The Burden of Medical Debt and the Impact of Debt Forgiveness

Date: September 18, 2018

Raymond Kluender[1], Neale Mahoney[2], Francis Wong[3], and Wesley Yin[4]

Abstract

This analysis plan pre-specifies our planned approach and main analysis of a randomized controlled trial of medical debt forgiveness. Collaborating with RIP Medical Debt, a non-profit that buys and abolishes medical debt, medical debt will be forgiven for randomly chosen "treated" individuals, while collections efforts will proceed as normal for "control" individuals. Outcomes will be measured using credit bureau data.

Departments

1. Department of Economics, Massachusetts Institute of Technology
2. Booth School of Business, University of Chicago
3. Department of Economics, University of California Berkeley
4. Department of Public Policy and Anderson School of Management, University of California Los Angeles

Introduction

Medical debt is potentially a large burden for many Americans—with 44 million individuals holding an aggregate $75 billion in medical debt. While these nominal amounts are staggering, it is unclear to what extent medical debt harms financial well-being. Medical debt recovery rates are low, suggesting that the pure "balance sheet" cost of medical debt is modest for most individuals. Yet medical debt may harm individuals through lower credit scores, higher interest rates, and reduced access to credit—impairing economic opportunities and perhaps even locking individuals in "debt traps."

Experimental Design

We are conducting a randomized-control trial (RCT) to estimate the impact of medical debt forgiveness on measures of financial well-being. Individuals are randomly assigned to one of two arms:

1) Control: No contact is made with debtors, and no action is taken on medical debt, by researchers; ongoing collections, to the extent that they occur, proceed as normal. The control arm captures status quo interactions with debt collections.
2) Treatment: For these subjects, RIP Medical Debt will purchase all observed medical debt in collections held by First Financial Asset Management (FFAM), and then retire that debt from collections. We will mail subjects letters informing them of the source and dollar amount of medical debt that was abolished.

Stratification

We stratify the sample along the following dimensions:

1) *Location of debtor*. Due to the expenditure allocation needs of our partner organization, debt forgiveness is randomized within specific purchase areas (i.e. metropolitan areas and states).
2) *Age of the debtor*. Credit scores, credit supply, and credit demand exhibit strong lifecycle patterns. Older individuals are defined as those aged 45 or older as of March 2018, while younger individuals are those younger than age 45.
3) *Age of the debt*. Recovery rates are decreasing in the age of the debt, suggesting that the mechanisms by which debt impacts financial outcomes may depend on the age of debt, whether because incentives to pay down debt differ by age of debt, or selection effects of consumers who carry older debt. Individuals are grouped by the age of their most recent medical debt in our debt sample. Recent debt is classified as any debt that was originated during or after 2012.
4) *Amount of the debt*. Debt forgiveness may yield larger impacts when the face value of the debt is large. High-debt individuals are defined as those whose total debt is strictly greater than $1,000, while low-debt individuals have total debt less than or equal to $1,000.

Primary outcome measure

1. Number of accounts past due
   Definition: Count of tradelines with payments 30+ days past due

Secondary outcome measures

1. Collections domain ("first stage")

    a. Number of debts in collection
       Definition: Count of collections items

    b. Amount of debt in collections
       Definition: Sum of debt in collections

2. Non-medical financial distress domain

    a. Number of accounts in default
       Definition: Count of (non-medical) tradelines with payments 90+ days past due

    b. Account balances past due
       Definition: Sum of account balances for (non-medical) tradelines with payments 30+ days past due

    c. Account balances in default
       Definition: Sum of account balances for (non-medical) tradelines with payments 90+ days past due

3. Public records domain

    a. Chapter 7 bankruptcy in last 12 months
       Definition: Indicator for Chapter 7 filing in last 12 months

    b. Chapter 13 bankruptcy in last 12 months
       Definition: Indicator for Chapter 13 filing in last 12 months

    c. Bankruptcy in last 12 months (aggregate measure)
       Definition: Indicator for bankruptcy filing in last 12 months

4. Credit supply domain

    a. Credit score (conditional on having one)
       Definition: Vantage Score (or if not available most widely used credit score available)

    b. Credit score indicator
       Definition: Indicator for having credit score used in 4a

    c. Credit limits
       Definition: Sum of credit limits across credit cards

5. Borrowing domain

   a. Number of credit cards
      Definition: Number of credit card accounts

   b. Number of mortgages
      Definition: Number of mortgages

   c. Number of auto loans
      Definition: Number of auto loans

   d. Credit card balances
      Definition: Sum of credit card balances

   e. Mortgage balances
      Definition: Sum of mortgage balances

   f. Auto loan balances
      Definition: Sum of auto loan balances

   g. Number of loans (aggregate measure)
      Definition: Number of loan accounts

   h. Total balances (aggregate measure)
      Definition: Sum of balances on all loan accounts

Empirical specification

We will estimate the treatment effect of debt abolishment (combined with the information intervention) with the following ordinary least squares equation:

$$Y_i = \beta_1 \text{Treatment}_i + \beta_{2,s(i)} + X_i{}'\beta_3 + \epsilon_i$$

In the above equation, i indexes individuals, Treatment_i is an indicator for whether the individual was assigned to the treatment group, $\beta_{2,s(i)}$is a fixed effect corresponding to stratification groups, and X_i is a vector of observable baseline individual characteristics ("controls") in the FFAM and credit bureau data.

The coefficient β_1 captures the average treatment effect of debt abolishment on outcome Y. We control for strata fixed effects because treatment status is only randomly assigned conditional on strata. While controlling for individual characteristics is not necessary for causal interpretation of the estimates, their inclusion should increase the precision of our estimates.

Inference

For our analysis of primary outcome, we will cluster our standard errors at the individual level. For our analysis of secondary outcomes, we will adjust our standard errors to account for multiple testing by domain using two approaches. First, within each domain, we will report family-wise p-values that adjust for multiple outcomes using the free step-down resampling method of Westfall and Young (1993). We will also report unadjusted p-values for reference. See Anderson (2008) for details on this approach and Finkelstein et al. (2012) for an application.

Second, as indicated in the outcomes section above, in some domains we will construct aggregate outcomes variables. These aggregate measures are similar to the summary index outcomes used in, for example, Anderson (2008) and allow us to measure whether there are aggregate effects for the domain without raising concerns about within-domain multiple testing. We will naturally not include these aggregate measures (denoted as such in the secondary outcomes) when we use the free step-down resampling method to adjust for multiple inference.

Heterogeneity

We are not only interested in the average effect of debt forgiveness but also in heterogeneity in the effect across individuals. We will conduct a heterogeneity analysis by estimating our baseline specification on samples split one-by-one (i.e., not fully interacted) on the following dimensions:

1. Location of the debtor. Because we do not know the geographic areas for which debt will be available, we cannot pre-specify exactly how we will geographically split our sample.
2. Age of the individual (aged 45 or older vs. younger than 45)
3. Age of the debt (2012 or more recent vs. 2011 or older)
4. Amount of the debt (less than or equal to $1,000 versus greater than $1,000)

The effects of debt forgiveness may vary based on if the medical debt appears on the individual's credit report and the extent to which the individual has other debt in collections. Examining heterogeneity on these dimensions requires baseline (i.e., time of intervention) credit report information. If we are able to obtain baseline credit reports, we will conduct a heterogeneity analysis by estimating our baseline specification on samples split one-by-one (i.e., not fully interacted) on the following dimensions:

5. Whether the debtor has at least one "other" debt in collections (i.e. collections that are not observed in our FFAM medical debt data).
6. Whether the medical debt we observe in the FFAM medical debt data appears on the credit report.

We will also investigate heterogeneity using the machine learning techniques proposed in Athey and Imbens (2016) to identify groups for whom medical debt forgiveness is particularly beneficial. For these analyses, we will use all available information from the FFAM data and variables from the baseline credit reports (if available) as input variables.

References

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, *103*(484), 1481-1495.

Athey, S., and Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, *113*(27), 7353-7360.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J.P., Allen, H., Baicker, K. and Oregon Health Study Group (2012). The Oregon health insurance experiment: evidence from the first year. *The Quarterly Journal of Economics*, *127*(3), 1057-1106.

Westfall, P. H., and Young, S. S. (1993). Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment (Wiley Series in Probability and Statistics).