Health and Health Care Utilization Effects of Medical Debt Forgiveness

Date: April 15, 2021

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

Brief Summary

The goal of this study is to estimate the direct, causal impact of medical debt on health care utilization, mental health, and wellbeing of patients. The investigators will conduct a survey to measure the impact of the debt forgiveness on health care use, mental health, and wellbeing. The survey will be administered to approximately 17,000 subjects of a recent medical financial intervention. In that prior intervention, a non-profit charity, RIP Medical Debt, purchased and abolished medical debt for a randomly selected about 6,000 (out of the 17,000) individuals. In this current protocol, the investigators will administer the survey, and will compare surveyed outcomes of subjects who received and did not receive the intervention.

Departments

1. Harvard Business School
2. Department of Economics, Stanford University
3. National Bureau of Economic Research
4. Department of Public Policy and Anderson School of Management, University of California Los Angeles

Introduction

This study will estimate the direct, causal impact of medical debt on health care utilization, mental health, and wellbeing of patients. To do so, the investigators will administer a survey to approximately 17,000 subjects of a recent medical financial intervention. In that intervention, a non-profit charity, RIP Medical Debt, purchased and abolished medical debt for a randomly selected about 6,000 (out of the 17,000) study subjects. In this current protocol, the investigators will compare surveyed outcomes of subjects who received and did not receive the medical debt abolishment intervention. Because debt abolishment was randomized, comparing surveyed outcomes of treated and control subjects in the cross-section will allow the study to estimate the causal impact of the medical debt abolishment. The survey will measure the effects of medical debt on three sets of outcomes: (i) health care utilization, as measured by medical care visits, prescription drug utilization and adherence, and unmet need for medical care; (ii) mental health, as measured by validated screens for depression and anxiety; and (iii) subjective wellbeing, as measured by self-reported health, forgone consumption, and financial strain. This study would be the first to provide a direct, causal connection between the rising personal debt associated with U.S. health care and the health outcomes of its recipients.

Groups and Interventions

Group: Treatment

> Subjects in this "treatment" group had their medical debt forgiven by a non-profit charity, RIP Medical Debt. This protocol will administer a survey to measure subjects' health care utilization, mental health, and subjective well-being.

Group: Control

> No intervention was given to subjects in this "control" group. This protocol will administer a survey to measure subjects' health care utilization, mental health, and subjective well-being.

Intervention: Medical debt forgiveness

> A non-profit charity, RIP Medical Debt, bought and retired medical debt for individuals that were randomly assigned to the treatment group.

Study Population and Exclusions

Study Population:

> The study population is a probability sample drawn from a population of individuals who owe medical debt held by First Financial Asset Management (FFAM), a debt collection agency.

Exclusion Criteria:

1) Excluded individuals who owed less than $500 in medical debt to FFAM
2) Excluded individuals with missing Social Security numbers

Stratification

We stratify the sample along the following dimensions, as was requested by FFAM:

1) *Location of debtor*. To maintain balance across geographic areas, we stratify by state, combining states with a small number of debtors to ensure sufficient sample size.
2) *Amount of debt*. Debt forgiveness may yield larger impacts when the face value of the debt is large, therefore we also stratify by amount of debt FFAM holds for the individual.
3) *Debt collector recovery score.* The debt collector assigns an internally generated score that is indicative of the expected recovery rate. We further stratify on this dimension.
4) *Insurance status.* Finally, we stratify on whether individual accounts are labeled as self-pay or self-pay after insurance, as this dimension may also be predictive of the expected recovery rate.

Primary Outcome Measure

Outcome 1

Title: 8-item Patient Health Questionnaire (PHQ-8) Depression Scale

Description: Scores on the 8-item Patient Health Questionnaire depression scale range from 0 to 24, with higher scores indicating greater severity of depression.

Time frame: An average of 12 months after the intervention.

Secondary Outcome Measures:

Health Care Utilization Domain

Outcome 1:

Title: Received Needed Health Care

Description: Binary response to: "If you needed medical care in the last 12 months, did you get ALL the medical care you needed?"

Time Frame: An average of 12 months after the intervention.

Outcome 2:

Title: Received Needed Rx

Description: Binary response to: "If you needed prescription medications in the last 12 months, did you get all the prescription medications you needed?"

Time Frame: An average of 12 months after the intervention.

Mental Health Domain

Outcome 3:

Title: 7-item Generalized Anxiety Disorder (GAD7) Scale

Description: Scores on the 7-item Generalized Anxiety Disorder Scale range from 0 to 21, with higher scores indicating greater severity of anxiety.

Time Frame: An average of 12 months after the intervention.

Outcome 4:

Title: Stress

Description: Binary response to: "Stress means a situation in which a person feels tense, restless, nervous or anxious or is unable to sleep at night because his/her mind is troubled all the time. Do you feel this kind of stress these days?"

Time Frame: An average of 12 months after the intervention.

General Health Domain

Outcome 5:

Title: General Health

Description: Response to: "In general, would you say your health is: Excellent, Very Good, Good, Fair, or Poor?"

Time Frame: An average of 12 months after the intervention.

Subjective Wellbeing Domain

Outcome 6:

Title: Happiness

Description: Response to: "Taken all together, how would you say things are these days - would you say that you are Very Happy, Pretty Happy, or Not Too Happy?"

Time Frame: An average of 12 months after the intervention.

Financial Stress Domain

Outcome 7:

Title: Problems paying other bills

Description: Response to: "Besides medical bills, have you had problems paying other types of bills in the past 12 months?"

Time Frame: An average of 12 months after the intervention.

Outcome 8:

Title: Changes in spending due to medical debt

Description: Indexed response to: "As a result of medical bills have you cut back on spending in the past 12 months on i) Basic necessities (like food, heat or housing, or other basic household items), ii) Big-ticket items (like cars, furniture, or appliances); iii) business investments?"

Time Frame: An average of 12 months after the intervention.

Outcome 9:

Title: Changes in borrowing due to medical debt

Description: Indexed response to: "As a result of medical bills, in the past 12 months, have you i) Increased your credit card debt, or charge card debt ii) Borrowed money from a payday lender; iii) Borrowed from friends and family; iv) Used up all or most of your savings; v) Increased debt on other lines of credit?"

Time Frame: An average of 12 months after the intervention.

Empirical Specification

We will estimate the treatment effect of debt abolishment with the following ordinary least squares regression:

$$y_i = \beta T_i + \alpha_{w,v} + \varepsilon_i$$

In the above regression, i indexes individuals, T_i is an indicator for whether the individual was assigned to the treatment group and $\alpha_{w,v}$ are fixed effects for fully interacted treatment waves and survey waves. Because the probability of treatment assignment differed across treatment waves and the probability of surveying treated and control individuals differed across survey waves, controlling for the full interaction is necessary and sufficient for recovering unbiased treatment effects.

The coefficient β captures the treatment effect of debt abolishment on outcome y. While controlling for individual characteristics is not necessary for a causal interpretation of the estimates, their inclusion should increase the precision of our estimates. In addition to a specification with only the treatment-wave-by-survey-wave fixed effects, we will estimate a specification that controls for observables (from FFAM and baseline credit report data if available) using the "post-double-selection" method outlined in

Belloni, Chernozhukov and Hansen (2014) and recommended by Duflo (2018) in her NBER Summer Institute Master Lecture.[1]

Inference

For our analysis of the 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. Within each domain, we will report *p*-values that adjust for multiple outcomes (e.g., using the free step-down resampling method of Westfall and Young (1993) or the false discovery rate method of Benjamin and Hochberg (1995)). 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.

In some domains, we may also construct summary indexes, similar to those in Anderson (2008). These summary indexes would allow us to measure whether there are aggregate effects for the domain without raising concerns about within-domain multiple testing. Naturally, we would not include these summary indexes when adjusting for multiple inference.

Heterogeneity

In addition to the average effect of debt forgiveness, we are also interested in treatment effect heterogeneity across individuals. We will conduct heterogeneity analyses by estimating our baseline specification on our sample split one-by-one (i.e., not fully interacted) on the following dimensions:

1. Age of the individual.
2. Age of the debt.
3. Amount of the debt.

The effects of debt forgiveness may vary based on the extent to which the individual has other debt in collections. If we are able to merge in credit bureau data, we will conduct an additional heterogeneity analysis by estimating our baseline specification on a sample split on the following dimension:

4. Whether the debtor has at least one "other" debt in collections (i.e. collections that are not observed in our FFAM medical debt data).

We may 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.

[1] http://conference.nber.org/conf_papers/f114791.slides.pdf

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

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, *81*(2), 608-650.

Benjamini, Y, and Yosef H. (2995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57.1 (1995): 289-300.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., ... & 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., & Young, S. S. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment* (Vol. 279). John Wiley & Sons.