b
UNIVERSITÄT
BERN

Faculty of Business, Economics
and Social Sciences

Department of Economics

Life after Default: Credit Hardship and its Effects

Giacomo De Giorgi, Costanza Naguib

22-06

August, 2022

DISCUSSION PAPERS

Schanzeneckstrasse 1
CH-3012 Bern, Switzerland
http://www.vwi.unibe.ch

Life after Default: Credit Hardship and its Effects

Giacomo De Giorgi* Costanza Naguib†

Abstract

We analyze the impact of credit default on individual trajectories. Using a proprietary dataset for the years 2004-2020, we find that after default individuals relocate to cheaper areas. Importantly, default has long-lasting negative effects on income, credit score, total credit limit, and home-ownership status.

JEL codes: J61, G51, D12

Keywords: mobility, bankruptcy, default, credit, income

1 Introduction

In this paper we analyze what happens to an individual after default. We show that defaults, bankruptcies, and delinquencies trigger individual relocation to cheaper ZIP codes in the US both within and outside the initial commuting zone. Further, default has long-lasting negative consequences on credit score, total credit limit, home-ownership status, and income. Given that the occurrence of default isn't typically random, we adopt several strategies to recover the causal impact of default, such as event studies, a double/debiased lasso estimation technique proposed by Chernozhukov et al. [2018] coupled with an instrumental variable(-like) approach.

Default and delinquencies are not uncommon, in particular focusing on a 1% snapshot of the population with credit reports in 2010, we note that about 1.5% of the US population of interest had a new episode of harsh default (foreclosure, chapter 7 or 13) in their records, while new soft defaults in 2010 (delinquencies of more than 90 or 120 days) affected almost 5% of the population. In that given year, the population affected by both soft and hard default is 0.7%. These figures are compatible with roughly 4-6 million people experiencing such an episode in a given year. So what happens to these people afterwards?

*GSEM-University of Geneva, BREAD, CEPR, and IPA., Uni-Mail, 40 Bd. du Pont d'Arve, CH-1211 Geneva 4, Switzerland.

†University of Bern

Our interest in this paper is on the tracing of individuals' lives after default and we do not aim to distinguish between different default motives (see for example Ganong and Noel [2020] and the literature cited therein).

More generally our work is related to two strands of the literature, one strand which focuses on the individual determinants of default, whereas the other focuses on the individual and social costs of default and on the analysis of debt relief policies. For example, Lawrence [1995] belongs to the first strand, she builds a theoretical model of consumer's choices with default as a possible option over the life-cycle. Similarly, Guiso et al. [2009] study the determinants of strategic default, i.e. when one does not repay the mortgage, even if she would be able to do so, because the house value has fallen far below the value of the mortgage. Giliberto and Houston Jr [1989] develop a theoretical model of residential mortgage default when borrowers face beneficial as well as costly relocation opportunities. Mnasri [2018] finds that both income and geographical mobility are main trigger factors of default. In fact, households with a higher mobility rate (i.e. young households) are more likely to default. In general, all these studies agree that default is more likely if the individual is not married, not homeowner, and if she has already migrated far from her birthplace.

To the second strand of literature belong, for example, Collinson and Reed [2018], who investigate the impact of eviction on low income households in terms of homelessness, health status, labor market outcomes and long-term residential instability. Similarly, Currie and Tekin [2015] show that foreclosure causes an increase in unscheduled and preventable hospital visits. Albanesi and Nosal [2018] investigate the impact of the 2005 bankruptcy reform, which made more difficult for individuals to declare either Chapter 13 or Chapter 7. They find that the reform hindered an important channel of financial relief. Diamond et al. [2020] analyze the negative impact of foreclosures on foreclosed-upon homeowners. They find that foreclosure causes housing instability, reduced home-ownership, and financial distress. Finally, Indarte [2022] analyzes the costs and benefits of household debt relief policies.

The current paper attempts to quantify the impact of default on individuals' mobility and residential choices, credit access and utilization, and income in the short and long run. Our thought experiment is that of comparing the life-trajectories of two statistically identical individuals where only one is subjected to default. Ultimately, we quantify the impact of soft and harsh defaults as hampering the ability of the individual to borrow for several years, at the same time harsh default increases the probability of moving within one's commuting zone by 50% in the first couple of years after the episode, and half that for the next 5 years, The mobility outside CZ increases only in the year of harsh default when it more than doubles. What is noticeable is that income falls by about 15,000-20,000USD (about 33%), the probability of having a low total credit limit, below 10,000USD, increases by almost 30pp after a harsh default, while revolving balance goes steadily down 5,000 to 10,000USD (around 70%), while we find no significant effect on home-ownership.

The effects of soft default are similar to the ones above, somewhat smaller in magnitude, yet fairly substantial.

It is worth noting that harsh default, or the ability to charge off previously contracted debt, is thought of as a form of social insurance which allows individuals in financial hardship to reboot with a clean slate. Our analysis suggests that it is not the case as those filings result in substantial negative outcomes in the short and more importantly in the long-run. It appears that the relief provided by bankruptcy filings isn't able to turn lives around, suggesting that more permanent components of economic distress should be the policy target.

Similarly to Diamond et al. [2020], we find evidence that a harsh default, such as a foreclosure, has negative, large and long-lasting negative impact on a series of outcomes. Hence foreclosure mitigation policies, bankruptcy and default averting policies, should be expandend/supported, in order to reduce monetary and non-monetary costs for the individuals involved. Further, we find that the largest effects are concentrated among individuals with higher then median credit scores (higher than 731) in 2010, perhaps this suggest that while these individuals had better overall conditions before default they lack the means to counteract such events ex-post.[1]

In the rest of the paper we present the data in Section 2, an event study based exploratory analysis in Section 3, a causal machine learning approach in Section 4, and some remarks in Section 7.

2 Data

We rely on a unique proprietary dataset from Experian (see for example De Giorgi et al. [2021] or for a similar dataset Albanesi and Vamossy [2019]). This dataset includes information on the credit scores, and more than other 400 credit variables, plus basic socio-demographic such as data of birth, ZIP codes of residence, and imputed incomes for a 1% sample of the total population of the US with valid credit score in 2010. Hence, we can

[1] A credit score between 700-749 is considered a good one, above 750 an excellent one, while 650-699 is a fair one, 649-550 a poor one, and lower than 549 a very poor one (see `https://www.experian.com/blogs/ask-experian/what-is-a-vantagescore-credit-score/`).

rely on a large sample size for our statistical analyses, as much as a panel of 1.5 million individuals per year for the period 2004-2016 and then a final wave for 2020.

To be more precise, the data include a series of variables on the number of trades made and several variables measuring credit behavior (i.e. number of bankruptcies, number of credit delinquencies, number of credit cards, average amount of credit and so on). Beyond the rich credit information (more than 400 variables) which of course include mortgages and car loans, we have information on individual's age and a measure of imputed income, which has been computed by Experian based on W2's. The exact procedure for imputation is property of Experian, however in Figures B1 and B2 in Appendix C we check data reliability in terms of both data representativeness and validity of the income imputation and the data provided from Experian appear to be fully consistent with those obtainable from the Census or from the IRS.

The reliability of these income imputation measures has been also checked by comparing them with the PSID Labor Income for the 2009/11 wave, and the results were plausible (see Table B1 and Table B2 in the Appendix). Further, Lee and Van der Klaauw [2010] have shown validity and representativeness of the New York Fed Consumer Credit Panel, which is constructed on the basis of similar data to ours but from Equifax credit bureau.

In Table A1, we report a few summary statistics for our main variables of interest, including the house price index (HPI).[2]. It is useful here to distinguish between different types of default. For example, in the US it is possible to declare either Chapter 7 or Chapter 13 bankruptcy.[3].

In Appendix A we briefly discuss how credit bureaus should treat different types of defaults, in particular we know that Experian keeps Chapter 7 flags for up to 10 years, and Chapter 13 (and other delinquencies) for up to 7 years, this in accordance to the Fair

[2]Data on the House Price Index are taken from the Federal Housing Finance Agency. `https://www.fhfa.gov/DataTools/Downloads/Pages/House-Price-Index-Datasets.aspx`

[3]A chapter 7 bankruptcy implies the liquidation of assets: *...the bankruptcy trustee gathers and sells the debtor's non-exempt assets and uses the proceeds of such assets to pay holders of claims (creditors) in accordance with the provisions of the Bankruptcy Code...* (see https://www.uscourts.gov/services-forms/bankruptcy/bankruptcy-basics/chapter-7-bankruptcy-basics). A chapter 13 bankruptcy is also called a wage earner's plan. It enables individuals with regular income stream to develop a plan to repay all or part of their debts. Under this chapter, debtors propose a repayment plan to make instalments to creditors over three to five years. If the debtor's current monthly income is less than the applicable state median, the plan will be for three years unless the court approves a longer period "for cause." (1) If the debtor's current monthly income is greater than the applicable state median, the plan generally must be for five years. In no case may a plan provide for payments over a period longer than five years. 11 U.S.C. 1322(d). During this time the law forbids creditors from starting or continuing collection efforts. This chapter discusses six aspects of a chapter 13 proceeding: the advantages of choosing chapter 13, the chapter 13 eligibility requirements, how a chapter 13 proceeding works, making the plan work, and the special chapter 13 discharge. https://www.uscourts.gov/services-forms/bankruptcy/bankruptcy-basics/chapter-13-bankruptcy-basics

Credit Reporting Act.[4]

To get a sense of the sample we quickly scan through the main variables of interest for our base year (2010) in Table 1 (a similar table for the entire available period (2004-20) is in the Appendix, Table A1). For example the average credit score is 722, with an average (nominal) income (W2 derived) of 56,430USD, and an average House Price Index of 341 points (the index is set to 100 in 1985 for ZIP code 01001, Agawam, Massachusetts). The average age is 53. On average, in our sample around 12% of the individual moved ZIP code in 2010 and around 5% moved to a different commuting zone. The share of homeowners in the sample is around 69%. The share of individuals experiencing a new 90+ days delinquency in year 2010 are around 4.5%, whereas those experiencing a new 120-day delinquency are slightly less (i.e. around 4%). Individuals experiencing a new foreclosure in 2010 are around 0.7%, those declaring a new Chapter 7 default are around 0.6% and those declaring a new Chapter 13 default in 2010 are sensibly less, i.e. around 0.2%.

Around 30% of individuals have a credit card balance equal to zero. Among those with a positive credit card balance, the average amount is 14,126USD, whereas in the whole population it is 12,015USD. For mortgages, the average in the population is slightly less than 100,000USD, and around 40% of individuals in the sample have a mortgage. Among those with a positive mortgage amount, the average is 189,327USD. Average total credit limit on all open trades is around 168,000USD. No individual in our sample has a credit limit equal to zero (that is mechanical due to the fact that in order to be in the credit bureau a credit line is needed). The same holds true for the total credit limit on open revolving trades, which has an average of around 45,000USD in our dataset. Finally, the total balance on revolving trades has an average of around 14,500USD in the whole dataset, and around 89% of the individuals in the data have a total balance on revolving trades greater than zero. Among those, the average recorded is around 16,290.28USD.

[4] https://www.ftc.gov/system/files/documents/statutes/fair-credit-reporting-act/545a_fair-credit-reporting-act-0918.pdf.

Variable	Mean	Std. Dev.	Min.	Max.	N
Credit Score	722.101	95.1456	300	839	1,548,988
Income	56,430.4953	27,639.3451	1000	331,000	1,548,988
HPI	341.369	222.6032	52.45	1602.03	1,548,988
Mortgage Bal.	99,807.1558	188,801.9173	0	8,463,378	1,548,988
Age	52.9571	15.5809	18	112	1,548,964
Credit Card Balance (w/zeros)	12,015.2758	21,344.1088	0	1,036,857	1,548,988
Move ZIP	0.1236	0.3291	0	1	1,544,550
Move CZ	0.046	0.2094	0	1	1,531,739
Home_own	0.6944	0.4607	0	1	1,548,988
New Delinq. (90+)	0.0464	0.2104	0	1	1,548,988
New Delinq. (120+)	0.0399	0.1957	0	1	1,548,988
New Forecl.	0.0073	0.085	0	1	1,548,988
New Ch. 7	0.0059	0.0764	0	1	1,548,988
New Ch. 13	0.0017	0.0409	0	1	1,548,988
Total credit limit on open trades (all)	167,698.0195	235,476.1313	1	9,548,471	1,548,988
Total balance on revolving trades	14,532.32	34,123.4	0	5,077,251	1,548,988
Total credit limit on open rev. trades	45,300.7532	75,005.0938	1	5,566,654	1,548,988
Harsh default	0.0145	0.1197	0	1	1,548,988
Soft default	0.047	0.2117	0	1	1,435,257

Table 1: Summary statistics of our main variables, 2010. Special codes credit scores lower than 300 have been trimmed. Similarly,the top 1% of total credit limit, total balance on revolving trades and total revolving limit have been trimmed.

3 Event study analysis

In this section we provide event-study evidence on the effect of different types of default on individual relocation probabilities, income, home-ownership, total credit and revolving credit limits, and balances on revolving credits. We limit this analysis to the years 2004-2016, as we don't have years 2017 and 2018 in our data. In Figures 1 and 2, we report the result of an event study, in which the event is, respectively, a harsh default or a soft default. We define harsh default as any of the following: Chapter 7 declaration, Chapter 13 declaration, and foreclosure. In our data, 76% of what we define "harsh default" is a Chapter 7 declaration. Around 12% is Chapter 13 declarations and about 17% are foreclosures. Notice that the previous shares exceed 1 as in some cases there is overlap, e.g. both Chapter 7 is declared and there is a foreclosure. Such cases are less than 5%. Overall, harsh defaults represent around 24% of all the defaults recorded in the dataset.

In Figure 2, we consider 90-days and 120-days delinquencies as the defaults of interest, i.e. we focus on forms of soft defaults.

In the event studies for harsh default we exclude those individuals who experience a soft default in 2010, while in the soft default event study we exclude those with harsh defaults in 2010. For both types of events we condition on no other default episodes between 2004 and 2010, i.e. we only keep individuals who had no defaults prior to 2010.

Only around 5% of individuals with 90-day delinquency do not also have 120-day delinquency. In essence, as often used in the literature and by industry standard, a 90 days delinquent credit is a defaulted one.

The reason why we consider all episodes in 2010, conditional on no previous episode as our baseline is because our panel data is representative of the US population with valid credit score in 2010. Further, 2010 is somewhat in the midpoint of our time dimension for the dataset and this makes for a more meaningful before comparison. Also focusing on one specific period events allows us to be more computationally parsimonious as applying the procedures suggested in Callaway and Sant'Anna [2021] would be quite demanding in our context, and this structure circumvents the issues raised in Borusyak et al. [2021].

Ultimately we estimate the following event-study equation:

$$y_{i,t} = \alpha + \sum_{l=2}^{L} \beta_l Default_{i,2010-l} + \sum_{k=0}^{K} \gamma_k Default_{i,2010+k} + \delta_1 age_{i,t} + \delta_2 age_{i,t}^2 + \eta_i + \varepsilon_{i,t}$$

with i being the individual, and t the year. y_{it} are the outcomes of interest, e.g. the probability of moving ZIP code, the probability of moving commuting zone, a dummy for being home-owner, the total credit limit, the total revolving credit amount and (imputed) income. η_i is an individual fixed effect and $Default_{it}$ is, either a soft or a harsh default. $\varepsilon_{i,t}$ is a mean 0 error, and we center the events at time 0 (of course that is the year 2010), so that in the Figures below the horizontal axis have a simple interpretation where 0 stands for 2010, -1 for 2009, 1 for 2011 and so on. Given our design, $\sum_{l=2}^{L} \beta_l$ will speak to pre-trends, and $\sum_{k=0}^{K} \gamma_k$ to the effects of interest.

3.1 Harsh Default

Let's focus on Figure 1, i.e. the impact of a harsh default on our outcome variables of interest.

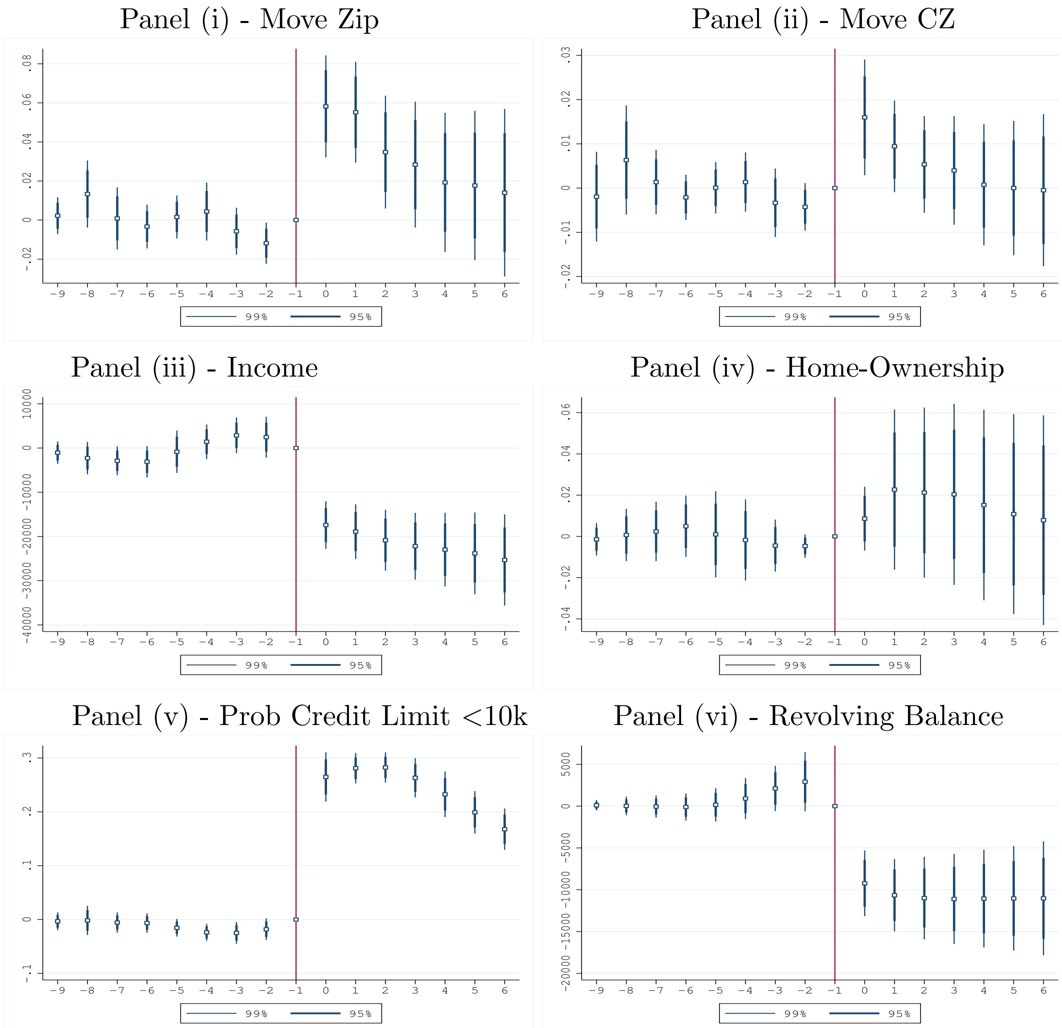

Figure 1: Event study dependent variable is: (i) move ZIP code, (ii) move commuting zone, (iii) income, (iv) home-ownership dummy, (v) the probability that total credit limit is lower than 10,000usd, i.e. very close to zero, (vi) total revolving credit amount. In all panels the events take place in 2010. The event considered is a harsh default, defined as either a Chapter 7 or a Chapter 13 declaration or a foreclosure. Individual fixed effects are included in all estimations. Other controls are age and age squared. 95 and 99% confidence intervals around the point estimates.

First let us notice that all the pre-event coefficients are very close to 0, and for the most part statistically insignificant. On impact, a harsh default entails an increase by around 6pp in the probability of moving ZIP code (Panel (i)). This impact is non-negligible if we consider that around 12% of all individuals in our sample move ZIP code from one year to the following one. This impact decreases to 3-5pp after one and two years from the event and becomes statistically insignificant afterwards. A harsh default is also associated with an increase by around 1.5pp (or a 40% effect size) on impact in the probability of moving commuting zone (Panel (ii)). This impact becomes immediately statistically insignificant afterwards. These are both very large effects, of course foreclosure implies some form of relocation but the effects here are not just coming from foreclosed upon individuals. If we only consider the impact of Chapter 13 or Chapter 7 declarations and we exclude foreclosures, the probability of moving ZIP code has a jump on impact (the increase is around 2pp, results not reported for brevity), but the effect becomes insignificant immediately afterwards. Further, after a Chapter 13 or Chapter 7 declaration, again excluding foreclosures, the probability of changing commuting zone increases by 1pp on impact. The effect becomes immediately insignificant afterwards (results not reported for brevity). Foreclosure gives raise to one to two thirds of the harsh default effects on mobility.

In the mid-panels of Figure 1, we assess the impact of a harsh default on income (Panel (iii)) and the probability of being homeowner (Panel (iv)). A harsh default is associated with a negative impact of around 20,000-25,000USD on annual (imputed) income, this is a persistent and statistically significant effect, i.e. it is still statistically significant six years after the event. Given that average imputed income is around 56,000USD, a harsh default is associated with a fall of around one third of individual annual income. However, the impact of a harsh default on the probability of being a homeowner is never statistically different from zero, neither at a 99, nor at a 95% confidence level.

We suspect that the lack of a significant effect is due to the way foreclosure proceedings are administered where it is common to have a long lag between the initiation and the finalization of the process, in some states the process can take up to 4-5 years and it is unclear to us at which stage the home-ownership passes to the lender.[5]

Finally, in the bottom panels of Figure 1, we study the impact of a harsh default on the probability that total credit limit is lower than 10,000usd, i.e. very close to zero (Panel (v)) and on revolving credit balance (Panel (vi)). Harsh default relates to a substantial increase in the probability of having a credit limit close to zero, i.e. around 25pp, declining to around 15/20pp after five/six years from the event. Further, a harsh default is associated with a drop by around 10,000USD in the revolving balance amount, this effect also remains statistically signficant over time, up to six years after the event. Given that average total revolving balance is around 14,500USD, the default is reflected in a fall in balance of around 70% of the amount.

[5]https://www.attomdata.com/news/most-recent/top-10-states-with-longest-foreclosure-timeline/

3.2 Soft Default

In Figure 2 we analyze the impact of a soft default, i.e. 90-days or 120-days delinquencies only. We note that all the pre-event coefficients are again very close to 0, and for the most part statistically insignificant. Differently from Figure 1, we add an extra Panel (vii) where we investigate the relationship between soft and harsh defaults overtime.

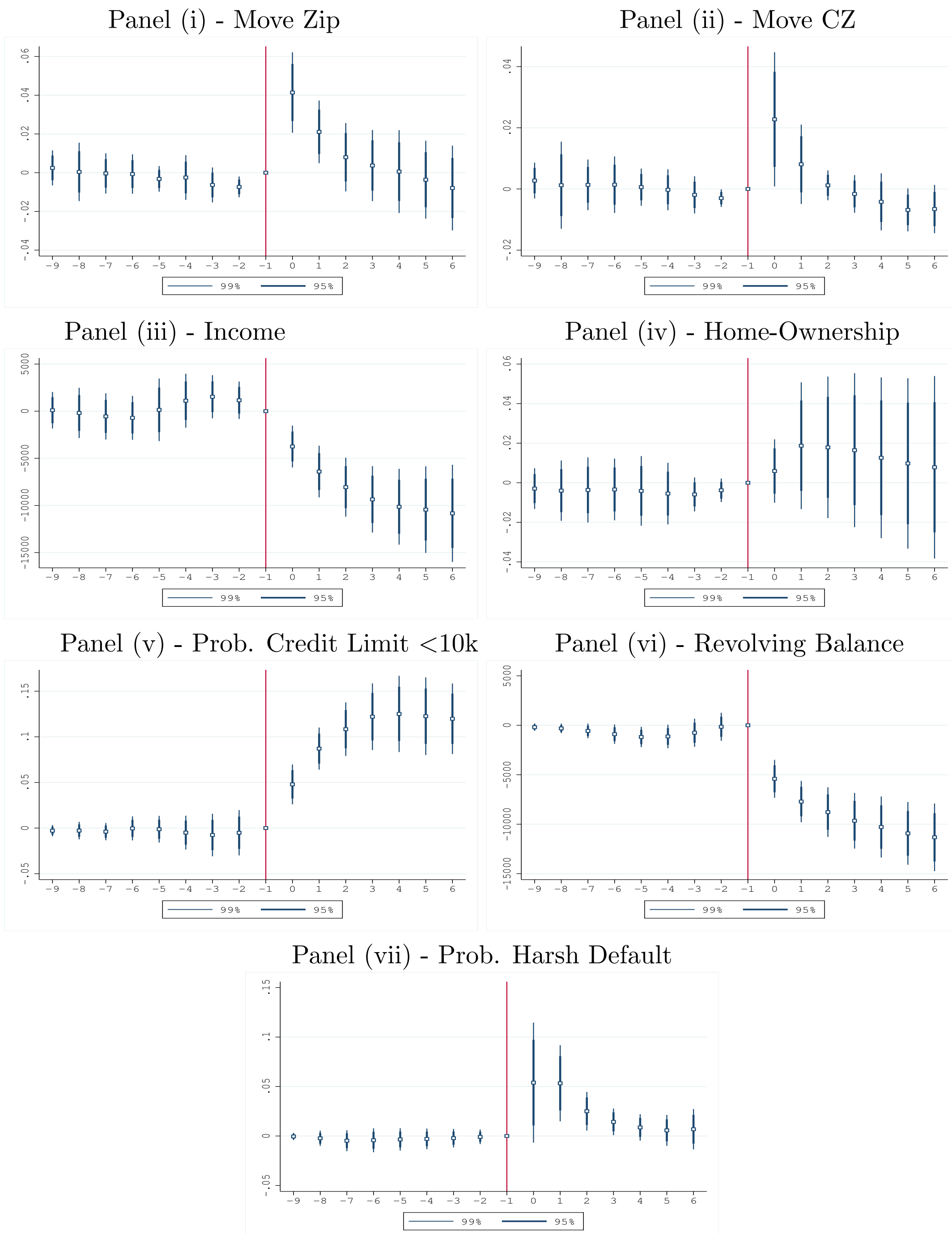

Figure 2: Event study: dependent variable is: (i) move ZIP code, (ii) move commuting zone, (iii) income, (iv) home-ownership dummy, (v) probability that total credit limit is lower than 10,000USD, (vi) total revolving credit amount, (vii) probability of having a harsh default in subsequent years. In all panels the events take place in 2010. The event considered is a soft default, i.e. a 90-day or a 120-day delinquency but no Chapter 7, Chapter 13 or foreclosure taking place in the same year. Individual fixed effects are included in all estimations. Other controls are age and age squared. 95 and 99% confidence intervals around the point estimates.

On impact, a soft default entails an increase by around 4pp in the probability of moving ZIP code (Panel (i)). After one year, the effect is still 2pp, but it becomes essentially indistinguishable from zero afterwards. These are large magnitudes, considering that the year-to-year zip code mobility is about 12%. A soft default is also associated with an increase by around 2pp on impact in the probability of moving commuting zone (Panel (ii)). This impact becomes statistically insignificant one year after the event, and even turns negative afterwards, i.e. by year 6 after the event we have a 1pp drop in commuting zone mobility. In terms of mobility, the impact of a harsh default on zip code mobility is larger than that of a soft default, whereas the impacts of the two types of default on commuting zone mobility are similar. The effects on mobility are large and persistent in the medium to long run, especially in the case of a soft default. Both soft and harsh defaults seem to trigger higher mobility across zip codes on impact, but a soft default appears to hamper it in the longer run. It is plausible that on impact people need to relocate, but then due to the persistent effect on credit score they find hard to relocate further due to relocation costs and most likely to credit score checks for new rentals (and of course purchases).

In the middle panels of Figure 2, we study the impact of a soft default on income ((Panel (iii)) and probability of being homeowner ((Panel (iv)). A soft default has a negative impact of around 5,000-10,000USD on annual imputed income, i.e. around one third/one half of the negative impact of a harsh default. The impact is statistically significant up to six years after the event. A soft default wipes away around 9-18% of annual average imputed income. The impact of a soft default on the probability of being a homeowner is never stastically different from zero.

Next, in Figure 2, we study the impact of a soft default on the probability that total credit limit is lower than 10,000USD, i.e. very close to zero ((Panel (v)) and on revolving credit balance ((Panel (vi)). For the probability of a low credit limit, a soft default is associated with an increase of 5pp to around 12pp (overtime) in the probability of having a total credit limit close to zero. This impact is increasing over time, and remains statistically significant up to six years after the event. For the revolving balance, a soft default entails a drop by around 5,000USD on impact and between 10,000 and 12,000USD, increasing over time. This impact is between half and 100% of that found previously for a harsh default on revolving credit amount.

Finally, we look at the impact of a soft default on the probability of a harsh one ((Panel (vii)) and indeed we find that the probability of a harsh default following a soft one increases by around 5pp, after one and two years from the event, diminishing and becoming statistically insignificant over time. This is a non-negligible magnitude on impact and is in line with models of opportunistic default on all loans such as Parlour and Rajan [2001].

4 The double/debiased lasso estimation method

Despite the fact that our event studies are meant to capture the causal impacts of the different types of default on the outcomes of interest we recognize that the causal impacts might be better identified using causal machine learning methods in our context (for an overview see Athey and Imbens [2019]). We therefore employ a double/debiased machine learning procedure in the spirit of [Chernozhukov et al., 2018], as described below.

Lasso is a so-called regularization method, in the family of machine learning techniques. It allows to deal with a large number of covariates (in our case, socio-demographic, as well as credit variables for all the pre-event years, i.e. from 2004 to 2009) by adding to the standard multivariate regression equation a penalization terms. The presence of the penalization term ($L1 \equiv \lambda \sum_{j=1}^{p} |\beta_j|$) pushes the coefficients towards zero, so that only the most relevant covariates are kept in the model. Lasso minimizes the following quantity:

$$\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j X_{ij}\right)^2 + \lambda\sum_{j=1}^{p}|\beta_j| \tag{1}$$

where y_i is the outcome variable (in our case: the probability of moving ZIP code, the House Price Index of the ZIP code of residence, or the credit score, the annual (imputed) income, or the total revolving credit limit, and so on). X_i is a vector of covariates that includes age, age squared, commuting zone fixed effects, year dummies, the amount of mortgage balance and the amount of car leasing open, measured in each of the pre-event years. In our case, the treatment is the default which, as in the previous sections, can be harsh or soft, depending on the model specification. The second part of the equation, the $L1$ penalty, actually forces some of the coefficients to be exactly equal to zero when λ is sufficiently large (James et al. [2013]). Hence, Lasso performs variable selection, resulting in sparse and hence also easier to interpret but less flexible models. Altogether, this implies that Lasso is expected to perform better in a setting where a relatively small number of predictors have substantial coefficients, and the remaining predictors have coefficients that are very small or that are equal zero (James et al. [2013]), that is, when only a few predictors actually influence the response. Finally, given that Lasso reduces some of the coefficients to zero and hence excludes certain variables, the coefficients of correlated variables are expected to be zeroed, except for one of the correlated predictors when that is large enough. As λ increases, the relative importance of the $L1$ penalty increases and thus more and more coefficients are set equal to zero.

The problem when applying Lasso regression to economic research questions is that all the estimated coefficients are pushed to zero by construction. Hence, their size cannot be directly interpreted. This is why a post-Lasso approach is often adopted, in which a second stage OLS with only the relevant variables selected by the Lasso in the first stage is performed. However, another issue is that, if the variable representing treatment is not randomly assigned, the post-Lasso estimation (as well as the Lasso) will provide biased results. In our framework, the treatment variable is a default (harsh or soft). This treatment variable is very likely not to be randomly assigned, but to depend on individual characteristics and behavior, some of them observable (e.g. age) and some of them not (i.e. individual preferences, level of commitment on the financial engagements taken and so on). Therefore to handle the potential endogeneity of the default variables, in Section 4 we perform all of our estimations by applying the double/debiased Lasso estimation technique proposed by Chernozhukov et al. [2018]. This method allows for robust inference in presence of many covariates and potentially many instruments. Their proposed procedure is as follows. Consider the model:

$$y_i = d_i\tau + X_i'\beta + u_i \tag{2}$$

where y_i and X_i are defined as above, d_i is the treatment variable (i.e. the occurrence of a default), and u_i is the error term. The first stage is defined as:

$$d_i = X_i'\delta + v_i \tag{3}$$

the double/debiased estimation procedure consists in the following steps: (i) Predict y_i and d_i using X_i with separate Lasso regressions and obtain $\hat{\beta}$ and $\hat{\delta}$, (ii) residualize: $\hat{u}_i = y_i - X_i'\hat{\beta}$ and $\hat{v}_i = d_i - X_i'\hat{\delta}$, (iii) the debiased estimator is:

$$\hat{\tau} = \left(\frac{1}{n}\sum_{i=1}^{n}\hat{v}_i d_i\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\hat{v}_i\hat{u}_i \tag{4}$$

4.1 Double/debiased Lasso baseline results

In this Section we report our baseline double/debiased lasso results for a series of outcome variables (probability of moving ZIP code, probability of moving commuting zone, House Price Index of the ZIP code of residence, credit score, income, total credit limit, revolving credit limit, home-ownership status, revolving credit utilization rate (balance over limit)).

In all the estimations presented in this Section, the controls are: age, age squared, year dummies, commuting zones fixed effects, the amount of mortgage balance open and the amount of car loan open. Except for age, age squared and year fixed effects, that are for sure exogeneous, all the other controls are measured in years 2004 to 2009, i.e. pre-event.

Another control included is the maximum interest rate allowed by the anti-usury laws in each state in each year of our dataset.[6] Since in some states this maximum rate depends on the current rate paid by the treasury bills, our instrument exhibits some variation over time, albeit limited.[7] Consistently with what we did in Section 3, in all the estimates presented in this Section we restrict the analysis to those affected by a harsh or a soft default in year 2010, compared to those who were not affected by the same type of default up to the same year. This means that both for the treatment and the control group, we drop from the sample those individuals who have a default between 2004 (the first year in our data) and 2009 in their records. When studying the impact of a soft default, we drop from the dataset those who were affected by a harsh default in the same year 2010.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	move	move_cz	hpi	cs	incomeW2	home_own	ratio	log rev bal	log tot bal
Harsh def	0.103***	0.0334***	-39.80***	-125.2***	-21664.5***	-0.0202***	0.0148***	-0.721***	-1.372***
	(87.06)	(42.88)	(-41.77)	(-423.00)	(-255.40)	(-13.07)	(16.55)	(-83.35)	(-211.96)
N	16204104	16198910	17416959	17416959	17416959	17416959	17416959	15070047	17415378

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Controls: age, age squared, year dummies, commuting zones fixed effects,
the amount of open mortgages and car loan pre-event (years 2004-2009), plus State max interest rate.

Table 2: Post lasso results - impact of a harsh default

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	move	move_cz	hpi	cs	incomeW2	home_own	ratio	log rev bal	log tot bal
Soft def	0.0575***	0.0253***	-12.93***	-106.7***	-5896.1***	-0.0128***	0.0991***	0.332***	-0.138***
	(79.49)	(53.03)	(-21.93)	(-609.76)	(-112.56)	(-13.46)	(183.68)	(75.76)	(-34.95)
N	15783731	15778721	16842842	16842842	16842842	16842842	16842842	14617983	16841370

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Controls: age, age squared, year dummies, commuting zones fixed effects,
the amount of open mortgages and car loan pre-event (years 2004-2009), plus State max interest rate.

Table 3: Post lasso results - impact of a soft default

From Tables 2 and 3 we deduce that, for most outcomes, the impact of a harsh default is larger than that of a soft default. A harsh default is associated with an increase in the

[6] Source: https://www.findlaw.com/state/consumer-laws/interest-rates.html

[7] An alternative would be to use bankruptcy fees. This alternative control variable has more geographical detail (94 US judiciary districts vs the 50 US states), however, to the best of our knowledge, it is only available for a snapshot in time (Lupica [2011]).

probability of moving to a different ZIP code by around 10pp (Column (1)), whereas a soft default only by an increase of around 6pp. Both types of defaults are associated with roughly the same probability of moving out of the current commuting zone (Column (2)), i.e. around 3pp increase. Individuals move to cheaper ZIP codes (Column (3)) after a harsh default (i.e. minus 40 points in the house price index), whereas this negative impact on the HPI is more moderate, i.e. around minus 13 points, in case of a soft default.

As expected, a harsh default causes a drop by around 125 points in the credit score (Column (4)), that is around 107 points for a soft default. Also income drops (Column (5)) , consistently with Diamond et al. [2020], the fall is substantial, around 22,000USD in case of a harsh default and around 6,000USD for a soft default. A harsh default causes lower probability of being homeowners by 2pp (Column (6)), whereas a soft default is associated with a decrease by around 1pp in the same probability.

Further, a harsh default entails an increase by around 1.5pp in the utilization rate of revolving credit (Column (7)), i.e. in the ratio between the open revolving balance and the total revolving credit limit. This increase is even higher, i.e. around 10pp in case of a soft default. Finally, a harsh default causes a decrease in both revolving credit balances (Column (8)) and total credit balance (Column (9)), respectively by around -0.72 log points and by -1.37, that is because the credit lines dried out. On the contrary, a soft default entails an increase in the log revolving credit balance (+0.33 log points) but a decrease in the total credit amount (-0.14 log points).

5 The Causal Effects of Default in the Long-Run (2020)

In this section we analyze the long-run impact of both a soft and a harsh default by means of the double/debiased lasso estimation techniques. This means that all of our outcome variables are measured in 2020 (the last year we have data for). We use the same control variables as in the previous sub-section. From Table 4 and 5 we deduce that, ten years after the event, both a soft and a harsh default cause an increase of around 2-4pp in the probability of moving ZIP code (Column (1)) and with an increase of around 1pp in the probability of moving commuting zone (Column (2)). Further, ten year after the event, a harsh default, the movers live in zip codes with lower property values (Column (3)), the House Price Index is lower by around 22 points, whereas the drop is only by 4 points in case of a soft default. The credit score drops in the long run is still of around 78 points in case of a harsh default and by around 53 points in case of a soft default (Column (4)).

Moreover, a harsh default in 2010 causes a drop in income by around 26,000USD still in 2020 (Column (5)), the long run impact of the soft default is lower, around 11,000USD. The long run impact on income is even larger than the short-term one.

In the long run, a harsh default decreases the probability of being homeowner (Column (6)) by around 1pp, i.e. the impact is smaller (but similar) in the long run than in the short term. A soft default appears to have an even larger effect.

Finally, in the long run both a harsh and a soft default generate an increase by around 5-6pp in the utilization rate of revolving credit (Column (7)). A harsh default decreases revolving credit balances in the long term by around 0.3 log points (Column (8)), whereas a soft default produces, in the long term, an increase by around 0.08 log points in the same variable. Further, harsh defaults increase the total credit balance (Column (9)), i.e. 0.06 log points in case of a harsh default, and 0.17 log points in case of a soft default.

Tables 4 and 5 show that the negative impact of both types of default on variables such as credit score and income, as well as the positive impact on relocation probabilities is long lasting and still statistically significant up to ten years after the event.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	move	move_cz	hpi	cs	incomeW2	home_own	ratio	log rev bal	log tot balance
Harsh def	0.0409***	0.00914***	-21.97***	-77.69***	-25749.2***	-0.0115***	0.0542***	-0.316***	0.0635***
	(37.95)	(12.70)	(-16.41)	(-242.23)	(-212.52)	(-7.11)	(61.43)	(-38.02)	(5.38)
N	14832053	14819987	14832053	14832053	14829579	14832053	14832053	11794940	14825666

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Controls: age, age squared, year dummies, commuting zones fixed effects,
the amount of open mortgages and car loan pre-event (years 2004-2009),plus max interest rate.

Table 4: DML long-run impact of a harsh default

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	move	move_cz	hpi	cs	incomeW2	home_own	ratio	log rev bal	log tot balance
Soft def	0.0206***	0.00636***	-4.207***	-52.89***	-10785.8***	-0.00876***	0.0559***	0.0798***	0.168***
	(30.71)	(14.20)	(-5.02)	(-268.61)	(-142.90)	(-8.71)	(102.78)	(15.91)	(22.92)
N	14349406	14337833	14349406	14349406	14347176	14349406	14349406	11450630	14343725

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Controls: age, age squared, year dummies, commuting zones fixed effects,
the amount of open mortgages and car loan pre-event (years 2004-2009),plus max interest rate.

Table 5: DML long-run impact of a soft default

6 Heterogeneity of the Effects

In this final subsection we analyze potential sources of heterogeneity in the long-run impact of a harsh default. We consider two sources or natural dimensions of potential heterogeneity: age and initial credit score and we simply split the sample below vs. above median. For age that is 50 years old which might not be ideal to distinguish between more or less mobile populations (in Appendix D.3 we also consider a split for younger than 30), while the median score is 731 (scores above 700 are considered as good or very good). Of course younger individuals have shorter credit histories and are typically more subject to job transitions and shocks as they are climbing the job ladder and typically more likely to move while less well insured against shocks as they are in the "borrowing" part of the lifecycle. We also know that the existing credit scores are crucial to access credit so that default can be more common but also potentially insured against or somewhat prevented by the availability of multiple/larger credit sources. In the following, we define high credit score, as credit score that was above the median in 2010 and low credit score as credit score that was below the median in the sample in the same year. Similarly, we define young individuals as those whose age is below the median and older individuals as those whose age is above the median.

6.1 By Age

From Tables 6 and 7 we deduce that the impact of a harsh default on relocation probabilities (both ZIP code and commuting zone) is roughly the same both for young and older individuals. However, the drop in House Price Index is wider for the young (i.e. -27 points vs -15 points). Also, whereas the drop in credit score is quite similar in the two cases (-71 points for the young, -85 points for the older individuals), the drop in income is substantially higher for the young people (i.e. -28,000USD vs -21,000USD). The negative impact on the probability of being homeowner is also stronger for the young (-2pp vs around -1pp for the older ones), whereas the increase in the utilization rate of revolving credit is around +5pp for both groups. The log revolving credit balance also decreases more for the young (-0.36 log points vs only around -0.2 log points). Finally, the log total credit balance decreases by around -0.3 log points for the young, whereas in the long run after a harsh default it increases by around 0.6 log points for older individuals. In general, we can conclude that younger individuals suffer more, in the long run, for a harsh default.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	move	move_cz	hpi	cs	incomeW2	home_own	ratio	logrevbalance	logallbalance
harshdef	0.0412***	0.00779***	-27.05***	-71.18***	-28062.3***	-0.0214***	0.0537***	-0.362***	-0.269***
	(27.97)	(8.27)	(-15.28)	(-172.49)	(-156.61)	(-11.22)	(42.54)	(-35.76)	(-24.40)
N	6999895	6992907	6999895	6999895	6999112	6999895	6999895	6024331	6996484

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: DML long-run impact of a harsh default for individuals below median age in 2010

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	move	move_cz	hpi	cs	incomeW2	home_own	ratio	logrevbalance	logallbalance
harshdef	0.0385***	0.0108***	-14.54***	-85.21***	-21325.7***	-0.00958***	0.0533***	-0.195***	0.551***
	(22.87)	(9.11)	(-6.71)	(-163.68)	(-130.42)	(-3.34)	(41.43)	(-12.91)	(23.18)
N	7477326	7472427	7477326	7477326	7475711	7477326	7477326	5467387	7474398

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: DML long-run impact of a harsh default for individuals above median age in 2010

6.2 By Credit Score in 2010

In this subsection we analyze heterogeneity in the long-run impact of a harsh default in 2010 along different initial credit scores. From Tables 8 and 9, we deduce that the positive long-run impact of a harsh default on relocation probabilities (both ZIP code and commuting zone) is higher for those with a higher initial credit score (respectively +3pp vs +6pp for the ZIP code mobility and +1pp vs +2pp for commuting zone mobility). The drop in the house price index is also higher for those with an initially higher credit score (-35 points vs -16 points), as well as the drop in income (-25,000USD vs -22,000USD, approximately).

The impact on the probability of being homeowner in the long run is also notably higher for those with an initial credit score above the median (i.e. -9pp vs -5pp), and the utilization rate of the revolving credit also increases far more for this group than for those with an initial credit score below the median (+6pp vs +2pp). Finally, the drop in the log revolving credit balance is rather similar across the two groups (i.e. -4-5 log points), and the log of total credit amount only decreases for individuals with an initial low credit score (-0.14 log points), whereas in the long run it increases for those with an initial high credit score (+0.55 log points). In general, however, we notice from these results that individuals with an initial high credit score suffer a lot more, in the long run, from the consequences of a harsh default.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	move	move_cz	hpi	incomeW2	home_own	ratio	logrevbalance	logallbalance
harshdef	0.0330***	0.00624***	-16.38***	-22416.7***	-0.0513***	0.0157***	-0.488***	-0.139***
	(25.80)	(7.51)	(-11.18)	(-157.81)	(-29.29)	(13.23)	(-52.39)	(-11.04)
N	6600352	6594325	6600352	6598292	6600352	6600352	5262273	6596667

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: DML long-run impact of a harsh default for individuals below median cs in 2010

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	move	move_cz	hpi	incomeW2	home_own	ratio	logrevbalance	logallbalance
harshdef	0.0573***	0.0189***	-34.46***	-25352.1***	-0.0944***	0.0630***	-0.436***	0.550***
	(22.22)	(10.67)	(-9.97)	(-88.56)	(-22.59)	(39.43)	(-20.15)	(17.71)
N	8174072	8168083	8174072	8173658	8174072	8174072	6485268	8171370

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: DML long-run impact of a harsh default for individuals above median cs in 2010

7 Conclusion

In this paper we aim at assessing the impact of different types of default on outcomes such as credit score, income, probability of moving ZIP code and probability of moving commuting zone. To do this, we avail ourself of a unique dataset, from Experian credit bureau, covering 1% of the US population for the years 2004-2020. We adopt multiple approaches to study the impact of harsh and soft default on individual trajectories. In particular our preferred approach is the double/debiased lasso estimation technique (as proposed by Chernozhukov et al. [2018]).

Our findings are that the impacts of both a harsh and a soft default are relevant and long-lasting, up to 10 years after the event. The negative impact of a soft default on our outcome variables of interest is somehow muted, but still substantial. After a default, the individual experiences an increase in the probability of moving ZIP code by around 5%, and an increase in the probability of moving commuting zone by around 2%. Yet, those who actually move relocate to cheaper areas, as the House Price Index of residence drops by 13-40 points, depending on the type of default.

A default is also associated with a substantial income loss (betwee 15,000 and 20,000 USD, which is around one third of total annual income). This impact is long lasting and statistically significant up to 10 years after the event, the long-term impact being even larger than the effect on impact. Also, a (harsh) default leads to a drop in the credit score by around 120 points, i.e. substantial, since the median credit score is 731 points. Finally, individuals recording a (harsh) default witness a surge by around 25% in the probability of having a very low credit limit (i.e. lower than 10,000USD), as well as a 5,000-10,000USD drop in their revolving balance and an increase in their utilization rate of revolving credit, meaning that their revolving credit limit drops even more.

All the impacts described above are statistically significant and long-lasting, up to 6-10 years after the event. The only exception is the impact of a default on home-ownership, which does not appear to be significant from our event study. This is likely because foreclosure procedures can take up to 5-6 years to be completed, i.e. for the property to the switched from the old owner to the new one. However, our double/debiased lasso results show that, in the long term, a default entails a (slight) decrease in the probability of being home-owner.

Finally, we find evidence that the impact of a harsh default is heterogeneous at least across two dimensions, i.e. age in the year of the event and credit score in the year of the event (2010). Indeed, the effects of a default on our outcome variables of interest are more marked for those who were younger than the median age in the sample, as well as for those who had a credit score higher than the median in the sample in 2010, when they were hit by the default.

Our findings are policy relevant, as knowing the cost of each type of default for the individual is essential in order to design adequate debt relief policies for individuals being in a situation of financial distress. For example, one could think of better shielding young individuals, as well as individuals with a higher initial credit score from a harsh default, in order to prevent the higher welfare losses recorded for these groups. With this paper we contribute to the ongoing debate on the topic.

References

Stefania Albanesi and Jaromir Nosal. Insolvency after the 2005 bankruptcy reform. Technical report, National Bureau of Economic Research, 2018.

Stefania Albanesi and Domonkos F Vamossy. Predicting consumer default: A deep learning approach. Technical report, National Bureau of Economic Research, 2019.

Susan Athey and Guido W. Imbens. Machine learning methods that economists should know about. *Annual Review of Economics*, 11(1):685–725, 2019. doi: 10.1146/annurev-economics-080217-053433. URL `https://doi.org/10.1146/annurev-economics-080217-053433`.

Kirill Borusyak, Xavier Jaravel, and Jann Spiess. Revisiting event study designs: Robust and efficient estimation, 2021. URL `https://arxiv.org/abs/2108.12419`.

Brantly Callaway and Pedro H.C. Sant'Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021. ISSN 0304-4076. doi: https://doi.org/10.1016/j.jeconom.2020.12.001. URL `https://www.sciencedirect.com/science/article/pii/S0304407620303948`. Themed Issue: Treatment Effect 1.

Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1), 2018. doi: 10.1111/ectj.12097.

Robert Collinson and Davin Reed. The effects of evictions on low-income households. 2018.

Janet Currie and Erdal Tekin. Is there a link between foreclosure and health? *American Economic Journal: Economic Policy*, 7(1):63–94, 2015.

Giacomo De Giorgi, Matthew Harding, and Gabriel F. R. Vasconcelos. Predicting mortality from credit reports. *FINANCIAL PLANNING REVIEW*, 4(4):e1135, 2021. doi: https://doi.org/10.1002/cfp2.1135. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/cfp2.1135`.

Rebecca Diamond, Adam Guren, and Rose Tan. The effect of foreclosures on homeowners, tenants, and landlords. Technical report, National Bureau of Economic Research, 2020.

Peter Ganong and Pascal J Noel. Why do borrowers default on mortgages? a new method for causal attribution. Working Paper 27585, National Bureau of Economic Research, July 2020. URL http://www.nber.org/papers/w27585.

S Michael Giliberto and Arthur L Houston Jr. Relocation opportunities and mortgage default. *Real Estate Economics*, 17(1):55–69, 1989.

Luigi Guiso, Paola Sapienza, and Luigi Zingales. Moral and social constraints to strategic default on mortgages. Technical report, National Bureau of Economic Research, 2009.

Sasha Indarte. The costs and benefits of household debt relief. 2022.

Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.

Emily C Lawrence. Consumer default and the life cycle model. *Journal of Money, Credit and Banking*, 27(4):939–954, 1995.

Donghoon Lee and Wilbert Van der Klaauw. An introduction to the frbny consumer credit panel. *FRB of New York Staff Report*, (479), 2010.

Lois R Lupica. The consumer bankruptcy fee study. *Available at SSRN 2132913*, 2011.

Ayman Mnasri. Downpayment, mobility and default: A welfare analysis. *Journal of Macroeconomics*, 55:235–252, 2018.

Christine A. Parlour and Uday Rajan. Competition in loan contracts. *American Economic Review*, 91(5):1311–1328, 2001.

A Harsh and Soft Defaults in Credit Reports

Positive and negative credit events are recorded by credit bureaus, i.e. Experian, Equifax, Transunion, in individuals' credit reports. Such events stay in the report for some time depending on the event type. For example, Experian keeps soft defaults for up to 7 years, chapter 7 for 10 years and Chapter 13 for 7 years.[8] This means that other things equal, an individual with a negative episode will have a lower credit score, of course the other things equal is not a plausible situation as the negative episode will immediately lower the score and diminish the ability of that individual to participate in credit operations. Therefore while the flag for Chapter 7 will stay on for at most 10 years, its impact will typically diminish overtime and the individual credit score could in principle recover in a much shorter time span, e.g. opening secured credit cards (security deposit for a given credit line).

Variable	Mean	Std. Dev.	Min.	Max.	N
Credit score	722.9473	93.1651	300	839	23,081,323
Income	56,099.5197	28,036.1698	1000	343,000	23,081,323
HPI	385.5708	276.6767	41.35	2641.8501	23,081,323
Mortgage Bal.	92,129.7948	179,450.3356	0	9,569,957	23,081,323
Age	53.2163	15.7251	12	130	23,072,889
Credit Card Balance (w/zeros)	12587.2825	22002.297	0	7364334	23,081,323
Move ZIP	0.1232	0.3287	0	1	21,363,071
Move CZ	0.0479	0.2135	0	1	21,179,058
Home_own	0.7155	0.4512	0	1	23,081,323
New Delinq. (90+)	0.0504	0.2189	0	1	23,081,323
New Delinq. (120+)	0.0463	0.2101	0	1	23,081,323
New Forecl.	0.0039	0.0619	0	1	23,081,323
New Ch. 7	0.0074	0.0859	0	1	23,081,323
New Ch, 13	0.0017	0.0408	0	1	23,081,323
Total credit limit on open trades (all)	163,638.4387	228,168.5649	1	9,981,334	23,081,323
Total balance on revolving trades	13,216.62	30,453.91	0	5,111,476	23,081,323
Total credit limit on open rev. trades	47,529.5838	73,527.3312	1	7,529,100	23,081,323
Harsh default	0.0125	0.1111	0	1	23,081,323
Soft default	0.0481	0.2139	0	1	21,495,710

Table A1: Summary statistics of our main variables, 2004-2020. Top 1% of total credit limit, total balance on revolving trades and total revolving credit limit have been trimmed for readability. Credit scores lower than 300 have also been trimmed.

B Pooled Credit Data

C Data Reliability

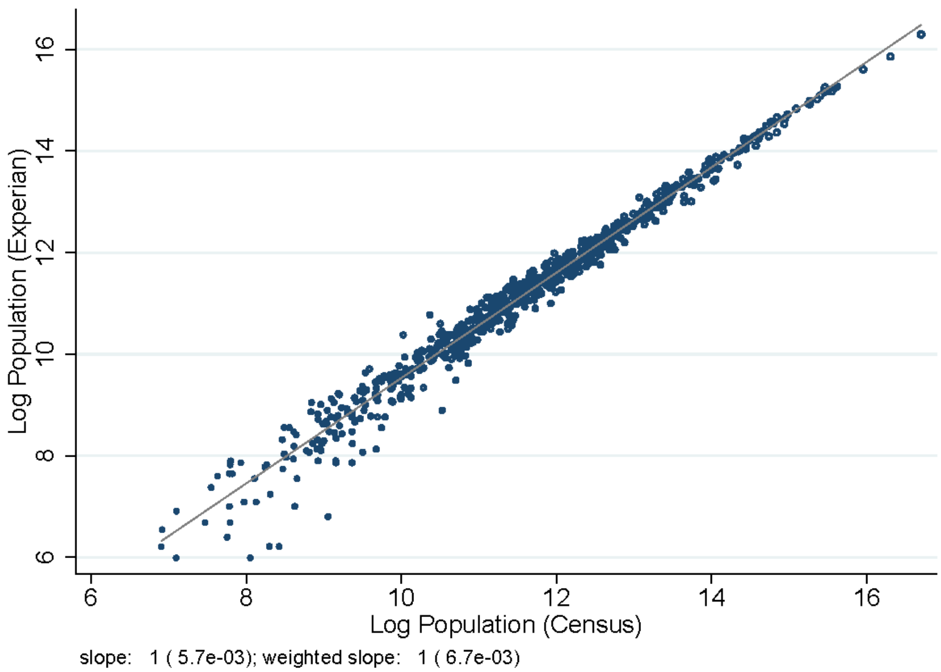

Figure B1: Data representativeness

[8]In the case of Experian see `https://www.experian.com/blogs/ask-experian/how-long-does-it-take-information-to-come-off-your-report/`, for Equifax `https://www.Equifax.com/personal/education/credit/report/how-long-does-information-stay-on-credit-report/`.

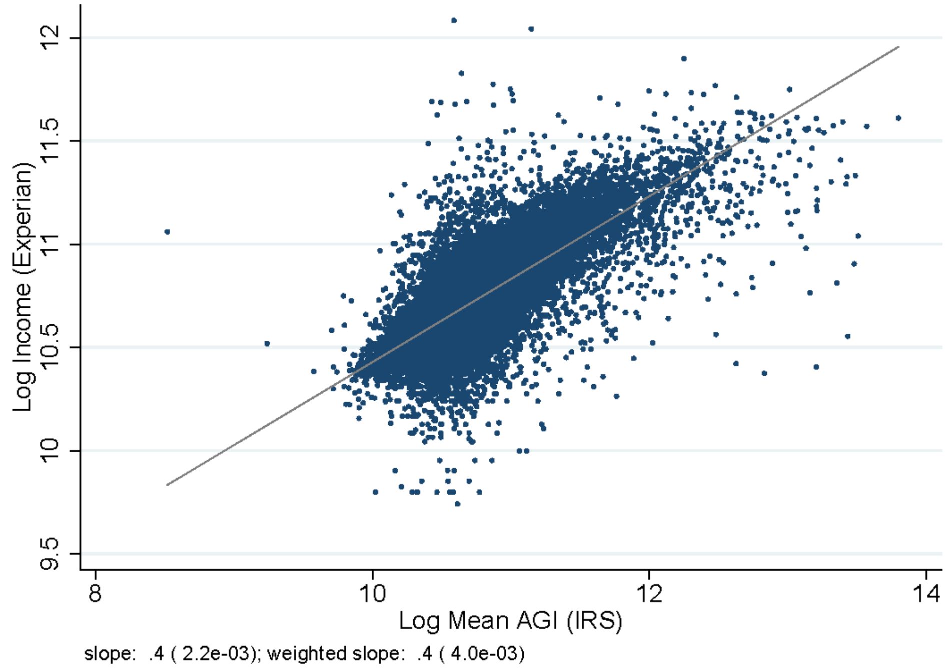

Figure B2: Validity of income imputation

	Labor Income
Mean	48378.05
Sd	82146.32
Min	10
Max	5210000

Table B1: PSID Labor Incomes for the 2009/11 Waves. Negative or zero incomes dropped

	Labor Income PSID	Labor Income Experian
Age	.099	.067
Sd	(.011)	(.0003)
Age Squared	-.001	-.001
Sd	(.0001)	(.0000)
N	10,302	4,390,814

Table B2: (Log)Labor Income Age Profiles in PSID and Experian. We use Labor incomes in the PSID and the W2 imputed income in Experian, we restrict the sample to the years between 2009 and 2011 and to individuals age 25 to 60.

D Additional Evidence in the Long-Term

D.1 Harsh Default in the Long-Term

As an additional exploratory analysis, in Figure B3, we present a snapshot in 2020 of the distribution of key variables such as credit score, income, total credit limit and home-ownership, for individuals who experienced a harsh default in 2010 vs for individuals who did not.

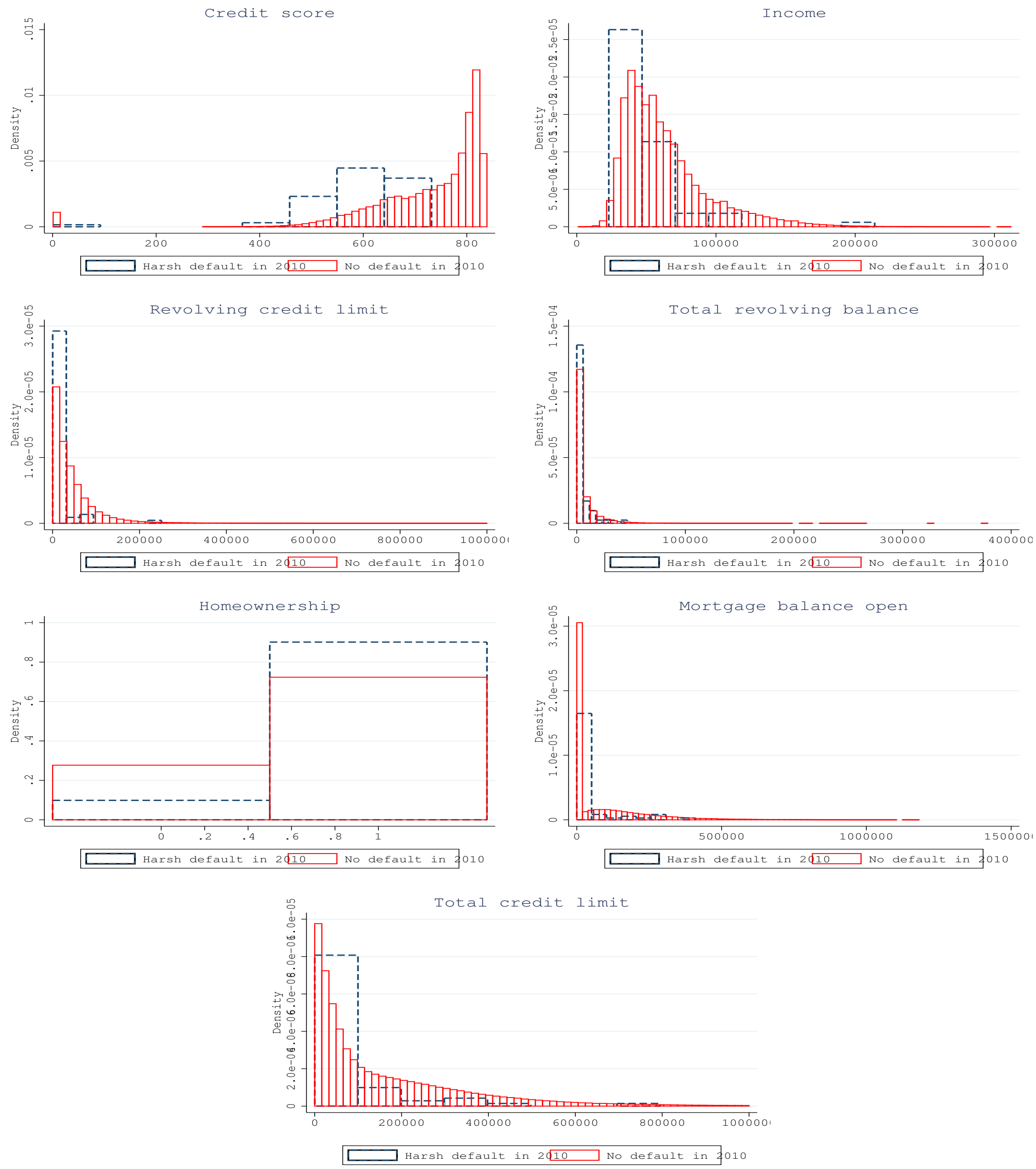

Figure B3: Comparison of the density (histogram) in 2020 of (i) credit score (ii) imputed income W2 (iii) revolving credit limit (iv) total balance of open revolving trades (v) homeownership status (vi) mortgage (amount of balance open) and (vii) total credit amount (i.e. open balance) for individuals who had a harsh default in 2010 vs for those who didn't. Top 1% of total credit amount, income, total revolving balance, revolving credit limit and mortgage balance open have been trimmed for readability of the graphs.

D.2 Soft Default in the Long-Term

In Figure B4, we repeat the same exercise as in Figure B3, but this time comparing those who experienced a soft default in 2010 with those who did not.

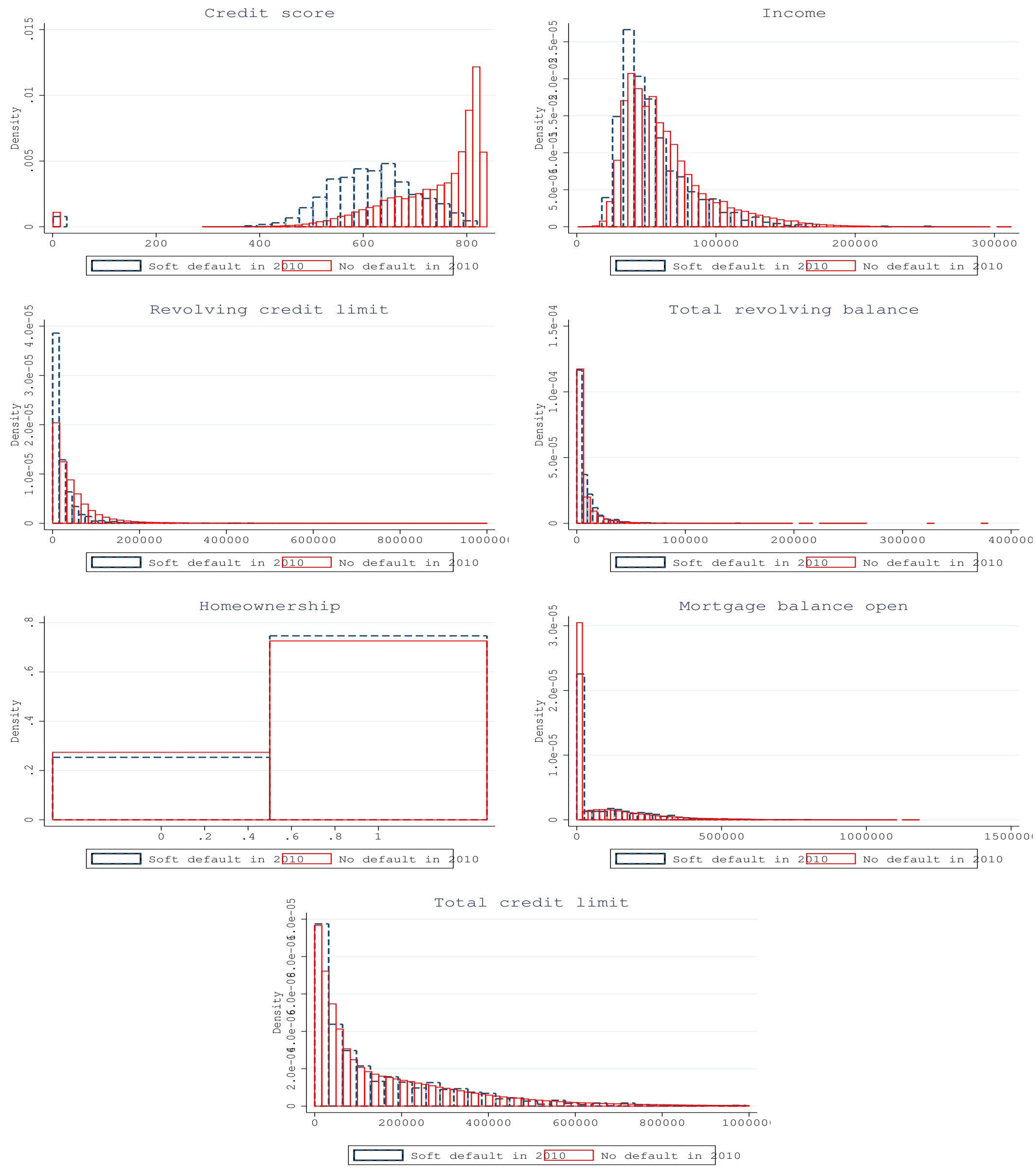

Figure B4: Comparison of the density (histogram) in 2020 of (i) credit score (ii) imputed income W2 (iii) revolving credit limit (iv) total balance of open revolving trades (v) homeownership status (vi) mortgage (amount of balance open) and (vii) total credit amount (i.e. open balance) for individuals who had a soft default in 2010 vs for those who hadn't. Top 1% of total credit amount, income, total revolving balance, revolving credit limit and mortgage balance open have been trimmed for readibility of the graphs.

D.3 Younger vs Older than 30 Years Old in 2010

	(1) move	(2) move_cz	(3) hpi	(4) cs	(5) incomeW2	(6) home_own	(7) ratio	(8) logrevbalance	(9) logallbalance
Harsh def	0.0380^{***}	0.00661^{*}	-36.42^{***}	-67.46^{***}	-23557.2^{***}	-0.0150^{*}	0.0468^{***}	-0.373^{***}	-0.330^{***}
	(7.73)	(2.13)	(-7.20)	(-57.12)	(-45.39)	(-2.38)	(12.68)	(-13.63)	(-10.91)
N	1003013	1001538	1003013	1003013	1002974	1003013	1003013	857245	1002186

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B3: DML long-run impact of a harsh default for individuals younger than 30 in 2010

	(1) move	(2) move_cz	(3) hpi	(4) cs	(5) incomeW2	(6) home_own	(7) ratio	(8) logrevbalance	(9) logallbalance
Harsh def	0.0419^{***}	0.00971^{***}	-19.95^{***}	-78.54^{***}	-25949.9^{***}	-0.00708^{***}	0.0550^{***}	-0.310^{***}	0.0897^{***}
	(38.24)	(13.19)	(-14.37)	(-236.26)	(-209.34)	(-4.27)	(60.76)	(-35.53)	(7.16)
N	13829040	13818449	13829040	13829040	13826605	13829040	13829040	10937695	13823480

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B4: DML long-run impact of a harsh default for individuals older than 30 in 2010

D.4 Counties Above vs Below the Median of Non-White Residents in 2010

	(1) move	(2) move_cz	(3) hpi	(4) cs	(5) incomeW2	(6) home_own	(7) ratio	(8) logrevbalance	(9) logallbalance
Harsh def	0.0375^{***}	0.00582^{***}	-55.45^{***}	-72.50^{***}	-27760.7^{***}	-0.0271^{***}	0.0522^{***}	-0.293^{***}	-0.0313^{*}
	(24.70)	(5.97)	(-26.81)	(-162.32)	(-152.05)	(-12.27)	(42.61)	(-26.25)	(-1.97)
N	7258778	7250495	7258778	7258778	7257528	7258778	7258778	5785083	7254495

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B5: DML long-run impact of a harsh default for individuals living in counties with below median share of non-Whites in 2010 (Census data)

	(1) move	(2) move_cz	(3) hpi	(4) cs	(5) incomeW2	(6) home_own	(7) ratio	(8) logrevbalance	(9) logallbalance
Harsh def	0.0443^{***}	0.0133^{***}	0.526	-83.30^{***}	-23520.9^{***}	-0.00385	0.0559^{***}	-0.352^{***}	0.169^{***}
	(28.83)	(12.44)	(0.40)	(-180.89)	(-150.43)	(-1.65)	(43.78)	(-28.24)	(9.55)
N	7573275	7569492	7573275	7573275	7572051	7573275	7573275	6009857	7571171

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B6: DML long-run impact of a harsh default for individuals living in counties with above median share of non-Whites in 2010 (Census data)