# 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

December 2024

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 they were no longer 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 develop a 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 negative indirect effects, but also did not benefit consumers whose information was removed. Finally, we show that large medical collections are also not meaningfully predictive of default, suggesting that the CFPB's proposal to delete remaining medical collections from credit reports would be similarly ineffective

*We thank Ray Kluender and Neale Mahoney for helpful comments and discussions.

# 1 Introduction

One in seven Americans had a medical account in collections between 2019 and 2022. Most medical collections arise when third-party debt collectors report unpaid medical debts to credit bureaus.[1] Policymakers have increasingly scrutinized these reporting practices, expressing concerns that making medical debt visible to lenders could impair access to credit following unexpected medical shocks. In response to these concerns, 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 CFPB proposed a rule in June 2024 to remove all remaining medical collections from credit reports, arguing that eliminating this information would enhance credit access and improve terms for consumers burdened with medical debt. To address concerns that deleting this information could harm consumers who had no medical debt, the CFPB cited evidence that medical collections are poor predictors of future default, and thus have limited relevance for credit origination decisions (Consumer Financial Protection Bureau, 2024). In other credit contexts, however, information deletion has been shown to reduce borrowing opportunities for some groups (Liberman et al., 2019). Moreover, if medical collections truly lack predictive power, policies aimed at deleting them are unlikely to improve credit access.

Against this backdrop, this paper studies the direct and indirect effects of deleting medical collections from credit reports. Specifically, we focus on the direct benefits to consumers whose medical collections are deleted, as well as the potential negative spillovers on other consumers. To assess how this deletion affects lenders' ability to predict default, we employ a machine learning model trained on two sets of 2019–2021 data. The first set uses data on borrower credit history, including information on medical collections below $500, while the second set excludes these collections. Comparing the output of these two credit scoring models reveals that excluding small medical collections has no meaningful effect on default prediction, indicating they have minimal relevance for credit origination decisions. Thus, the April 2023 intervention that deleted this information is unlikely to affect credit access,

[1] Hospitals can also report medical collections, but this is less common (Brevoort and Kambara, 2014).

either directly or indirectly.

We confirm this insight using two distinct research designs. First, we employ a Regression Discontinuity (RD) approach centered around the $500 information deletion threshold. Comparing individuals just above and below this threshold, we find no evidence that information deletion improved credit access or financial health. Our 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 our two default prediction models—one with and one without information on medical collections below $500—to identify consumers potentially harmed by the removal of this information. Specifically, we focus on 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. We use a difference-in-differences research design to compare them to consumers whose predicted probability falls by at least 2 percentage points. We show that the two groups are observationally similar across a wide range of characteristics and evolved in close parallel prior to the 2023 information deletion.[2] Consistent with our RD results, we find no evidence of negative spillover effects from the information deletion, with precise estimates that again rule out small effects.

We conclude that the 2023 decision to delete small medical collections had no negative indirect effects, but also provided no benefits to those whose information was removed. Additionally, we show that medical collections above $500 are also not predictive of default behavior, suggesting that the CFPB's 2024 proposed rule is unlikely to significantly affect credit access. Our findings suggest that policies promoting information deletion do not have the intended benefit for consumers burdened by medical debt.

Our paper builds on a broad literature studying information deletion across different contexts. Liberman et al. (2019) study the deletion of default information in Chile, finding

[2]Both groups consist primarily of low-income consumers with thin credit files. As we show in the main text, when reliable information is scarce, even randomly generated variables can influence default predictions. Thus, a noisy predictor like medical collections effectively partitions this population into two randomly assigned groups.

that it increases borrowing for consumers whose predicted probability of default decreases post-deletion, but reduces borrowing for those classified as higher default risks. Jansen et al. (2024) estimate that bankruptcy flag removals directly benefit affected consumers through lower interest rates, while raising rates for those with no history of bankruptcy, resulting in a small decline in social surplus.[3] Agan and Starr (2017) study the removal of 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. We contribute to this literature by studying information deletion in the context of medical collections, a topic of ongoing policy debate. Unlike most prior studies, we find that the deleted information has minimal predictive value, indicating there should be no direct or indirect effects. We confirm this prediction using two separate research designs, providing strong evidence that policies promoting the deletion of small medical collections have no effect on credit access and financial health. Unlike other works in this literature, our setting consists of information deletion for accounts below a cut-off, allowing us to use regression discontinuity as one of our research designs.

We also relate to work on medical debt forgiveness, which is often considered a policy alternative to deleting information about medical debt. Kluender et al. (2024) conduct two large-scale randomized experiments forgiving medical debt and find no impact of debt relief on credit access, utilization, and financial distress. Taken together, the findings of our two papers suggest that neither deleting information about medical debt nor forgiving it are effective at alleviating financial distress for people burdened by medical debt. These results stand in contrast to studies on other types of debt forgiveness, which generally find positive results (Dobbie and Song, 2015; Di Maggio et al., 2020; Cespedes et al., 2024).

This paper also contributes to the literature using machine learning techniques for credit scoring and, more broadly, to better understand credit markets. Our finding that removing medical collections from credit scoring models leads to non-random sorting of consumers by predicted default probabilities despite their lack of predictive power is closely related to

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

Blattner and Nelson (2022), who show that default probability estimates for low-income consumers with thin credit files are noisier than those of other consumers. Similar to our paper, Liberman et al. (2019) uses machine-learning algorithms to study the effects of information deletion. Also related are works using machine learning for credit scoring (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). We contribute to these works by using machine learning credit scoring models to evaluate the predictive power of medical collections and identify consumers at risk of negative spillover effects following the deletion of this information.

Finally, we also add 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.

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

We use the Gies Consumer and Small Business Credit Panel (GCCP), a panel dataset with anonymized credit record data on 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.[4] The dataset spans 2004–2024, with annual snapshots of credit records as of the end of the first quarter of each year. 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 assigned and out of the panel through death or inactivity, mirroring 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.[5] In Table A.1, we benchmark the share of people with medical collections according to our classification against other sources.

We restrict our sample to the years 2019–2024 and to consumers aged 18 or older. We exclude people with missing age, credit score, or income data, 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 our sample is female. On average, consumers in the full sample 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 focus on individuals with medical collections. Consumers with medical collections have, on average, lower credit scores, lower income, and lower balances, and are substantially more likely to have a subprime credit record. They

[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.

have 2.44 medical collections on average, with 1.45 medical collections below $500.

### 2.1 Regression Discontinuity Sample

In our RD analysis, we further restrict the sample to consumers with a medical collection in 2022 and a non-missing credit score between 2022 and 2024. The resulting sample has 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 accounts are medical collections and 1.56 are medical collections below $500 on average.

## 3 Do Medical Collections Predict Default?

The effect of deleting medical collections hinges on their predictive power for future default. If medical collections predict default, some lenders will opt to use this information in their in-house credit scoring models, and thus deleting this information may alter their origination decisions and restrict the access to credit of some consumers. On the other hand, if medical collections are not meaningfully predictive of future default, deleting this information should not have consequences for risk pricing, including for consumers for whom the information is deleted.

Note that not all lenders have in-house credit scoring models, with some opting to rely solely on the output of models 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 directly impacts credit scores. VantageScore stopped including medical collections below $500 in its model in January 2023, followed by FICO a few months later.[6] This implies that the deletion of small medical collections should not affect the origination decisions of lenders that rely solely on these models.

[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.

For lenders that develop proprietary credit scoring models, whether they include medical collections in these models or not will depend on whether this information helps predict defaults. We thus start our analysis by investigating the predictive power of medical collections. To mimic the consequences of the April 2023 decision to delete medical collections below $500, we train two credit scoring models: one that includes data on medical collections below $500 and one that does not. While our models are likely not identical to those used by any particular lender, they rely on similar information and algorithms and perform well compared to other models in the literature. The assumption behind this analysis is that the output of our models is correlated to those of other well-designed models.

As we describe 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 and terms, and 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 remaning medical collections should also not lead to significant direct or indirect effects on credit access and terms.

## 3.1 Credit scoring with and without medical collections

Traditional credit scoring models, such as the FICO and the VantageScore, are typically logit models estimated using person-level data (Federal Reserve Board, 2007). The typical model's objective is to predict "default," defined as any account that becomes 90 days or more past due within the next 18–24 months, based on observed variables relating to payment history, amounts owed, length of credit history, new credit, and credit mix (Federal Reserve Board, 2007). Consistent with these traditional models, we select 48 features with 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] Unlike traditional models, our baseline algorithm is XGBoost, a state-of-the-art machine learning algorithm for classification problems.

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

We train two person-level credit scoring models: one including the number of medical collections below $500 and one excluding this information. To mimic the effect of the 2023 intervention, both models include information on medical collections above $500 and non-medical collections. We train both models using data from 2019 to 2021, a period preceding the removal of information on medical collections below $500. Specifically, we use features in 2019 data on over 2.8 million consumers to predict defaults occurring between 2020 and 2021, reserving 90% of observations to train the model and 10% to assess model performance out-of-sample.

We report model performance across a range of metrics in Table 3. The first column shows results for models with medical collections below $500, and the second column without. In our context, the Accuracy score—measured as the share of correct predictions—is not particularly meaningful because default is a relatively uncommon outcome. For example, a model that predicts that consumers never default would have an Accuracy score of 87%, equal to the share of consumers in the 2019 sample that do not default in 2020–2021. Similarly, the area under the Receiver Operating Characteristic curve (AUC), which measures the probability that the model will assign a higher default probability to a true defaulter than to a non-defaulter, is also not as informative as the precision or recall scores in this setting (Davis and Goadrich, 2006).

Precision and Recall scores 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 positive identifications (in our case, predicted defaults) that were correctly classified, while recall measures the proportion of actual positives that were correctly classified as such. We use a threshold of 50% when assigning individuals as positive or negative based on predicted probabilities. Because both precision and recall matter, we also compute the F1 score, which is a 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 does a good job of predicting default. We obtain an F1 Score of 0.557, which is a harmonic mean of recall and precision. Our recall of 0.448 is among the highest in the literature, which typically has values ranging from 0.35 to 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 when predicting default 3 months and 6 to 9 months ahead, respectively.[8] Predicting default in a 3-, 6-, or 9-month window typically yields higher precision and recall than a 24-month window, which likely contributes to the difference in performance between our model and those in Khandani et al. (2010) and Chioda et al. (2024).

Our precision score of 0.736 also compares favorably to the literature, where values often range from 0.06 to 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 0.853 precision when predicting default 3 months ahead. Precision-recall is a more appropriate performance metric than AUC in rare occurrence settings such as default prediction (Davis and Goadrich, 2006), but our AUC of 0.712 is within the typical range of 0.66 to 0.88 obtained in the literature.

Comparing the first two columns in Table 3 shows that removing medical collections has no impact on model performance. All metrics are unchanged up to third decimals except for Accuracy, which *increases* by 0.001 when small medical collections are removed. This provides strong evidence that medical collections below $500 are not meaningfully predictive of default. In the third column, we show that deleting information on all medical collections, including those exceeding $500, also has no impact on model performance. We thus expect that the CFPB's proposed rule to erase all remaining medical collections will also have little or no meaningful effect on consumers.

To further support the notion that medical collections are not an important feature, we report SHAP feature importance values in Figure 1, with variables in the y-axis ranked

[8] We compute recall and precision for Khandani et al. (2010) from the confusion matrix for the December 2008 3-month forecast using the classifier threshold of 50%. Chioda et al. (2024) use a threshold of 20% to classify positive identifications, which raises recall and decreases precision relative to a 50% threshold by increasing the number of true and false positives.

in order of importance. Small medical collections rank very low in importance with an average SHAP value of 0.0011, compared to an average value of 0.278 across the 10 most important features. Additional analysis, illustrated in Figure 2, shows that predicted default probabilities change little when removing information on medical collections below $500, with only approximately 10% of consumers seeing a change greater than 2 p.p. in absolute value.[9] Next, we argue that these changes are likely driven by noise rather than meaningful differences in default risk.

We start by categorizing 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 medical collections were meaningfully predictive of default, we would expect significant differences between those consumers whose predicted probabilities increase versus those whose probabilities decrease when medical collections are removed. Specifically, we would expect consumers with medical debt to see lower predicted default probabilities once that information is 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.

Table 4 shows summary statistics for all three groups. Negatively and positively treated consumers are similar in demographics and credit characteristics, but have lower credit scores,

[9]For comparison, 2 p.p. is the difference in 2022Q2–2024Q1 90 days or more delinquencies between consumers with a VantageScore of 300–500 and those with scores of 501–520 (https://www.vantagescore.com/lenders/risk-ratio/).

lower income, and lower balances than unaffected consumers. Figure 3 visualizes these results using balancing regressions. All variables are standardized and each dot is a regression coefficient of the variable labeled on the y-axis on either the positive (blue) or negatively treated flag (red). While both the positively and negatively treated groups are meaningfully different from the unaffected group, they are similar to each other. If medical collections were strongly predictive of default, we would expect a pronounced sorting effect along that variable. Figure 3 shows no evidence of sorting among positively and negatively treated consumers, who 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 in non-random sorting of consumers. This sorting arises because default probability estimates for low-income consumers with thin credit files are noisy (Blattner and Nelson, 2022). Consequently, even a feature with little to no predictive power such as small medical collections may receive positive loadings and influence predictions for these consumers. Excluding this predictor 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 them across the two groups randomly. To confirm this explanation, we next introduce a randomly generated variable into the model and demonstrate that excluding it reproduces the same pattern observed when excluding medical collections under $500, underscoring their limited predictive value (see Section 3.2).[10]

## 3.2 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,

[10]In theory, models estimated using state-of-the-art algorithms such as XGBoost should learn to disregard variables with no predictive value during training. In practice, these variables can still have non-zero importance due to limitations of finite sample sizes.

which does not include this random variable. We show performance metrics in Table A.2 in the Appendix. While removing small medical collections doesn't impact model performance, removing a random variable slightly improves performance, marginally increasing the F1 Score 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. In Figure A.1, we overlay the histogram of probability differences from Figure 2 with the analogous histogram obtained from 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 (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 flag (red). In panel B, we present an analogous plot by comparing two models, one with and without the noise variable, and re-sorting consumers into three groups as described in Section 3.1. We again find that "treated" consumers have, on average, lower credit scores, lower income, and lower balances, reinforcing the idea that estimates of default probabilities 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 effect of medical debt deletion on consumer credit outcomes. Accounts with less than $500 of medical debt were deleted. For analytical convenience, we recenter the medical debt running variable by measuring its value as distance (in dollars) from the $500 cutoff. Although treatment occurs at the account level, our main analysis is conducted at the individual level, which necessitates aggregating the running variable across an individual's accounts. We aggregate by taking the maximum debt value among all

accounts. Alternative methods of aggregation and their interpretations are discussed below.

Our main identifying assumption is that assignment to either side of the medical debt cutoff is as good as random. We believe this assumption is reasonable: it is difficult for individuals to precisely manipulate their medical debt, which depends on fixed prices that often are unknown in advance to the individual. Our data come from Experian's administrative records, which minimizes concerns about sample selection bias.

We estimate the following model at the individual level:

$$Y_{it} = \alpha DEBT_i^{2022} + \beta ABOVE_i^{2022} + \gamma(ABOVE_i^{2022} \times DEBT_i^{2022}) + \epsilon_{it} \tag{4}$$

The dependent variable, $Y_{it}$, is an outcome for individual $i$ in year $t$. The running variable, $DEBT_i^{2022}$, is measured as distance of the individual's maximum-medical-debt account from the \$500 cutoff, and $ABOVE_i^{2022}$ is an indicator equal to one if $DEBT_i^{2022} \geq 0$. 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.

We use a triangular kernel in all 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).

Our focal parameter of interest is $\beta$, 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.

## 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%).[11] 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 of deleting medical collections below $500 from credit reports. To do so, we restrict our sample to consumers whose predicted probability of default increases when medical collections below $500 are removed from our credit scoring model (negatively treated) and those whose probability of default decreases (positively treated), as described in Section 3. Because these groups are sorted based on their 2019 characteristics, we restrict our sample for this analysis to 2020–2024.

[11] 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.

If small medical collections are not predictive of default and the changes in predictive default that we document in Section 3 correspond to noise, we would expect no differential effect of the 2023 deletion on these two groups of consumers. As we discuss in Section 3 and show in Figure 3, positively and negatively treated groups are similar across a wide range of observables, making our key identifying assumption—that outcomes for the two groups would have evolved similarly in the absence of the information deletion—more plausible.

We estimate the following regression model at the individual level:

$$Y_{ict} = \alpha + \beta TREATED_i \times POST_t + \lambda_i + \delta_{ct} + \epsilon_{it}, \tag{5}$$

where $Y_{ict}$ is an outcome for consumer $i$, residing in county $c$, in year $t$; $TREATED_i$ is a binary variable equal to one if the consumer belongs to the negatively treated group and zero if they belong to the positively treated group; $POST_t$ equals zero before 2023 and one after; and $\lambda_i$ and $\delta_{ct}$ are vectors of consumer and county-year fixed effects, respectively. Because we assign treatment based on differences in the predicted probability of default with and without medical collection, 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. The negative and positively treated groups correspond to approximately 10% of the full sample, and we are thus left with 100 clusters.

Our coefficient of interest is $\beta$ which, under the assumption of parallel trends, is an estimate of the average effect of deleting medical collections on outcomes of consumers whose predicted probability of default increases when medical collections are removed, relative to consumers whose predicted probability of default decreases. Note that, to the extent that removing medical collections affects the credit outcomes that we measure, the control group—consumers whose predicted probability of default decreases when medical collections are deleted—would benefit. That means that $\beta$ is potentially an overestimate of what we would obtain by comparing consumers in our negatively treated group to a control group that is unaffected by the information deletion. Thus, the true indirect of deleting medical collections may be even smaller than the estimates we report below, which rule out with 95% confidence even a small impact.

To show evidence in favor of the parallel trends assumption, we also estimate event-study versions of Equation (5):

$$Y_{ict} = \alpha + \sum_{\kappa=2020}^{2024} \beta_\kappa TREATED_i \times \mathbb{I}_{\kappa=t} + \lambda_i + \delta_{ct} + \epsilon_{it}, \tag{6}$$

where $\mathbb{I}_{\kappa=t}$ is a binary variable equal to one in year $t$ and zero otherwise. We omit the binary variable for the year prior to the intervention (2022), which serves as the reference year.

## 5.2 Results

Figure 7 plots coefficient estimates and 95% confidence intervals for Equation (6) for 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). Across all outcomes, we find that consumers in the negative and positively treated groups evolved in close parallel before 2023, supporting the parallel trends assumption. After 2023, we see no differential effects between the two groups in these measures.

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.[12] Taken together, these results suggest that the positive treated group is a good counterfactual for the negatively treated group and that deleting medical collections from credit reports does not have negative spillover effects on other consumers.

We report pooled estimates from Equation (5) in Table 6, for the same set of outcomes as Figures 7 and 8. All but one estimate are statistically insignificant and all are small relative to the sample mean. The one significant estimate is a 0.721 increase in credit scores, which corresponds to 0.12% of the sample mean of 617.13 and is thus not economically significant. Overall, our estimates are precise and rule out even small negative spillovers from the deletion

[12] We show additional outcomes in Appendix Figure A.5.

of small medical collections with 95% confidence.

# 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 random-

ized 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 (2024). 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.

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.

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. *Journal of Finance Forthcoming*.

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). The effects of medical debt relief: Evidence from two randomized experiments. *Working paper*.

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

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

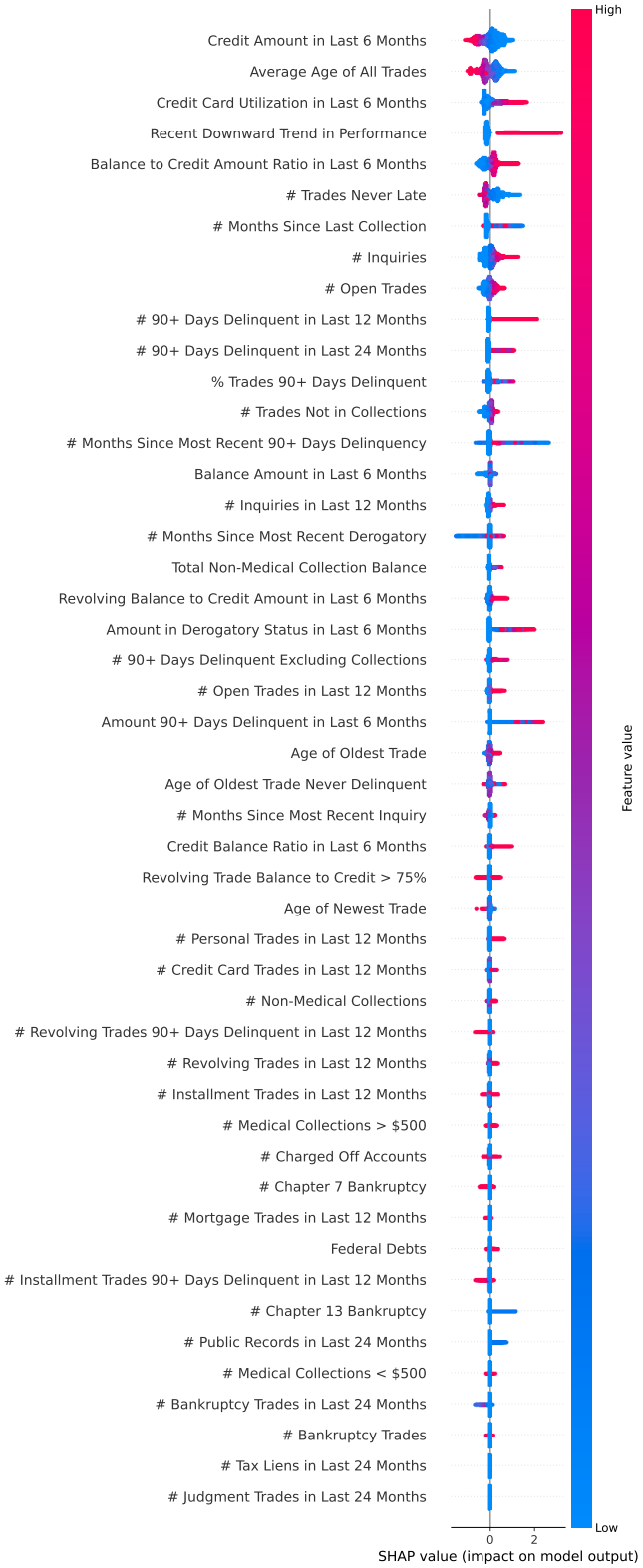
Meursault, V., D. Moulton, L. Santucci, and N. Schor (2022). The time is now: Advancing fairness in lending through machine learning. *FRB of Philadelphia Working Paper No. 22-39*.

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.

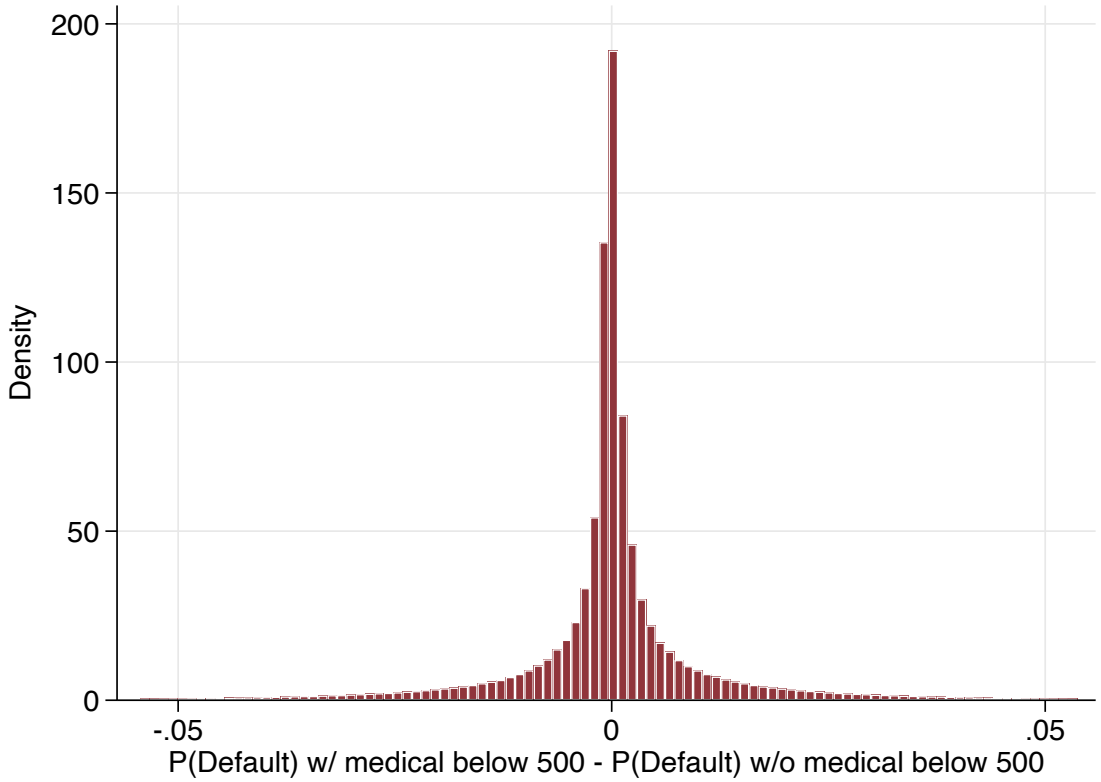
Sandler, R. and L. Nathe (2022). Paid and low-balance medical collections on consumer credit reports. *Consumer Financial Protection Bureau Office of Research Reports Series* (22-5).

**Figure 1:** SHAP Values for XGBoost Model With Medical Collections



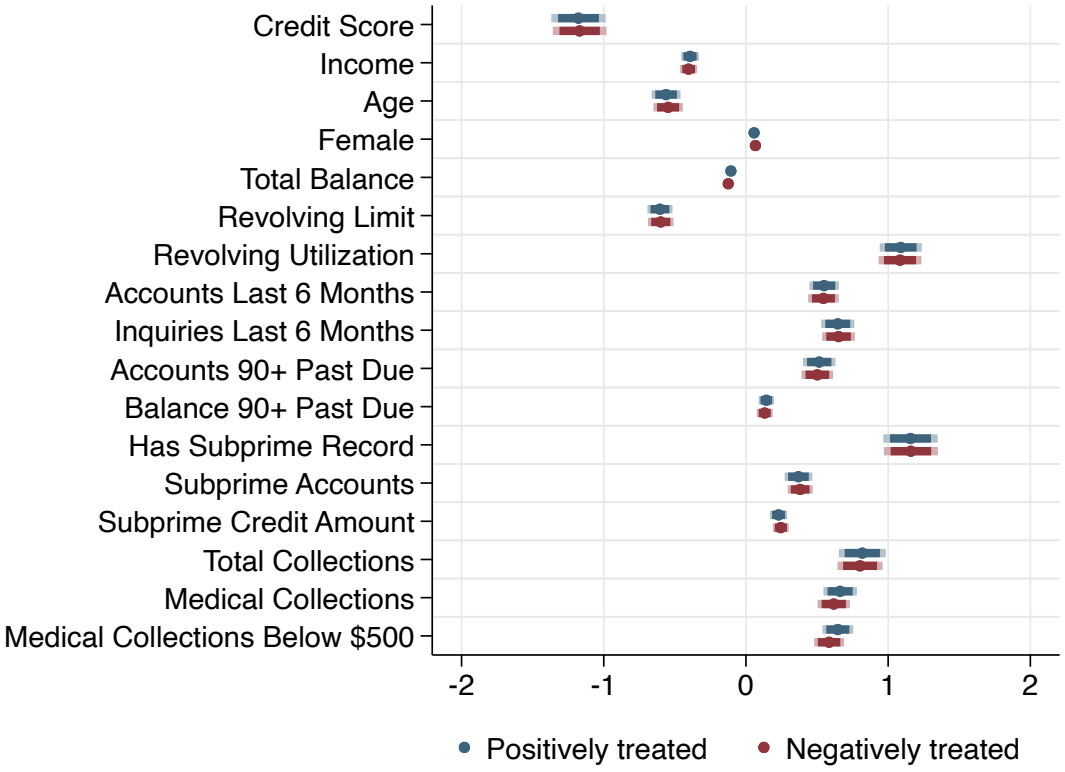

This figure shows variable importance measures (SHAP values) for our baseline XGBoost model including medical collections below $500. The features labeled on the y-axis are ranked according to their average importance and the distribution of SHAP values is shown in each row.

**Figure 2:** Changes in Predicted Default Probabilities from Removing Medical Collections




This figure shows a histogram of the difference in predicted probability of default under our baseline credit scoring model with medical collections below $500 and in the model without this information. In both versions of the model, default is predicted using 2019 features.

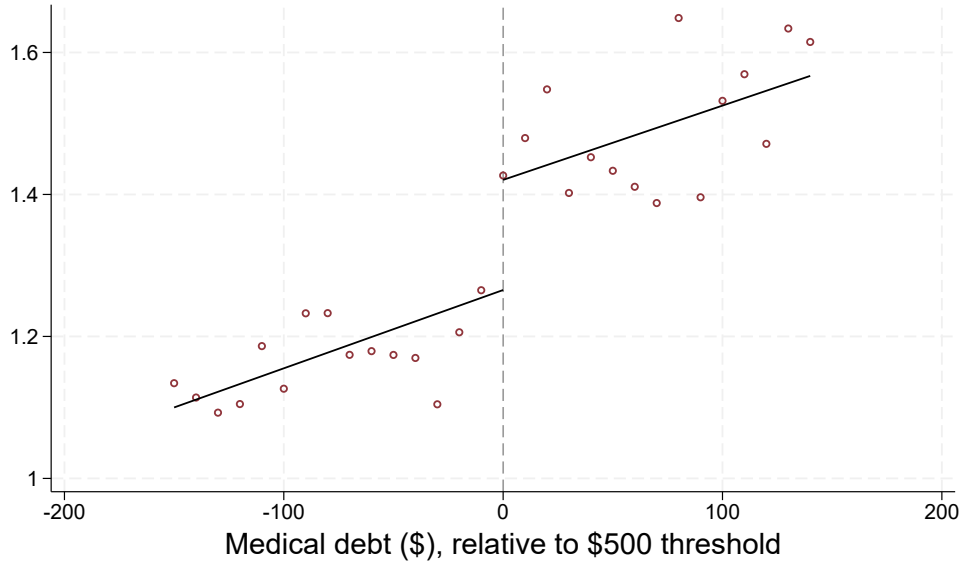
**Figure 3:** Covariate Balance by Changes in Default Probabilities



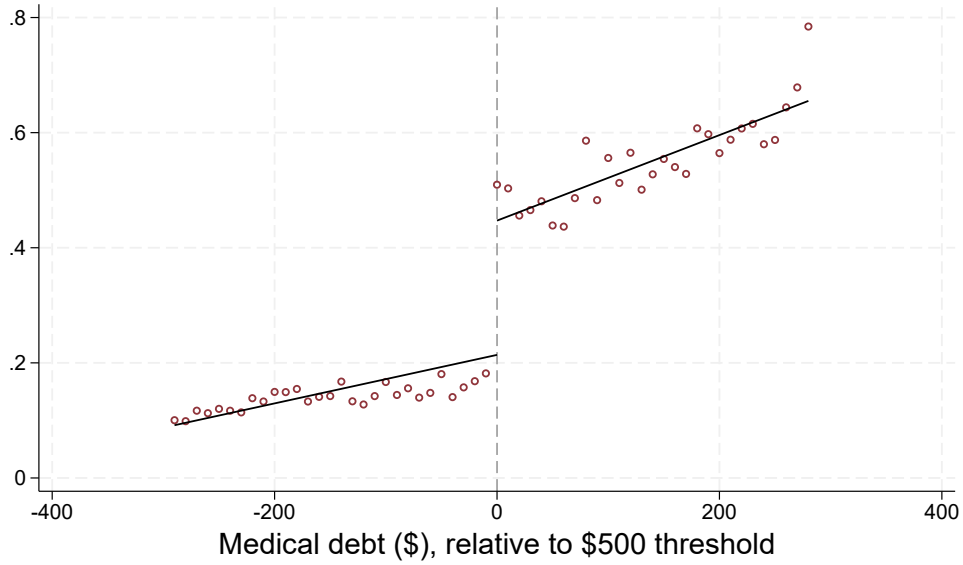

This figure shows coefficients of balancing regressions for a range of covariates, denoted in the y-axis. Negatively Treated refers to consumers above the 95th percentile in the distribution of probability difference, whose predicted probability of default increases by 2 p.p. or more when medical collections 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 2 p.p. or more when medical collections below $500 are removed from our baseline credit scoring model. The omitted category consists of consumers between the 25th and 75th percentiles, who experience a change in predicted default probabilities (in absolute value) of 0.0015 p.p. or less. We construct 100 equal-sized bins of the difference in predicted probability of default and cluster standard errors at the bin level.

**Figure 4:** Number of Accounts in Collections, 2024

**(a)** Total Collections (#)

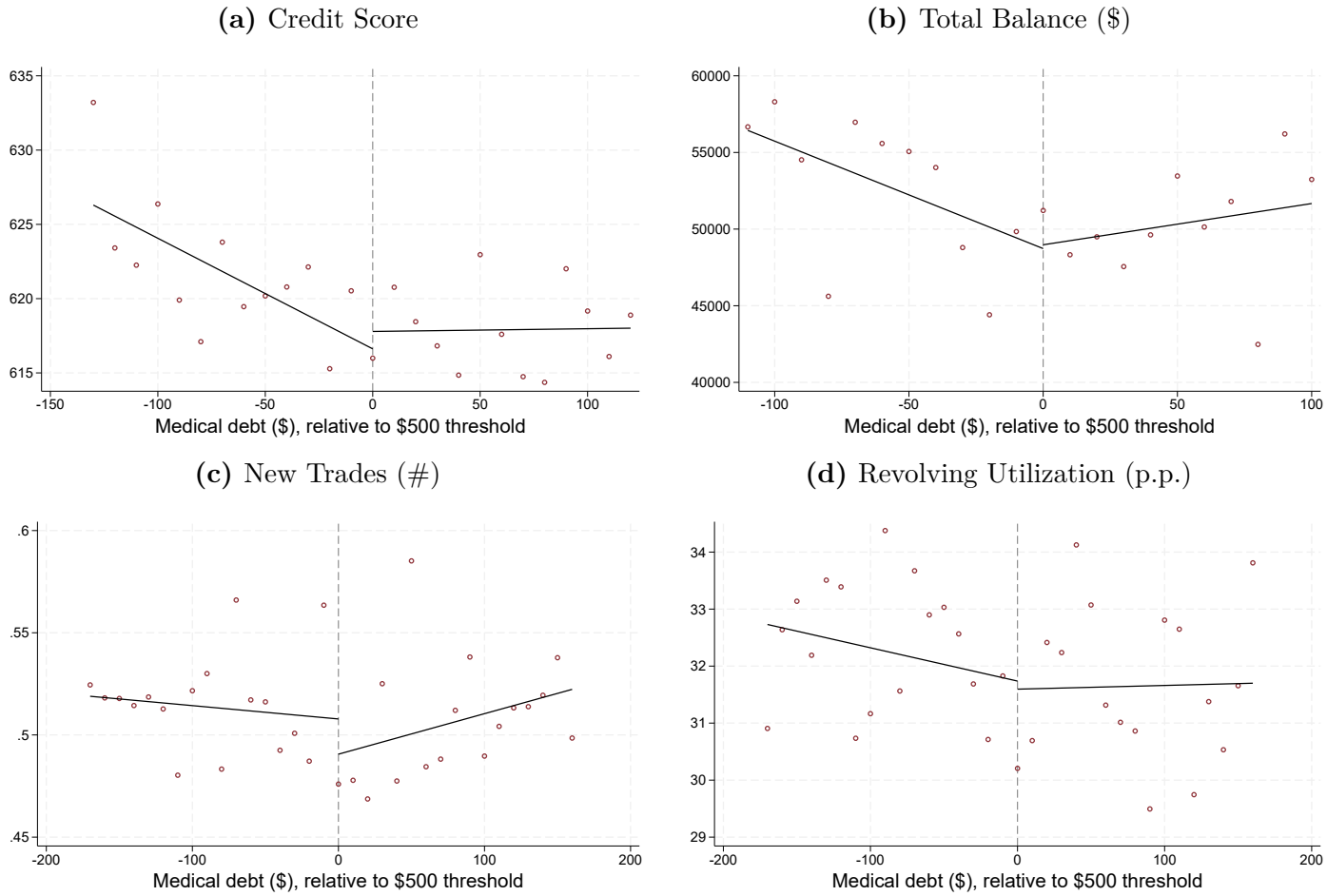



**(b)** Medical Collections (#)




Notes: This figure shows the average number of consumer accounts in collections 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. Total Collections is the total number of trades in collection on a consumer's credit report. Medical Collections is the total number of trades in collection under Medical or Healthcare category. The fitted lines are estimated using Equation (4).

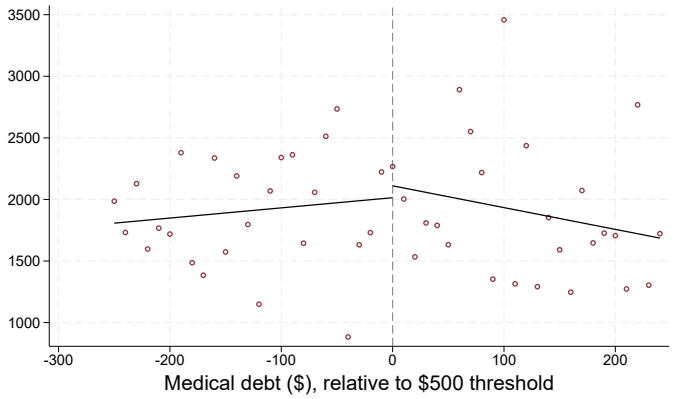
**Figure 5:** Credit Access and Utilization, 2024



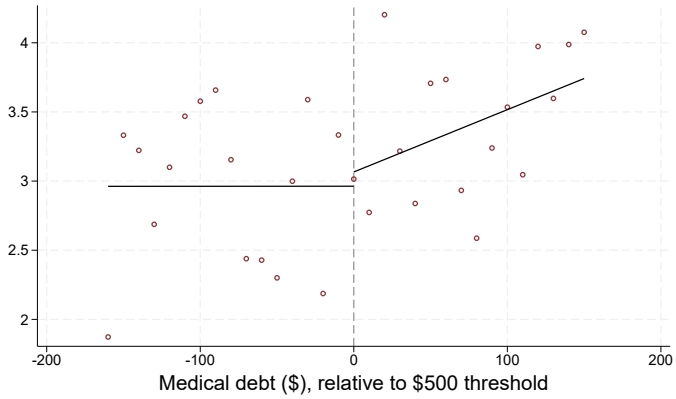

Notes: This figure shows the average consumer credit score, total balance, revolving limit, and revolving utilization by medical debt, relative to the $500 threshold. The running variable for medical debt corresponds to the highest debt amount across the consumer's medical collections accounts. Credit 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 (4).

**Figure 6:** Derogatories and Alternative Credit, 2024

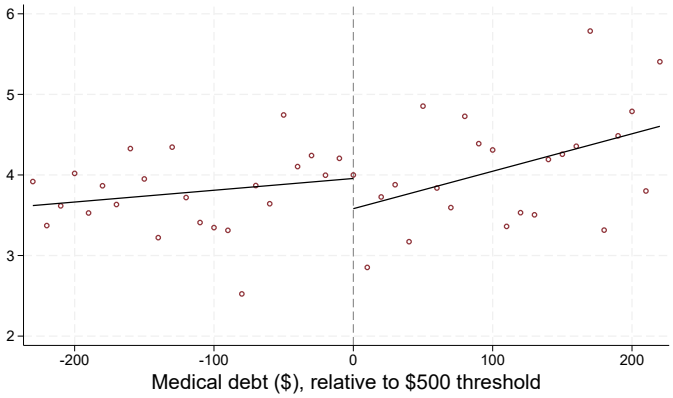
**(a)** Balance 90 days+ Past Due ($)



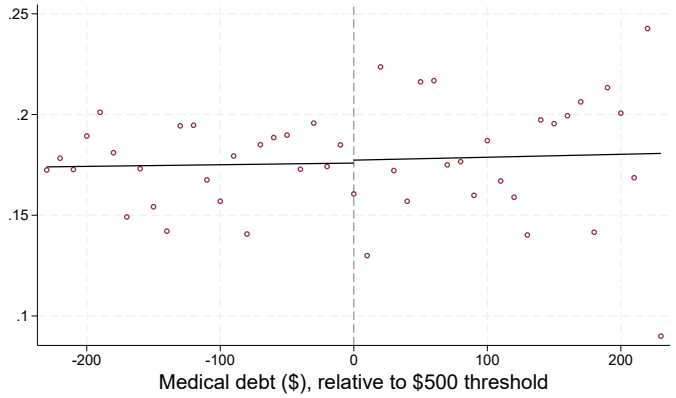

**(b)** Has Bankruptcy Flag (p.p.)




**(c)** Has Alternative Credit Balance (p.p.)

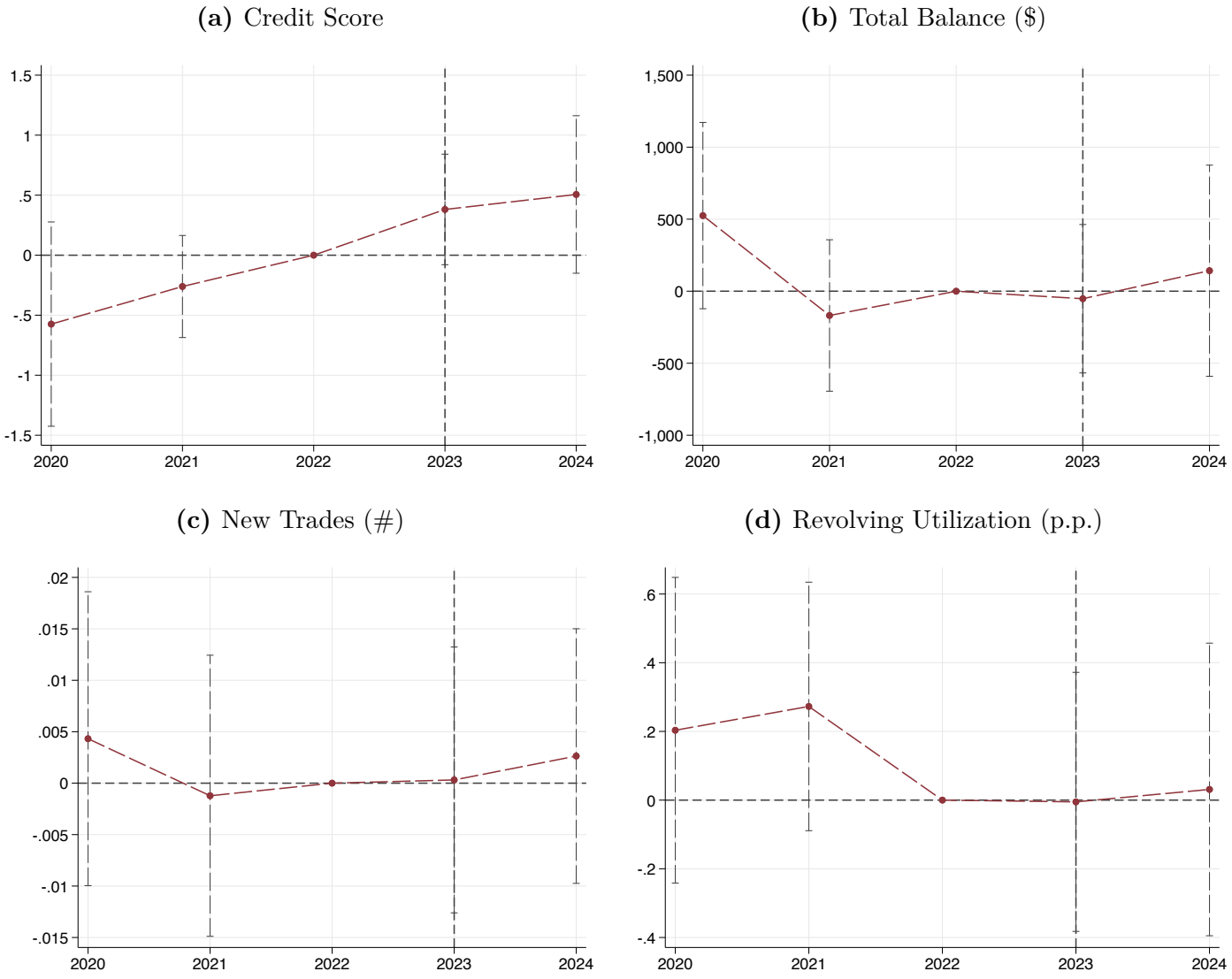



**(d)** Alternative Credit Trades (#)



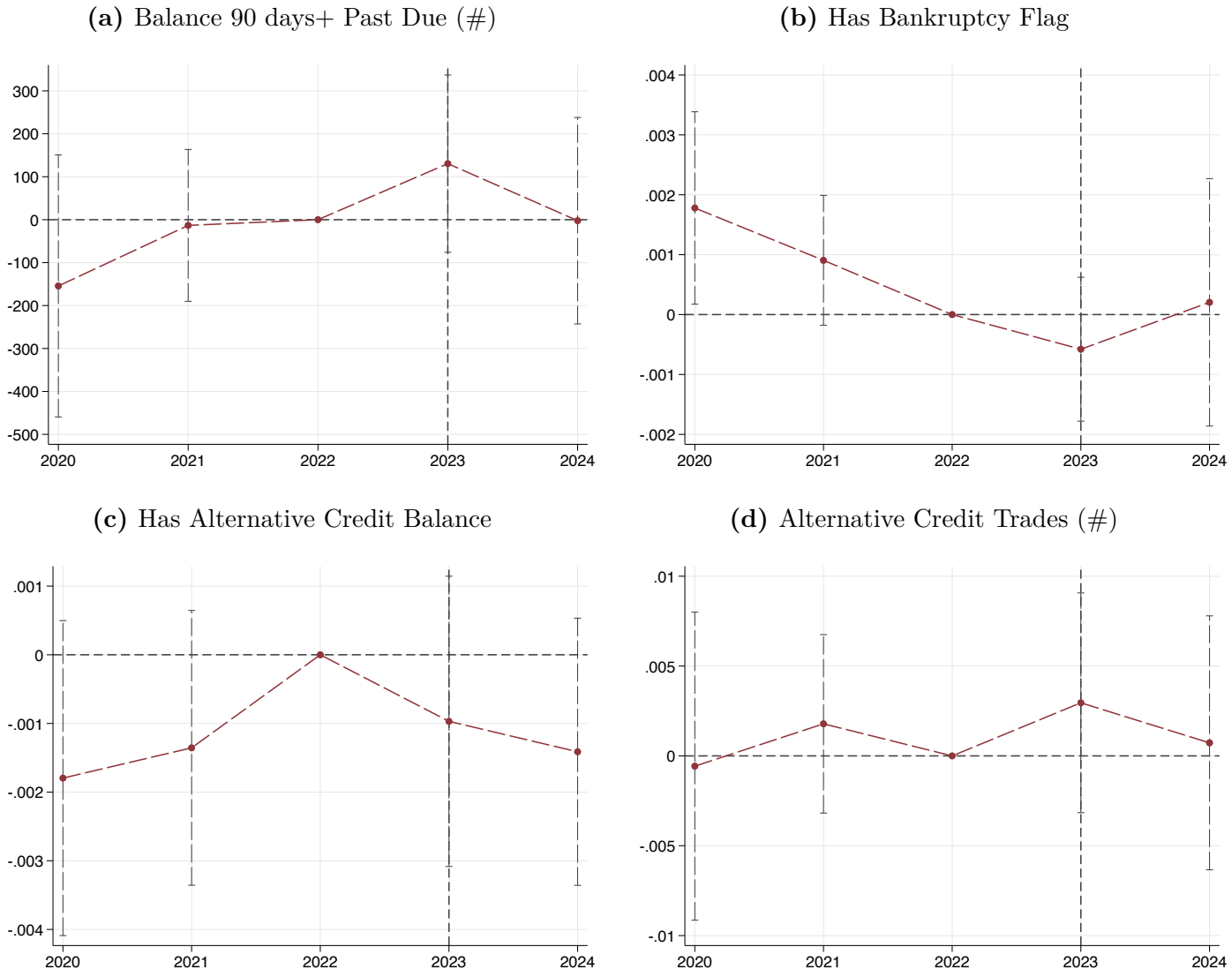

Notes: This figure shows the average consumer number of past-due trades, alternative trades, and alternative credit by medical debt, relative to the $500 threshold. The running variable for medical debt corresponds to the highest debt amount across the consumer's medical collections accounts. Balance 90 days+ Past Due is the total balance on trades 90 days or more past due. Has Alternative Credit Balance is a binary variable that equals one if the consumer has a positive alternative credit amount. Alternative Credit Trades is the total number of alternative credit accounts in the consumer's Clarity report. The fitted lines are estimated using Equation (4).

**Figure 7:** Event Study: Credit Access and Utilization




This figure shows coefficient estimates and 95% confidence intervals of Equation (6) for the outcomes denoted in panel captions. 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 months. Revolving Utilization is the balance-to-limit ratio across all revolving credit accounts. 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.

**Figure 8:** Event Study: Derogatories and Alternative Credit




This figure shows coefficient estimates and 95% confidence intervals of Equation (6) for the outcomes denoted in panel captions. Balance 90 days+ Past Due is the total balance on trades 90 days or more past due. Has Alternative Credit Balance is a binary variable that equals one if the consumer has a positive alternative credit amount. Alternative Credit 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 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 1:** Summary Statistics

| | Full Sample | | | With Medical Collection | | |
|---|---|---|---|---|---|---|
| | 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 shows descriptive statistics between 2019 and 2024 for our full sample in the first three columns and of the sample conditional on at least one medical collection on the next three columns. All variables come from the Gies Consumer and Small Business Credit Panel.

**Table 2:** Summary Statistics for Consumers with Medical Collections, 2022

| | 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 shows summary statistics from the Gies Consumer and Small Business Credit Panel (GCCP). The statistics are calculated using 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 in 2022–2024 with at least one medical collection account.

**Table 3:** Credit Scoring Metrics

| 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 the performance of our two baseline credit scoring models: with and without information of medical collections below $500. In the last column, we also report the performance of a model without any medical collections, above or below $500. 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 4:** Summary Statistics by 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:** Direct Effect: Regression Discontinuity Design

| 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) |
| 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 ($\beta$) estimates and 95% confidence intervals of Equation (4). 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

| Dependent Variable: | 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 (5). 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

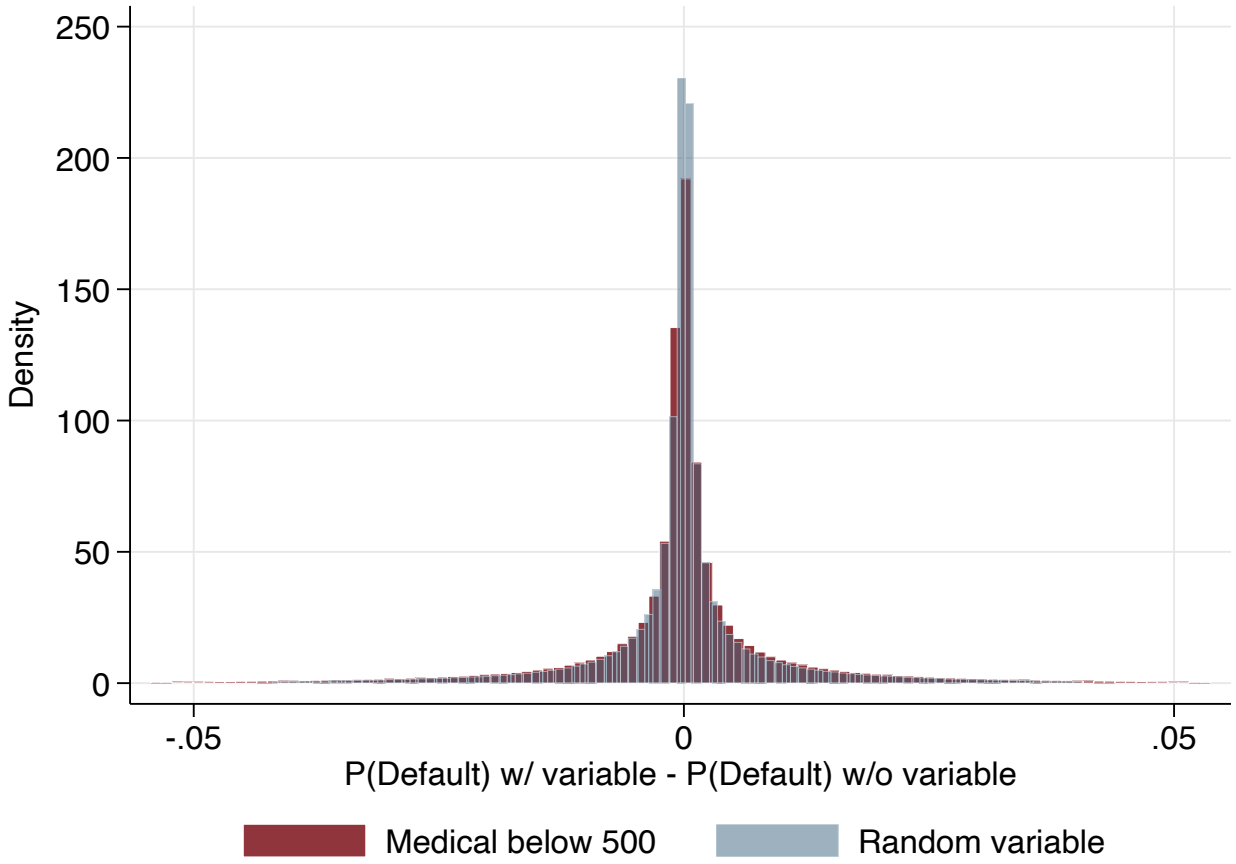
# Online Appendix

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

Victor Duarte, Julia Fonseca, Divij Kohli, Julian Reif

## A Additional Figures and Tables

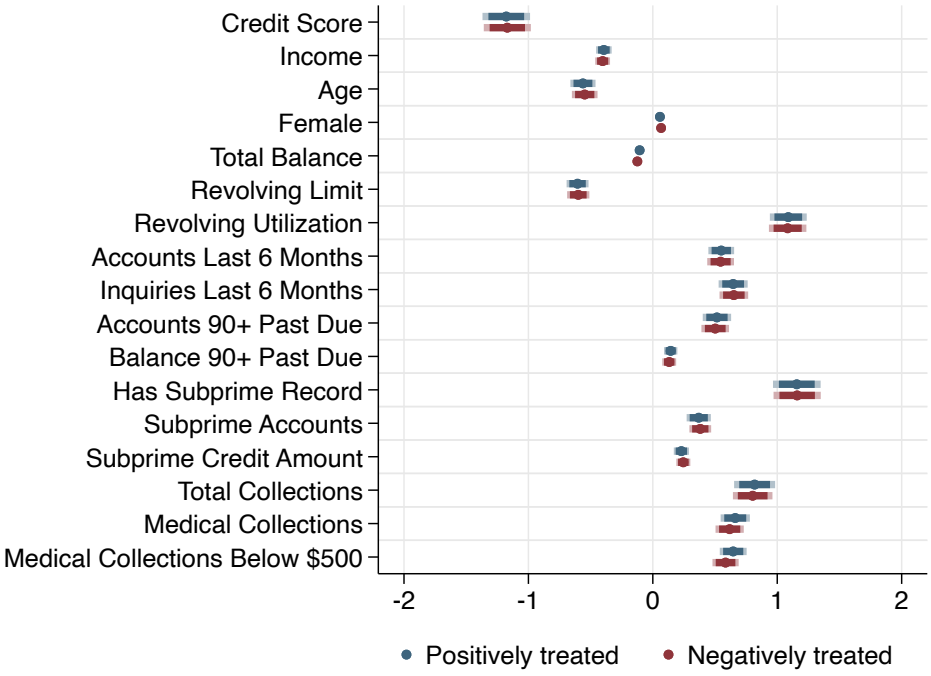
**Figure A.1:** Changes in Predicted Default Probabilities: Medical vs. Noise



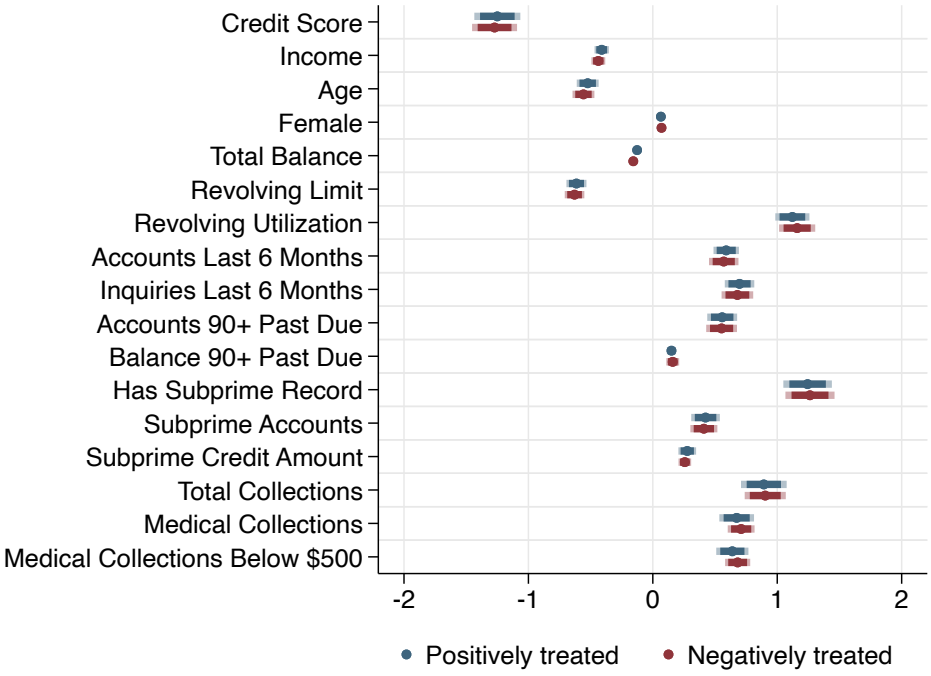

This figure shows a histogram of the difference in predicted probability of default under our baseline credit scoring model with medical collections below $500 and in the model without this information in red. In blue, it shows an analogous histogram for credit scoring models with and without a feature drawn randomly from a uniform distribution. In all versions of the model, default is predicted using 2019 features.

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

**(a)** Medical Collections Below $500



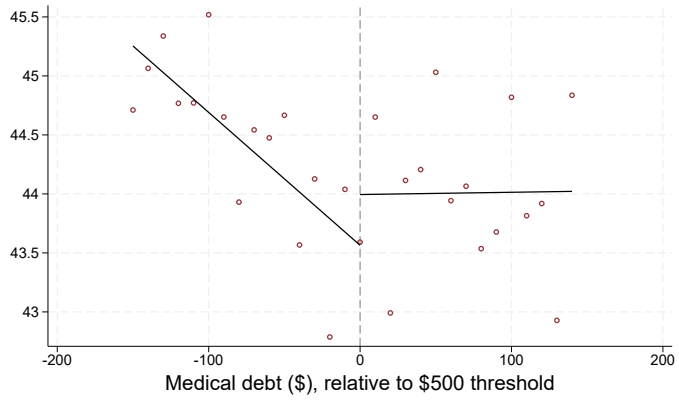

**(b)** Random Variable



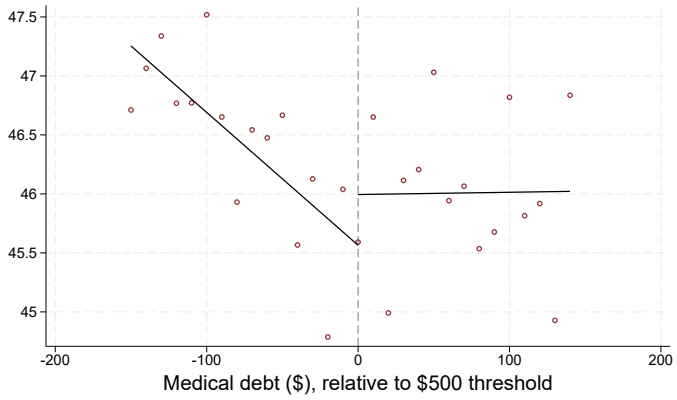

This figure shows coefficients of balancing regressions for a range of covariates, denoted in the y-axis. In each panel, we train credit scoring models with and without a variable and compute the difference in predicted probabilities of default under each model using 2019 data. In panel A, the variable i smedical collections below $500 and, in panel B, the variable is randomly drawn from a uniform distribution. 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. The omitted category consists of consumers between the 25th and 75th percentiles. We construct 100 equal-sized bins of the difference in predicted probability of default and cluster standard errors at the bin level.

**Figure A.3:** Covariates Smoothness

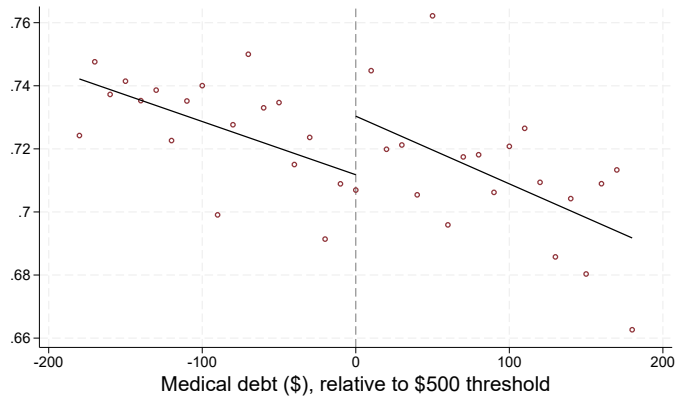
**(a)** Age in 2022



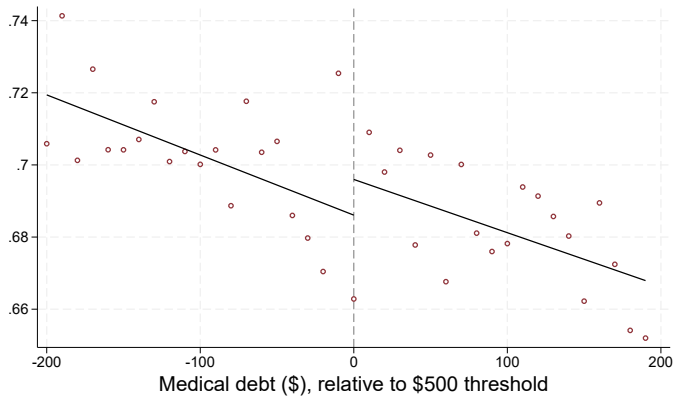

**(b)** Age in 2024



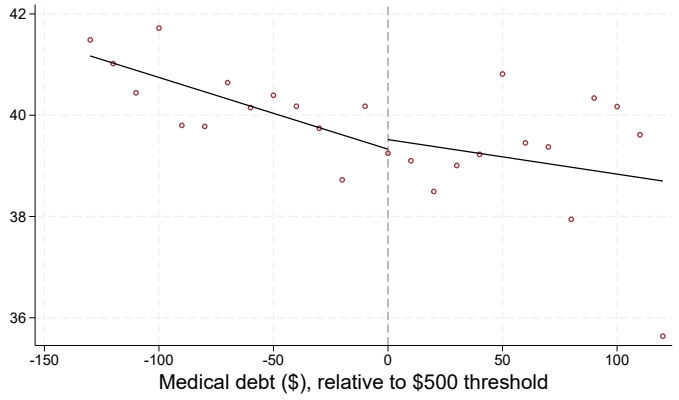

**(c)** Married in 2022



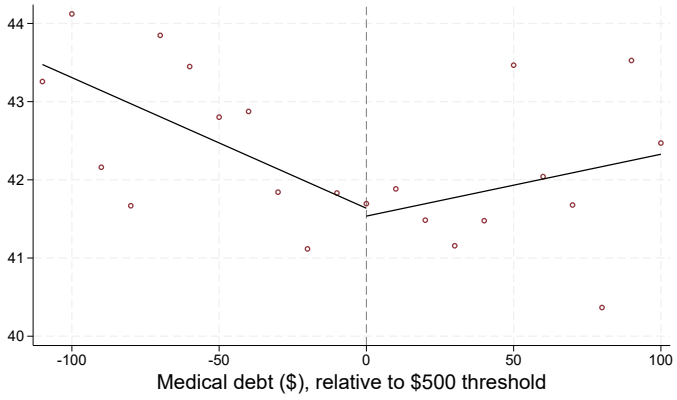

**(d)** Married in 2024




**(e)** Income in 2022



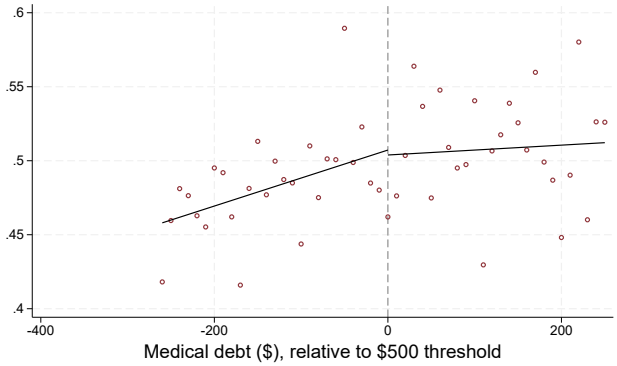

**(f)** Income in 2024



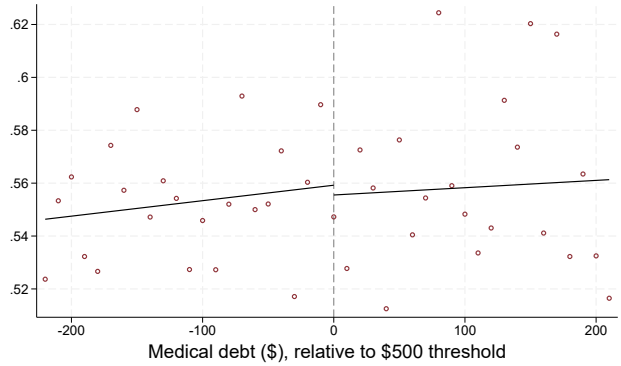

Notes: This figure shows the average Age, Marital Status and Income by medical debt, relative to the $500 threshold. The running variable for medical debt corresponds to the highest debt amount across the consumer's medical collections accounts. The fitted lines are estimated using Equation (4).

**Figure A.4:** RD Design: Additional Outcomes

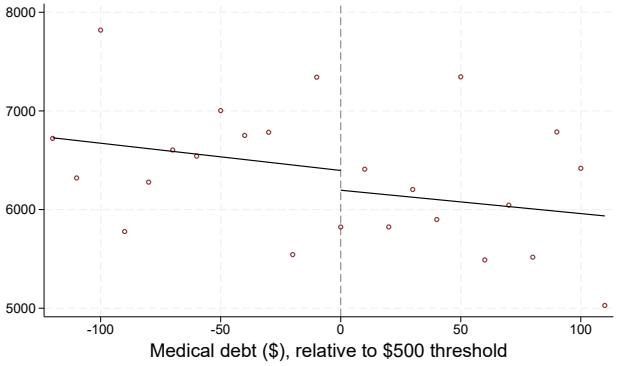
**(a)** Trades 90 days+ Past Due ($)



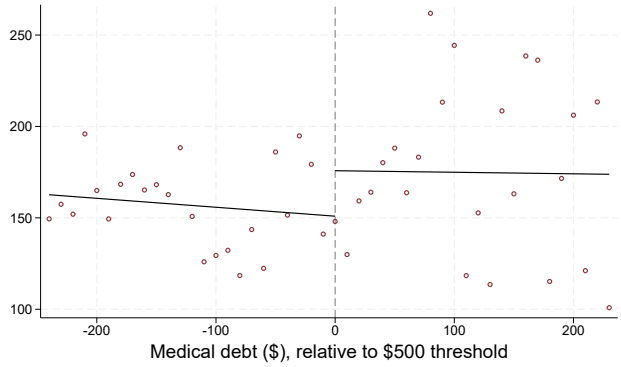

**(b)** New Inquiries (#)



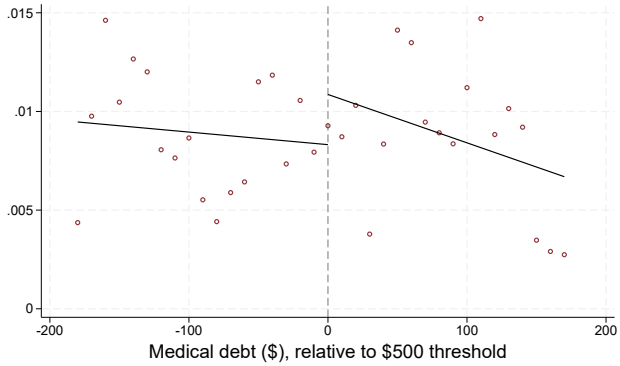

**(c)** Revolving Limit ($)




**(d)** Alternative Credit ($)



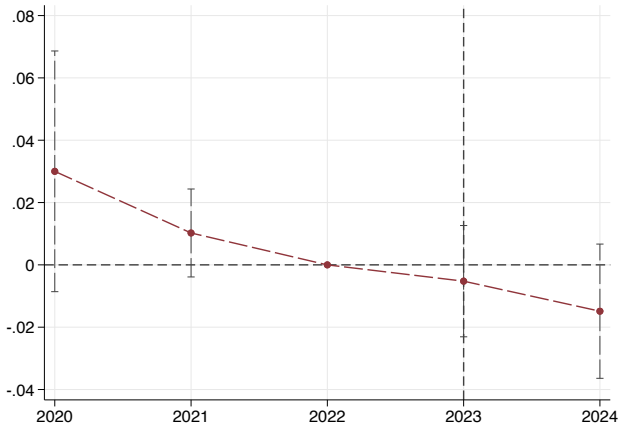

**(e)** New Mortgage Trades (#)



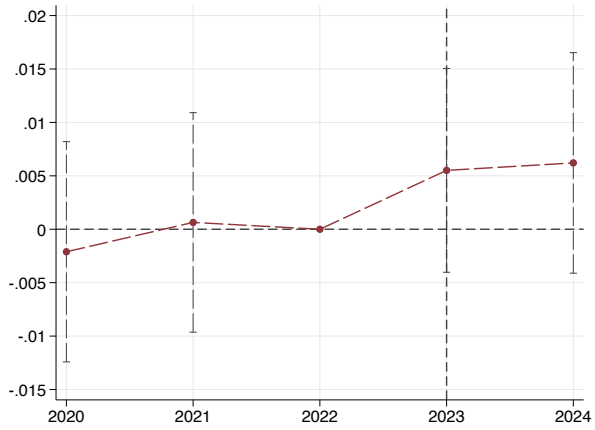

Notes: This figure shows coefficient estimates and 95% confidence intervals of Equation (4) for the outcomes denoted in panel captions. Trades 90 days+ Past Due is the number of trades 90 days or more past due. New Inquiries is the number of inquiries in the prior 6 months. Revolving Limit is the total limit across all revolving accounts. New Mortgage Trades is the number of mortgage trades opened in the prior 6 months.

**Figure A.5:** Event Study: Additional Outcomes

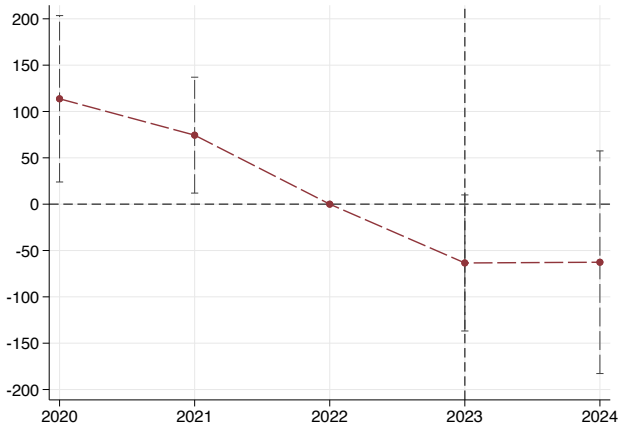
**(a)** Trades 90 days+ Past Due ($)



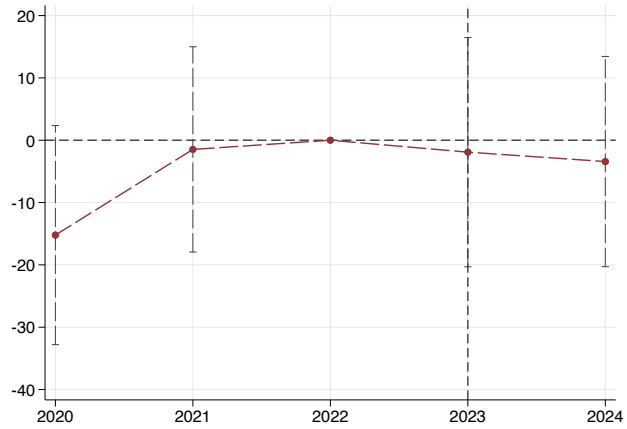

**(b)** New Inquiries (#)



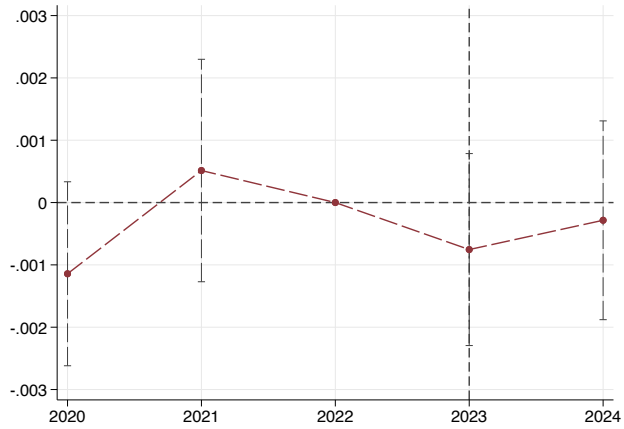

**(c)** Revolving Limit ($)




**(d)** Alternative Credit ($)




**(e)** New Mortgage Trades (#)




This figure shows coefficient estimates and 95% confidence intervals of Equation (6) 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.