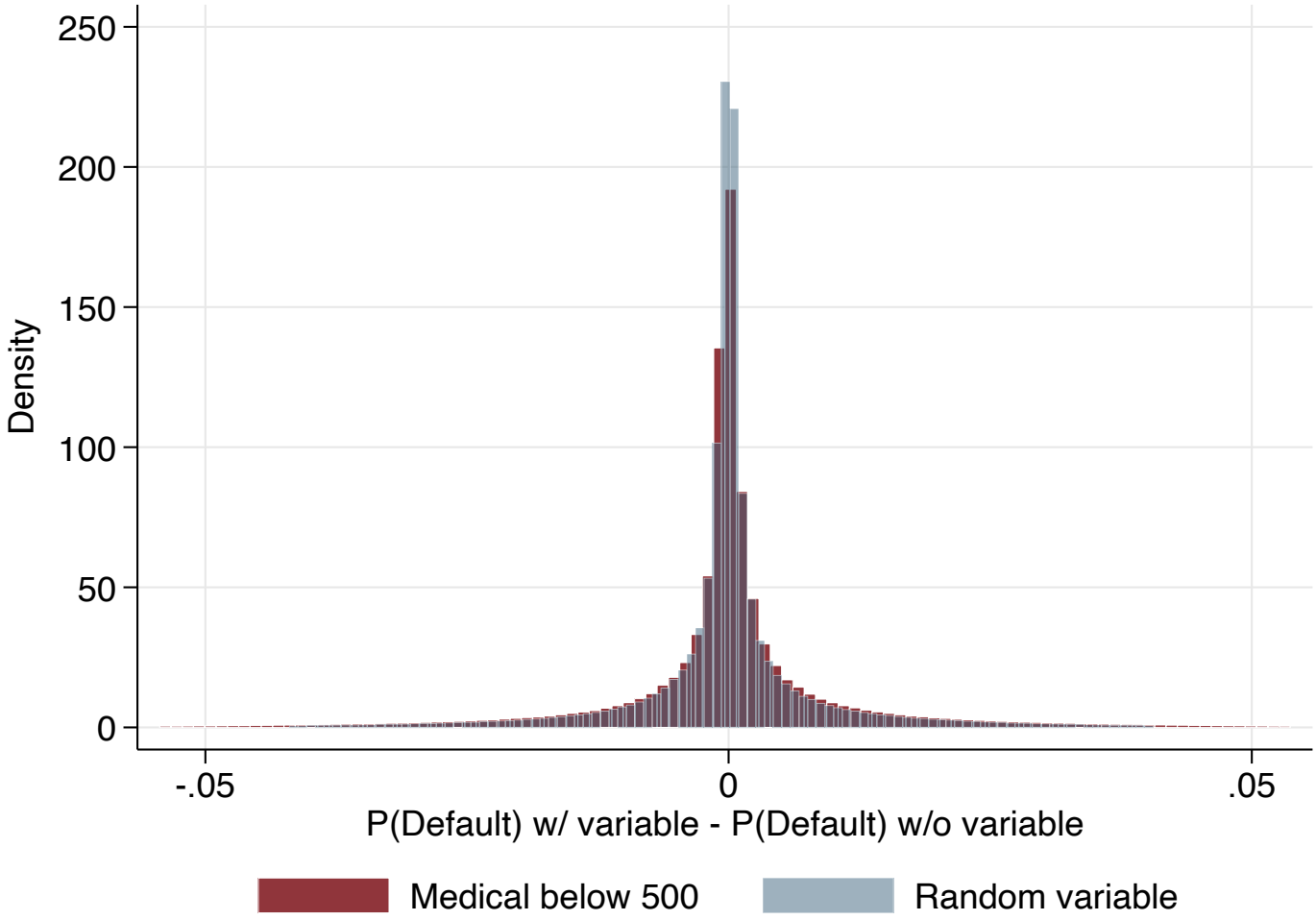
1. Treatment occurs only if all of an individual's medical collections are deleted.
2. Treatment occurs if any (i.e., at least one) medical collection is deleted.
3. Treatment scales with the proportion of deleted accounts.

The first scenario corresponds to requiring that the maximum value of all medical collection accounts is below the cutoff. The second scenario corresponds to requiring that the minimum value is below the cutoff. The third scenario can be estimated using an account-level regression.

The estimating equation in (5) aligns with the first scenario. In this appendix, we present estimates for the second and third scenarios.

B Additional Figures and Tables

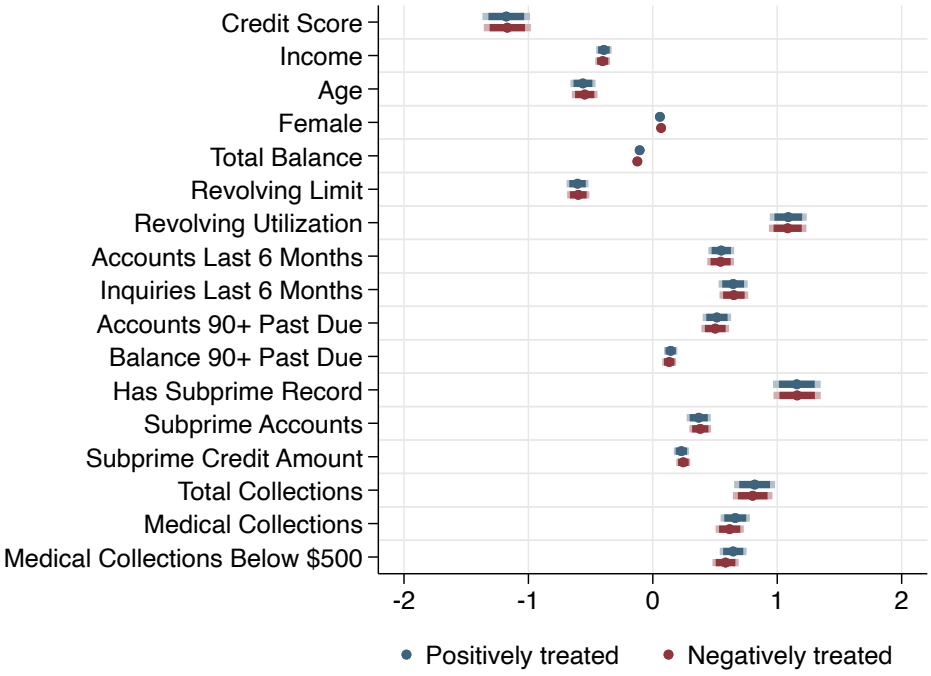
Figure A.1: Effect of Removing Small Medical Collections vs. Removing Noise



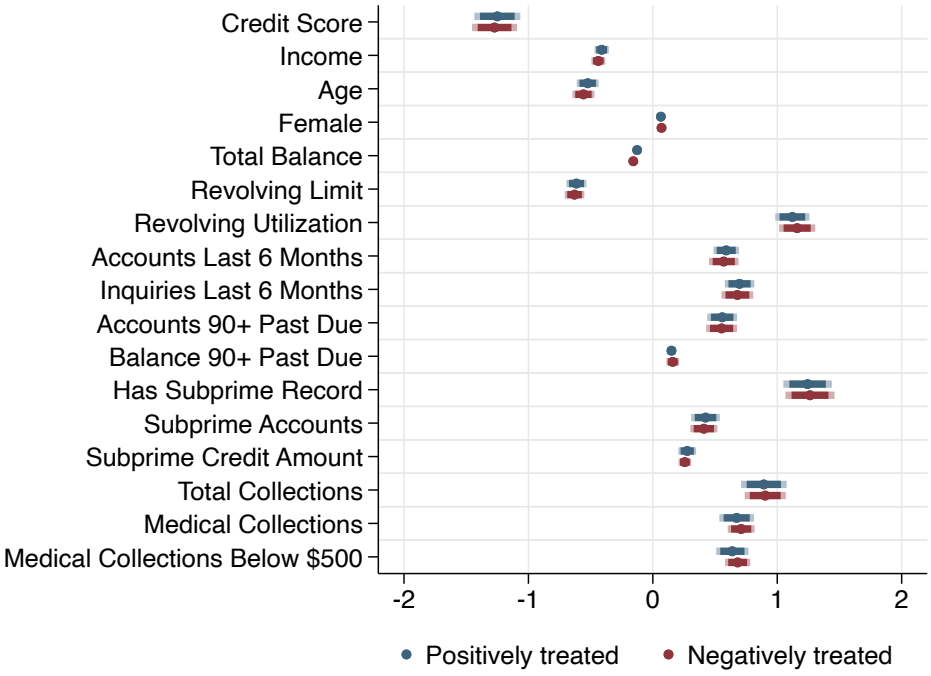
Notes: This figure shows the change in the predicted probability of default over 24 months following the removal of a variable from the credit scoring model described in Section 3. The red histogram reproduces the plot from Figure 2, showing the effect of removing small (< \$500) medical collections. The blue histogram shows the effect of removing a random noise predictor that was drawn randomly from a uniform distribution.

Figure A.2: Covariate Balance by Changes in Default Probabilities: Small Medical Collections vs. Noise

(a) Small (< $500) Medical Collections

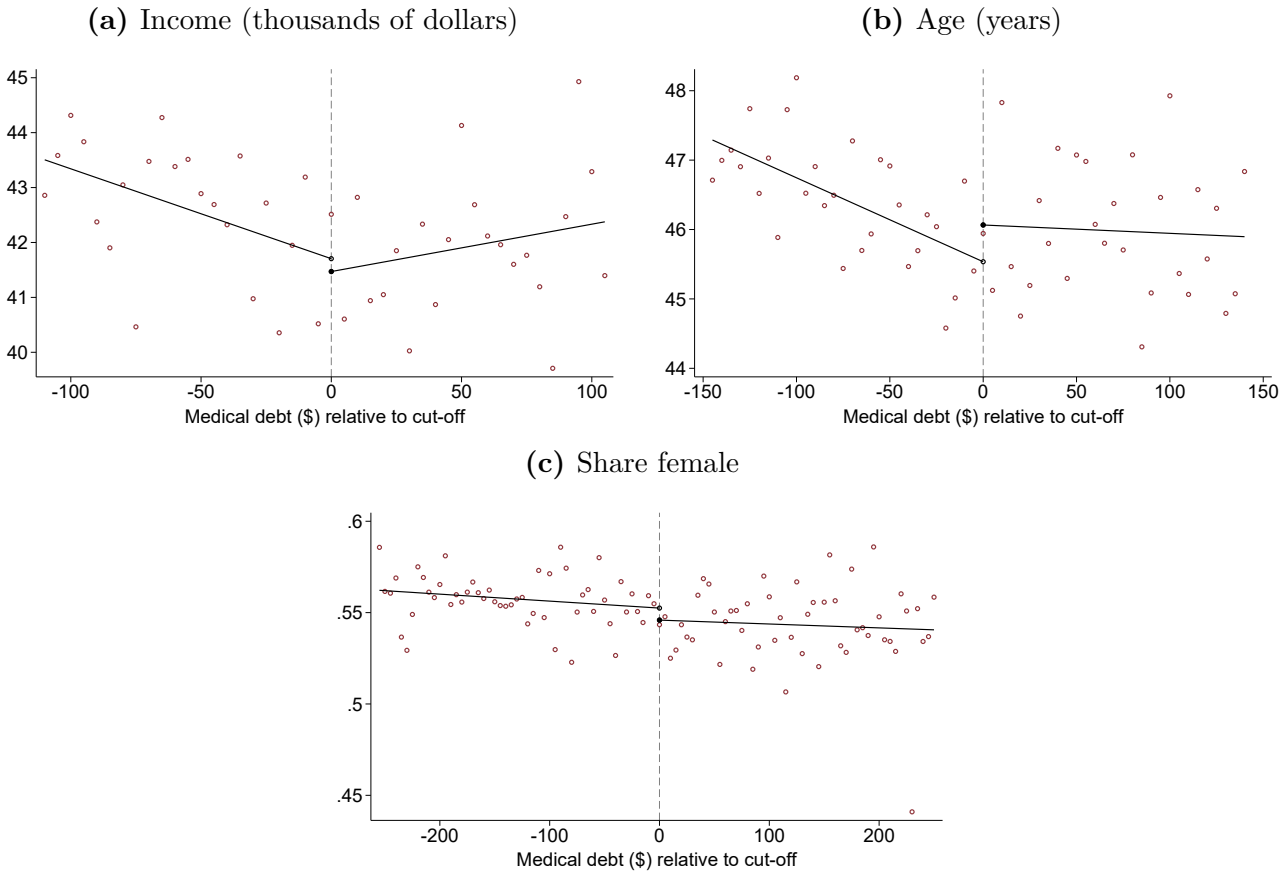


(b) Random Variable



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 0.0015 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.

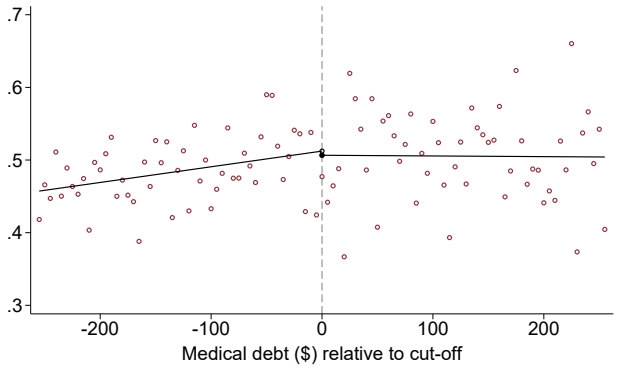
Figure A.3: Covariate Smoothness Tests



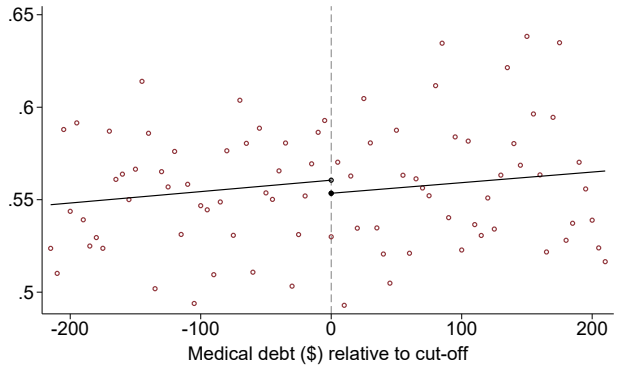
Notes: This figure shows average income, age, and share female by medical debt amount, where the amount is measured as distance from the $500 threshold. The running variable is equal to the maximum value of the consumer's medical collections accounts in 2022. The covariates are measured in 2022. The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

Figure A.4: Additional Credit Outcomes for RD Analysis, 2024

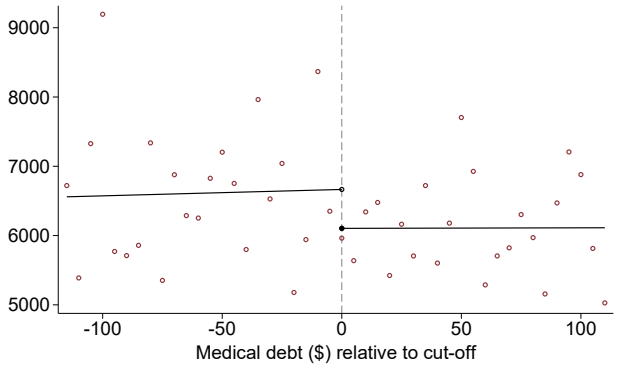
(a) Number of Accounts 90+ days Past Due



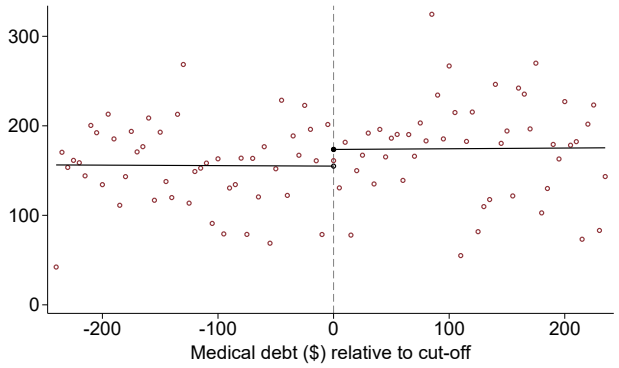
(b) Number of New Inquiries



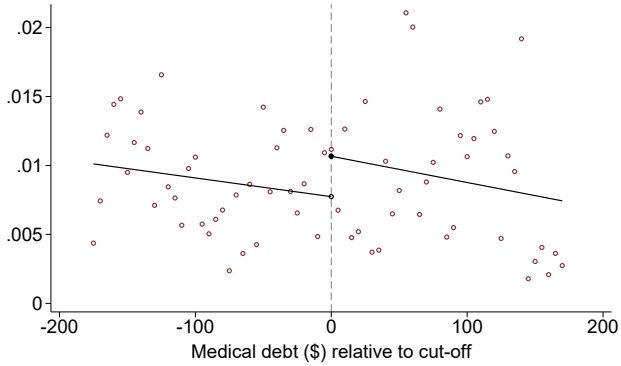
(c) Revolving Limit ($)



(d) Alternative Credit Balance ($)

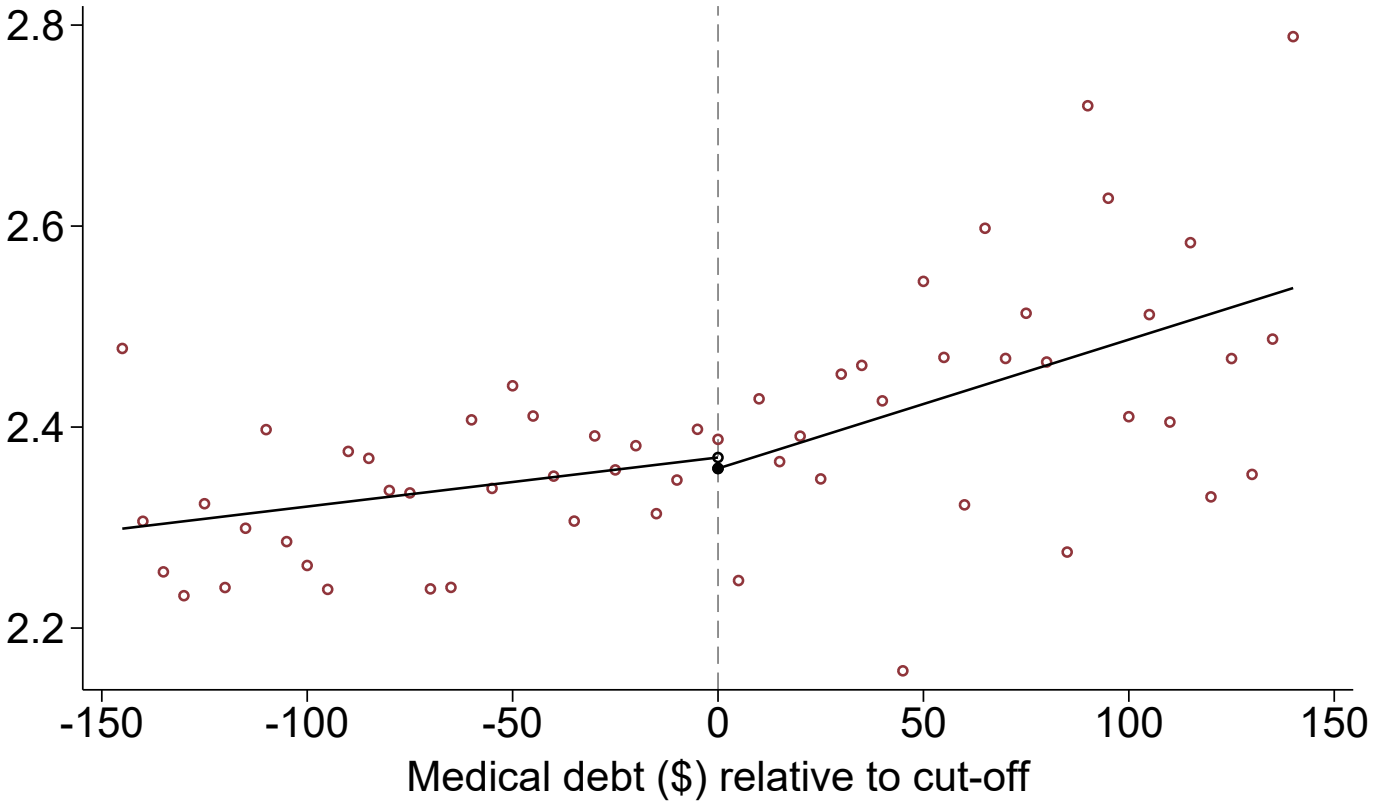


(e) Number of New Mortgage Accounts



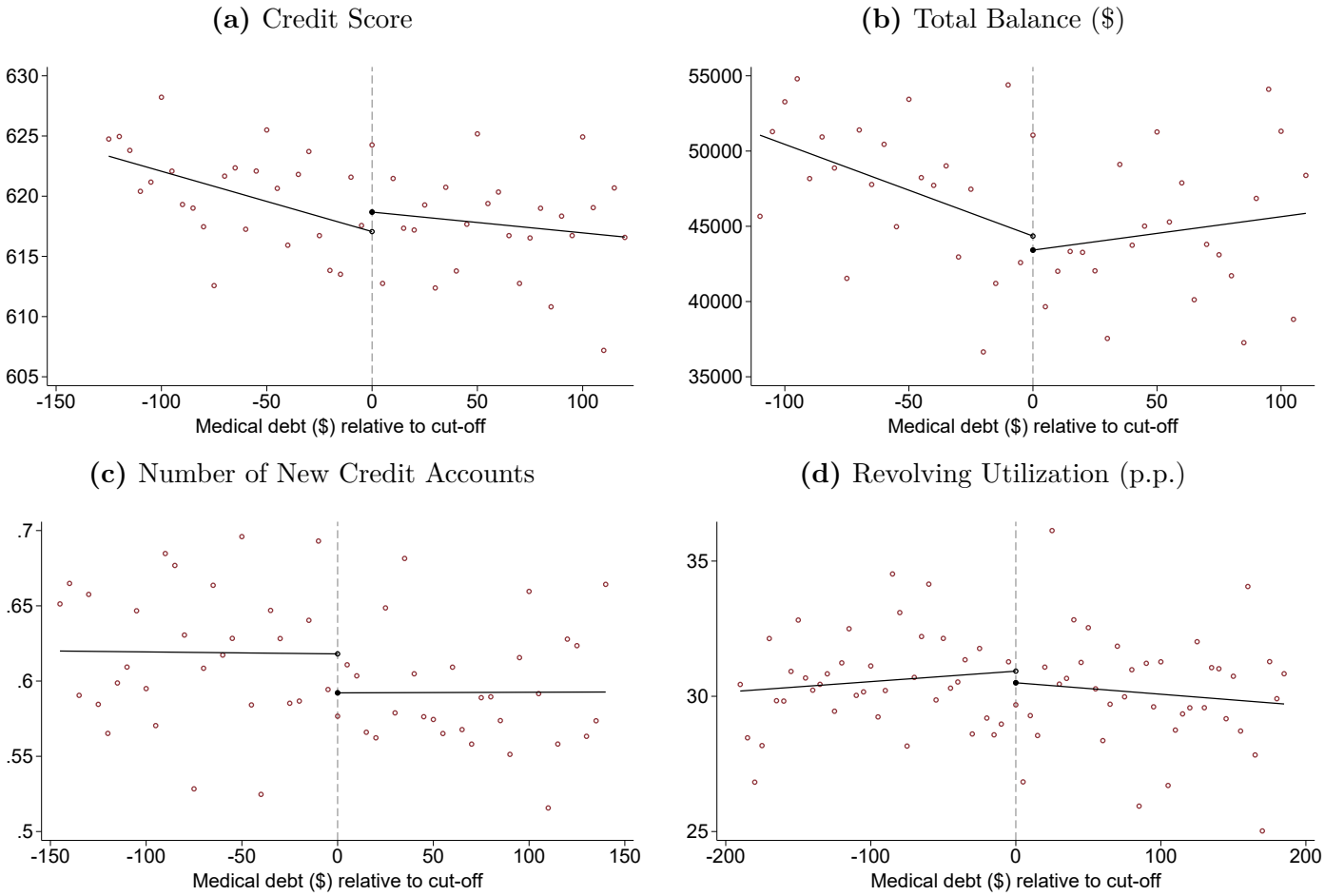
Notes: This figure shows the relationship between 2022 medical debt and five supplementary credit outcomes. Medical debt is defined as the maximum value of the consumer's 2022 medical collections accounts, measured relative to the $500 threshold. Number of Accounts 90+ days Past Due is the number of 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. The fitted lines are estimated using Equation (5).

Figure A.5: Falsification test: Number of Accounts in Medical Collections, 2022



Notes: This figure shows the relationship between the medical debt running variable and the number of medical debt accounts 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. The fitted lines are estimated using Equation (5).

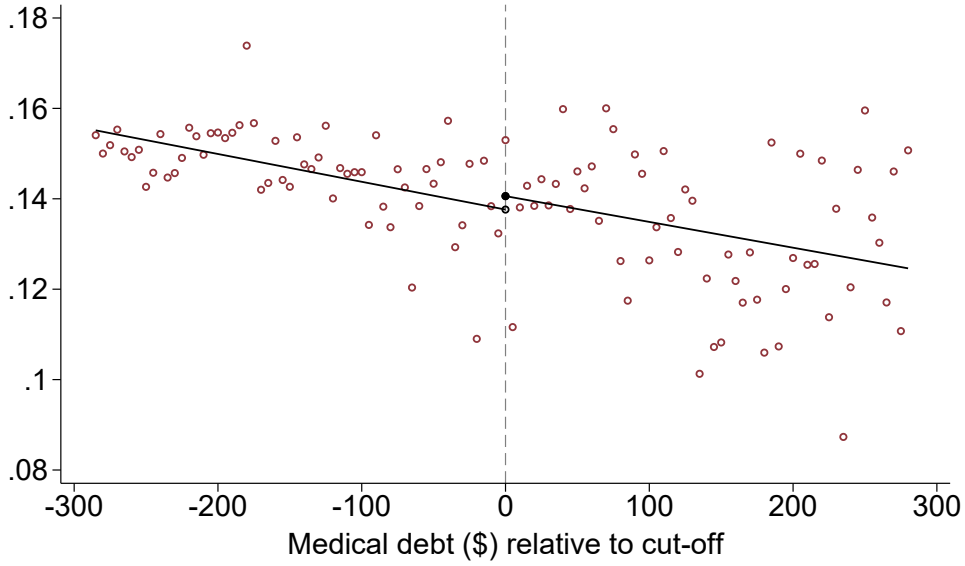
Figure A.6: Falsification test: Credit Access and Utilization, 2022



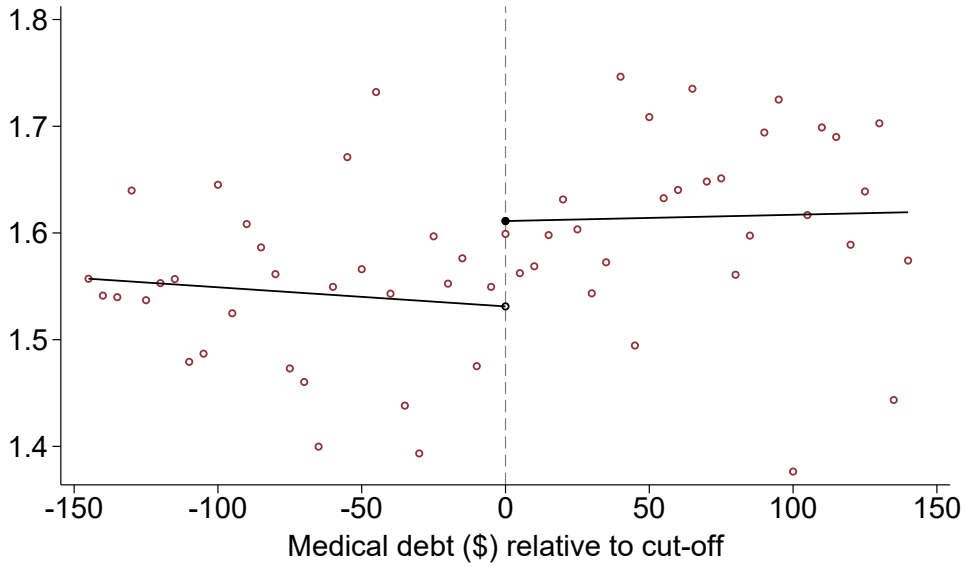
Notes: This figure shows the relationship between 2022 medical debt and four credit measures in 2022: consumer credit score, total balance, number of new credit accounts, and revolving utilization. Medical debt is defined as the maximum value of the consumer's 2022 medical collections accounts, measured 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. The fitted lines are estimated using Equation (5).

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

(a) Share of Surviving Accounts

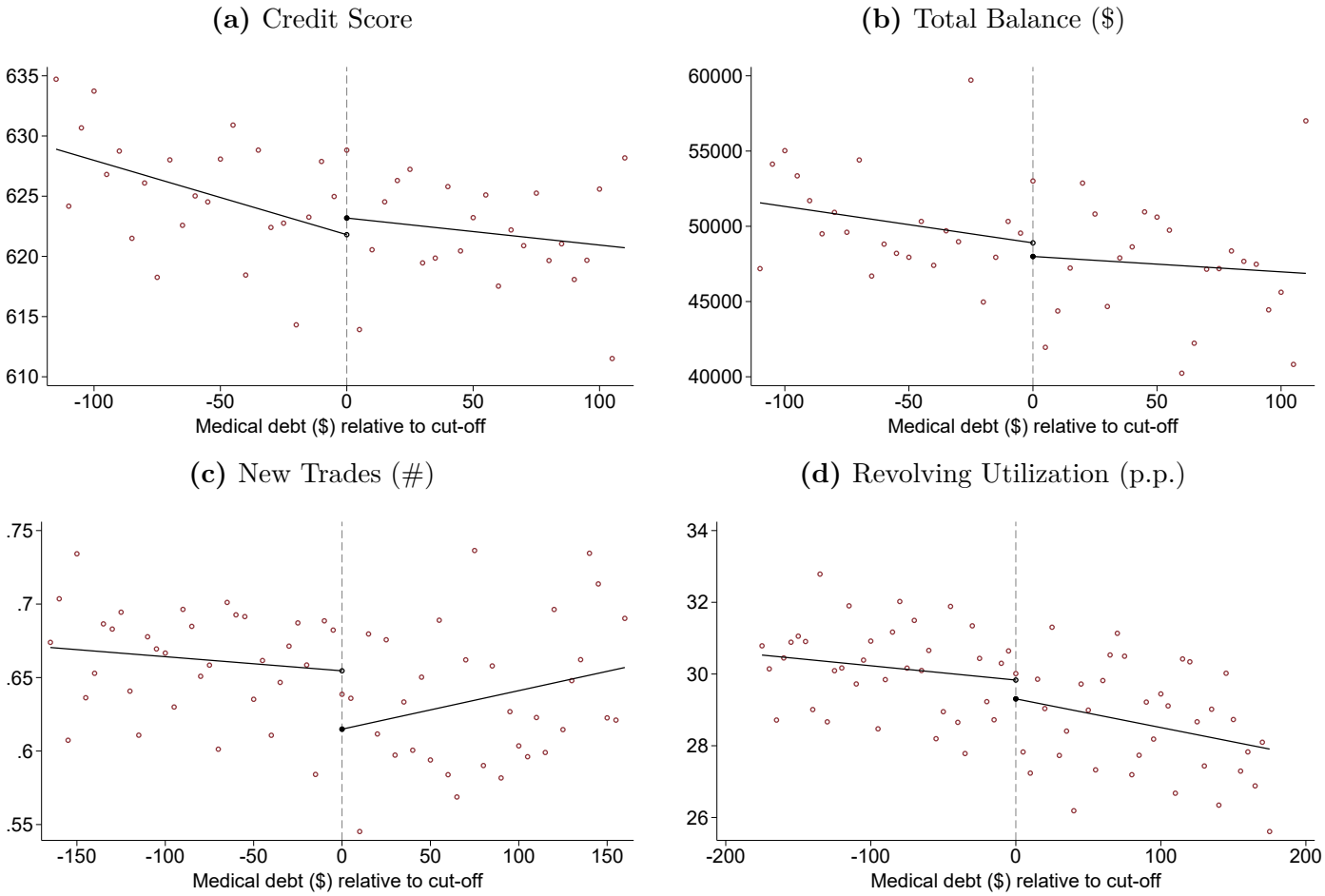


(b) Average Number of Accounts per Person



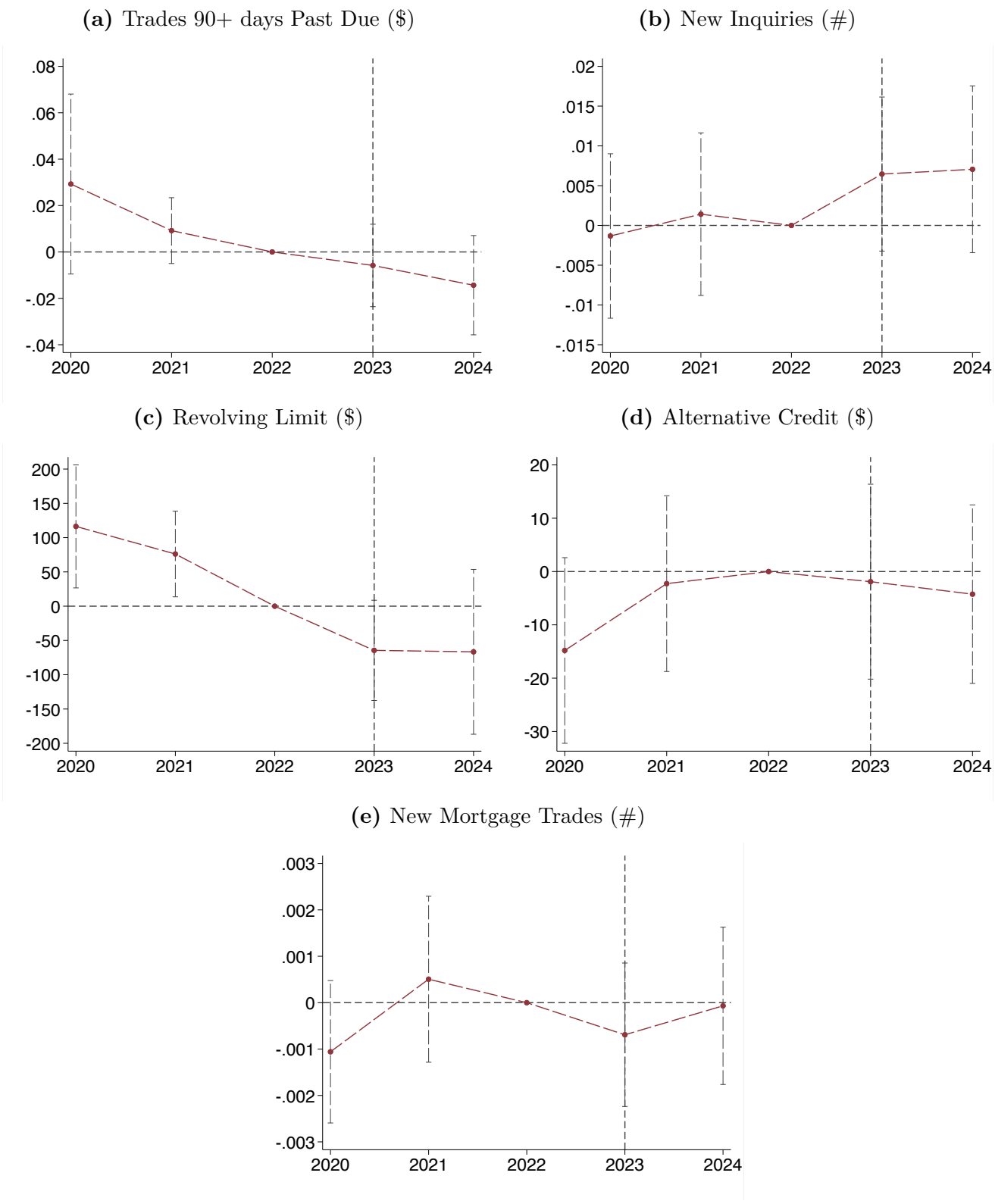
Notes: Panel (a) shows the proportion of 2020 medical collection accounts which remain present on 2022 credit reports by account amount, where the amount is measured as distance from the $500 threshold. Panel (b) shows the average number accounts per person. In panel (b), the running variable is equal to the maximum value of the consumer's medical collections accounts. The fitted lines are estimated using Equation (5).

Figure A.8: Placebo test: Credit Access and 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 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. The fitted lines are estimated using Equation (5).

Figure A.9: Effect of Removing Small Medical Collections on Additional Outcomes



This figure shows coefficient estimates and 95% confidence intervals of Equation (7) 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.1: Medical Collections Data Benchmarking

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 benchmarks the share of individuals with medical collections in our data against other sources. The second column reports the share of consumers with at least one account in medical collections compared in GCCP sample. The third and fourth columns report the share of consumers with medical collections from Blavin et al. (2023) and Sandler and Nathe (2022), respectively. GCCP data is from March of each year, while Urban Institute data is from August. Sandler and Nathe (2022) use the CFPB's Consumer Credit Panel containing monthly snapshots. We report approximate values as of January of each year.

Table A.2: Credit Scoring Metrics: Medical vs. a Random Variable

Metric	Baseline	Without Medical < $500	With 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

This table reports the performance of our two baseline credit scoring models—with and without information of medical collections below $500—as well as the performance of a model including a feature randomly drawn from a uniform distribution. Accuracy score is the share of correct predictions. Precision is the proportion of predicted defaults that were correctly classified. Recall is the proportion of actual defaults that were correctly classified. F1 score is a harmonic mean of Precision and Recall. AUC if the area under the Receiver Operating Characteristic curve, which measures the probability that the model will assign a higher default probability to a true defaulter than to a non-defaulter.

Table A.3: Direct Effect: Alternative RD Specifications

Dependent Variable:	Total Collections (#)	Medical Collections (#)	Credit Score	Balance ($)	New Trades (#)	Utilization (%)	Del. Balance ($)	Bankruptcy (%)	Has Alternative (%)	Alt. Trades (#)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					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]
Sample Mean	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

This table shows the coefficient (β) estimates and 95% confidence intervals of Equation (5). The running variable in first row for medical debt corresponds to the highest debt amount across the consumer's medical collections accounts. Whereas, the running variable in second row for medical debt corresponds to the smallest 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. Total collections is the total number of accounts in collections on a consumer's credit report, Medical Collections is the total number of trades in collection in Medical or Healthcare categories. Credit Score are credit scores according to the VantageScore model. Balance is the total balance across all credit accounts, New Trades is the total number of trades opened in prior 6 months. Utilization is the balance-to-limit ratio across all revolving credit accounts (scaled by 100). Del. Balance is the total balance on trades 90 days or more past due. Bankruptcy is a binary variable that equals one if the consumer has any bankruptcies. Has Alternative is a binary variable that equals one if the consumer has a positive alternative credit amount. Alternative Trades is the total number of alternative credit accounts in the consumer's Clarity report. A ***, **, and * indicate significance at the 1%, 5%, and 10% level respectively, using conventional inference.