University of Alberta

Essays on Personal Bankruptcy and Mortgage Foreclosure

by

Vyacheslav Mikhed

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Economics

Fall 2013
Edmonton, Alberta

Abstract

This thesis consists of four chapters. Chapter 1 tests the hypothesis that income shocks cause bankruptcy. Using a difference-in-difference specification, we exploit an exogenous fiscal payment, paid to Albertans, and find that this payment causes a decrease in bankruptcies, as predicted by the income shock hypothesis. Using insolvent's balance sheet data, we find that the financial benefits of bankruptcy (liabilities discharged minus wealth forgone) are higher for those filers who received the payment. This is consistent with those potential filers, with smaller advantages from bankruptcy, being dissuaded from filing by the payment.

Chapter 2 examines the effect of income inequality on debt and financial distress. Following the 2008 crisis, several authors have argued that growing inequality increases debts of the poor, who attempt to match the consumption of the rich; and that these debts lead to bankruptcy. We test this argument using a unique database of essentially every personal bankruptcy filing in Canada from 2005 to 2010. Our main finding is that increased income inequality is associated with higher levels of debt in bankruptcy; in particular, larger unsecured and credit card debt and increased risk of bankruptcy.

Chapter 3 explores the impact of the distance between filers and bankruptcy professionals on bankruptcy filing costs. We test if longer distances between debtors and their closest bankruptcy professionals, implying higher transactions costs, lead to debtors demanding larger financial benefits from their bankruptcy to make the bankruptcy worthwhile. We show that distance related

costs are particularly important in rural areas, where distances to the closest bankruptcy professionals are typically large.

Chapter 4 examines the impact of government policies on US mortgage foreclosures. Before the 2008 financial crisis, the US government encouraged mortgage lending to low income borrowers designated as a special ‘under-served’ group by the Community Reinvestment Act. We explore whether this law influenced mortgage foreclosures in 2003-2010. We exploit the 80 percent threshold discontinuity embedded in the law to identify the causal effect of the law on foreclosures. We find that regions with relatively faster and less expensive non-judicial foreclosure process experienced an increase in foreclosures due to the Community Reinvestment Act.

Acknowledgement

I am very grateful to my thesis supervisors Dr. Barry Scholnick and Dr. Melville McMillan for all their continuous help, support, and encouragement. I also would like to thank Dr. Haifang Huang for his insightful comments and suggestions.

Financial support from the Office of the Superintendent of Bankruptcy (OSB) Canada to conduct the research on which this thesis is based is gratefully acknowledged. The views expressed in this thesis are not necessarily those of the Office of the Superintendent of Bankruptcy, of Industry Canada or of the Government of Canada. Particular thanks to Janice Jeffs, Stephanie Cavanagh, Gord Kelly and Lynne Santerre for their help with the OSB data. I also thank the Government of Alberta for providing data on the 2006 Alberta Resource Rebates.

Table of Contents

List of Tables

List of Figures

Abbreviations

ATM	Automated Teller Machine
BLS	Bureau of Labor Statistics
CANSIM	Canadian Socio-Economic Information Management System
CD	Census Divisions
CMA	Census Metropolitan Area
CPI	Consumer Price Index
CRA	Community Reinvestment Act
CSD	Census Sub Divisions
DA	Census Dissemination Area
DB	Census Dissemination Blocks
DD	Differences-in-Differences estimator
FBB	Financial Benefits of Bankruptcy
GIS	Geographic Information System
GSE	Government Sponsored Enterprises
HUD	US Department of Housing and Urban Development
IALSS	International Adult Literacy and Skills Survey
IPO	Initial Public Offerings
JCHS	Joint Center for Housing Studies
LMI	Low and Moderate Income
MIZ	Metropolitan Influence Zone
MSA	Metropolitan Statistical Area
NCLC	National Consumer Law Center
OLS	Ordinary Least Squares
OSB	Office of the Superintendent of Bankruptcy
PCCF	Postal Code Conversion File
PSID	Panel Study of Income Dynamics
PUMA	Public Use Microdata Area
RD	Regression Discontinuity
ZCTA	ZIP code tabulation area

Chapter 1: Do Income Shocks Cause Bankruptcy? Evidence from Exogenous Fiscal Payments[1]

1.1 Introduction

The income shock explanation for bankruptcy implies that negative income shocks e.g. job loss, health shocks, divorce etc. should lead to increases in bankruptcy filings. This explanation is at the center of a longstanding theoretical debate over whether it, or various alternative explanations such as strategic default, financial benefits or lowered bankruptcy stigma, are the main causes of bankruptcy (see e.g., Fay, Hurst and White, 2002; Gross and Souleles, 2002; White, 2011 and many others).

Given the importance of the income shock hypothesis in the literature, statistical evidence supporting it is quite rare. Recent empirical research has tested this hypothesis by examining the impact of plausibly exogenous events on bankruptcy filings. Gross and Notowidigdo (2011) examine exogenous increases in U.S. state level Medicaid coverage and find that there is indeed a causal relationship between increased state level Medicaid coverage (which would reduce the financial impact of unexpected health shocks) and *reduced* bankruptcy. This finding is thus consistent with the negative income shock hypothesis of bankruptcy.

Other research, however, has found that exogenous shocks do not impact bankruptcy filings in ways predicted by the negative shocks hypothesis. Gross, Notowidigdo and Wang (2013) exploit the randomized timing of US tax rebate checks and directly examine their impact on bankruptcy filings. They find that receipt of US tax rebates actually causes *increased* bankruptcies, which is the opposite to what would be predicted by the income shock hypothesis. Their proposed, institutionally based, explanation is that once liquidity constrained

1 This Chapter is a joint work with Dr. Barry Scholnick (School of Business, University of Alberta).

individuals receive the fiscal cash transfer, they can afford to pay the administrative costs (filing fees) required to file for bankruptcy.

Hankins, Hoekstra and Skiba (2011) exploit the random differences of small and large lottery winnings, and examine their impact on bankruptcy. They conclude that winning the lottery does not reduce, but only postpones bankruptcy. They propose various behavioral type explanations for their findings, including high discount rates or mental accounting, which may lead to lottery winners increasing consumption rather than paying down debt.

Despite these conflicting findings, the first aim of this paper is to provide new empirical evidence on this question. For exogenous variation, we exploit a politically motivated, one time only, fiscal cash transfer paid to every resident of one Canadian province in one specific year (the so called "Ralph-bucks" payments to every resident of Alberta in 2006) but not to any other Canadians. Using a difference-in-difference methodology, we examine the impact of these fiscal cash transfers on every bankruptcy filing in Canada. We find that these exogenous fiscal cash transfers lead to a significant *reduction* in bankruptcies. Even though the nature of the exogenous shocks examined by Gross and Notowidigdo (2011) and the shocks examined in this paper are quite different, both of these papers are consistent with the income shock hypothesis of bankruptcy, in that both find that positive shocks *reduce* bankruptcy. This is in distinction to Hankins, Hoekstra and Skiba (2011) and Gross, Notowidigdo and Wang (2013) who find that positive shocks do not reduce and may even increase bankruptcy.

Issues of endogeneity are central to empirical tests of the income shock hypothesis of bankruptcy because bankruptcy can be both a cause of, and a consequence of, negative income shocks. It is in order to overcome this endogeneity issue that Gross and Notowidigdo (2011), Hankins, Hoekstra, and Skiba (2011), Gross, Notowidigdo, and Wang (2013), as well as the current study, all seek to exploit plausibly exogenous income shocks, and examine how these

shocks impact bankruptcy. In essence, all these tests distinguish between the standard income shock hypothesis of bankruptcy, which predicts that positive income shocks should reduce bankruptcies, and the various behavioral (e.g. high discount rates, mental accounting proposed by Hankins, Hoekstra, and Skiba, 2011) or institutional (e.g. high filing fees proposed by Gross, Notowidigdo and Wang, 2013) explanations which predict that positive shocks should not reduce, or should even increase, bankruptcies.

The key identification strategy we use in this paper is that we can distinguish exactly which individuals received this unexpected fiscal payment (residents of the Province of Alberta in 2006) and which did not, based on their province of residency (approximately 10.5 percent of Canadians are resident in the province of Alberta). We examine the impact of this Alberta specific income shock on every Canadian bankruptcy as measured by data provided to us by the Canadian bankruptcy regulator, the Office of the Superintendent of Bankruptcy (OSB). These OSB data are the result of special runs of the OSB data extraction system conducted specifically for this project. These data contain counts of every Canadian bankruptcy filing in all Canadian Census Dissemination Areas (DAs) in every year.

Specifically, we find that the positive income shock from the 2006 Alberta fiscal cash transfer lowered consumer bankruptcies in our treatment group (Albertans in 2006) by about 7 percent compared to consumer bankruptcies in our control group (Canadians not residents of Alberta in 2006, and residents of all provinces in all other years). This finding is thus consistent with standard predictions of the negative income shock model, which implies that negative income shocks increase bankruptcy filings, while positive shocks reduce bankruptcy counts.

While the first part of the paper documents that the positive income shock reduces the number of bankruptcies, as predicted, the second part of the paper provides new evidence on *why* income shocks impact the number of bankruptcies.

Our argument is that receipt of the exogenous income shock will affect the weighting of the benefits of bankruptcy relative to the costs of bankruptcy of the marginal potential bankruptcy filer. Specifically, we propose the hypothesis that receipt of the income shock will persuade marginal potential filers that they can use the money from the transfer to avoid bankruptcy, thus lowering the number of filers after receipt of the income shock.

A testable implication of this argument is that those individuals who actually do file for bankruptcy, even after receipt of the income shock (the treatment group) will have higher expected benefits from bankruptcy compared to the benefits of filing for bankruptcy of the control group. In other words, those who file for bankruptcy even after receipt of the positive income shock, will be those who are most motivated to file, in spite of the shock. This hypothesis has not previously been developed or tested in the literature.

In order to test the hypothesis that those who proceed to file for bankruptcy, even after receipt of the income shock, are those with the most to gain from bankruptcy, we use detailed balance sheet data from individual bankruptcy filings. In addition to providing count data of every bankruptcy filing in Canada, the OSB has also furnished us with detailed balance sheet data of every Canadian bankruptcy filed electronically (rather than using paper based filings). In the period under study more than 75 percent of bankruptcy filings were completed electronically, thus providing us with an extremely large sample of very detailed bankruptcy filings, containing the full balance sheet of each bankrupt.

Access to these detailed balance sheet data allows us to address the issue of how the receipt of an exogenous income shock impacts the balance sheet of filers who received the shock (the treatment group) compared to those that did not (the control group). Specifically, we examine what Fay, Hurst and White (2002) define as the *financial benefit* of bankruptcy. These authors define the financial benefit of bankruptcy as being the benefits to the filer from the amount of unsecured debt discharged in bankruptcy (e.g. credit card debt) minus the value of

the assets forgone in bankruptcy (i.e. wealth net of province/state level exemptions). The larger this financial benefits amount, the greater will be the benefits to the individual from filing.

Because we have access to the complete balance sheet of every electronic filing, we are able to precisely calculate the financial benefits of filing for bankruptcy of each of these filers. We use the same formula as Fay, Hurst and White (2002), to calculate the financial benefits to every bankruptcy filer in our sample (appropriately adjusted for differences in exemptions between Canadian provinces). We use this calculated financial benefits amount for each bankruptcy filer as our dependent variable in a difference-in-difference specification on the exogenous income shock payable to Albertans in 2006. We are then able to examine whether the average level of financial benefit from bankruptcy of our treatment group (who received the income shock), is higher than the average benefit of our control group. Our results support this hypothesis, in that the financial benefits of our treatment group are significantly higher than the financial benefits of our control group. Thus, for the first time in the literature we can show that the balance sheet characteristics (specifically the financial benefits of bankruptcy) of individuals who proceed to file for bankruptcy, even after receipt of the income shock are different from those who file without access to the shock.

Our total count data (which counts every bankruptcy in Canada, irrespective of whether it is a paper or electronic filing) are available at annual frequencies. A major advantage of our balance sheet data (which are only limited to electronic filings in Canada), is that we can identify the exact month of every filing. Thus, as an additional robustness test, we can examine the timing of the response to the income shock in the difference-in-difference framework. We find that there is a significant impact on the financial benefits of bankruptcy of the treatment group relative to the control group in the first three months after the payment of the exogenous income shock. However, there is no significant response in financial benefits in the subsequent nine months. These timing

specifications are thus consistent with the exogenous income shock having a causal impact on the response of the financial benefits of bankruptcy of the treatment group in the three month period after the income shock.

Our second robustness check exploits the specific characteristics of the Alberta payment to examine the impact of income shocks of different magnitudes. The 2006 Alberta rebate payment consisted of payments of C$400 to every resident of Alberta, including every adult and every child. Thus, a household of 2 received C$800, while a household of 4 received C$1600. Our detailed individual bankruptcy filing data allow us to observe exactly how many individuals live in the bankrupt's household. We can thus compare the impact of income shocks of different magnitudes. As predicted, we find that Alberta Rebate had a significantly greater impact on the balance sheets (specifically financial benefits) in households with more than 3 individuals compared to households containing 3 or fewer individuals.

1.2 The Exogenous Income Shock - The Alberta 2006 Resource Rebate

Our main independent variable of interest is the exogenous income shock that was received by our treatment group (residents of Alberta in 2006) but not our control group (residents of other provinces in Canada in 2006 and all Canadians in all other years). On 12th of September 2005, the government of Alberta announced that it would mail a check of C$400 to every resident of Alberta. Children were also entitled to receive the C$400 fiscal transfer, but their amounts were added to the checks of their parents. The only Alberta residents not eligible for the transfer were prison inmates as of September 1, 2005. The transfer was exempt from taxes and it did not alter eligibility for other government programs. Based on data made available to us privately by the Government of Alberta, 92.2 percent of all payments made by the Alberta Government were done

in January 2006. Furthermore, it is estimated by the Government of Alberta that 97 percent of all individuals eligible to receive the payments had been paid by July 2006. The actual distribution of the checks was performed by the Canada Revenue Agency, the federal government body responsible for tax collection and tax refund dissemination among other tasks.

There was much discussion in the popular press at the time that the motivation for this one-off payment was a "vote grab" designed to increase the popularity of the then Premier of Alberta, Ralph Klein. The politically motivated nature of these cash transfers is indicated by the fact that this kind of payment has never occurred either before or since in the recent history of the province of Alberta,[2] or indeed in any other Canadian province. Because of the perceived political motivation for these one-off payments, Albertans almost universally referred to them as "Ralph-bucks". We follow a variety of authors (e.g. Levitt, 1997) who argue that politically motivated actions of politicians are a good source of exogenous variation.

The magnitude of the Alberta cash transfer was C$400 for each and every member of the household (adults and children). The relative impact of this amount on bankruptcy filers can be seen by comparing this amount to the median assets of bankruptcy filers, which in 2005-2008 were C$5,200. Thus, to take as an example a household of three individuals (who would receive a cash transfer of C$1,200), the magnitude of this cash transfer is 23% of the median assets of bankruptcy filers. Furthermore, the magnitude of the fiscal cash transfers in this study are similar to the magnitudes of the exogenous fiscal cash transfers (US tax rebates) examined by Gross, Notowidigdo and Wang (2013) in their bankruptcy study, which typically fell between US$300 and US$1200 per household.

The nature of the fiscal cash transfer examined in this study allows us to

2 The only, somewhat similar, event we are aware of were the "Social Credit" payments in the 1950s when the Government of Alberta paid around $20 to each Alberta resident. Since these "Social Credit" payments occurred in the distant past, it is unlikely that there was any widespread expectation that a similar payment would be disbursed in 2006.

test the income shocks explanation of bankruptcy vs. various other explanations of bankruptcy including strategic default, the filing fees hypothesis, and behavioral explanations. The transfer provided C$400 to each resident of Alberta, thus generating a positive income shock which according to the income shock explanation of bankruptcy should reduce bankruptcies. All alternative explanations of bankruptcy imply that the transfer should not change or should even increase bankruptcies.

1.3 The Impact of the Shock on Annual Bankruptcy Count Data

1.3.1 OSB Bankruptcy Count Data

In this section, we test whether a positive and exogenous income shock from the Alberta 2006 Resource Rebate had any effect on the number of personal bankruptcy filings in Canada.[3] Our administrative bankruptcy data are provided to us uniquely by the Canadian bankruptcy regulator, the Office of the Superintendent of Bankruptcy (OSB). Under Canadian law, every bankruptcy filing has to be submitted by a licensed bankruptcy trustee (typically a certified lawyer or accountant) to the OSB. Thus, the OSB has a record of every bankruptcy filing in Canada. We have access to two separate databases from the OSB.[4] The first database includes a full count of *all* Canadian bankruptcy filings

3 Consumer bankruptcy in Canada relieves an insolvent individual from some of his/her debts in exchange for non-exempt assets. After bankruptcy filing, all debt collection efforts of creditors against the debtor must be stopped. Secured creditors are given priority in receiving payments from the sale of the assets which stand as a collateral of debts owed to them. Unsecured creditors may receive the remaining portion of the payments only after secured debts are completely repaid. These assets form the bankruptcy estate and they are sold to pay off the debts of the bankrupt.

4 The second database contains balance sheets of all bankruptcy filers as well as their individual income and demographic information. We describe this dataset in detail in the next section. We use these data to test whether the fiscal transfer had any effect on financial benefits of bankruptcy as measured by balance sheet data.

by year and by Canadian postal code.[5] We label this database "annual count data". This database is a very accurate and complete count, but it does not contain any other details from the bankruptcy filings.

These Canadian bankruptcy count data are extremely rich compared to US bankruptcy data. Unlike the United States, there is a single bankruptcy regulator for the whole of Canada, the Office of the Superintendent of Bankruptcy (OSB), which forms a part of Industry Canada (the Federal industry ministry). Every Canadian bankruptcy petition has to be filed with the OSB. This centralized filing process is different from that used in the United States, where bankruptcy filings are recorded at the level of the Bankruptcy Court District (of which there are 94 in the United States). Bankruptcy researchers using US data thus have to acquire individual bankruptcy filings data from each of the separate Bankruptcy Court Districts, and in some cases some courts have refused to divulge these data (e.g. the bankruptcy data used in Gross, Notowidigdo and Wang, 2013 contained 74 percent of bankruptcies in the United States).

While the original data from the OSB are available at the level of Canadian six-character postal codes, we aggregate these annual bankruptcy count data to the level of Census Dissemination Areas (DAs). This aggregation allows us to match these data with Census and other data containing many control variables available at the DA level. There were 54,626 dissemination areas in Canada during the 2006 Census.[6] Census dissemination areas are geographic spaces used by Statistics Canada. On average, they have between 400 and 700

5 These data are specifically constructed in this way so as to count family bankruptcies only once. This way of counting does not capture secondary but non-joint spousal bankruptcies of individuals who live at a different address than the primary filer. Such filings are extremely rare.

6 The OSB provided us with the bankruptcy count data at the six character postal code level. Canadian postal codes are extremely small geographic units containing only 15 households on average. There are approximately 805,640 active postal codes in Canada. We matched postal codes to Dissemination Areas using the Postal Code Conversion File developed by Statistics Canada and Canada Post. Matching postal code and census data is quite common in Canada. Some of the postal codes have very large counts of bankruptcies. These large values may be due to typos or data entry errors. We drop observations with more than 20 personal bankruptcies in a postal code per year. This amounted to 144 observations or about 0.098 percent of the total sample size being omitted.

inhabitants and can be considered neighborhoods or smaller suburbs. The first obvious advantage of these very fine grained dissemination area level bankruptcy counts is that we can determine exactly which province each region falls in, and thus identify filers who were Alberta residents and who therefore received the 2006 Alberta specific fiscal cash transfer.

1.3.2 Controlling for Stigma and Information Effects - Block Level Spillovers

One of the most prominent explanations for bankruptcy (besides the negative shock hypothesis discussed above) concerns stigma and information spillover effects between individuals in social contact with each other (e.g. Gross and Souleles, 2002; Fay, Hurst and White, 2002; Livshits, MacGee and Tertilt, 2010; White, 2011). This argument states that if individuals are in social contact with other individuals who have previously filed for bankruptcy, then they will be more likely to file themselves. This is either because this social interaction lowers the perceived stigma of bankruptcy, or because knowledge about the procedures of bankruptcy will be transferred. Existing empirical research on these social contact (stigma/information) effects has defined an individual's set of social contacts based on geographic location of previous bankruptcy filers. Fay, Hurst and White (2002) for example, examine stigma/information effects by using aggregate bankruptcy filing data at the US bankruptcy court district level (of which there are 94 in the US), while Gross and Souleles (2002) examine these effects using aggregate bankruptcy filing data at US State level.

In this paper we can control for these social interaction stigma/information effects by including data on lagged aggregate bankruptcy measured at the very small Dissemination Area level. We construct an indicator variable equal to 1 if there were one or more bankruptcies in the DA in the five year period 2000-2004.[7]

7 All our estimates of interest are virtually unchanged with either the dummy variable or the actual count of past bankruptcies.

This period falls before the data used in our main diff-in-diff specification which includes data from 2005 to 2008. It has been argued in the literature (e.g. Scholnick, 2012) that stigma and/or information flows are slow moving and long lasting processes within a social network, thus it is appropriate to include five lagged years in this specification. We argue that because of the very small geographic areas involved (around 500 persons on average) our definition of this variable, which we call the lagged neighborhood effect variable, is much more likely to capture actual social interactions compared to US state level or US bankruptcy court level aggregations (both of which contain many millions of individuals). The stigma/information flows effect predicts a positive coefficient on this variable because past bankruptcies in the social circle should have a positive effect on new bankruptcies in that circle.

1.3.3 Controlling for Observable Income Shocks – Neighborhood and Suburb Level

We include a variety of control variables, measured at different levels of geographic aggregation to control for income levels, income shocks and other demographics. Our first group of variables is taken from 2006 Canadian Census data, which are made available by Statistics Canada at the DA level. We include census measures of average family income in the DA. We also use DA level data to capture 2006 census data on demographic issues such as marital status, age, education, gender and the proportion of homeowners. One concern with the DA level census data is that all data are taken from a single census year (2006), thus while we can use these data to control for the level of average family income in the DA in 2006, we cannot use these data to examine the rate of change in income - i.e. income shocks.

In order to examine income shocks in specific geographic areas we use data from the Canadian tax authorities, the Canada Revenue Agency (CRA). In particular we use local code statistics for personal income tax returns filed. These

data are measured at the level of Census Sub Divisions (CSD). Critically for our purposes, these CSD level CRA data are available annually, thus we are able to capture annual percentage rates of change of income within the specific geographic area. In particular, these data have measures of total income (taxable and not taxable) for all individuals in the CSD, and also the total number of individual tax returns filed in the CSD. We can thus calculate the average personal income of individuals in the CSD by dividing total income across all filers in the CSD by total number of returns filed in the CSD.

In terms of the predicted signs on individual control variables, based on the discussion of the negative shock hypothesis above, we expect that positive income shocks should have a negative effect on bankruptcies, while divorce should have a positive effect on bankruptcies.

1.3.4 Controlling for Financial (Numerical) Literacy

There is a large literature linking issues such as bankruptcy with levels of financial literacy (see e.g. Lusardi, 2012 and many others). Furthermore, Lusardi (2012) argues that a central element of financial literacy is numeracy - i.e. the capacity to conduct relatively complex calculations. A unique feature of this paper is that we can employ a measure of numerical literacy, available at the DA geographic level, to control for the possible influence of financial literacy on personal bankruptcy. Our numerical literacy data were developed by Murray (2011).[8] This variable is computed using the 2003 International Adult Literacy and Skills Survey (IALSS) and the 2006 census. IALSS evaluated numerical skills for a very large sample of the Canadian population and collected various demographic data. The average level of numerical literacy for each DA was estimated, based on the demographic characteristics of that DA. DA numerical literacy scores vary from 100 to 500, where higher levels of the score correspond to higher levels of numerical literacy. Our expectation is that greater levels of

8 We are grateful to Scott Murray for providing us with these data.

numerical literacy (as a proxy for financial literacy) should result in lower levels of bankruptcy.

1.3.5 Number of Persons in Each Dissemination Area

Our final control variable provides the number of residents in each DA. A direct count of actual inhabitants in each DA is done by Statistics Canada during the census. We use this variable to account for dissimilar sizes of regions we study. Some of these regions may have more bankruptcies than other regions just because they have more persons living in them who might go bankrupt.

Matching data from all these sources resulted in a decrease in the number of Dissemination Areas available from 54,626 to 37,563. This decrease was due to the following reasons. Certain DAs have no match in all data sources or have missing values for the variables of our interest, thus, these regions are not included into our sample. In addition, we limit our sample to the ten provinces of Canada, and delete DAs in Yukon, Nunavut, and Northwest Territories as there are very few people living in these regions. A small number of CSD regions in the CRA tax data are reported multiple times with different values of incomes. We omit these CSDs from further analysis to avoid errors.

1.3.6 Annual Count Model Specification

We use the following model to examine the impact of the exogenous income shock on bankruptcy counts:

$$Y_{it} = \delta FiscalTransfer_{it} + \beta_1' IncShocks_{it} + \beta_2' CenCont_{it} + \beta_3' OtherCont_{it} + \varepsilon_{it} \qquad (1)$$

As the dependent variable, Y_{it}, we use DA level bankruptcy count. The main independent covariate of interest is *Fiscal Transfer* which is an indicator variable for DAs in Alberta in 2006, i.e., this variable is equal to 1 for DAs in

Alberta in 2006 and it is equal to 0 otherwise. This specification is a standard Differences-in-Differences estimator (DD) used in the literature (e.g. Angrist and Pischke, 2009; Bertrand, Duflo and Mullainathan, 2004; Pischke, 2007). This estimator measures the impact of the intervention (income shock from the transfer) as the difference between the outcomes, in our case bankruptcy counts, in the treatment group, DAs in Alberta in 2006, and the control group, DAs in the rest of Canada and Alberta not in 2006.

We include only one year in our treatment group because the transfer was paid in 2006 and because of the possibility of autocorrelation in multiple years of a DD specification as highlighted by Bertrand, Duflo, and Mullainathan (2004). Our data and estimation run from 2005 to 2008. We do not examine data after 2008, because of the financial crisis, which significantly impacted the rate of personal bankruptcies.

DA level average income and CSD level changes in personal income are included in Income Shocks controls, (labeled *IncShocks,* in equation (1)). The vector of Census controls, (labeled *CenCont* in eq (1)), contains DA level proportions of homeowners, proportions of males, age distribution (proportions of population of 20-39 years old, proportion of population aged 40-64, and over 65), proportions of divorced, separated and widowed individuals, and proportions of population with five levels of educational attainment (high school, apprenticeship, college, university, graduate). Other controls, *OtherCont*, include numerical literacy, lagged neighborhood effect, and DA population. All regressions include year and province fixed effects to account for time trends and differences between provinces. As can be seen from Table 1-1, treatment and control groups are similar in terms of the DA characteristics as measured by these control variables.[9] We use a Negative Binomial Model to estimate equation (1) because our

9 The prior default variable included in Table 3-4 is available at the individual level only. It shows whether a particular bankrupt has filed for bankruptcies or filed a proposal before. No such variable can be constructed at the DA or other level of aggregation. However, the lagged neighborhood effect variable available at the DA level partially captures the impact of past bankruptcies in a region on current consumer defaults.

dependent variable is a count.[10]

1.3.7 Annual Count Difference-in-Difference Results

Table 1-2 provides a summary of our results by only reporting the key diff-in-diff fiscal transfer variable of interest,[11] with the rest of coefficients reported in appendix A, see Table A2. Each cell in this Table represents a separate regression, and these results are reported in percentage terms. The coefficient on consumer bankruptcies shows that the expected value of bankruptcies in Alberta in 2006 declined by 7.88 percent compared to the rest of Canada. This coefficient is estimated using pooled data with all controls and standard errors clustered at the DA level. With DA level fixed effects and no controls (they are excluded because they are mostly time-invariant), the coefficient is -6.8 percent and it is statistically significant. The finding for consumer bankruptcies indicates that a *positive* income shock from the Alberta-2006 fiscal cash transfer led to a significant *reduction* in bankruptcies, as predicted by the standard negative income shock explanation for bankruptcies.

The various alternative explanations of the impact of income shocks on bankruptcy imply that a positive shock should have increased bankruptcies, as people would have more money to pay bankruptcy filing fees and therefore declare bankruptcy (Gross, Notowidigdo, and Wang, 2013), or that the shock should have no effect on bankruptcy because individuals file for bankruptcy irrespective of income shocks (Hankins, Hoekstra and Skiba, 2011). Our results do not support these alternative explanations as we show that bankruptcies decreased after a positive shock.

In order to control against the possibility of an unobservable province-year (specifically Alberta-2006) shock driving our results, we conduct several

10 We also tried a Poisson Model and obtained similar results.

11 To simplify interpretation of the coefficients on the *Fiscal Transfer*, we transformed them into percentage terms using the following formula from Long and Freese (2001): [exp(δ)-1]*100%.

falsification tests. Firstly, we examine the impact of the exogenous shock on both small business bankruptcies and consumer bankruptcies. The major distinction between these two types of bankruptcies is that consumer filers have more than 50 percent of debts which are consumer related, while small business filers have more than half of liabilities business related.[12]

We argue that exogenous income shocks payable to all individuals are more likely to have a significant impact on consumer rather than small business bankruptcies because small businesses have larger balance sheets (both assets and liabilities) than consumers. On the other hand, if our results were being driven by an unobservable province-year shock then this unobserved shock should impact *both* small business bankruptcies as well as consumer bankruptcies. The second line of Table 1-2 shows that there is no significant impact of the transfer on small business bankruptcies, which provides support for our argument that the change in consumer bankruptcies are being driven by the exogenous fiscal payment rather than unobservable province-year shocks.

Our second falsification test exploits legal distinctions in Canadian bankruptcy law between filing for bankruptcy and filing a proposal to creditors for a reduction and/or delay in debt repayment, which can be considered a "haircut". A proposal to creditors (haircut) requires the consent of creditors, while a bankruptcy filing does not, thus, the financial situation of proposal filers will tend to be superior to the financial situation of bankruptcy filers (our bankrupts' balance sheet data show that the median assets of a bankruptcy filer in 2005-2008 were C$5,200, while the median assets of a proposal filer were C$17,251). The fiscal cash transfer (C$400 per household member) is larger relative to the size of household assets of bankruptcy filers compared to the assets of proposal filers. Thus, if what we are finding is due to the transfer, the impact on consumer

12 Small business bankruptcies used in this paper are different from bankruptcies of large firms and corporations, which may operate in multiple provinces. Small business are more likely to be local ventures headed by individual entrepreneurs. We argue that a count of small business bankruptcies is better for falsification tests than a count of large business bankruptcies because small businesses are local and more likely to be influenced by province specific shocks.

bankruptcy filers should be significantly larger than on consumer proposal filers. Alternatively, if unobserved macroeconomic shocks are driving our results, the effect of Alberta 2006 on consumer bankruptcies will be not necessarily larger than this effect on consumer proposals. The last line of Table 1-2 shows that effects of the transfer on consumer proposals are not different from zero, thus confirming our conjecture that our results are due to the transfer but not due to Alberta-specific economic shocks.

1.4 Financial Benefits of Bankruptcy

The second part of this paper uses balance sheet data from electronically filed bankruptcies to test the hypothesis that the individuals who proceed to file for bankruptcy in spite of receipt of the exogenous income payment will be filers with higher financial benefits to filing (as defined by Fay, Hurst and White, 2002). Our argument is that filers with high net financial benefits of filing will proceed to declare bankruptcy in spite of receiving the fiscal payment. However, we surmise that those marginal filers who are considering filing but have relatively low financial benefits of bankruptcy may use the fiscal transfer to avoid bankruptcy.

1.4.1 OSB Balance Sheet Data

The second database we use includes the full balance sheet and other filing information provided by the bankruptcy trustee to the OSB (and thus to the bankruptcy court) at the time of the filing. We label this database "balance sheet data". Unlike the bankruptcy count data used in the previous section, this database is not a complete listing of every bankruptcy filing, but it is limited to those filings made electronically using an E-Filing system, rather than filed using paper. The OSB instituted the E-Filing system in 2002, and by 2007 essentially all filings were filed electronically. In the years of interest in this study (before and after the Ralph-bucks payments of 2006) the percentages of electronic filings

were: 2005 - 62.2%, 2006 - 77.4%, 2007 - 97.7%, 2008 - 98.9%.

Thus, while the balance sheet database contains a large majority of bankruptcy filers in Canada, it is not exhaustive and is limited to electronic rather than paper filers. We argue that selection of electronic rather than paper filing should not be a concern for this study because the decision as to whether to file electronically or using paper is a decision of the bankruptcy trustee rather than the individual debtor. Furthermore, the transition to electronic filing was essentially made by all trustees in Canada by the latter part of our study, indicating that there is no systematic reason for why paper or electronic should be preferable compared to the other.

The annual bankruptcy count database used in the previous section has the advantage that it includes a complete count of all bankruptcies, thus we can test the impact of the exogenous shock on the total count of bankruptcies. Because the individual balance sheet data are not complete, i.e., they contain only e-filings but not paper bankruptcy filings, we cannot use these data to calculate counts of bankruptcy events per month or compute numbers of different types of bankruptcy. However, we can use these balance sheet data to explore how the income shock impacts the balance sheet characteristics of filers in our treatment and control groups.

Hankins, Hoekstra and Skiba (2011) and Gross, Notowidigdo and Wang (2013) also use balance sheet data (balance sheet data for individual bankruptcy filers). The balance sheet data used by these authors, however, are limited to small samples, thus these balance sheet data are not part of the main specifications in their papers. Hankins, Hoekstra and Skiba (2011) collect data on the balance sheets of approximately 200 filers. Because of the small nature of this balance sheet sample, these authors do not run formal statistical tests on their data, but only use them to examine differences in means, etc. Gross, Notowidigdo and Wang (2013) collect a somewhat larger sample of approximately 6500 bankruptcy balance sheets. They use that data to examine the impact of their exogenous shock

on the characteristics of bankruptcy filers balance sheets, but the statistical significance of their results is limited.

The main advantage of our balance sheet data is that we have access to every electronically filed bankruptcy balance sheet in Canada, which amounts to a database of many hundreds of thousands of balance sheets. Our main empirical specification for analyzing our balance sheet data closely follows that of Gross, Notowidigdo and Wang (2013) in that the main independent variable is the exogenous income shock, but the dependent variable of interest to us is the financial benefits of bankruptcy.

1.4.2 Financial benefits of bankruptcy

We define individual debtor's financial benefits of filing for bankruptcy in the same way as Fay, Hurst, and White (2002):

$$FinBen_{it} = max\left[D_{it} - max\left[W_{it} - E_{it}, 0 \right], 0 \right] \quad (2)$$

where D_{it} is unsecured liabilities of filers eliminated in bankruptcy, W_{it} is total wealth of bankruptcy filers minus all secured debts, and E_{it} represents bankruptcy exemptions available to filers in a particular year and province. All bankruptcy exemptions allowed in Canada during our study period are described in Table 1-3. Most of the exemptions are related to particular assets such as principal residence, car, furniture, or pensions accounts. Equation (2) captures the idea of bankruptcy which discharges unsecured liabilities of filers in exchange for non-exempt filer's assets. If assets minus secured debts and exemptions are less than equal to zero, then there is nothing to distribute among unsecured creditors and all the bankrupt's unsecured debts are discharged. Following Fay, Hurst and White (2002), we set this formula to be non-negative because individuals are assumed not to file for bankruptcy if their wealth is larger than unsecured liabilities. In this

case, it may be easier for the debtor to sell assets and pay down unsecured debts to avoid bankruptcy altogether.

In order to integrate the 2006 Alberta income shock into this formula, we utilize the very specific ruling made by the OSB (the official bankruptcy regulator in Canada) as to how these payments should be dealt with in bankruptcy. The OSB stated very explicitly that the Alberta 2006 transfer payments were exempt from seizure in bankruptcy. Specifically, the OSB ruled that "the rebate amounts are exempt from execution or seizure, and cannot be assigned…The rebates are considered property of the bankrupt that is *not* divisible amongst the creditors. "[13]

This ruling is of key importance in determining how these payments should be dealt with in the context of the Fay, Hurst and White (2002) formula. Because the exogenous payment increased wealth (cash on hand) of each individual, W (wealth) increased. At the same time, however, the OSB ruling that this cash payment was "exempt" from creditors in bankruptcy, implies that E (exemptions) also increased, by exactly the same amount as the payment. In other words, the change in W-E equals zero, thus the exogenous payment, by definition, had no impact on the financial benefits as defined by equation (2). Intuitively, the Alberta payment should not increase the financial benefit from bankruptcy because the payment does not impact how much additional advantage the individual gets in bankruptcy. The individual keeps the amount of this payment irrespective of whether or not the individual files for bankruptcy.

It is important to note that equation (2) is simply an accounting identity – it shows how much each individual will benefit from filing for bankruptcy, compared to the individual not filing for bankruptcy. This accounting identity, however, is distinct from the specific theoretical hypothesis proposed by Fay, Hurst and White (2002) which is that the larger the amount of financial benefit in equation (2), the greater the probability that the individual will file. Because the specifics of the Alberta 2006 payment had a zero impact on the accounting

13 The full OSB ruling on the Alberta 2006 Rebates is at http://www.ic.gc.ca/eic/site/bsf-osb.nsf/eng/br01567.html

identity (because W and E both increased by the same amount, thus the change in W – E = 0), the financial benefits hypothesis of Fay, Hurst and White (2002) predicts that the Alberta payments should have a zero impact on the decision to file for bankruptcy. This is in distinction to the income shock hypothesis, which specifically predicts that income shocks should indeed impact bankruptcy choices.

We can also compare the income shock and financial benefits hypotheses in terms of their predictions as to the effect of the Alberta income shock on financial benefits of bankruptcy filers. Because the Alberta income shock does not change the amount of financial benefits in bankruptcy (because change in W = change in E) the prediction of the Fay, Hurst, and White (2002) financial benefits hypothesis is that there should be no significant difference in the financial benefits of bankruptcy filers in the treatment and control groups. The income shock hypothesis, on the other hand, predicts that the Alberta payment will lead to a reduction in individuals filing for bankruptcy, with those who proceed to file in spite of receipt of the income shock being those with the most to gain from bankruptcy.[14] Thus, the income shock hypothesis predicts that the financial benefits of the treatment group will be higher than that of the control group.

1.4.3 The Impact of Exogenous Payments on Financial Benefits of Bankruptcy

In this section of the paper we argue that a testable implication of the income shock hypothesis is that those individuals who proceed to file for bankruptcy in spite of receipt of the exogenous payment (the treatment group) will be those with the greatest benefit from bankruptcy. Receipt of the payment should reduce the probability of filing if individuals use the payment to avoid

14 It is important to note that what filers gain from bankruptcy is a discharge of debts, i.e. they do not receive any income from filing, rather some of their financial liabilities are reduced or eliminated. As we describe above, the transfer does not change the bankruptcy calculation, it does not alter debts or assets. Hence, filers in Alberta in 2006 are not better or worse off by the amount of the transfer if they decide to file. Their financial position with regard to the transfer is exactly the same in bankruptcy and out of bankruptcy.

bankruptcy. However, those individuals who proceed to file for bankruptcy, even after receiving the payment, will be those who are most motivated to file – i.e. those with the highest benefits from filing. Thus our testable hypothesis in this section is that the level of financial benefits of the treatment group (filers who received the payment) will be significantly higher than the level of financial benefits of the control group (filers who did not receive the payment).

We can test this argument by comparing average benefits of filing in the treatment and control groups. If individuals with low financial benefits after accounting for bankruptcy costs drop out of the bankruptcy filing pool, the average value of the financial benefits of those who file after the transfer will increase compared to the benefits in the sample of filers without the transfer. Alternatively, if the financial benefits of filling play no role in bankruptcy decision, we will observe no change in the average benefits of filers in the treatment and control groups.

To test this hypothesis, we use the specification presented in equation (1) with the financial benefits of bankruptcy as the dependent variable. As our independent variables, we use both individual level and neighborhood level controls. Individual balance sheets provide us with information on filer's age, marital status (divorce), self-employment, presence of a car, prior defaults, and household size. DA level controls consist of numerical literacy, average personal income, change in personal income (annual CSD data), and educational attainment. We also use postal code level controls such as bankruptcy stigma/information effect's indicator variable and population. All regressions include year and province fixed effect to account for possible time trends and differences among provinces in terms of bankruptcy exemptions and other factors. We run simple OLS on these data with standard errors clustered at the postal code level. We omit individuals with more than 100,000 in financial benefits to remove outliers.

Table 1-4 provides key summary statistics for our dependent and

independent variables of interest. This Table includes several independent variables, e.g., prior defaults, derived from individual bankruptcy filings. These variables are not available at an aggregated level such as DA and so they cannot be used with the bankruptcy count data described in Table 1-1. A comparison of characteristics of the treatment and control groups described in Table 1-4 allows for a conclusion that, with a few exceptions, these two groups are similar in their observable attributes.

As can be noted from Table 1-4, the treatment group constitutes around 1.82 percent of the balance sheet data sample.[15] On the other hand, Table 1-1 demonstrates that the treatment group represents 2.82 percent of the overall bankruptcy count data sample (4143/(142761+4143)*100%). The difference in the sizes of the treatment group in these two samples can be explained by the fact that the bankruptcy count data record DAs irrespective of whether any bankruptcy was filed in them. However, the balance sheet data contain balance sheets of actual bankrupts only. Thus all individuals in the treatment group of the balance sheet data who decide not to file for bankruptcy will not be included in the sample. Our results for the count data in Table 1-2 suggest that the number of bankruptcy filings was indeed reduced in the treatment group. Hence we observe less bankruptcy filers in the treatment group of the bankruptcy balance sheet data.

Figure 1-1 presents densities of financial benefits of filers in the treatment and control groups. As can be seen from this graph, densities of benefits in these two groups are remarkably similar in shape. This similarity may show that bankruptcy filers in these two groups are similar in their characteristics, but for receipt of the transfer. In addition, the density of the financial benefits in the treatment group seems to be shifted slightly to the right which shows that bankrupts in the treatment group derived slightly more benefits from filing compared to the debtors in the control group.

15 The number of observations in the treatment group is 3819, and the total number of observations is 3819+206021=209840. Hence, the percentage of observations in the treatment group is 3819/209840*100%=1.82%.

While Figure 1-1 provides some intuitive evidence on the influence of the income shock on the financial benefits of filing for bankruptcy, Table 1-5 summarizes results of formal statistical tests. This Table shows that financial benefits of filers in Alberta in 2006 increased by 12.13 percent as measured by annual balance sheet data. This finding suggests that on average filers in the treatment group benefited more from bankruptcy than bankrupts in the control group, thus supporting the income shock hypothesis of bankruptcy.

The balance sheet data allow us to split bankruptcy filers by month. Monthly data provide additional insights as to when exactly the change in the filing behavior occurred. Most (92%) of the 2006 Alberta resource rebate was paid in January 2006. Hence, if the increase in the filing benefits of bankrupts in Alberta in 2006 is due to the transfer, the bulk of this increase will happen in the months immediately after the receipt of the transfer, i.e. in February-April of 2006. However, if our findings are due to some other factors, e.g. macroeconomic shocks, then changes in the financial benefits may be detected in some other periods as well.

Table 1-5 shows that financial benefits in the treatment group (Alberta 2006) increased significantly only in February-April 2006, and they were not different from zero in the three quarters after that. The coefficient itself implies that the receipt of the transfer on average increased financial benefits of filers by 25.6 percent in February-April of 2006. Taken together, results in Table 1-5 are consistent with the hypothesis that filers with lower financial benefits of filing compared to the additional costs of filing drop out of the pool of bankruptcy filers and thus drive the average financial benefits in the treatment group up. This finding supports the conjecture that debtors take into account financial benefits of bankruptcy when making their filing decision.

The balance sheet data of individual bankrupts allow us to examine another aspect of the transfer: the impact of the size of the rebate on the financial benefits of filing. These balance sheet data record household size of every

bankruptcy filer. The transfer provided C$400 per each household member, including children. Thus, the size of the transfer received by each household was proportional to the number of family members, with larger households receiving more money. Therefore, we compare response to the transfer among smaller households and larger households. Our hypothesis here is that the increase in the average benefits of filing of larger households will be higher than the increase in the average benefits of smaller households. Since larger households receive more money from the transfer, those with small financial benefits of filing will be even more likely to avoid bankruptcy and drop out of our sample of bankrupts. This tendency will increase average benefits of filing in this group of larger households even more.

Table 1-6 depicts results of the impact of the transfer on the financial benefits of filing for the smaller and larger households. We define smaller households as those with not more than 3 members, while larger households have more than 3 members.[16] For the annual data, financial benefits of both smaller and larger households increased after the transfer. However, benefits of the smaller households on average jumped by 9.76 percent, while benefits of larger households on average grew by 26.55 percent. Hence, the effect of the transfer on the financial benefits of larger families was almost three times larger than the effect for smaller families.

When we split our sample by months after receipt of the transfer, we find similar patterns as in Table 1-5. Financial benefits of both small and large families increased significantly only in February-April of 2006, with the increase in benefits for large families being greater than the increase in benefits for smaller families. These results show that our findings for the impact of the transfer on average financial benefits of bankruptcy filers are robust.

16 We are not able to split our data by household size into smaller groups because some groups have very few observations.

1.5 Conclusion

This paper shows that a fiscal cash transfer reduces the number of personal bankruptcies. This finding differs from a number of other recent papers in the literature which have found that plausibly exogenous shocks (US tax rebates and lottery winnings) do not reduce bankruptcies, but increase or have no impact on bankruptcies. Our finding is consistent with Gross and Notowidigdo (2011) who find that increased Medicaid coverage, which is akin to a negative shock to health care expenditures or a positive income shock, reduces personal bankruptcies.

We use a difference-in-difference estimator by exploiting the "Ralph-bucks" fiscal payments, which were a once only, politically motivated transfer, payable to every resident of Alberta but not payable to any other Canadian. Because we can identify the province of residence of each bankruptcy filer in Canada, we can identify exactly our treatment group of bankruptcy filers who received the transfer (those resident in Alberta in 2006) as well as our control group of bankruptcy filers who did not receive the transfer (those filers resident in the rest of Canada in 2005-2008). Our bankruptcy data are particularly rich, in that we can identify the Census Dissemination Area of every bankruptcy filer in Canada, thus we can include a large variety of DA level controls to account for observable income shocks and other demographics.

In addition to our count data, which include the complete set of every bankruptcy filing in Canada, we also have access to detailed balance sheet data on every bankruptcy filing in Canada submitted electronically. We use this balance sheet data to calculate the financial benefit from bankruptcy as defined by Fay, Hurst and White (2002). Broadly, the financial benefit from bankruptcy is the advantage to the bankrupt from the discharge of unsecured liabilities minus the loss of wealth from bankruptcy (which is positive equity in assets net of bankruptcy exemptions). We find that those bankruptcy filers in the treatment group who received the income shock had significantly larger financial benefits from bankruptcy relative to the bankruptcy filers in the control group who did not

receive the income shock. This finding is consistent with debtors who proceeded to file for bankruptcy, in spite of receiving the income shock, being those with the most to gain from bankruptcy – i.e. those with the highest financial benefits from filing (liabilities discharged minus wealth forgone). In other words, those potential bankruptcy filers who receive the income payment and then choose not to file will be those debtors whose benefit from bankruptcy will be the lowest.

Our results are important for policy, as well as theoretical reasons. In terms of policy, our results are inconsistent with the policy argument that fiscal payments may not be effective in reducing bankruptcy. We find that such payments do indeed reduce bankruptcy filings. In terms of theory, our results contribute to the longstanding theoretical debate as to whether bankruptcy is driven by negative income shocks or by various other explanations including stigma or financial benefit, etc. Our finding that a positive income shock reduces bankruptcy is consistent with the negative income shock hypothesis for bankruptcy.

Figure 1-1: Distribution of the financial benefits of bankruptcy in the treatment group (Alberta 2006, n = 3819) and control group (rest of Canada and Alberta not in 2006, n = 206021)

This figure shows distributions of the financial benefits of bankruptcy in the treatment and control groups after receipt of the transfer. We hypothesize that after receipt of the income payment, those debtors with higher benefits of bankruptcy will be those most likely to file. Thus this hypothesis predicts that benefits of bankruptcy in the treatment group will be higher than benefits in the control group.

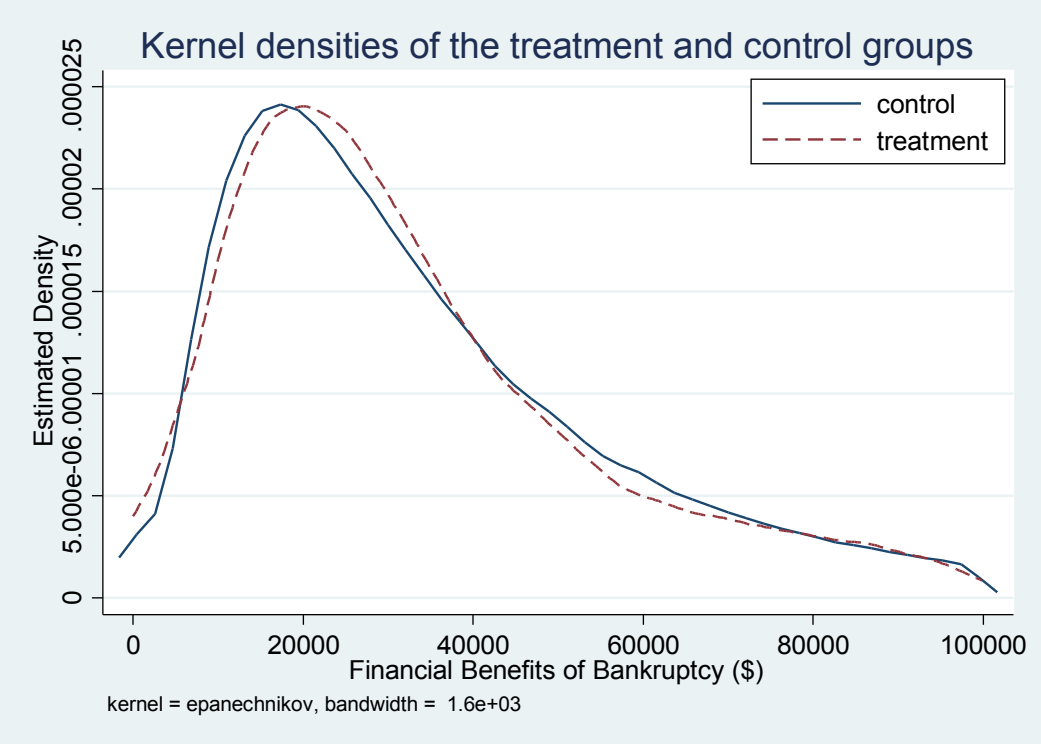


Table 1-1: Summary Statistics of bankruptcy count data and control variables, Treatment Group (DA, Alberta in 2006) and Control Group (all Canadian DAs but for Alberta in 2006)

	Treatment Group			Control Group		
Variable	**Obs**	**Mean**	**Std. Dev.**	**Obs**	**Mean**	**Std. Dev.**
Consumer bankruptcy (DA)	4143	1.154	1.669	142761	1.687	2.296
Small business bankruptcy (DA)	4143	0.130	0.400	142761	0.101	0.360
Consumer proposal (DA)	4143	0.235	0.577	142761	0.416	0.923
Fiscal transfer (DA)	4143	1.000	0.000	142761	0	0
Lagged neighborhood effect (DA)	4143	0.922	0.268	142761	0.919	0.272
Population (DA)	4143	670.8	585.9	142761	629	449
Average personal income (DA)	4143	42337	25161	142761	36089	18412
Change in average personal income (CSD)	4143	11.078	2.925	138614	3.568	6.165
Homeowners (DA)	4143	0.743	0.249	142761	0.725	0.264
Males (DA)	4143	0.499	0.031	142761	0.489	0.032
Age 20-39 (DA)	4143	0.291	0.103	142761	0.262	0.084
Age 40-64 (DA)	4143	0.342	0.069	142761	0.359	0.062
Age over 65 (DA)	4143	0.114	0.090	142761	0.140	0.092
Divorced (DA)	4143	0.080	0.034	142761	0.078	0.036
Separated (DA)	4143	0.029	0.016	142761	0.031	0.018
Widowed (DA)	4143	0.049	0.047	142761	0.060	0.047
High school (DA)	4143	0.235	0.074	142761	0.237	0.079
Apprenticeship (DA)	4143	0.122	0.063	142761	0.113	0.066
College (DA)	4143	0.199	0.069	142761	0.187	0.073
University (DA)	4143	0.183	0.111	142761	0.177	0.107
Graduate (DA)	4143	0.064	0.071	142761	0.077	0.079
Numerical literacy (DA)	4143	276.7	11.8	142761	268.5	13.8

Table 1-2: The impact of the Exogenous Income Shock on bankruptcies

These tests examine the hypothesis that debtors file for bankruptcy after declines in income due to adverse events such as job loss, health problems or divorces. A finding of a reduction in bankruptcies after a positive income shock from the fiscal transfer indicates support for income shock explanation of bankruptcy. These tests use a Negative Binomial model with standard errors clustered at the DA level.

	pooled data robust s.e.	fixed effects
Consumer bankruptcies	-7.879***	-6.779***
Small business bankruptcies	-8.412	-6.984
Consumer proposals	-0.380	0.108

Notes: This Table summarizes the full results reported in an on-line appendix. Each cell reflects one regression and only reports the estimated coefficient on the Fiscal Transfer (Differences-in-Differences) term. Results are reported in percentage terms. *** indicates significance at 1%, ** - significance at 5%, * - significance at 10%. The fixed effect specifications have only DA fixed effects and no control variables. Further details are provided in the appendix.

Table 1-3: Bankruptcy exemptions by Canadian provinces

	Exemptions					
Provinces	**House**	**Car**	**Pension**	**Personal Effects**	**Furniture**	**Land**
Alberta	40000	5000	No	4000	4000	All if rural
British Columbia	12000	5000	All	up to 4000 together		No
Manitoba	2500	3000	All	All	4500	No
New Brunswick	No	6500	All	No	5000	No
Newfoundland and Labrador	10000	2000	All	4000	4000	No
Nova Scotia	No	6500	All	All	All	No
Ontario	No	5650	All	5600	11300	No
Prince Edward Island	No	3000	All	All	2000	No
Quebec	No	No	All	up to 6000 together		No
Saskatchewan	50000	10000	All	7500	All	No

Notes: We use bankruptcy exemption as described by
http://www.bankruptcycanada.com/bankruptcyexemptions.htm
All amounts are in Canadian dollars and apply to equity in the asset. These amounts represent maximum values of assets protected from seizure by creditors in bankruptcy.

Table 1-4: Summary Statistics of balance sheet data and control variables, Treatment Group (Alberta in 2006) and Control Group (all Canadian filers but for Alberta in 2006)

	Treatment Group			Control Group		
Variable	**Obs**	**Mean**	**Std. Dev.**	**Obs**	**Mean**	**Std. Dev.**
Financial benefits	3819	32506.5	21424.5	206021	33131	21576.35
Log of financial benefits	3819	9.8	2.7	206021	10.01	1.924744
Fiscal transfer	3819	1	0	206021	0	0
Age	3819	41.572	13.563	206021	42.698	13.34532
Self-employment	3819	0.050	0.219	206021	0.047	0.2109524
Car ownership	3819	0.686	0.464	206021	0.616	0.4863578
Divorce	3819	0.152	0.359	206021	0.132	0.3389461
Numerical literacy	3819	274.009	11.179	206021	264.621	12.86064
Lagged neighborhood effect	3819	0.662	0.473	206021	0.658	0.4742882
Average income (CSD)	3819	48.289	9.679	206021	37.709	7.973957
Change in income (CSD)	3819	10.911	3.089	206021	3.069	6.712775
Prior defaults	3819	0.168	0.374	206021	0.174	0.3788863
High school (DA)	3819	0.245	0.069	206020	0.242	0.0732812
Apprenticeship (DA)	3819	0.130	0.060	206020	0.126	0.0654541
College (DA)	3819	0.194	0.065	206020	0.184	0.0712781
University (DA)	3819	0.156	0.100	206020	0.146	0.0936073
Graduate (DA)	3819	0.049	0.053	206020	0.058	0.0614403
Postal code population	3819	36.613	53.318	206021	36.489	49.78954
Household size	3819	1.950	1.290	206021	2.025	1.301871

Table 1-5: The effect of the exogenous income shock on the financial benefits of bankruptcy

These tests show which debtors do not file for bankruptcy after receipt of a positive income shock. We test the hypothesis that filers who avoid bankruptcy are those with smaller benefits of bankruptcy compared to other costs of bankruptcy such as no access to credit or bankruptcy stigma. A finding of a positive impact of the transfer on the financial benefits indicates that filers with smaller benefits are dropping out of the pool of bankruptcy filers, thus increasing the average benefits of the remaining filers in the treatment group. These tests use OLS with postal code clustered standard errors. The event study methodology indicates that the effect of the transfer should occur in the months immediately after the payment of the transfer (which was paid in January 2006). A positive coefficient in February-April 2006 is thus consistent with the expected effect of the transfer. No significant effect of the transfer after April 2006 shows that these findings are due to the income shock, but not due to other factors

		Time periods			
Dependent variable	**2006**	**February-April**	**May-July**	**August-October**	**November-January**
log of financial benefits	12.125**	25.622***	5.555	13.270	4.535

Notes: This Table summarizes the full results reported in an on-line appendix. Each cell reflects one regression and only reports the estimated coefficient on the Fiscal Transfer (Differences-in-Differences) term. Results are reported in percentage terms. *** indicates significance at 1%, ** - significance at 5%, * - significance at 10%. Further details are provided in the appendix.

Table 1-6: Magnitude of the transfer's effect on financial benefits of bankruptcy

These tests show which debtors do not file for bankruptcy after receipt of the transfer. We test the hypothesis that filers who receive a larger transfer (those with more than 3 household members) will file only if their benefits are particularly high. This hypothesis implies a larger increase in the financial benefits after the transfer for big households compared to small households. These tests use OLS with postal code clustered standard errors. The event study methodology indicates that the effect of the transfer should occur in the months immediately after the payment of the transfer (which was paid in January 2006). A positive coefficient in February-April 2006 is thus consistent with the expected effect of the transfer. No significant effect of the transfer after April 2006 shows that these findings are due to the income shock, but not due to other factors.

	Time periods				
Sample	**2006**	**February-April**	**May-July**	**August-October**	**November-January**
HH size<=3	9.764*	24.543***	4.248	14.019	3.390
HH size>3	26.553**	30.614*	13.543	9.140	15.663

Notes: This Table summarizes the full results reported in an on-line appendix. Each cell reflects one regression and only reports the estimated coefficient on the Fiscal Transfer (Differences-in-Differences) term. Results are reported in percentage terms. *** indicates significance at 1%, ** - significance at 5%, * - significance at 10%. Further details are provided in the appendix.

Bibliography

Angrist, J., and J. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton: Princeton University Press.

Bertrand, M., Duflo, E., and S. Mullainathan (2004). "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119(1), 249-275.

Fay, S., Hurst, E., and M. White (2002). "The Household Bankruptcy Decision," *The American Economic Review, 92*(3), 706-718.

Gross, T., and M. Notowidigdo (2011). "Health insurance and the consumer bankruptcy decision: Evidence from expansions of Medicaid," *Journal of Public Economics, 95*, 767-778.

Gross, T., Notowidigdo, M., and J. Wang (2013). "Liquidity Constraints and Consumer Bankruptcy: Evidence from Tax Rebates," *The Review of Economics and Statistics*, forthcoming.

Gross, D., and N. Souleles (2002). "An Empirical Analysis of Personal Bankruptcy and Delinquency," *The Review of Financial Studies, 15*(1), 319-347.

Hankins, S., Hoekstra, M. and P. M. Skiba (2011). "The Ticket to Easy Street? The Financial Consequences of Winning the Lottery," *The Review of Economics and Statistics, 93*(3), 961-969.

Levitt, S. (1997). "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *The American Economic Review, 87*(3), 270-290.

Livshits, I., MacGee, J., and M. Tertilt (2010). "Accounting for the Rise in Consumer Bankruptcies," *American Economic Journal: Macroeconomics, 2*(2), 165-193.

Long, S., and J. Freese (2001). *Regression models for categorical dependent variables using Stata.* College Station, Tex.: Stata Press.

Lusardi, A. (2012). "Numeracy, Financial Literacy, and Financial Decision-

making," National Bureau of Economic Research Working Paper No. 17821.

Murray, S. (2011). "Financial Literacy: A Conceptual Review," Research paper prepared for the Task Force on Financial Literacy. Canadian Council on Learning.

Pischke, J. (2007). "The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years," *The Economic Journal, 117*, 1216-1242.

Scholnick, B. (2012). "Bankruptcy, Public Disclosure and Neighborhood Effects," University of Alberta School of Business Working Paper.

White, M. (2011). "Corporate and Personal Bankruptcy Law," *Annual Review of Law and Social Science, 7*, 139-164.

Appendix A

Table A1. Data sources, levels of aggregation, and variables

Variables	Aggregation	measurement	# of units	Data Source
Consumer bankruptcies	Dissemination Areas (DAs)	counts per DA	54,626	Office of the Superintendent of Bankruptcy (OSB)
Population	Dissemination Areas (DAs)	households	54,626	2006 Canada Census
Average personal income		dollars		
Homeowners		proportion		
Males		proportion		
Age 20-39		proportion		
Age 40-64		proportion		
Age over 65		proportion		
Divorced		proportion		
Separated		proportion		
Widowed		proportion		
High school		proportion		
Apprenticeship		proportion		
College		proportion		
University		proportion		
Graduate degree		proportion		
Numerical literacy	Dissemination Areas (DAs)	score between 100 and 500	54,626	Murray (2011)
Change in average personal income	Census Sub Divisions (CSDs)	percent	5,418	Canada Revenue Agency (CRA)

Notes: Number of units as reported by Statistics Canada.

Table A2. The effect of the exogenous income shock on the number of bankruptcies (pooled data, robust standard error). Full results for all variables and untransformed raw coefficients which are used to calculate results reported in Table 1-2.

Independent Variables	Con. Bank.	Bus. Bank.	Con. Proposal
Fiscal transfer (DA)	-0.082***	-0.088	-0.004
	(0.019)	(0.056)	(0.041)
Lagged neighborhood effect (DA)	0.987***	0.793***	0.802***
	(0.026)	(0.053)	(0.038)
Average income (DA)	-0.000***	-0.000**	-0.000***
	(0.000)	(0.000)	(0.000)
Change in average personal income (CSD)	0.002***	0.008***	0.004***
	(0.001)	(0.002)	(0.001)
Numerical literacy (DA)	0.006***	0.011***	-0.015***
	(0.001)	(0.002)	(0.001)
Divorced (DA)	3.694***	2.014***	3.195***
	(0.183)	(0.428)	(0.258)
Separated (DA)	3.413***	-0.216	1.977***
	(0.351)	(0.778)	(0.497)
Widowed (DA)	0.473**	0.128	0.388
	(0.227)	(0.498)	(0.311)
Homeowners (DA)	-0.743***	-0.072	-0.220***
	(0.024)	(0.059)	(0.034)
Age 20-39 (DA)	0.391***	0.265	1.605***
	(0.101)	(0.233)	(0.138)
Age 40-64 (DA)	0.093	0.423	-0.347**
	(0.121)	(0.265)	(0.159)
Age over 65 (DA)	0.154	1.021***	-1.021***
	(0.115)	(0.247)	(0.159)
High school (DA)	-0.265***	-0.748***	1.176***
	(0.082)	(0.178)	(0.110)
Apprenticeship (DA)	0.458***	0.469**	1.469***
	(0.109)	(0.228)	(0.143)
College degree (DA)	-0.463***	-0.739***	0.924***
	(0.096)	(0.199)	(0.128)
University degree (DA)	-1.774***	-1.585***	0.742***
	(0.086)	(0.186)	(0.116)
Graduate degree (DA)	-1.708***	-1.291***	0.280*
	(0.105)	(0.235)	(0.146)
Males (DA)	1.448***	3.154***	-0.291
	(0.183)	(0.408)	(0.251)
Population (DA)	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)
Constant	-2.012***	-8.092***	-0.615**
	(0.225)	(0.455)	(0.305)
Observations	142,757	142,757	142,757

Notes: Raw coefficients from the Negative Binomial Model are reported. ***, **, * denote statistical significance at 1, 5, and 10 percent respectively. Standard errors in parentheses. Data include all Canadian DAs in 2005-2008. The treatment variable is defined to be equal 1 in 2006 and 0 otherwise.

Table A3. The effect of the exogenous income shock on the number of bankruptcies (DA fixed effects). Full results for all variables and untransformed raw coefficients which are used to calculate results reported in Table 1-2.

Independent Variables	Con. Bank.	Bus. Bank.	Con. Proposal
Fiscal transfer (DA)	-0.070***	-0.072	0.001
	(0.019)	(0.056)	(0.040)
Constant	2.152***	-0.596	-0.680
	(0.150)	(0.561)	(0.426)
Observations	133,476	40,142	88,851

Notes: Raw coefficients from the Negative Binomial Model are reported. ***, **, * denote statistical significance at 1, 5, and 10 percent respectively. Standard errors in parentheses. Data include all DAs in 2005-2008. The treatment variable is defined to be equal 1 in 2006 and 0 otherwise.

Table A4. The effect of the exogenous income shock on the financial benefits of bankruptcy
Full results of Table 1-5

Independent Variables	2006	Time periods: February-April	May-July	August-October	November-January
Fiscal transfer (DA)	0.121**	0.256***	0.056	0.133	0.045
	(0.054)	(0.077)	(0.103)	(0.091)	(0.091)
Age	-0.001*	-0.001*	-0.001*	-0.001*	-0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Self-employment	0.253***	0.253***	0.253***	0.253***	0.253***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Car ownership	0.131***	0.131***	0.131***	0.131***	0.131***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Divorce	-0.013	-0.013	-0.013	-0.013	-0.013
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Numerical literacy	-0.005***	-0.005***	-0.005***	-0.005***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged neighborhood effect	-0.051***	-0.051***	-0.051***	-0.051***	-0.051***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Average income (CSD)	0.006***	0.006***	0.006***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Change in income (CSD)	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Prior defaults	-0.054***	-0.054***	-0.054***	-0.054***	-0.054***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
High school (DA)	0.128*	0.129*	0.130*	0.129*	0.130*
	(0.077)	(0.077)	(0.077)	(0.077)	(0.077)
Apprenticeship (DA)	-0.066	-0.067	-0.067	-0.067	-0.067
	(0.092)	(0.092)	(0.092)	(0.092)	(0.092)
College (DA)	0.146*	0.145	0.146*	0.146*	0.146*
	(0.089)	(0.089)	(0.089)	(0.089)	(0.089)
University (DA)	0.842***	0.842***	0.843***	0.842***	0.843***
	(0.074)	(0.074)	(0.074)	(0.074)	(0.074)
Graduate (DA)	0.735***	0.734***	0.737***	0.737***	0.737***
	(0.106)	(0.106)	(0.106)	(0.106)	(0.106)
Postal code population	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Household size	0.041***	0.041***	0.041***	0.041***	0.041***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Constant	10.958***	10.956***	10.955***	10.956***	10.955***
	(0.147)	(0.147)	(0.147)	(0.147)	(0.147)
Observations	209,839	209,839	209,839	209,839	209,839

Notes: Raw coefficients from the OLS are reported with logarithm of benefits as the dependent variable. ***, **, * denote statistical significance at 1, 5, and 10 percent respectively. Standard errors in parentheses. Data include all Canadian bankruptcy filers in 2005-2008, we drop observations with more than $100,000 in benefits.

Table A5. The magnitude of the income shock and financial benefits of bankruptcy. Households with no more than 3 members. Full results of Table 1-6

		Time periods			
	2006	**February-**	**May-**	**August-**	**November-**
Independent Variables		**April**	**July**	**October**	**January**
Fiscal transfer (DA)	0.098*	0.245***	0.042	0.140	0.034
	(0.059)	(0.086)	(0.113)	(0.096)	(0.100)
Age	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Self-employment	0.270***	0.271***	0.270***	0.270***	0.270***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Car ownership	0.141***	0.141***	0.141***	0.141***	0.140***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Divorce	-0.028**	-0.028**	-0.028**	-0.028**	-0.028**
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Numerical literacy	-0.005***	-0.005***	-0.005***	-0.005***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged neighborhood effect	-0.057***	-0.057***	-0.057***	-0.057***	-0.057***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Average income (CSD)	0.005***	0.005***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Change in income (CSD)	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Prior defaults	-0.047***	-0.047***	-0.047***	-0.047***	-0.047***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
High school (DA)	0.165**	0.165**	0.166**	0.166**	0.166**
	(0.081)	(0.081)	(0.081)	(0.081)	(0.081)
Apprenticeship (DA)	-0.130	-0.131	-0.131	-0.130	-0.131
	(0.098)	(0.098)	(0.098)	(0.098)	(0.098)
College (DA)	0.203**	0.203**	0.204**	0.204**	0.204**
	(0.093)	(0.093)	(0.093)	(0.093)	(0.093)
University (DA)	0.849***	0.850***	0.850***	0.850***	0.850***
	(0.078)	(0.078)	(0.078)	(0.078)	(0.078)
Graduate (DA)	0.745***	0.744***	0.747***	0.747***	0.747***
	(0.110)	(0.110)	(0.110)	(0.110)	(0.110)
Postal code population	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	10.999***	10.998***	10.997***	10.998***	10.997***
	(0.151)	(0.151)	(0.151)	(0.151)	(0.151)
Observations	176,746	176,746	176,746	176,746	176,746

Notes: Raw coefficients from the OLS are reported with logarithm of benefits as the dependent variable. ***, **, * denote statistical significance at 1, 5, and 10 percent respectively. Standard errors in parentheses. Data include all Canadian bankruptcy filers in 2005-2008, we drop observations with more than $100,000 in benefits.

Table A6. The magnitude of the income shock and financial benefits of bankruptcy. Households with more than 3 members. Full results of Table 1-6

		Time periods			
	2006	**February-**	**May-**	**August-**	**November-**
Independent Variables		**April**	**July**	**October**	**January**
Fiscal transfer (DA)	0.266**	0.306*	0.135	0.091	0.157
	(0.129)	(0.166)	(0.245)	(0.277)	(0.186)
Age	0.000	0.000	0.000	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Self-employment	0.179***	0.179***	0.179***	0.179***	0.179***
	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)
Car ownership	0.097***	0.097***	0.096***	0.096***	0.096***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Divorce	0.068	0.067	0.068	0.068	0.068
	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)
Numerical literacy	-0.007***	-0.007***	-0.007***	-0.007***	-0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Lagged neighborhood effec	-0.018	-0.018	-0.018	-0.018	-0.018
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Average income (CSD)	0.009***	0.009***	0.009***	0.009***	0.009***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Change in income (CSD)	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Prior defaults	-0.087**	-0.087**	-0.087**	-0.087**	-0.087**
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
High school (DA)	0.007	0.008	0.011	0.011	0.012
	(0.217)	(0.217)	(0.217)	(0.217)	(0.217)
Apprenticeship (DA)	0.309	0.307	0.311	0.307	0.312
	(0.233)	(0.233)	(0.233)	(0.233)	(0.233)
College (DA)	-0.035	-0.040	-0.035	-0.036	-0.035
	(0.242)	(0.242)	(0.242)	(0.242)	(0.242)
University (DA)	0.831***	0.828***	0.833***	0.831***	0.832***
	(0.217)	(0.217)	(0.217)	(0.217)	(0.217)
Graduate (DA)	0.550*	0.548*	0.548*	0.548*	0.550*
	(0.322)	(0.322)	(0.322)	(0.322)	(0.322)
Postal code population	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	11.449***	11.434***	11.437***	11.436***	11.437***
	(0.404)	(0.404)	(0.404)	(0.404)	(0.404)
Observations	33,093	33,093	33,093	33,093	33,093

Notes: Raw coefficients from the OLS are reported with logarithm of benefits as the dependent variable. ***, **, * denote statistical significance at 1, 5, and 10 percent respectively. Standard errors in parentheses. Data include all Canadian bankruptcy filers in 2005-2008, we drop observations with more than $100,000 in benefits.

Chapter 2. Inequality, Debt and Bankruptcy: Evidence from Insolvent's Balance Sheets[17]

2.1. Introduction

Following the 2008 crisis, a number of authors have linked income inequality and debt to financial distress (e.g. Rajan, 2010; Kumhof and Ranciere, 2011; Galbraith, 2012). The basic argument is that increasing inequality leads to higher debt levels of the poor, so that they can attempt to match the consumption of the rich, which in turn leads to bankruptcy. These authors have linked income inequality and financial distress in an attempt to explain the well-known stylized fact from Piketty and Saez (2003 and updated data) that income inequality in the US, as measured by the share of income to the top 1% or top 0.1%, peaked in the periods before the 1929 and 2008 financial crises, which were associated with very high levels of personal debt.

Despite a large amount of discussion on the links between inequality, debt and bankruptcy, individual micro based empirical evidence on these relationships is very rare. Bertrand and Morse (2013) provide micro evidence on one element of the story by showing empirically that the consumption levels of the rich induce the non-rich to consume more. Their individual level evidence does not, however, examine levels of debt, or of personal bankruptcy. Another element of the story is explored by Georgarakos, Haliassos and Pasini (2013) who use a Dutch individual level data set to demonstrate that persons with lower than average income tend to accumulate more in secured and unsecured liabilities.

Bertrand and Morse (2013) do examine the links between inequality and bankruptcy, but use cross sectional US state level regressions which show that higher income inequality leads to higher rates of aggregate bankruptcies in the

17 This Chapter is a joint work with Dr. Barry Scholnick (School of Business, University of Alberta).

state. Frank, Levine and Dijk (2010) use essentially the same methodology of regressing income inequality on aggregate bankruptcies, but rely on data from the 100 most populous counties in the US, and also find that higher levels of income inequality lead to more bankruptcies.[18]

While the existing literature explores particular elements of the link between inequality and bankruptcy, we aim to specifically examine how inequality affects the debt composition of bankrupts, the reasons for their decision to file, and the number of bankruptcy filings. We use a new and unique data base containing detailed balance sheet data of essentially *every* personal bankruptcy filing in Canada from 2005 to 2010. These individual balance sheet data consist of the dollar amounts of all debts of each bankruptcy filer, including all mortgages, credit cards and other debts, etc. These data were provided to us by the Canadian Bankruptcy regulator, the Office of the Superintendent of Bankruptcy (OSB). There are almost half a million individual bankruptcy filings in the database.

We use these data to address three related research questions in this paper. First, we examine whether cross sectional differences in income inequality impact the levels of debt of each bankruptcy filer compared to other bankruptcy filers. Second, we test whether bankruptcy filers in unequal regions are more likely to declare bankruptcy due to excessive borrowing compared to other reasons such as health shocks and divorce. Third, we explore whether higher levels of income inequality are associated with more bankruptcies. These three questions all examine various elements of the story in the literature (e.g. Rajan, 2010; Kumhof and Ranciere, 2011; Galbraith, 2012) that income inequality leads to more debt, which in turn leads to more financial distress.

Our first finding is that income inequality is indeed associated with higher levels of debt across bankruptcy filers, and in particular higher levels of unse-

18 Bordo and Meissner (2012) use country level data from 1920 to 2000 and find that while credit booms lead to financial crises, increases in top income shares do not lead to credit booms. They thus conclude that while debt leads to financial distress, there is no empirical link between inequality and debt.

cured and credit card debt. This finding is of particular interest in the context of the 2008 crisis because we can compare the impact of inequality on both mortgage debt and credit card debt of Canadian bankruptcy filers. It is possible that Canada did not face a very severe crisis in 2008, largely because mortgage lenders in Canada were far more restricted in the amounts of mortgages they could provide relative to US mortgage lenders. The specific role of the mortgage market in the US is central to much of the discussion linking income inequality to financial distress.[19] However, mortgage debt seemed to be less important for bankrupts in Canada as only 23.5 percent of bankruptcy filers had mortgage debt outstanding compared to 89 percent of all bankruptcy filers having credit card debt outstanding.

One way to interpret our results is that bankrupts living in high inequality areas will have had greater pressures to increase their debts before bankruptcy in order to match their consumption with that of their richer neighbors. Thus, these individuals are likely to have relatively higher levels of debt when they file for bankruptcy. However, bankrupts living in low inequality areas will have had fewer pressures to increase debt to match their neighbors' consumption because their neighbors would not be that much richer. Thus, these individuals will be more likely to declare bankruptcy for reasons other than excessive debt and will have lower levels of debt when they file for bankruptcy.

Our direct evidence for the heterogeneity of motivations for bankruptcy across filers forms the basis for our second main finding. This heterogeneity can be seen by using our own database of Canadian bankruptcy filings. Each bankruptcy filer in Canada is required to respond to an open-ended question in the filing form, which asks *"Give reasons for your financial distress."* (OSB Form 79, Question 19). We have access to the complete textual responses to this question

19 Frank (2011) describes mortgage borrowing and expenditure cascades as follows (p. 61). "*Top earners build bigger mansions simply because they have more money….But the larger mansions of the rich shift the frame of reference of the near-rich...So the near rich build bigger, too, and that shifts the relevant framework for others just below them, and so on, all the way down the income scale."*

from every bankruptcy filer. Using textual analysis software, we classify all these responses into a large number of different categories – one of which is "overuse of credit". Approximately 60 percent of all bankruptcy filings listed "overuse of credit" in their response to this question. Other possible reasons include divorce, gambling, unemployment, etc. The fact that 60 percent of bankrupts indicated that excessive borrowing was a reason for their financial distress demonstrates that while debt is clearly an important rationale for bankruptcy, it is not the only reason. We use this heterogeneity in possible reasons for bankruptcy across individual filers to form our second testable hypothesis, which is that income inequality influences the rationale for declaring bankruptcy.

Thus, our second finding is that income inequality increases the probability of filing for bankruptcy due to overuse of credit as opposed to other reasons such as unemployment, health problems, gambling, etc. The existing literature on personal bankruptcy (e.g. Fay, Hurst, and White, 2002; Gross and Souleles, 2002; Livshits et al., 2010; White, 2011, and many others) has emphasized a wide variety of possible reasons for bankruptcy – only one of which is excessive debt. Thus, simply running a cross section regression of income inequality on *total* bankruptcies in a geographic area, as done by Bertrand and Morse (2013) and Frank, Levine and Dijk (2010), fails to account for the fact that not all bankruptcies are driven by issues of excessive borrowing. However, Rajan (2010), Kumhof and Ranciere (2011), Galbraith (2012) emphasize that the possible influence of income inequality on personal bankruptcy runs through excessive debt. Thus we test the hypothesis that inequality increases the probability of declaring bankruptcy due to excessive borrowing and find support for this hypothesis.

Our third section examines related issues, using an alternative database. This database contains data on all Canadian bankruptcy counts at the level of six digit postal codes, which are very small geographic areas with only 15 households on average. These bankruptcy count data are complementary to our balance sheet data. While our balance sheet data contain information on individual bankruptcy

filers, these data do not have information on persons that might have been under financial duress but had not filed for bankruptcy. Thus, with these data alone we cannot examine whether income inequality increases probability of filing for bankruptcy as opposed to not filing for bankruptcy. In order to overcome this obstacle, we use data on counts of bankruptcy and proposal events in a geographic area. This methodology is common to the bankruptcy literature (e.g. Gross and Notowidigdo, 2011; Bertrand and Morse, 2013; Frank, Levine, and Dijk, 2010). If inequality affects the probability of an individual's filing for bankruptcy, it will also change counts of bankruptcies in a region where this individual resides. Thus, using these new count data, we examine whether financial distress induced by inequality and excessive borrowing results in larger counts of bankruptcies. We test this hypothesis and find that income inequality has positive effects on the counts of consumer defaults, and, especially, on the number of bankruptcies due to overuse of credit.

The remainder of the paper is structured as follows. Bankruptcy balance sheet data and liabilities of bankrupts are examined in Section 2. Section 3 provides evidence of the impact of income inequality on the motivation to file for bankruptcy. Section 4 describes our findings for the effect of income inequality on the counts of personal bankruptcies. A summary and conclusion are provided in Section 5.

2.2 Income Inequality and Debt

2.2.1 Bankruptcy Balance Sheet Data

All bankruptcies in Canada must be filed with the Office of the Superintendent of Bankruptcy (OSB). Bankruptcy filers are required to provide full details of their assets and liabilities (on OSB Form 79)[20] as well as their

20 Form 79 is at http://www.ic.gc.ca/eic/site/bsf-osb.nsf/eng/br02196.html.

current income and expenses (on OSB Form 65).[21] Our data include most of the information from these two forms. All of these forms are required by law to be filed by a Bankruptcy Trustee, who is typically a Lawyer or Accountant certified by the OSB to file bankruptcies. Bankruptcy petitions are legal documents, thus appropriate levels of proof are required by the Trustee from the bankruptcy filer, before petitions can be filed with the OSB.

The OSB was able to provide us with all files submitted to them electronically by Trustees, but not any paper based filings. A significant shift to electronic based filings was implemented by the OSB in the period after 2004. Our data run from 2005 to 2010, and in those years the percentage of electronic (versus paper) filings was 2005 - 62.2%, 2006 - 77.4%, 2007 - 97.7%, 2008 - 98.9%, 2009 - 98.6% and 2010 - 99.6%.

The OSB provided us with data on balance sheets from consumer bankruptcies and consumer proposals. Bankruptcies and proposals are two types of personal insolvencies allowed in Canada. Bankruptcies permit borrowers to extinguish their debts but require them to forgo their non-exempt assets. Proposals allow filers to keep their assets but force them to repay part of their debts to creditors. Canadian personal bankruptcy and proposal are roughly comparable to US Chapter 7 and Chapter 13 bankruptcies respectively. The OSB defines consumer bankruptcies and proposals as those where more than half of the debts are consumer related.

Consumer insolvencies which we use in this paper are primary filings only and they do not include secondary or corollary bankruptcies.[22] Those corollary bankruptcies are of spouses or other relatives who file for bankruptcy as a result of a default of the primary filer. The corollary bankruptcies are excluded in order

21 Form 65 is at http://www.ic.gc.ca/eic/site/bsf-osb.nsf/eng/br02190.html.

22 Secondary or corollary bankruptcies are defaults triggered by the bankruptcy of the primary bankruptcy filer. For instance, filer's spouse or children may also file for bankruptcy because their financial affairs worsen significantly after the bankruptcy of the primary filer. These secondary bankruptcies are in essence family bankruptcies, so they should not be counted as two or more separate defaults.

to record family bankruptcies only once and to avoid repeat usage of the same files. The exclusion of secondary bankruptcies does not allow us to account for extremely rare defaults where a corollary bankruptcy is filed by a relative (not spouse) with a different address than that of the primary bankruptcy filer.

Our balance sheet data of each bankruptcy filer (from form 79) include separate amounts for a variety of assets (e.g. house, furniture, car, personal effects, securities, and cash). In terms of liabilities, we are able to observe different types of liabilities including: real property mortgages, non-mortgage bank loans, credit cards from banks, credit cards from non-banks (e.g. retailers), student loans, loans from individuals (e.g. pawn shops) and other liabilities. For each of these categories of liabilities we are able to distinguish between secured and unsecured liabilities.

We can also observe the current income and expenses of each debtor from form 65. Income categories include employment income, self-employment income, pension, child/spousal support, etc. Expenditure categories include child/spousal support, housing expenses (rent or mortgage), transport, insurance, and other personal expenses.

Our data include some but not all demographic information from these forms. We have data on the age, marital status and number of individuals in the household of each bankruptcy filer, but we were not provided with data on the gender, race, official Canadian language or education level. In terms of geographic location we do not have data on the exact street address of the filer, but we do have data on the Canadian six digit postal code of each debtor. Canadian six digit postal codes are extremely small geographic areas containing on average only 15 households, thus these postal codes provide us with a very precise indicator of each filer's geographic location. These six digit postal codes are central to our matching of individual bankruptcy filing data with income inequality data described below.

2.2.2 Data on Income Distributions and Inequality Measures

Canadian six digit postal codes of individual bankruptcy filers are used in matching of our individual bankruptcy data with the data on personal income distributions at two levels of aggregation. These levels are Census Subdivisions (CSDs) (5,000 inhabitants) and Census Divisions (117,000 residents). Income inequality measured at the CSD level may capture local income inequality more precisely than data for US counties (used in Frank, Levine, and Dijk, 2010), which on average have 100,000 residents (the 2000 US Census), US Public Use Microdata Areas (PUMAs) with an average population of 127,000 (Luttmer, 2005), or US states with millions of inhabitants (used by Bertrand and Morse, 2013).

Another unique feature of our income data is that they come from two different sources. In particular, data on income distributions at the Census Division level are from the 2006 Canada Census. While this source covers only one year (2006) and it relies on a survey, it supplements our second data base which is based on Tax returns. The second source of data on income distributions is the Canada Revenue Agency (CRA), which is Canadian tax authority. The CRA data are for Census Subdivisions (CSDs) and they are available for 2005-2008. These data are drawn from Personal Income Tax returns. False reporting on tax returns is punished by law; hence, these data may have fewer mistakes than the Census data.

We use income distributions from these two sources to compute income inequality as Gini coefficients for each region in our database. Gini coefficients for CSDs and CDs are computed using the method suggested by Mark Burkey and R-library *ineq*.[23] We match these coefficients to the rest of our data using the Postal Code Conversion File developed by Statistics Canada and Canada Post. Matching postal code level data with Census data at the CSD, CD and other

23 A description of this method can be found here http://www.ncat.edu/~burkeym/DOCS/Gini%20coefficients%20census%20data.doc

geographic levels is common in studies with Canadian data.

Taxes and transfers may alter income distributions, thus Gini coefficients computed using incomes before taxes tend to be larger than Gini coefficients based on incomes after taxes. However, we only have access to income distributions before taxes for Census Subdivisions. Hence, we decided to use income inequality measures before taxes throughout this paper. Our average before-tax CSD Gini coefficient of 0.448 (see Table 2-1) is very similar to before-tax Gini coefficient for entire Canada computed by Statistics Canada and equal to 0.436.[24]

2.2.3 Controls

Our most important control variables are average incomes at the very local level (Census Dissemination Areas) and income shocks at the provincial level. These covariates are important because income levels and shocks may have strong impacts on bankruptcies. DA average incomes come from the 2006 Canada Census and therefore, they measure only geographic variation but no temporal variation. Since provincial economies could have performed differently during our study period, we use provincial level data on changes to income from Statistics Canada to capture shocks to income across time and space.

Recent literature has argued that numerical literacy is crucial in financial decisions (e.g. Lusardi, 2012). We are able to control for the impact of this factor on debt accumulation using a new numerical literacy variable capturing the ability to use mathematical tools in consumer finance.[25] Financial literacy scores rely on the data available in the 2003 International Adult Literacy and Skills Survey (IALSS) and the 2006 Census. IALSS provides demographic and financial literacy data for a sample of Canadian population, which are used to estimate a

24 Canadian Socio-Economic Information Management System (CANSIM) table 202-0709.
25 We are grateful to Scott Murray for these data. See Murray (2011) for further details.

model of how the demography influences financial literacy. The coefficients from this model and data from the Census are combined to predict financial literacy scores for all Dissemination Areas (DAs) in Canada.

We are also able to control for bankruptcy stigma/information effects using six digit postal code level data on past bankruptcies. The existing literature argues that past local bankruptcies may lower perceived bankruptcy costs or might provide information about the bankruptcy filing process, thus encouraging more debtors to file (e.g. Fay, Hurst and White, 2002; Scholnick, 2012). We use data from the OSB on all bankruptcies in a six digit postal code in 2000-2004 to construct a variable capturing the influence of bankruptcy stigma/information at the postal code level. This covariate is an indicator variable equal to 1 if a given postal code had at least one bankruptcy in 2000-2004 and equal to zero otherwise.

Data from individual bankruptcy filings allow us to construct several control variables at the individual level. In particular, we control for filer's age, household size, divorce, car ownership, mortgage (which proxies for homeownership), self-employment, total assets, and prior personal defaults. It has been argued in the previous literature on personal bankruptcies that individual level shocks to income or expenditures may increase bankruptcies (e.g. Gross and Notowidigdo, 2011). We control for such shocks using variables for divorce, age, household size, and self-employment. Prior personal defaults allow us to account for repeat bankruptcy filers who could be driven in the filings not so much by shocks as by strategic motive (i.e. debt elimination and asset retention). Finally, car ownership, mortgages, and total assets may capture the effect of assets on debt accumulation and use of credit to buy consumption goods.

While census data and financial literacy data are measured at the Dissemination Area level, bankruptcy stigma is available at the postal code level. We matched these data together using the Postal Code Conversion File provided by Statistics Canada. The matching reduced the number of observations available, but it allowed us to add controls for important factors influencing personal

bankruptcies identified in the previous studies of this subject and described above. After the matching, we have around 500,000 personal bankruptcy balance sheets available for analysis. Table 2-2 provides summary statistics for these data.

2.2.4 Estimation Strategy

Our main econometric specification has the following form:

$$Y_{it} = \delta Inequality_{it} + \beta_1' Income_{it} + \beta_2' Individual_{it} + \beta_3' Neighborhood_{it} + Province_{it} + Year_{it} + \varepsilon_{it} \quad (1)$$

As the dependent variable, Y_{it}, we use one of the various measures of financial distress experienced by an insolvent before filing for bankruptcy or filing a proposal. These measures include the dollar value of total liabilities, unsecured liabilities, credit card debt, and mortgage liabilities. We consider additional dependent variables such as the ratios of unsecured liabilities to total liabilities and credit card debt to total liabilities.

We use total liabilities of bankrupts to examine the hypothesis that bankruptcy filers in more unequal regions declare bankruptcy with more debt than defaulters in less unequal localities. Our additional dependent variables such as unsecured liabilities and credit card debts allow us to explore whether inequality influences debt structure of bankruptcies and whether certain types of debts are more responsive to inequality than other types of debt. Ratio of unsecured debt to total liabilities and credit card debt to total liabilities also reveal the changes in the debt structure of bankrupts induced by income inequality.

We estimate equation (1) using a Tobit model with the lower limit set to zero for the levels and ratios of liabilities described above. The lower limit is set to zero because these liabilities and their ratios cannot assume negative values for

bankrupts. For the ratios of unsecured liabilities to total liabilities and credit card debt to total liabilities, we also set the upper limit to 1 because these types of liabilities cannot exceed total liabilities.

Total assets of bankrupts serve as a control in equation (1) instead of being a denominator of our dependent variables, and thus forming debt-to-assets ratios. This approach allows us to save many observations as total assets are equal to zero for many insolvents and the debt-to-asset ratio is not defined. In addition, we use total family income as a control in our robustness checks for the same reason. This control does not change our results in any significant way. However, inclusion of total family income and total assets together as controls do not allow some Tobit models to converge. Therefore, we keep only total assets as a control in our main regressions.

We measure income inequality by the Gini coefficient at either Census Subdivision (CSD) or Census Division (CD) level. As described in the previous section, data used to compute these coefficients come from two sources and these coefficients are either constant over time or vary both with time and space. We use these inequality metrics one by one to check the robustness of our results to alternative ways of measuring inequality. Lower values of these inequality coefficients imply less inequality, while higher values correspond to more inequality.

As argued by Grinblatt et al. (2008), unobserved attributes or shocks may occur simultaneously in small and large neighborhoods. For instance, some individuals may dislike both inequality and debt. They may self-select into neighborhoods with lower inequality and on average have less debt and bankruptcy. This unobserved selection may bias coefficients on regular Gini measures. Alternatively, local shocks such as plant closures may increase income inequality due to layoffs or depressed wages and force individuals to borrow more because of unemployment.

Because we measure income inequality in the two concentric rings, we are

able to implement the methodology developed by Grinblatt, Keloharju, and Ikaheimo (2008), which examines the impact of neighborhood characteristics on individuals. The basic idea of Grinblatt et al. (2008) is to compare the impact of near neighbors from "the inner ring" neighborhoods with the impact of distant neighbors from "outer ring neighborhoods" on the individual. Essentially the methodology of Grinblatt et al. (2008) involves creating a new variable by subtracting the characteristics of the outer ring neighborhood from the characteristics of the inner ring neighborhood (in our case the subtracting of the Gini measures of a larger concentric ring from the Gini measure of a smaller concentric ring). Grinblatt et al. (2008) use this "net neighborhood" measure as a new independent variable of interest capturing neighborhood effects, which is then regressed on the individual level dependent variable (in our case levels of debt etc. of individual bankruptcy filers).

There are both interpretative and econometric reasons for implementing the Grinblatt et al. (2008) methodology in our context. In terms of interpretation, the net neighborhood measure (smaller inner ring Gini minus larger outer ring Gini) indicates the relative importance of the inequality in smaller areas (the inner ring) relative to larger regions (the outer ring) on the individual debt and bankruptcy. In this paper we provide empirical evidence of the relative importance of different consumption reference groups, (as measured by the Gini coefficient in different concentric rings) on the individual bankruptcy filer.

The econometric advantages of using the net neighborhood measure as an independent variable is emphasized by Grinblatt et al. (2008). Essentially, the subtraction of outer-ring neighbor effects from inner-ring neighbor effects controls for omitted common attributes that are shared by residents of both the inner- and outer-rings. This allows us to control for possible endogeneity that could arise from the self-selection of individuals into inner neighborhoods, at the level of the outer-ring neighborhood. Grinblatt et al. (2008) argue that "more distant neighbors are an instrument for omitted control variables that might generate spurious

inferences about near neighbor influences" (p. 735). They also note that "instrumenting with more distant neighbors purchases (in our case Gini coefficients) also controls for other variables (observed and unobserved) that are common to a larger community" (p. 736). Therefore, in addition to regular Gini coefficients, we use differences in these coefficients.

Vector of income control variables (*Income*) in equation (1) includes both DA level average incomes and percent change in income measured at the Provincial level. Individual controls (*Individual*) consist of filer's age, household size, dummy for divorce, dummy for self-employment, dummy for prior personal bankruptcies, total assets, dummy for mortgage, and dummy for car ownership. DA numerical literacy and bankruptcy stigma are included among neighborhood controls (*Neighborhood*). Table 2-2 provides summary statistics on these controls. Year and province fixed effects are added to these models. We cluster standard errors at the Dissemination Area level which is the most disaggregated level of clustering possible in this case.[26]

2.2.5 Results

A summary of our results are presented in Tables 2-4 and 2-5. Each cell in these Tables represents one regression and reports the coefficient on an inequality measure along with its standard error in parentheses. The dependent variable for the corresponding regression is displayed in the first row of each Table, and the inequality measure used in the regression is shown in the first column of the appropriate Table. Each regression is estimated using the Tobit model with control variables and clustered standard errors as described in the previous section. We report estimates of untransformed coefficients from corresponding regressions.

26 We can cluster standard errors at the postal code level, but there are too few observations for each postal code. There are only 15 households in a postal code on average and bankruptcies are relatively infrequent events. Hence, postal code level clustering is not possible practically. We also tried CSD level clustering and no clustering, and got similar results.

Table 2-4 reveals that income inequality at both neighborhood level (CSD) and town level (CD) is associated with higher levels of total liabilities, unsecured, credit card and mortgage debt of bankruptcy filers. Ratios of unsecured liabilities to total liabilities and credit card debt to total liabilities also seem to be higher in more unequal regions. In addition, income inequality at a local level such as Census Subdivision (5000 residents) seems to have a slightly smaller effect on debt accumulation compared to inequality at a more aggregated level (Census Divisions, 117,000 inhabitants on average).

In terms of interpretation, a move from the complete equality at the CSD level (Gini = 0) to the complete inequality (Gini=1) with all other variables constant increases total liabilities of insolvents by $88,827, unsecured liabilities by $78,189, credit card liabilities by $75,886, and it adds $170,989 to mortgage debt. The same change in inequality brings up the ratio of unsecured liabilities to total liabilities by 0.517 and the ratio of credit card debt to total liabilities by 0.95. These findings suggest that individuals in more unequal communities carry much more unsecured and credit card debt when they file for bankruptcy compared to bankrupts in more equal regions.

One standard deviation change in the CSD Gini coefficient, which is equal to 0.043, adds $3819.56 to total liabilities, $3362.13 to unsecured liabilities, $3263.10 to credit card liabilities, and $7352.53 to mortgage debt of bankrupts. Therefore, we can conclude that income inequality increased liabilities of insolvent individuals and families in Canada in 2005-2010. As shown in Table 2-3, only 23.5 percent of insolvents have mortgage, but all bankrupts have unsecured liabilities and around 89 percent of them have credit card debt. This debt structure and the positive impact of inequality on the ratio of credit card debt to total liabilities allow for a conclusion that credit cards and unsecured debt were primary channels of debt accumulation before bankruptcy in Canada.

This conclusion is in contrast to the case of the USA, where, as Rajan (2010) argues, mortgage credit was the primary way of debt accumulation in

response to income inequality before the 2008 financial crisis. However, US and Canadian mortgage lending practices were different in the years before the crisis. While mortgage lending standards were relaxed substantially in the US in the early 2000s, they remained rather strict in Canada. This could be a reason why, unlike its US counterpart, Canadian real estate market did not experience a surge in mortgage foreclosures and a fall in house prices during and after the crisis.

As argued by Grinblatt et al. (2008), common unobservable neighborhood shocks or characteristics may be taken into account by using a difference between two measures of neighborhood effects, one more local and the other more distant. In our analysis, we use Gini coefficients at the CSD and CD levels to construct a difference between them, which could eliminate the impact of some unobservable neighborhood factors on our estimates. Results of Table 2-4 demonstrate that even after we take this difference, the effect of income inequality on debt accumulation is positive, but smaller in magnitude.

While Table 2-4 presents results for all insolvency filers (both proposals and bankruptcies), Table 2-5 summarizes results for these two groups of insolvent individuals separately. Overall, income inequality increases liabilities of both bankruptcy filers and proposal filers. However, estimated effects of inequality on liabilities of proposal filers are smaller than inequality's effects on debts of bankruptcy petitioners.

It can be argued that income inequality is important in debt accumulation only for individuals who are behind their neighbors or other wealthier peers. This is because only those who are behind have motivation to catch up and use credit when they cannot earn enough income. While there are many ways to designate those who are behind, people with incomes below median or average values are natural candidates. Our data on family incomes of bankrupts confirm that 97 percent of them have incomes below DA average incomes, and 95 percent of insolvents are with incomes below DA median incomes. Incomes of individuals filing for bankruptcy may be depressed for some time before filing and so their

value at the time of filing may not capture the complete history of the relative position of these bankrupts in income distribution during debt accumulation (which may take years). Despite this caveat, our results in Tables 2-4 and 2-5 suggest that most bankrupt individuals, who are behind their neighbors in terms of income, also have larger liabilities in more unequal localities.

2.3 Inequality and Overuse of Credit

2.3.1 Why do people file for bankruptcy?

While the previous section explored the question of whether income inequality leads to higher levels of debts of bankrupts and whether any particular liabilities are disproportionately affected by inequality, this section examines which types of bankruptcy filings are influenced by income inequality.

Previous research on personal bankruptcy has emphasized a variety of reasons for bankruptcy including shocks to income or expenses, declining bankruptcy stigma, and strategic defaults (e.g., Fay, Hurst, and White, 2002; Gross and Souleles, 2002; Livshits et al., 2010; White, 2011). Although almost all bankrupts file for bankruptcy because they are unable to make timely debt payments, these individuals may differ in reasons why they cannot make timely payments. One of these reasons is overuse of credit, while other reasons could be health problems, unemployment, gambling, etc. We argue that individuals, who use credit to assuage income inequality and catch up with consumption of their wealthier peers, are more likely to file for bankruptcy citing overuse of credit as the reason for filing rather than other reasons such as health problems, unemployment, etc. Hence, we hypothesize that income inequality affects bankruptcy filings through excessive debt accumulation, but not through other channels. In this section we test this hypothesis using our individual bankruptcy balance sheet data.

2.3.2 Data

Our OSB data on individual bankrupts' balance sheets include the full textual answer to the open ended question posed on the Bankruptcy filing form "*Give reasons for your financial difficulties"* (OSB Form 79, Question 14). Using textual analysis software, these open-ended answers are coded into 17 categories described in Table 2-3. One of these answers is "overuse of credit". We use this answer to separate bankrupts who experienced financial distress because of excessive credit use from insolvents who filed for bankruptcy due to financial shocks and other reasons. Thus, our data allow us to test whether income inequality indeed increases probability of filing due to excessive borrowing compared to other reasons.

2.3.3 Estimation strategy

The econometric specification which we use in this section is very similar to equation (1) and has the following form:

$$Y_{it} = \delta Inequality_{it} + \beta_1' Income_{it} + \beta_2' Individual_{it} + \beta_3' Neighborhood_{it} + Province_{it} + Year_{it} + \varepsilon_{it} \quad (2)$$

Based on the answer to the question "*Give reasons for your financial difficulties"*, we construct a binary variable for filing due to overuse of credit. This variable is equal to 1 when a bankruptcy or proposal filer answers the question by stating that overuse of credit was among their reasons for declaring bankruptcy. This variable is zero otherwise. We estimate the effect of income inequality on the probability to file for bankruptcy due to overuse of credit using the specification shown in equation (2) and the overuse of credit binary variable as our dependent variable in a Logit regression model. The overuse of credit

dependent variable permits to test the hypothesis that individuals in more unequal regions are more likely to file for bankruptcy due to overuse of credit compared to persons in more equal areas.

The same income, individual, and neighborhood level control variables are used in these regressions as in the previous section to account for other factors influencing individual bankruptcy decision.

2.3.4 Results

Table 2-6 summarize results for the overuse of credit specification. This Table shows that income inequality increases probability of filing for bankruptcy as a result of overuse of credit. These findings hold for both CSD and CD Gini coefficients in the whole sample as well as subsamples of bankruptcy and proposal filers. The coefficients imply that one standard deviation growth in CSD Gini coefficient increases the probability of filing for bankruptcy or proposal because of overuse of credit by 0.043*0.324*100%=1.39 percent.

As the last row of Table 2-6 shows, income inequality increases bankruptcy and proposal filings due to overuse of credit even after we account for possible unobserved local attributes or shocks using our methodology similar to Grinblatt et al. (2008). As argued in the previous section, there is no clear theoretical prediction about the relative importance of local and global inequality on overuse of credit, debt accumulation, and bankruptcy. Therefore, from theoretical perspective, differences in Gini coefficients may be either positive or negative. However, the coefficient on the difference between CSD Gini and CD Gini implies that local inequality is more important than global inequality for bankruptcy filings due to overuse of credit.

2.4 Inequality and Personal Bankruptcy

2.4.1. Bankruptcy Count Data

While the previous sections described the effect of income inequality on liabilities of bankrupts and the probability to declare bankruptcy due to overuse of credit, this section explores the impact of inequality on total counts of bankruptcies and proposals and counts of bankruptcies and proposals due to overuse of credit.

As we described in Section 2, there are two types of personal insolvency in Canada: proposal to creditors and bankruptcy. Proposals are roughly similar to Chapter 13 bankruptcy in the USA and allow proposal filers to keep their assets, but require them to repay part of their debts to creditors over several years. Canadian bankruptcies, which are similar to US Chapter 7 bankruptcy, eliminate most debts of insolvents and force them to relinquish their non-exempt assets to creditors. The Office of Superintendent of Bankruptcy (OSB) provided us with counts of these consumer defaults per six digit postal code per year in 2005-2010.

In our analysis of counts of consumer bankruptcies and proposals per six digit postal code per year, we are not able to use individual controls on insolvent and solvent individuals and households. Therefore, we once again turn to the 2006 Canada Census to obtain additional controls. In particular, we attempt to account for local (i.e. postal code and neighborhood) economic and demographic conditions using proportion of homeowners, gender composition, age distribution, number of divorces and family break-ups, and educational attainment (e.g. high school, college, university).

According to Frank, Levine, and Dijk (2010), population density is a very important part of the relation between income inequality and personal bankruptcy. In particular, persons in highly populated urban areas may easily observe each other's consumption, and therefore they are more likely to replicate their neighbors' higher expenditures by borrowing. Higher borrowing may eventually

lead to bankruptcy. That is why Frank, Levine, and Dijk (2010) limit their sample to the 100 most populous US counties. We control for this effect directly using variable capturing region's level of urbanization. To this end, we use the concept of the Metropolitan Influence Zone (MIZ) introduced by Howatson-Leo et al. (1996). This variable ranges from 1 to 8, and it measures the proportion of area's residents commuting to an urban center. Hence, MIZ captures the characteristic of a region as an urban area (closer to 1) or a rural one (near 8). MIZ variable is provided to us at the Census Subdivision (CSD) level in 2006. Since population density is highly correlated with urbanization, the MIZ covariate controls for the influence of this factor on personal bankruptcies.

Our last control is the number of dwellings per each postal code. This control may be important since larger postal codes might have more bankruptcies just because of their size. The number of households is not available directly from any source, but we estimate it using Census data. For this purpose, we use household counts for Census Dissemination Blocks (DBs) which are Census geographic spaces even smaller than Dissemination Areas (DAs). Each DB contains around 5 postal codes on average. So, after we match postal codes to DBs, we divide the number of households in the appropriate DB by the number of postal codes in this DB. Thus, we generate estimates of households for all postal codes in this DB. We are able to match postal codes to DBs for around 85 percent of postal codes. The number of households in the remaining 15 percent of postal codes is imputed using the same method but household counts for DAs.

While aggregating individual cases of bankruptcies and proposals to Canadian six digit postal codes, we find that the majority of postal codes had no insolvency in 2005-2010. We set counts of both bankruptcies and proposals in these regions to zero. After matching counts of insolvencies to Census data, financial literacy data, and other data described above, we have around 670,000 postal codes for six years (2005-2010) remaining in our dataset. Summary statistics for these data are presented in Table 2-1.

Some observations in our dataset seem to be data entry errors or outliers. While this problem is not widespread and it does not affect our estimates in any significant way, we decide to omit certain outliers from further analysis. In particular, we delete all observations with more than 20 consumer bankruptcies or proposals per year. Regions with more than 20 bankruptcies or proposals of any type seem to be data entry errors since there are only 15 households in a postal code on average. The omission of these outliers resulted in 393 observations being deleted which are about 0.01 percent of the overall sample.[27] Our tax data have a few CSDs reported several times in the same year with various incomes. We exclude these regions to minimize data errors. We also delete postal codes with less than 3 households as reported by the Canada Census because these regions have mostly business addresses and a very few private residences.

2.4.2 Estimation Strategy

$$Y_{it} = \delta Inequality_{it} + \beta_1^{'} Income_{it} + \beta_2^{'} Neighborhood_{it} + Province_{it} + Year_{it} + \varepsilon_{it} \quad (3)$$

We model the relation between income inequality and personal bankruptcy using the specification outlined in equation (3). As the dependent variable, we use either count of personal bankruptcies, proposals, insolvencies (bankruptcies and proposals combined), or bankruptcies, proposal or insolvencies due to overuse of credit. We use these six dependent variable to examine whether income inequality has different effect on various types of consumer defaults. Bankruptcy and proposal offer different paths to solvency and impose dissimilar requirements on bankrupts. In addition, the usage of counts of bankruptcies and proposals due to overuse of credit as dependent variables allows us to test whether such defaults

27 We omit only one observation where a code had more than 20 bankruptcies per year. All other years for this postal code remained in the sample provided that their bankruptcy counts did not exceed 20.

are more prevalent in more unequal regions.

All these six dependent variables are counts of events per year; hence, we use a Negative Binomial regression model to estimate the effect of income inequality on these variables. While our data is a panel, some inequality measures and most of the control covariates are constant over time. Hence, we can use either a random effects model or a pooled regression with clustered standard errors. As suggested by Angrist and Pischke (2009), we opt for a pooled regression, and clustered standard errors in all regressions with these dependent variables at the postal code level.

With the counts of various bankruptcies and proposals on the left-hand side, we control for proportion of homeowners, gender composition, age distribution, number of divorces and separations, educational attainment, financial literacy, proximity to an urban center, bankruptcy stigma (which is an indicator variable equal 1 if a postal code had at least one bankruptcy in 2000-2004 and it is equal to 0 otherwise), and number of households. Table 2-1 reports summary statistics for these control variables. We also include year and province fixed effects in all regressions to account for any time specific shocks to bankruptcies and differences among provinces in terms of bankruptcy rules and exemptions. Standard errors are clustered at the postal code level in these regressions.

2.4.3 Results

Our results for the impact of income inequality on bankruptcies, proposal and other consumer defaults are summarized in Table 2-7. These results reveal that income inequality increases counts of proposals and insolvencies (bankruptcies and proposals), while it has no significant effect on bankruptcies alone. These coefficients may be interpreted as follows: if Census Subdivision Gini increases by one standard deviation, the average total insolvencies are

expected to grow by around 2 percent.

Looking at the bankruptcies, proposals, and insolvencies due to overuse of credit, we can conclude that income inequality is associated with higher numbers of all three types of default with most coefficients statistically different from zero. A comparison of total bankruptcies to overuse of credit (thereafter, overuse) bankruptcies shows that more unequal regions experience a higher number of consumer defaults due to overuse of credit than bankruptcies due to other reasons. Similar patterns can be observed when comparing total proposals to overuse proposals and total insolvencies to overuse insolvencies. Hence, income inequality plays a larger role for debtors who file for bankruptcy because of overuse of credit. These results support our conjecture that income inequality leads to bankruptcies through overuse of credit and excess debt accumulation.

After we control for unobservable neighborhood effects by taking the difference between CSD Gini and CD Gini coefficients, we still find positive and significant effects of inequality on total insolvencies and consumer defaults due to overuse of credit. These results demonstrate that income inequality at the local CSD level has a larger effect on bankruptcies and proposals than the inequality at the more aggregated (CD) level.

The increases in the number of consumer bankruptcies, proposals, and insolvencies due to income inequality are consistent with our hypothesis that income inequality leads to more borrowing, and, eventually, inequality may trigger additional bankruptcies. This idea is also supported by the finding that bankruptcies due to overuse of credit are more responsive to income inequality than total bankruptcies. While bankruptcies may occur due to various shocks, defaults involving overuse of credit are more likely in unequal neighborhoods.

2.5 Conclusion

Several recent studies have argued that excessive debt accumulation is the main mechanism through which income inequality increases personal bankruptcies. The peaks in income inequality before the Great Depression of 1929 and the Great Recession of 2008 are perceived to cause both crises. While this reasoning might sound plausible, there is very little individual level empirical research linking inequality, individual debt accumulation, and bankruptcy. This study attempts to fill this gap.

Our study uses a unique database of virtually all personal bankruptcies and proposals filed in Canada in 2005-2010. These data provide us with balance sheets of individual insolvents with very detailed assets, liabilities, incomes, demographic and socio-economic attributes. We are also able to identify a special group of bankrupts for whom excessive borrowing could be a problem. We define this group as those bankruptcy or proposal filers who indicated that overuse of credit led to their defaults. Finally, these unique data allow us to compute counts of bankruptcies and insolvencies for six digit Canadian postal codes. These postal codes are very tiny geographic units containing only 15 households on average.

Using these data, we show that income inequality is associated with higher total liabilities, unsecured liabilities, credit card debt, and mortgage liabilities of bankrupts. A small number of bankruptcy filers with mortgage debt, widespread accumulation of credit card liabilities and unsecured debts by bankruptcy filers, and positive effects of inequality on credit card to total liabilities ratio allow us to conclude that credit cards and unsecured credit were primary channels of debt accumulation used by bankrupts in more unequal communities. This finding is in contrast to arguments of Rajan (2010), who suggests that mortgage debt was used in the US by individuals to keep up with their wealthier peers. However, this discrepancy may be explained by more restrictive mortgage lending standards in Canada than in the US in the 2000s. Even though only 23.5 percent of insolvents

in our sample have mortgage liabilities at the time of filing, those bankrupts who have mortgages tend to have higher mortgage debts in more unequal regions as well.

Our findings suggest that debtors tend to file for bankruptcy due to overuse of credit rather than other reasons in more unequal regions. This result supports the idea that income inequality may induce bankrupts to accumulate more debt and file for bankruptcy because of overuse of credit. Total counts of bankruptcies, proposals, and insolvencies also increase in response to higher income inequality. In addition, we argue that income inequality as measured by Gini coefficients increases filings for bankruptcy and proposals due to overuse of credit.

Table 2-1: Summary Statistics of bankruptcy count data

Variable	Obs	Mean	Std. Dev.	Min	Max
Total bankruptcies (postal code)	3618274	0.102	0.481	0	20
Total proposals (postal code)	3618274	0.028	0.204	0	11
Total insolvencies (postal code)	3618274	0.130	0.572	0	20
Overuse bankruptcies (postal code)	3618274	0.059	0.327	0	17
Overuse proposals (postal code)	3618274	0.019	0.164	0	10
Overuse insolvencies (postal code)	3618274	0.078	0.395	0	18
Gini (CSD)	2391446	0.448	0.041	0.143	0.809
Gini (CD)	3618274	0.458	0.032	0.376	0.557
Gini CSD-CD	2391446	-0.010	0.025	-0.310	0.347
Neighborhood effect (postal codes)	3618274	0.220	0.414	0	1
Postal code population	3618274	14.461	17.254	3	897
Average income (DA)	3618274	37080	19674	9108	601418
Change in income (Province)	3618274	4.924	2.907	-6.738	25.40
Homeownership (DA)	3618274	0.732	0.244	0	1
Males (DA)	3618274	0.487	0.031	0.219	0.795
Age 20-39 (DA)	3618274	0.263	0.082	0	0.875
Age 40-64 (DA)	3618274	0.359	0.059	0.029	0.607
Age over 65 (DA)	3618274	0.144	0.089	0	0.952
Divorced (DA)	3618274	0.080	0.035	0	0.317
Separated (DA)	3618274	0.031	0.018	0	0.149
Widowed (DA)	3618274	0.063	0.046	0	0.564
High school (DA)	3618274	0.236	0.076	0	0.595
Apprenticeship (DA)	3618274	0.115	0.065	0	0.5
College (DA)	3618274	0.189	0.070	0	0.581
University (DA)	3618274	0.179	0.105	0	0.786
Graduate (DA)	3618274	0.078	0.079	0	0.744
Numerical literacy (DA)	3618274	269.3	13.638	213.8	323
Proximity to urban center (CSD)	3618274	1.721	1.363	1	8

Table 2-2: Summary statistics of individual bankruptcy data

Variable	Obs	Mean	Std. Dev.	Min	Max
Overuse of credit	492397	0.602	0.489	0	1
Total liabilities	492397	93929	119237	2	1000000
Unsecured liabilities	492397	49785	57967	0	1000000
Unsecured to total	492397	0.788	0.310	0	1
Credit cards total	492397	18321	24564	0	856694
Cards to total	492397	0.328	0.317	0	1
Unsecured to assets	487385	634	6896	0	1000000
Cards to assets	487385	228	2622	0	750000
Gini (CSD)	268189	0.449	0.043	0.143	0.751
Gini (CD)	492397	0.459	0.034	0.3759	0.557
Gini CSD-CD	268189	-0.010	0.024	-0.3099	0.283
Age	492397	43.4	13.0	18	90
Car	492397	0.649	0.477	0	1
Mortgage	492397	0.235	0.424	0	1
Self-employment	492397	0.062	0.242	0	1
Numerical literacy	492397	265.5	12.774	213.8	323.1
Bankruptcy stigma	492397	0.627	0.484	0	1
Household size	492397	2.119	1.344	1	12
Average DA income	492397	32164.1	11694	9273	601418
Change in income	492397	4.419	2.754	-6.74	25.40
Divorce	492397	0.123	0.329	0	1
Total assets	492397	54755.1	103093	0	2265000
Prior defaults	492397	0.184	0.388	0	1

Table 2-3: Reasons for bankruptcy and types of debt

Reason for bankruptcy	Proportion of bankrupts	Type of debt	Proportion of bankrupts	Average value of debt
Marital Breakdown	0.160	mortgage	0.235	166527.4
Unemployment	0.260	secured	0.470	93775.8
Insufficient Income	0.348	unsecured	1.000	49785.6
Business Failure	0.095	credit cards	0.894	20498.4
Health Concerns	0.187	bank credit cards	0.770	16284.8
Accidents / Emergencies	0.023	other credit cards	0.634	9110.3
Overuse of Credit	0.602			
Student Loans	0.007			
Gambling	0.022			
Tax Liabilities	0.043			
Loans to Friends	0.013			
Bad / Poor Investments	0.017			
Garnishee	0.014			
Legal Action	0.014			
Moving / Relocation	0.018			
Substance Abuse	0.017			
Supporting Relatives	0.054			

Table 2-4: Impact of income inequality on debt accumulation by bankrupts (all types) in 2005-2010

We examine the hypothesis that income inequality affects levels and ratios of various liabilities of bankruptcy filers. We conjecture that filers in more unequal communities have more pressure to borrow to compensate for their lack of income and catch up with the consumption of wealthier neighbors. Thus bankrupts in more unequal localities are hypothesized to have more debt compared to filers in more egalitarian regions.

Inequality measures	Total liabilities	Unsecured liabilities	Unsecured to total	Credit card liabilities	Cards to total	Mortgage debt
Gini (CSD)	88,827***	78,189***	0.517***	75,886***	0.950***	170,989***
	(3,393)	(3,197)	(0.0176)	(1,558)	(0.0198)	(10,423)
Gini (CD)	127,111***	103,481***	0.855***	97,821***	1.256***	267,024***
	(3,661)	(3,424)	(0.0186)	(1,670)	(0.0199)	(10,343)
Gini CSD-CD	39,471***	42,548***	0.241***	54,583***	0.673***	38,936***
	(6,244)	(5,697)	(0.0320)	(2,881)	(0.0377)	(12,280)

Notes: This Table summarizes the full results of Tobit regressions of various types of liabilities on inequality measures and controls as shown in equation (1) and text. Each cell reflects one regression and only reports the estimated coefficient on the income inequality measure indicated in the first column and liability type or ratio indicated in the first row as the dependent variable. Results are reported as raw coefficients, standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table 2-5: Impact of income inequality on debt accumulation by proposal and bankruptcy filers in 2005-2010

We examine the hypothesis that income inequality affects levels and ratios of various liabilities of bankruptcy filers. We conjecture that filers in more unequal communities have more pressure to borrow to compensate for their lack of income and catch up with the consumption of wealthier neighbors. Thus bankrupts in more unequal localities are hypothesized to have more debt compared to filers in more egalitarian regions.

Inequality measures	Total liabilities	Unsecured liabilities	Unsecured to total	Credit card liabilities	Cards to total	Mortgage debt
			Proposal Filers			
Gini (CSD)	55,593***	7,842*	0.379***	32,302***	0.750***	204,425***
	(5,950)	(4,290)	(0.0293)	(2,144)	(0.0359)	(22,270)
Gini (CD)	77,164***	-5,836	0.559***	33,930***	0.922***	274,876***
	(6,234)	(4,326)	(0.0278)	(2,445)	(0.0329)	(18,113)
Gini CSD-CD	36,563***	30,737***	0.335***	42,772***	0.792***	53,577**
	(12,293)	(9,299)	(0.0575)	(4,388)	(0.0716)	(24,444)
			Bankruptcy Filers			
Gini (CSD)	117,372***	112,567***	0.590***	89,378***	0.935***	167,490***
	(4,060)	(3,926)	(0.0203)	(1,897)	(0.0221)	(11,368)
Gini (CD)	178,580***	165,025***	1.027***	121,594***	1.244***	291,526***
	(4,441)	(4,291)	(0.0213)	(2,050)	(0.0224)	(12,381)
Gini CSD-CD	52,629***	55,707***	0.227***	57,894***	0.602***	37,470***
	(6,957)	(6,617)	(0.0361)	(3,278)	(0.0400)	(13,031)

Notes: This Table summarizes the full results of tobit regressions of various types of liabilities on inequality measures and controls as shown in equation (1) and text. Each cell reflects one regression and only reports the estimated coefficient on the income inequality measure indicated in the first column and liability type or ratio indicated in the first row as the dependent variable. Results are reported as raw coefficients, standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table 2-6: Impact of income inequality on the probability to file due to overuse of credit

We examine the hypothesis that income inequality increases probability to file for bankruptcy due to overuse of credit. We test whether filers in more unequal communities are more likely to file for bankruptcy because of overuse of credit rather than other reasons outlined in Table 2-3.

Inequality measures	All bankrupts	Bankruptcy filers	Proposal filers
Gini (CSD)	0.324***	0.157***	0.465***
	(0.0301)	(0.0331)	(0.0554)
Gini (CD)	0.574***	0.356***	0.598***
	(0.0320)	(0.0357)	(0.0511)
Gini CSD-CD	0.334***	0.176***	0.699***
	(0.0571)	(0.0614)	(0.113)

Notes: This Table summarizes full results of Logit regressions of the overuse of credit indicator variable on inequality measures and controls as shown in equation (1) and text. Each cell reflects one regression and only reports the estimated coefficient on the income inequality measure indicated in the first column. Results are reported as marginal effects, standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table 2-7: Effect of income inequality on the number of bankruptcies and proposals in Canada in 2005-2010

We test the hypothesis that higher income inequality increases instances of consumer bankruptcies. This hypothesis implies a positive effect of Gini measures on counts of bankruptcies and proposals and, especially, bankruptcies and proposals due to overuse of credit

Inequality measures	Total bankruptcies	Total proposals	Total insolvencies	Overuse bankruptcies	Overuse proposals	Overuse insolvencies
Gini (CSD)	0.232	13.554***	1.937***	1.518**	40.556***	5.388***
Gini (CD)	-0.614	26.842***	1.397***	0.436	70.32***	4.89***
Gini CSD-CD	0.834	1.444	1.37**	2.189***	8.417***	3.674***

Notes: This Table summarizes the full results of Negative Binomial regressions of personal defaults on inequality measures and controls as shown in equation (1) and text. Each cell reflects one regression and only reports the estimated coefficient on the income inequality measure indicated in the first column and default type shown in the first row as the dependent variable. Results are reported as the percent change in the dependent variable due to a standard deviation change in the inequality measure. Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Bibliography

Angrist, J., and J. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton: Princeton University Press.

Bertrand, M., and A. Morse (2013). "Trickle-Down Consumption," National Bureau of Economic Research Working Paper No. 18883.

Bordo, M., and C. Meissner (2012). "Does Inequality Lead to a Financial Crisis?" National Bureau of Economic Research Working Paper No. 17896.

Cowen, T. (2011). "Keeping Envy Local," New York Times Opinion Pages.

Fay, S., Hurst, E., and M. White (2002). "The Household Bankruptcy Decision," *The American Economic Review, 92*(3), 706-718.

Frank, R. (2011). *The Darwin Economy: Liberty, Competition, and the Common Good.* Princeton: Princeton University Press.

Frank, R., Levine, A., and O. Dijk (2010). "Expenditure Cascades," Social Science Research Network Working Paper No. 1690612.

Galbraith, J. (2012). *Inequality and Instability: A Study of the World Economy Just Before the Great Crisis.* New York: Oxford University Press.

Georgarakos, D., Haliassos, M., and G. Pasini (2013). "Household Debt and Social Interactions," Sustainable Architecture for Finance in Europe Working Paper No. 1.

Grinblatt, M., Keloharju, M., and S. Ikaheimo (2008). "Social Influence and Consumption: Evidence from the Automobile Purchases of Neighbors," *The Review of Economics and Statistics, 90*(4), 735-753.

Howatson-Leo, L., Earl, L., Puderer, H., Cunningham, R., Bollman, R., and W. Rea (1996). "A New Approach to Non-CMA/CA Areas," Statistics Canada, Agriculture and Rural Division, Working Paper No. 31.

Kumhof, M., and R. Ranciere (2011). "Inequality, Leverage and Crises," IMF working Paper 10/268.

Livshits, I., MacGee, J., and M. Tertilt (2010). "Accounting for the Rise in

Consumer Bankruptcies," *American Economic Journal: Macroeconomics, 2*(2), 165-193.

Lusardi, A. (2012). "Numeracy, Financial Literacy, and Financial Decision-making," National Bureau of Economic Research Working Paper No. 17821.

Luttmer, E. (2005). "Neighbors as Negatives: Relative Earnings and Well-Being," *The Quarterly Journal of Economics, 120*(3), 963-1002.

Murray, S. (2011). "Financial Literacy: A Conceptual Review," Research paper prepared for the Task Force on Financial Literacy. Canadian Council on Learning.

Piketty, T., and E. Saez (2003). "Income Inequality in the United States," *The Quarterly Journal of Economics 118*(1), 1-39.

Rajan, R. (2010). *Fault Lines: How Hidden Fractures Still Threaten the World Economy.* Princeton: Princeton University Press.

Saez, E. (2008). "Striking It Richer: The Evolution of Top Incomes in the United States," University of California, Berkeley Research Note.

Scholnick, B. (2012). "Bankruptcy Spillovers: Distance, Public Disclosure and Opaque Information," University of Alberta School of Business Working Paper.

White, M. (2011). "Corporate and Personal Bankruptcy Law," *Annual Review of Law and Social Science, 7*, 139-164.

Chapter 3: Distance as a Bankruptcy Filing Cost[28]

3.1 Introduction

A large literature in Finance has shown that geographic and distance related costs can impact the interaction between parties to a financial contract, in a variety of contexts. Geographic distance has been shown to have an impact on such diverse financial contracts as bank lending (Hauswald and Marquez, 2006) as well as corporate acquisition, commercial real estate trading, and securities investment (Hau, 2001). Other studies have found that banks lend more but at higher costs to borrowers located further away from their physical branches (Degryse and Ongena, 2005), that investors prefer shares of local firms (Coval and Moskowitz, 1999 and 2001), that proximate real estate is traded to decrease asymmetric information (Garmaise and Moskowitz, 2004), and that Initial Public Offerings (IPO) of firms provide different signals to nearby and more distant acquirers (Ragozzino and Reuer, 2011).

The aim of this paper is to propose and test a new hypothesis – that distance and geography matter in the context of personal bankruptcy. A central hypothesis in the bankruptcy literature (as summarized by White, 2007 and many others) is that the various *costs* of bankruptcy filing (e.g. social costs and economic costs) impact filing decisions. The impact on bankruptcy choices from social costs such as stigma have been discussed by Gross and Souleles (2002), Fay, Hurst and White (2002) and Scholnick (2012). The impact on bankruptcy choices from economic costs related to filing fees have been discussed by Gross, Notowidigdo, and Wang (2013). This paper is the first in the literature to examine a new kind of cost on the bankruptcy filing process - the costs imposed by geography and distance.

Specifically, we propose the new hypothesis that geography matters

28 This Chapter is a joint work with Dr. Barry Scholnick (School of Business, University of Alberta).

because bankruptcy filing is an interactive process that requires distressed debtors to interact with bankruptcy professionals (e.g. lawyers, trustees etc.) in order to file. A distressed debtor who lives in an area with a large number of bankruptcy professionals within close proximity will thus face lower geographically imposed filing costs compared to a distressed debtor who lives in areas with few proximate bankruptcy professionals.

Our study shows for the first time, that geographic distance between the bankruptcy filer and the trustee impacts the personal bankruptcy decision. The new hypothesis proposed in this paper follows a central element of the bankruptcy literature, which states that an individual will choose to file for bankruptcy if the benefits from bankruptcy exceed the costs imposed by the bankruptcy. This paper will examine a specific element of this cost-benefit trade-off by testing the hypothesis that an increase in the costs of bankruptcy filing imposed by geographic distance will increase the financial benefits required to make a bankruptcy filing worthwhile to the individual. Our specific testable hypothesis is that if the geographic costs of filing are higher (e.g. if the individual lives in an area that is not well served by local bankruptcy trustees), then that individual will require higher financial benefits from bankruptcy (FBB), in order to overcome these geographic costs of filing, and thus to be persuaded to file.

The concept of financial benefits of bankruptcy (FBB) is taken from Fay, Hurst and White (2002), and captures the net effect of the amount of unsecured debt that is discharged in bankruptcy (which is a benefit to the filer) *minus* the liquidated nonexempt assets which are used to repay creditors (which the filer loses in bankruptcy). In other words, we hypothesize that individuals who live in areas with high geographic costs of filing (i.e. areas that are underserved by bankruptcy trustees) will require larger amounts of financial benefits from bankruptcy (e.g. larger amounts of unsecured credit card debt written off in bankruptcy, and/ or smaller amounts of secured assets forgone in bankruptcy) in order to overcome the higher geographic costs and be persuaded to file.

In order to examine the impact of geography and distance on bankruptcy filings, we use a unique database of essentially every electronically filed bankruptcy in Canada provided to us by the Canadian bankruptcy regulator, the Office of the Superintendent of Bankruptcy (OSB). Our data consist of detailed balance sheet and location information, provided by the OSB, for *every* Canadian bankruptcy that was filed electronically from 2005 to 2010. In total we observe more than 386,000 bankruptcy filings, each containing the full balance sheet submitted by the filer to the OSB at the time of the bankruptcy. The key reason that we are able to obtain such detailed and extensive data is that Canada has a single bankruptcy regulator (the OSB), and every filing in Canada has to be made to that regulator. This differs from the US, where there is no central bankruptcy regulator, and where bankruptcy filings have to be made to individual bankruptcy court districts. The unique advantage of our Canadian data is that we are able to measure both key elements of our central hypothesis: (1) balance sheet data from individual bankruptcy filings, used to calculate the financial benefits of bankruptcy and (2) location data of filers and trustees, used to calculate geographic costs.

Our study exploits an important element of Canadian bankruptcy law, which is that *every* bankruptcy has to be filed by a bankruptcy professional specifically licensed by the OSB, which in Canada is called a bankruptcy trustee. The Canadian system of requiring bankruptcy trustees to be licensed is very different from the US bankruptcy system, where essentially any professional can be used to make a bankruptcy filing, or a filing can be made by the individual without any professional help. Because of this Canadian trustee licensing system, our data allow us to identify the specific location of the full universe of *every* licensed bankruptcy trustee operating in Canada. Combining these data with the information on the exact location of each E-filer, we are able to generate two different measures of the costs associated with the geographic distance between trustees and filers. First, we measure the distance between the filer and the

geographically closest trustee. Second, we measure how many trustees are located within a 10 km radius of each filer. These measures provide very significant exogenous variation, across the different geographic areas in Canada where the individual bankruptcy filers reside. As an example, our data show that the number of bankruptcy trustees who are located within a 10 km radius of Canadian bankruptcy filers ranges from a minimum of zero trustees, to a maximum of 94 trustees.

Our OSB database not only contains the location of filers and trustees but also it records the full balance sheet of every filer, including all assets and all liabilities at the date of bankruptcy. We use these data to calculate financial benefits of bankruptcy for every electronically filed bankruptcy in Canada in 2005-2010. The concept of the financial benefits of bankruptcy (benefits from debt discharged minus assets lost) was initially developed by Fay, Hurst and White (2002). These authors developed and tested the hypothesis that increased financial benefits (their independent variable) would predict larger number of bankruptcy filers (their dependent variable). In this paper, however, we test the very different hypothesis that increased geographic costs of filing (our exogenous independent variable) will impact the financial benefits of filers (our endogenous dependent variable).

In this paper we argue that the level of financial benefits accruing to the filer when the individual chooses to file, is endogenous. This endogeneity flows from the fact that the choice and timing of the decision to file for bankruptcy is that of the individual. Our main exogenous independent variable is the geographic costs of filing, as measured by the geographic context (e.g. supply of trustee services in the proximate geographic area) of the individual filer.

We use a variety of different specifications to relate our distance based measures (distance to nearest trustee and number of trustees within 10 km) to the geographic based transactions cost of interacting with a trustee. The simplest specification makes the assumption that these distance-based measures are

linearly related to geographic costs. That is the closer the nearest trustee and the more trustees within 10 km, the lower the geographic costs of filing.

Our results are different for bankruptcy filers in rural and urban areas. We find in rural areas that the more trustees within a 10 km radius of the individual (implying lower geographic costs of filing), the lower the average financial benefits from bankruptcy (debt discharged minus assets forgone) required to persuade the individual to file for bankruptcy. Similarly, we find that in rural areas, the larger the geographic distance between the filer and the closest trustee (implying larger geographic costs of filing), the larger the average financial benefits from bankruptcy (debt discharged minus assets forgone) required in order to persuade the individual to file.

While our results strongly support our hypothesis in rural areas, we do not find evidence to support our hypothesis in urban areas. We argue that the reasoning for these differences between rural and urban areas reflects the very different distribution of trustees in the different areas. In rural areas there are an average of 1.1 trustees within 10 km of each filer (standard deviation of 1.8), and the closest trustee is on average 47.6 km away from the filer (standard deviation 88 km). We argue that it is because of this sparse distribution of available trustees in rural areas, that the data support our geographic costs hypothesis. In a context where distances are large and there are very few proximate trustees, it seems plausible that distance costs will indeed impact the decisions of bankruptcy filers.

On the other hand, the provision of trustees is very different in urban areas. On average there are 18.6 trustees within a 10 km radius of each urban filer (standard deviation 21.3), and the average distance between an urban filer and the closest trustee is 3.3 km (standard deviation 4.6 km). There is thus a very dense coverage of available trustees within close geographic proximity of each filer in urban areas – which we argue explains why the data do not support our geographic costs argument in urban areas. In a context where distances are short and there are multiple proximate trustees, it is plausible that distance should not

play a significant role in the decision to file for bankruptcy.

3.2 Data

3.2.1 Bankruptcy institutional details

The source of our bankruptcy data, the Office of the Superintendent of Bankruptcy (OSB) regulates all bankruptcies in Canada, thus every bankruptcy filing has to be made to the OSB. In 2002 the OSB introduced E-filing, whereby trustees could submit filings electronically rather than in paper form. The OSB provided us with all individual level electronic filing data from Forms 79 (Statement of Affairs) and 65 (Monthly Income and Expense Statement of the Bankrupt/Debtor and the Family Unit and Information (or Amended Information) Concerning the Financial Situation of the Individual Bankrupt) for all Canadian consumer E-Filers from 2005 to 2010, but did not provide us with data from paper filings. The share of E-Filings for individual years was: 2005 - 62.2%, 2006 - 77.4%, 2007 - 97.7%, 2008 - 98.9%, 2009 - 98.6% and 2010 – 99.6%, thus the process of moving to E-Filing was essentially complete by 2007.

Every insolvency case in Canada must be filed by a Bankruptcy Trustee. Trustees are typically professionals who are certified by the OSB to process, prepare, and file bankruptcy petitions and other legal documents related to bankruptcy. Because of this trustee licensing process, our data include the complete universe of trustees in Canada, as well as their geographic location. The OSB provided us with data on the geographic location of *every* office for every trustee, thus we know the location of every office in multiple office trustee firms.

We are able to link OSB E-Filing data (in particular the postal code of the filer) with postal code data of individual trustees. This matching allows us to measure the geographic distance between the postal code of the filer and the

postal code of the nearest trustee. Canadian six-digit postal codes are extremely small geographic units containing just 15 households on average, and can be smaller than city blocks in urban areas. Postal codes may encompass larger regions in rural areas. These postal codes cover the whole territory of Canada, and as of 2006 there were 832,163 active postal codes. In this paper, postal codes are used to provide geographic location, and also for matching purposes because a large amount of demographic data (e.g. Census data) can be matched to postal code level data (more details below).

Under the Canadian insolvency system, a distressed debtor can choose to file for bankruptcy or file a proposal. A bankruptcy (which is somewhat similar to a Chapter 7 bankruptcy in the US) constitutes the debtor having unsecured debt discharged, but possibly having to liquidate assets (e.g. a house) in order to pay secured creditors. A proposal (which is somewhat similar to Chapter 13 bankruptcy in the US) constitutes the debtor and creditors negotiating a new payment stream (usually smaller amounts over a longer period), but does not involve the liquidation of assets. Because the main dependent variable in this study is the financial benefits of bankruptcy (i.e. unsecured debt discharged minus non-exempt assets liquidated as defined by Fay, Hurst and White, 2002) the focus of this paper is only on bankruptcies and not on proposals. This is because each proposal is uniquely negotiated between the distressed debtor and creditors, and concerns a new future pattern for the required stream of payments. Thus, while the concept of the financial benefits of bankruptcy (Fay, Hurst, White, 2002) is central to discussions of bankruptcy, it is not an appropriate categorization for discussion of proposals.

The bankruptcy filings provided by the OSB are primary filings only and they exclude corollary insolvencies. In most cases, corollary bankruptcies are filings of spouses of bankrupt individuals. Joint bankruptcy filings, i.e. when both spouses file for bankruptcy, are recorded twice in the original OSB data, once for the wife and once for the husband. However, these two files are identical as this is

actually one joint filing. In order to avoid double-counting and repeated usage of the same bankruptcy file, the OSB does not provide us with corollary insolvencies.

An important element of our study is the licensing process used to regulate bankruptcy trustees in Canada. In order to acquire a bankruptcy trustee license, a license seeker needs to undertake a multiple-year education program, pass exams, and complete a number of hours of in-service training. As is common in many regulated professions, the OSB regulates both the licensing of trustees as well as the maximum price a trustee can charge for a bankruptcy filing.

3.2.2 Balance Sheet Data and the Financial Benefits of Bankruptcy (Dependent Variable)

While some studies in the bankruptcy literature have attempted to utilize the balance sheets of bankrupts, such studies have been severely limited by data constraints. Fay, Hurst and White (2002) measure financial benefits from individual bankruptcy (FBB) filers using data from the Panel Study of Income Dynamics (PSID) survey, which captures various elements of individual balance sheets. Their data, however, consist of only 254 bankruptcy balance sheets because only a very small fraction of individuals surveyed for the PSID filed for bankruptcy. Hankins, Hoekstra and Skiba (2011) in their study of the impact of lotteries on bankruptcies hand collect about 250 bankruptcy balance sheets. While they use these balance sheets to examine issues such as total secured and unsecured debt, they do not calculate the financial benefits of bankruptcy (FBB) for each individual. Similarly, Gross, Notowidigdo and Wang (2013) in their study on liquidity constraints and bankruptcy are able to hand collect data on the balance sheets of approximately 6500 filers. They also do not examine FBB, but limit their analysis to balance sheet data such as total liabilities and total income, etc.

Our study makes use of balance sheet data from more than 386,000 bankruptcy balance sheets. Our study is thus the first to examine the concept of FBB, as developed by Fay, Hurst and White (2002), using detailed balance sheet data from many hundreds of thousands of bankruptcy filers, which are essentially the population of filers during the study period.

We define individual debtor's net financial benefits of filing for bankruptcy (FBB) in the same way as Fay, Hurst, and White (2002):

$$FBB_{it} = max\left[D_{it} - max\left[W_{it} - E_{it}, 0\right], 0\right] \qquad (1)$$

where D_{it} is unsecured liabilities of filers eliminated in bankruptcy, W_{it} is total wealth of bankruptcy filers minus all secured debts, and E_{it} represents bankruptcy exemptions available to filers in a particular year and province. Equation (1) captures the central idea of bankruptcy which discharges unsecured liabilities of filers in exchange for non-exempt filer's assets. If assets minus secured debts and exemptions are less or equal to zero, then there is nothing to distribute among unsecured creditors and all bankrupt's unsecured debts are discharged. The central advantage of our data is that we can use our detailed balance sheet information from each bankruptcy filing to calculate a dollar value of FBB for each bankruptcy filer. As can be seen from Table 3-1, the average value of FBB in our sample is $48,532. Total liabilities on average are equal to $86,204 and mean total assets are $43,741.

Our measure of unsecured liabilities (D) is the direct measure of total unsecured debt on the bankruptcy filer's balance sheet (including credit card and all other forms of unsecured debt). Our measure of wealth (W) is also taken directly from the filer's balance sheet and is calculated as total assets minus total secured debt. This is the amount of positive equity that will be liquidated in bankruptcy. Our measure of exemptions (E) is more complex because it entails various province-specific exemptions allowed to bankruptcy filers in different

provinces. All bankruptcy exemptions allowed in Canada during our study period are described in Table 3-2. Most of the exemptions are related to particular assets such as principal residence, car, furniture, or pensions accounts. A key advantage of our data is that we can observe all of these different categories of assets in the filer's balance sheet, and can thus calculate the exemptions for each individual. These bankruptcy exemptions apply to either equity in these assets (houses and cars) or assets value determined by the trustee (everything else). For instance, the province of Alberta allows the bankrupt to keep up to $40,000 of housing equity in bankruptcy. So, if house value net of mortgages secured by this house is lower than $40,000, the filer retains the whole house. If house equity is larger than the maximum house exemption, then this maximum value of the exemption is kept by the bankrupt. Hence, the value of house exemption in equation (1) is the lower of $40,000 or house equity of the debtor.

Exemptions on furniture and personal effects are calculated by comparing the value of these assets with the maximum allowed exemptions on them. If asset value is higher than the maximum exemption, then this maximum amount is used in equation (1). If the opposite is true, then asset value is used in equation (1). Some provinces bundle furniture and personal effects into one category and impose a cap on their joint value exempt in bankruptcy. We also include these features in our calculations.

Registered pension accounts are exempt from seizure by creditors in bankruptcy in all provinces but Alberta before October 1, 2009, which we take into account in our calculations. Canadian provinces have special exemption rules for farmers and non-farmers. Most of these rules include exemptions for assets which we cannot observe such as livestock, seeds, etc. Hence, we do not use these specific exemptions in our formula. However, Alberta allows up to 160 acres of land to be exempt in bankruptcy if the bankrupt is a farmer. The size of the land is unobserved to us, hence we exempt all land in bankruptcy for Albertan filers in rural areas, which is a proxy for farmers. For more details on differences in

bankruptcy exemptions across Canadian provinces, see Table 3-2.

In 2006 the government of Alberta distributed $400 to every resident of this province. In order to integrate the 2006 Alberta income shock into the FBB formula, we utilize the very specific ruling made by the OSB as to how these payments should be dealt with in bankruptcy. The OSB ruled very explicitly that the Alberta 2006 transfer payments were exempt from seizure in bankruptcy in 2006, thus we account for this exemption in our formula.

3.2.3. Geographic Costs (Independent Variable)

Our key independent variable is the geographic cost of filing for bankruptcy, which we argue is related to the distance between bankruptcy filers and trustees. A unique element of our data, which has not previously been available in the literature, is that they include the postal code of every filer and every trustee in the database. Canadian postal codes are extremely small areas, containing 15 households on average, and often less than the size of a city block.[29] These Canadian postal codes are thus orders of magnitude smaller than US ZIP Codes. We use the center point of the postal code as our basis of the geographic location of individual addresses in the postal code. The distance between the debtor and the trustee is calculated using the Haversine formula.[30] We use the Geographic Information System (GIS) software ARCGIS to estimate these distances.

Another unique element of our data is that we are able to identify the postal code of the *complete universe* of all bankruptcies filed electronically in Canada. To do this we exploit the fact that the only agents able to file bankruptcy petitions with the OSB are bankruptcy trustees that are specifically licensed by the

29 Rural postal codes may contain more population and cover larger areas. However, even in rural areas, Canadian postal codes are much smaller than five digit US ZIP codes.

30 For more details, see http://mathworld.wolfram.com/SphericalTrigonometry.html

OSB. This system is very different from that in the United States, where it is not necessary for a professional to file a bankruptcy with a court, and indeed a distressed debtor is able to make a bankruptcy filing without any professional assistance (called *pro se* filing).

Our first measure of the geographic cost of filing is the geographic distance between the bankruptcy filer and the licensed trustee that is geographically closest to that filer. This geographic distance is exogenous because it reflects the geographic area that the individual filer is located, rather than any choices specifically made by the filer. We exploit the large variation of this distance across all filers located in different areas of Canada. This exogenous variation is reflected in the data which show that across 386,000 observations the mean distance between the filer and closest trustee is 17.5 km with a standard deviation of 54.4 km.

Our second measure of geographic filing costs is the number of trustees within a 10 km radius of each individual filer. This measure allows us to provide data on the geographically proximate supply of trustees for each of the 386,000 individual bankruptcy filers in our database. We argue that this measure is also exogenous because it only reflects the geographically proximate supply of trustees, rather than any individual choices made by the bankruptcy filer. Once again we exploit the large variation in the count of trustees within the 10 km radius of the individual filer, with a mean of 13 trustees, and a standard deviation of 19 trustees across all the filers in our database.

We argue that there are two possible channels by which an increase in the number of trustees within a 10 km radius can reduce the costs of filing. Firstly, the greater the supply of proximate trustees within the radius, the lower the geographic costs of the individual accessing one of those trustees. Secondly, the larger supply of trustees may reduce bankruptcy filing fees. Recall, that in our discussion above we noted that the OSB regulates the maximum price that the trustee can charge for filing a bankruptcy – but it does not regulate the minimum

price. Thus it is possible that some price competition could exist among trustees. Based on our discussion with trustees, it would appear that because of the regulatory environment (i.e. a regulated price ceiling), little if any price competition actually does exist, and that most trustees charge the regulated maximum rate. However, it is possible that price competition could occur if trustees charge lower than the regulated rate. In the context of our study, we argue that an increase in the supply of proximate trustees within a radius would be one factor that caused trustees to charge less than the regulated rate. In this case more trustees would lead to lower filing fees. Thus the costs of filing would be lower, whether an increased supply of proximate trustees within a radius lowered (1) the geographic transactions costs or (2) the price charged by the trustee, or (3) both. We can thus test our main hypothesis that lowering the costs of filing (by increasing the supply of proximate trustees) would lead to a lower FBB required in order to persuade the individual to file.

3.2.4 Control Variables

We use a variety of control variables measured at both filer's individual level and filer's geographic area level. An important advantage of our OSB data is that they contain some important demographic variables about individual filers. The OSB provided us with a large amount of individual level demographic and economic data including: filer’s age, car ownership, self-employment status, household size, marital status, and prior insolvencies. All of these individual level data are included in all our regressions.

In addition to these individual level control variables, we also include a large variety of control variables measuring the characteristics of the geographic area of the filer. Because we know the exact postal code of each individual filer, we are able to match the postal code with Canadian census data. We match individual level and postal code level data from the OSB with Census and other

data using the Postal Code Conversion File (PCCF) developed by Statistics Canada and Canada Post. Matching of postal code data to Dissemination Area or Census Subdivision Data is common for papers involving Canadian data.

In order to capture neighborhood income, we use 2006 Census data on average personal incomes at the level of Census Dissemination Areas (DAs). DAs have between 400 and 700 inhabitants, and 500 persons on average. In addition to average incomes, we also control for shocks to income using changes to annual personal disposable income at the provincial level. Data on personal disposable income are from Statistics Canada. Both local levels of income and income shocks are potentially important factors in bankruptcies because they capture local effects on bankruptcy filings, for example, plant closures.

We also control for a neighborhood's level of financial literacy using data provided to us by Scott Murray (2011). Financial literacy and, in particular, numerical literacy such as ability to perform calculations required in consumer finance and comparisons of various financial products has been argued to matter in financial decisions and debt management (see e.g. Lusardi, 2012). We use numerical literacy measures estimated from the 2003 International Adult Literacy and Skills Survey (IALSS) and the 2006 Census. IALSS collected data on actual numerical literacy scores of a sample of Canadians as well as their demographic attributes. These data are used to estimate the relationship between numerical literacy and demographic variables. After that, coefficients on demographic characteristics from this relationship and 2006 Census demographics for all regions in Canada are used to calculate imputed numerical literacy scores for every DA in the country.

Fay, Hurst and White (2002), Gross and Souleles (2002) and Scholnick (2012) argue that past bankruptcies in an individual’s proximate geographic area could impact bankruptcies through stigma or information effects. These authors suggest that more neighborhood bankruptcies in the past could increase the probability of individual’s filing for bankruptcies, either through the lowering of

bankruptcy stigma or through the spread of information about the bankruptcy process. We control for past neighborhood bankruptcies using measures of all past bankruptcies in every postal code in our sample. To this end, we take counts of bankruptcies in 2000-2004, i.e. bankruptcies before the start of our sample, and create an indicator variable equal to 1 if a postal code experienced at least one bankruptcy during this period (2000-2004) and equal to 0 otherwise. Note that while our balance sheet data from the OSB are only available for electronic filers (and not paper filers), the OSB was able to provide us with full count data of bankruptcies per postal code of all filers including both electronic as well as paper filers.

3.2.5 Rural-Urban Index

In order to distinguish between urban and rural areas in Canada, we employ the widely used Statistics Canada definition of Census Metropolitan Areas (CMAs). We denote all CMAs as urban areas, and define all non CMA areas as rural regions. Statistics Canada designates Census Metropolitan Areas as geographic centers, primarily cities, with more than 100,000 inhabitants. As an additional measure of rural or urban geographic character, we use the concept of Metropolitan Influence Zone (MIZ) developed by Howatson-Leo et al. (1996) and used by Statistics Canada. This concept separates geographic areas based on the flows of their residents commuting to and from a central city. This is what the "influence" of a zone constitutes. There are eight categories of MIZ areas: Census Metropolitan Area, Tracted Census Agglomeration, Non-tracted Census Agglomeration, Strongly Influenced Zone, Moderately Influenced Zone, Weakly Influenced Zone, No Influenced Zone, and Territories.[31] Tracted Census

31 Statistics Canada defines Tracted Census Agglomerations as geographical spaces subdivided into Census Tracts, hence the name. Non-tracted Census Agglomeration are smaller units not divided into Census Tracts.

agglomerations are smaller units, i.e. towns, with between 50,000 and 100,000 total population. Non-tracted Census agglomerations are even smaller towns with more than 10,000 but less than 50,000 residents. A strongly influenced zone has at least 30% of its residents commuting to and employed in a metropolitan center. Areas are defined as moderately influenced if between 5% and 30% of their residents commute to the central city. Weakly influenced are those geographic spaces where the commuter flow is larger than zero but smaller than 5%. No influenced zones have no city commuters. Territories are areas in Yukon, the Northwest Territories, and Nunavut. We identify rural/urban postal codes using these eight categories of geographic areas. As MIZ scale changes from 1 (Census Metropolitan Areas) to 8 (Territories), regions get more rural and less urban. Table 3-3 summarizes main variables for urban and rural areas.

3.3 Estimation Strategy

We employ the following econometric specification:

$$FBB_{it} = \delta Geography_{it} + \beta_1' Income_{it} + \beta_2' Indcontrols_{it} + \beta_3' Neigborcontrols_{it} + Province_i + Year_t + \varepsilon_{it} \quad (2)$$

Our endogenous dependent variable FBB is the Financial Benefits of Bankruptcy as developed by Fay, Hurst and White (2002), and it is measured from individual balance sheets. The specific definition of FBB is provided in equation (1). Broadly speaking, FBB gauges the amount of unsecured debt discharged in bankruptcy minus the non-exempt assets lost in bankruptcy. Figure 3-1 demonstrates that FBB has a log-normal distribution. Therefore, we use the logarithm of FBB in our specifications.

The major independent variable of interest is $Geography_{it}$. This variable is

captured by one of the two measures of the supply of trustees in the filer's proximate geographic area. These are (1) the geographic distance between the filer and the filer's closest trustee, and (2) the number of trustees within a 10 km radius of the individual filer. We argue that both of these measures are exogenous in that they reflect the geographic area around the filer, rather than the specific choices made by the filer.

We control for various measures of income in the geographic area surrounding the filer to reflect idiosyncratic shocks in that geographic area. The set of variables $Income_{it}$ includes average DA income and province level percentage changes in personal disposable income.

Individual Controls consist of the filer's age, number of household members, marital status, presence of a mortgage, self-employment status and prior consumer defaults. *Neighborhood controls* comprise of financial literacy variable and bankruptcy stigma/information effect (neighborhood bankruptcies during 2000-2004). We also include province and year fixed effects to control for any time specific or province specific shocks and differences between provinces in terms of bankruptcy rules. All reported standard errors are clustered at the DA level.

3.4 Results

We present our results in Tables 3-4 to 3-8. Table 3-4 summarizes our main baseline specifications. Each cell in this Table reflects a single regression. We only report the coefficient on the geographic cost term (distance to the closest trustee or number of trustees within a certain radius) on a regression on FBB. In the attached appendix we report the full regressions from this Table as well as all other Tables, including all control variables, standard errors, and other statistics.

The key finding from Table 3-4 is that the results for rural regions (defined as all localities outside census metropolitan areas) are as predicted by our

hypothesis above. In rural areas we find that an increase in the number of trustees within a 10 km radius will lead to a reduction in the FBB of individual filers. In other words, as the number of proximate trustees increases the costs of bankruptcy faced by the individual filer decline. Thus these results are consistent with the hypothesis that when geographic filing costs are lowered, individuals will be persuaded to file even though their financial benefits of filing (debt discharged minus assets lost) are lower. In terms of economic magnitudes, our findings imply that one extra trustee within the 10 km radius will decrease the average financial benefits by 1.88 percent. If calculated at the mean of rural FBB ($43,372), this implies that one extra trustee in the 10 km radius of a rural filer will decrease FBB by 0.0188*$43372=$815.39. This coefficient is highly significant at 1%.

Our results for the distance in kilometers to the filer's closest trustee are also consistent with our predictions for filers in rural areas. The coefficient is highly significant at 1%. Our results show that one extra kilometer between the filer and the closest trustee (i.e. increasing geographic costs) will increase FBB for rural filers by 0.059 percent. Measured at the mean of rural FBB this implies that one extra kilometer to the closest trustee will increase FBB by 0.00059*$43372=$25.59.

While our results for rural areas are all consistent with our hypothesis that increased geographic costs will increase the FBB required to persuade the individual to file, our results for urban areas are of the opposite sign from this prediction, although the magnitudes of the coefficients are orders of magnitude smaller.

As we describe in the introduction, we argue that these results can be explained by the very different geographic distribution of trustees in rural and urban areas. As shown in Table 3-3, the nearest trustee in rural areas is on average 47.6 km from the average filer (sd = 88.6), compared to urban areas where the average distance is only 3.3 km (sd = 4.6). Our results show that our distance as a transaction cost hypothesis is strongly supported in areas (such as rural areas)

where distances are generally large. On the other hand, our distance as a transaction cost hypothesis is not supported in areas (such as urban areas) where distances are generally short. Consider the situation of an individual at the mean distance compared to an individual at the mean plus one standard deviation distance. In urban areas mean distance is 3.3 km, while the mean plus one standard deviation is 7.9 km. We argue that it is unlikely, for example, to make a significant difference to the FBB of the filer that the closest trustee is 3.3 km rather than 7.9 km away. This is consistent with our urban results not supporting our hypothesis. On the other hand, in rural areas, mean distance is 47.6 km and mean plus standard deviation distance is 136.2 km. Our rural results imply that moving from having to travel 47 km compared to 136 km will indeed significantly increase transactions costs to such an extent that this impacts the FBB of individual filers.

In order to test the argument that the influence of geographic distance should get stronger as areas get more rural, we run tests that interact geographic distance with Metropolitan Influence Zone (MIZ) covariate. We use interaction terms to examine whether issues related to geographic distance (closest trustee and trustees within a radius) become more acute as the area that the individual lives in becomes more rural and less urban. We interact the MIZ index with one of the individual level geographic cost measures (closest trustee and trustees within a radius).

Recall that Statistics Canada has categorized each postal code in Canada in eight MIZ zones where, Census Metropolitan Areas are coded 1, and areas further and further from the Metropolitan area are coded 2 to 8, with the most remote areas being coded 8. Our strategy is to interact the MIZ code variable with each of the two geographic cost variables (trustees in 10 km and distance to the closest trustee). As is standard when using interaction terms, all specifications also include the two components of the interaction term. Furthermore, we argue that both elements of the interaction term (MIZ code as well as geographic

location of trustees) are plausibly exogenous, in that neither reflects choices made by the individual filer. The specification of this equation is as follows:

$$FBB_{it} = \delta_1 \text{Geography}_{it} + \delta_2 \text{MIZ}_{it} + \delta_3 \text{Geography}_{it} * \text{MIZ}_{it} + \beta' Controls_{it} + Province_i + Year_t + \varepsilon_{it}$$

where the *Geography* variable is either the number of trustees within 10 km or the distance to the closest trustee. *MIZ* covariate is either the original MIZ scale (ranging from 1 to 8) or the rural indicator variable (equal to 0 for Census Metropolitan Areas (MIZ=1) and equal to 1 for all other regions).

The interaction term δ_3 captures the impact of both geographic costs as well as urban/rural nature of the area on FBB. In other words, using the interaction term we can test the hypothesis that the impact of the geographic cost (trustees within a certain radius or closest distance to trustee) should be higher the more rural the area. If it is true that the impact of geographic distance get stronger as the location of the filer moved from urban to more and more rural areas, then we would expect the estimated coefficient to be the same sign as that predicted in our main hypothesis – i.e. that more trustees within a 10 km radius of the filer would reduce FBB, and increased distance to the nearest trustee would raise FBB.

Our results for the interaction terms are reported in Table 3-5. The main results in this Table demonstrate that when MIZ is interacted with geographic location of trustees, these two trustee location measures (trustees within 10 km and distance to the closest trustee) are highly significant at 1% and have the predicted sign. Thus based on these results we can confirm the hypothesis that the impact of trustee location on FBB is stronger as the location of the filer becomes more rural and less urban. The results in Table 3-5 can be compared to those of Table 3-4. Table 3-5 suggests that trustee's proximity increases financial benefits of bankruptcy as regions becomes more rural, thus supporting our explanation that geographic distances reflect costs of accessing trustees in rural regions better than

in urban localities.

In addition to interacting trustee location with MIZ we also interact the trustee location variables (the number of trustees within 10 km, and the distance to the closest trustee) with a simple rural/urban indicator, which takes the value of 0 for Census Metropolitan Areas (MIZ=1) and it takes the value of 1 for all non-Metropolitan regions (MIZ between 2 and 8). The results from these interaction terms are the same as for the full MIZ interactions – i.e. highly significant coefficients with the predicted sign for trustees within 10 km and distance to closest trustee. These results also confirm the hypothesis that the impact of the trustee location on FBB is higher in rural areas.

One possible concern with our specification is that our results may be impacted by the financial crisis of 2008. Recall that our data run from 2005 to 2010 which includes the period before and after the financial crisis. In order to examine this possibility, we rerun all our results for the two periods January 2005 to August 2008 and separately the period September 2008 to December 2010. We run these models for all our various measures of geographic distance as well as for all the models with interaction terms (see Tables 3-6, 3-7 and 3-8). Our results are very robust to changing these time periods, which indicates that the financial crisis did not have a significant impact on our main conclusions.

3.5 Conclusion

A standard framework for analyzing bankruptcy filings is that individuals will file for bankruptcy when the benefits of filing outweigh the costs. This paper is the first to examine a different cost new to the bankruptcy literature – the costs associated with the geographic distance between the bankruptcy filer and the bankruptcy trustee. Our central argument is that geography matters because a bankruptcy filing typically involves interactions between the filer and the trustee. Thus we hypothesize that a distressed debtor who is located at some distance from potential trustees will face higher costs of filing compared to a distressed debtor

who is located close to trustees.

While our paper is the first in the literature to show that geographic distance matters in the context of personal bankruptcy filings, it forms part of a much larger literature demonstrating that distance matters in a large variety of other financial contracts that involve relationships between parties located some space apart. Distance has been shown to influence financial contracts as different as bank lending (Hauswald and Marquez, 2006), investment in local firms (Coval and Moskowitz, 1999 and 2001), and the purchase of Initial Public Offerings (Ragozzino and Reuer, 2011) amongst many others.

We test our hypothesis using Canadian bankruptcy data, and we exploit several unique elements of the Canadian bankruptcy system to develop our empirical strategy. Bankruptcy in Canada is federally regulated, thus every bankruptcy filing must be made to the regulator (the Office of the Superintendent of Bankruptcy or OSB). Furthermore, only bankruptcy professionals licensed by the OSB (called bankruptcy trustees) are allowed to make a bankruptcy filing. Our data are collected by the OSB, thus we are able to observe the exact geographic location of the full universe of bankruptcy trustees in Canada, as well as the location of every bankruptcy electronic filer. We use these data to measure geographic distances, which we use as proxies for the geographic cost of filings. These are our main independent variables.

Our main dependent variable is the Financial Benefit of Bankruptcy (FBB) as developed by Fay, Hurst and White (2002). FBB is simply the amount that the individual filer gains from bankruptcy (through the discharge of unsecured debt) *minus* the amount the filer loses (from the loss of non-exempt assets). The specific hypothesis we test is that as the geographic costs of filing increase (i.e. as it becomes more costly to interact with a more distant trustee) so the individual will have to have higher levels of benefits (FBB) in order to compensate for these increased geographic costs, and thus to be persuaded to file.

Our main result, based on a regression with almost four hundred thousand

individual filings, shows that our new hypothesis is supported in rural but not urban areas. In rural areas, distances are typically large, implying that issues related to distance will impact individual choices such as those related to bankruptcy. On the other hand, in urban areas, distances are typically much more compressed, implying that issues other than distance impact these individual choices. For rural bankruptcy filers, therefore, we find that increased costs of filing, caused by greater geographic distances to trustees, will result in higher required levels of financial benefit from filing (debt discharged minus assets lost in bankruptcy) in order to compensate for these increased geographic costs.

Figure 3-1. Financial Benefits of Bankruptcy (Bankruptcy Filers, 2005-2010, n=386k)

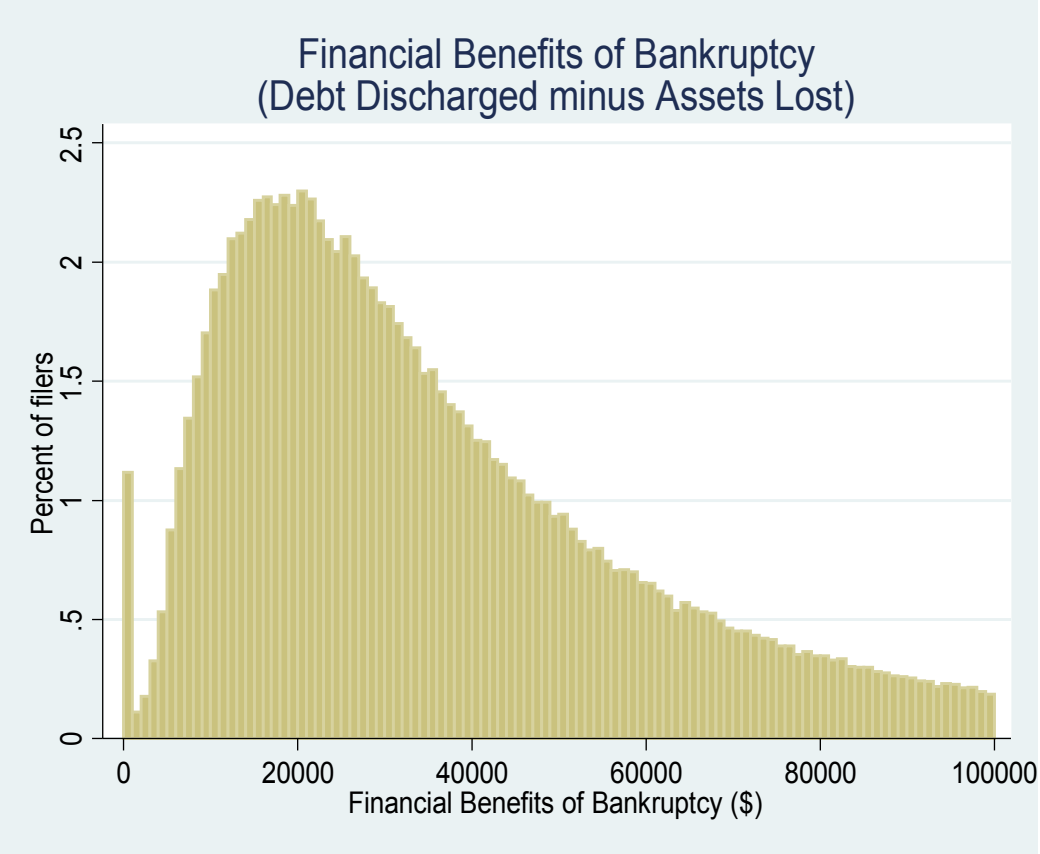


Table 3-1: Summary statistics for individual bankruptcy filing data

Variable	Obs	Mean	Std. Dev.	Min	Max
Number of trustees within 10 km	386770	13.032	19.427	0	94
Closest trustee-debtor distance	386768	17.537	54.398	0	1896.05
Financial benefits of bankruptcy	386770	48531.82	60286.59	0	1000000
Log of financial benefits	386770	10.2013	2.19	-9.21	13.82
Total liabilities	386770	86203.79	113327.9	1000	1000000
Total assets	386770	43741.35	90489.4	0	1876194
Secured liabilities	386770	35374.28	82842.54	0	955186
Unsecured liabilities	386770	50804.06	61405.54	0	1000000
Bankruptcy exemptions	386770	6704.79	15368.72	0	921001
Age	386770	43.56	13.29	18	90
Car	386770	0.628	0.48	0	1
Self-employment	386770	0.058	0.23	0	1
Numerical literacy	386770	265.47	12.81	213.827	323.07
Bankruptcy stigma	386770	0.64	0.48	0	1
Household size	386770	2.06	1.33	1	12
Average income	386770	31872.10	11492.45	9273	601418
Change in income	386770	4.46	2.88	-6.74	25.40
Divorce	386770	0.13	0.34	0	1
Prior defaults	386770	0.18	0.38	0	1
MIZ scale	386770	1.97	1.62	1	8

Table 3-2: Bankruptcy exemptions by Canadian provinces

	Exemptions					
Provinces	**House**	**Car**	**Pension**	**Personal Effects**	**Furniture**	**Land**
Alberta	40000	5000	No*	4000	4000	all if rural**
British Columbia	12000	5000	All	up to 4000 together		No
Manitoba	2500	3000	all	all	4500	No
New Brunswick	No	6500	all	No	5000	No
Newfoundland and Labrador	10000	2000	All	4000	4000	No
Nova Scotia	No	6500	All	All	All	No
Ontario	No	5650	All	5600	11300	No
Prince Edward Island	No	3000	All	All	2000	No
Quebec	No	No	All	up to 6000 together		No
Saskatchewan	50000	10000	All	7500	All	No

Notes: Bankruptcy exemptions are from http://www.bankruptcycanada.com/bankruptcyexemptions.htm
All amounts are in Canadian dollars and apply to equity in the asset. These amounts represent maximum values of assets protected from seizure by creditors in bankruptcy.

* Pension accounts are exempt in bankruptcy from October 1, 2009.

** See p. 87 for details.

Table 3-3: Summary statistics for urban and rural bankruptcy filers

	Urban			Rural		
Variable	**Obs**	**Mean**	**Std. Dev.**	**Obs**	**Mean**	**Std. Dev.**
Number of trustees within 10 km	262860	18.648	21.336	123910	1.118	1.851
Closest trustee-debtor distance	262858	3.355	4.603	123910	47.620	88.657
Financial benefits of bankruptcy	262860	50964.14	62820	123910	43372	54164
Log of financial benefits	262860	10.28122	2.028	123910	10.032	2.482
Age	262860	43.738	13.251	123910	43.180	13.374
Car	262860	0.591	0.492	123910	0.706	0.456
Self-employment	262860	0.064	0.245	123910	0.046	0.209
Numerical literacy	262860	267	12.856	123910	262.747	12.284
Bankruptcy stigma	262860	0.611	0.487	123910	0.693	0.461
Household size	262860	2.015	1.309	123910	2.163	1.355
Average income	262860	32860	12742.8	123910	29776	7831
Change in income	262860	4.412	2.650	123910	4.565	3.321
Divorce	262860	0.137	0.344	123910	0.116	0.320
Total assets	262860	42892	94888.4	123910	45542	80335
Prior defaults	262860	0.185	0.388	123910	0.171	0.377
MIZ scale	262860	1	0	123910	4.018	1.402

Table 3-4: Impact of bankruptcy trustees' proximity on financial benefits of bankruptcy

These tests examine the hypothesis that debtor-trustee distance has a positive impact on financial benefits of bankruptcy. Debtors with higher costs of filing as measured by the distance to a trustee will require a compensation in terms of bankruptcy benefits. This hypothesis predicts a negative coefficient for the number of trustees within 10 km from the debtor and a positive coefficient for the closest debtor-trustee distance.

Sample	Trustees within 10 km	Closest distance
Whole sample	0.00477***	-0.00010
	(0.00019)	(0.00010)
Rural	-0.01876***	0.00059***
	(0.00390)	(0.00013)
Urban	0.00415***	-0.00301**
	(0.00022)	(0.00130)

Notes: Each cell reports coefficients on the number of trustees within 10 km or closest debtor-trustee distance with standard errors from a regression with the financial benefits of bankruptcy as the dependent variable. We estimate these regressions on samples of bankruptcy filers described in the first column of this Table. Control variables as described in the text are included, but not reported. Full results for these regressions are presented in an on-line appendix. Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table 3-5: Impact of bankruptcy trustees' proximity on financial benefits of bankruptcy

Interactions of rural indicators and trustee-debtor distance

These tests examine the hypothesis that debtor-trustee distance has a positive impact on financial benefits of bankruptcy. Debtors with higher costs of filing as measured by the distance to a trustee will require a compensation in terms of bankruptcy benefits. This hypothesis predicts a negative coefficient for the number of trustees within 10 km from the debtor and a positive coefficient for the closest debtor-trustee distance.

Interactions	Trustees within 10 km	Closest distance
MIZ scale	-0.01026***	0.00018***
	(0.00258)	(0.00004)
Rural	-0.01961***	0.00439***
	(0.00383)	(0.00126)

Notes: Each cell reports coefficients on the interaction of the number of trustees within 10 km or closest debtor-trustee distance with the rural indicator or MIZ scale. In addition to the interaction term each regression includes number of trustees or distance and a rural indicator. Standard errors are reported in parentheses. The dependent variable is the financial benefits of bankruptcy. Control variables as described in the text are included, but not reported. Full results for these regressions are presented in an on-line appendix. Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table 3-6: Impact of bankruptcy trustees' proximity on financial benefits of bankruptcy before the financial crisis (January 2005 – August 2008)

These tests examine the hypothesis that debtor-trustee distance has a positive impact on financial benefits of bankruptcy. Debtors with higher costs of filing as measured by the distance to a trustee will require a compensation in terms of bankruptcy benefits. This hypothesis predicts a negative coefficient for the number of trustees within 10 km from the debtor and a positive coefficient for the closest debtor-trustee distance.

Sample	Trustees within 10 km	Closest distance
Whole sample	0.00479***	0.00001
	(0.00026)	(0.00014)
Rural	-0.02016***	0.00068***
	(0.00519)	(0.00018)
Urban	0.00428***	-0.00293
	(0.00032)	(0.00191)

Notes: Each cell reports coefficients on the number of trustees within 10 km or closest debtor-trustee distance with standard errors from a regression with the financial benefits of bankruptcy as a dependent variable. We estimate these regressions on samples of bankruptcy filers described in the first column of the Table. Control variables as described in the text are included, but not reported. Full results for these regressions are presented in an on-line appendix. Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table 3-7: Impact of bankruptcy trustees' proximity on financial benefits of bankruptcy after the financial crisis (September 2008 – December 2010)

These tests examine the hypothesis that debtor-trustee distance has a positive impact on financial benefits of bankruptcy. Debtors with higher costs of filing as measured by the distance to a trustee will require a compensation in terms of bankruptcy benefits. This hypothesis predicts a negative coefficient for the number of trustees within 10 km from the debtor and a positive coefficient for the closest debtor-trustee distance.

Sample	Trustees within 10 km	Closest distance
Whole sample	0.00472***	-0.00023*
	(0.00023)	(0.00014)
Rural	-0.01829***	0.00056***
	(0.00515)	(0.00019)
Urban	0.00394***	-0.00320**
	(0.00027)	(0.00144)

Notes: Each cell reports coefficients on the number of trustees within 10 km or closest debtor-trustee distance with standard errors from a regression with the financial benefits of bankruptcy as a dependent variable. We estimate these regressions on samples of bankruptcy filers described in the first column of the Table. Control variables as described in the text are included, but not reported. Full results for these regressions are presented in an on-line appendix. Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table 3-8: Impact of bankruptcy trustees' proximity on financial benefits of bankruptcy

Interactions of rural indicators and trustee-debtor distance before the 2008 economic crisis and after the crisis

These tests examine the hypothesis that debtor-trustee distance has a positive impact on financial benefits of bankruptcy. Debtors with higher costs of filing as measured by the distance to a trustee will require a compensation in terms of bankruptcy benefits. This hypothesis predicts a negative coefficient for the number of trustees within 10 km from the debtor and a positive coefficient for the closest debtor-trustee distance.

Interactions	Trustees within 10 km	Closest distance
	Before the crisis (January 2005 – August 2008)	
MIZ scale	-0.01083***	0.00016***
	(0.00280)	(0.00005)
Rural	-0.02070***	0.00445**
	(0.00503)	(0.00186)
	After the crisis (September 2008 – December 2010)	
MIZ scale	-0.01031***	0.00020***
	(0.00313)	(0.00005)
Rural	-0.01892***	0.00442***
	(0.00495)	(0.00142)

Notes: Each cell reports coefficients on the interaction of the number of trustees within 10 km or closest debtor-trustee distance with the rural indicator or MIZ scale. In addition to the interaction term each regression includes number of trustees or distance and a rural indicator. Standard errors are reported in parentheses. The dependent variable is the financial benefits of bankruptcy. Control variables as described in the text are included, but not reported. Full results for these regressions are presented in an on-line appendix. Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Bibliography

Agarwal, S., and R. Hauswald (2010). "Distance and Private Information in Lending," *The Review of Financial Studies, 23*(7), 2757-2788.

Allen, J., and E. Damar (2012). "Household Insolvency in Canada," Bank of Canada Review Winter (2011-2012), 43-54.

Coval, J., and T. Moskowitz (1999). "Home Bias at Home: Local Equity Preference in Domestic Portfolios," *The Journal of Finance, 54*(6), 2045-2073.

Coval, J., and T. Moskowitz (2001). "The Geography of Investment: Informed Trading and Asset Prices," *Journal of Political Economy, 109*(4), 811-841.

Degryse, H., and S. Ongena (2005). "Distance, Lending Relationships, and Competition," *The Journal of Finance, 60*(1), 231-266.

Fay, S., Hurst, E., and M. White (2002). "The Household Bankruptcy Decision," *The American Economic Review, 92*(3), 706-718.

Garmaise , M., and T. Moskowitz (2004). "Confronting Information Asymmetries: Evidence from Real Estate Markets," *The Review of Financial Studies, 17*(2), 405-437 .

Gross, D., and Souleles, N., (2002) "An empirical analysis of personal bankruptcy and delinquency," *The Review of Financial Studies, 15*(1), 319-347.

Gross, T., Notowidigdo, M., and J. Wang (2013). "Liquidity Constraints and Consumer Bankruptcy: Evidence from Tax Rebates," *The Review of Economics and Statistics*, forthcoming.

Hau, H. (2001). "Location Matters: An Examination of Trading Profits," *The Journal of Finance, 56*(5), 1959-1983.

Hauswald , R., and R. Marquez (2006). "Competition and Strategic Information Acquisition in Credit Markets," *The Review of Financial Studies, 19*(3), 967-1000.

Howatson-Leo, L., Earl, L., Puderer, H., Cunningham, R., Bollman, R., and W.

Rea (1996). “A New Approach to Non-CMA/CA Areas,” Statistics Canada, Agriculture and Rural Division, Working Paper No. 31.

Lusardi, A. (2012). “Numeracy, Financial Literacy, and Financial Decision-making,” National Bureau of Economic Research Working Paper No. 17821.

Murray, S. (2011). “Financial Literacy: A Conceptual Review,” Research paper prepared for the Task Force on Financial Literacy. Canadian Council on Learning.

Ragozzino, R., and J. Reuer (2011). “Geographic Distance and Corporate Acquisitions: Signals from IPO Firms,” *Strategic Management Journal, 32*, 876-894.

Scholnick, B. (2012). “Bankruptcy Spillovers: Distance, Public Disclosure and Opaque Information,” University of Alberta School of Business Working Paper.

White, M. (2007). “Bankruptcy Reform and Credit Cards,” *Journal of Economic Perspectives*, *21*(4), 175-199.

Appendix B

Table B1. The impact of bankruptcy trustees' proximity on financial benefits of bankruptcy. Full results with all controls reported. See complete description in Table 3-4.

Independent Variables	Whole sample	Rural	Urban
Number of trustees within 10 km	0.00477***	-0.01876***	0.00415***
	(0.00019)	(0.00390)	(0.00022)
Car	0.16327***	0.15831***	0.17166***
	(0.00693)	(0.01359)	(0.00801)
Self-employment	0.43125***	0.43580***	0.42011***
	(0.01428)	(0.03170)	(0.01578)
Numerical literacy	0.00130***	0.00330***	0.00180***
	(0.00041)	(0.00090)	(0.00046)
Age	0.00346***	0.00229***	0.00401***
	(0.00026)	(0.00049)	(0.00030)
Bankruptcy stigma	-0.11831***	-0.06317***	-0.13399***
	(0.00799)	(0.01559)	(0.00909)
Household size	0.09353***	0.08700***	0.09698***
	(0.00293)	(0.00548)	(0.00346)
Average income	0.00001***	0.00001***	0.00001***
	(0.00000)	(0.00000)	(0.00000)
Change in income	-0.00896***	-0.00791***	-0.01052***
	(0.00182)	(0.00252)	(0.00263)
Divorce	-0.00925	-0.03642*	0.00018
	(0.01038)	(0.02130)	(0.01171)
Total assets	-0.00000***	-0.00000***	-0.00000***
	(0.00000)	(0.00000)	(0.00000)
Prior defaults	-0.08983***	-0.02639	-0.11882***
	(0.00959)	(0.01859)	(0.01107)
Constant	9.18142***	8.68634***	9.04478***
	(0.10174)	(0.21558)	(0.12170)
Observations	386,770	123,910	262,860
R-squared	0.09445	0.18328	0.02841

We estimate these regressions on samples of bankruptcy filers described in the first row of this Table. Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table B2. The impact of bankruptcy trustees' proximity on financial benefits of bankruptcy. Full results with all controls reported. See complete description in Table 3-4.

Independent Variables	Whole sample	Rural	Urban
Closest trustee-debtor distance	-0.00010	0.00059***	-0.00301**
	(0.00010)	(0.00013)	(0.00130)
Car	0.13569***	0.16039***	0.15005***
	(0.00686)	(0.01359)	(0.00790)
Self-employment	0.43996***	0.43937***	0.42694***
	(0.01429)	(0.03168)	(0.01578)
Numerical literacy	0.00196***	0.00327***	0.00180***
	(0.00041)	(0.00090)	(0.00046)
Age	0.00383***	0.00229***	0.00433***
	(0.00025)	(0.00049)	(0.00030)
Bankruptcy stigma	-0.11269***	-0.05409***	-0.12112***
	(0.00802)	(0.01545)	(0.00906)
Household size	0.09176***	0.08717***	0.09560***
	(0.00294)	(0.00548)	(0.00347)
Average income	0.00001***	0.00001***	0.00000***
	(0.00000)	(0.00000)	(0.00000)
Change in income	-0.00886***	-0.00792***	-0.01059***
	(0.00182)	(0.00252)	(0.00263)
Divorce	-0.00699	-0.03692*	0.00054
	(0.01038)	(0.02129)	(0.01172)
Total assets	-0.00000***	-0.00000***	-0.00000***
	(0.00000)	(0.00000)	(0.00000)
Prior defaults	-0.09585***	-0.02652	-0.12679***
	(0.00957)	(0.01861)	(0.01106)
Constant	9.03793***	8.55063***	9.07374***
	(0.10578)	(0.21956)	(0.12233)
Observations	386,768	123,910	262,858
R-squared	0.09305	0.18330	0.02716

We estimate these regressions on samples of bankruptcy filers described in the first row of this Table. Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table B3. The interaction of rural indicators and trustee-debtor distance. See complete description in Table 3-5.

Independent Variables	Interactions	Interactions
MIZ scale	-0.01257***	
	(0.00332)	
Trustees* MIZ scale	-0.01026***	
	(0.00258)	
Trustees within 10 km	0.01427***	0.00360***
	(0.00254)	(0.00021)
Rural indicator		-0.08022***
		(0.01207)
Trustees*Rural indicator		-0.01961***
		(0.00383)
Car	0.16439***	0.16540***
	(0.00692)	(0.00692)
Self-employment	0.43020***	0.42833***
	(0.01428)	(0.01428)
Numerical literacy	0.00101**	0.00093**
	(0.00041)	(0.00041)
Age	0.00349***	0.00347***
	(0.00025)	(0.00026)
Bankruptcy stigma	-0.11536***	-0.11464***
	(0.00800)	(0.00793)
Household size	0.09366***	0.09381***
	(0.00293)	(0.00293)
Average income	0.00001***	0.00001***
	(0.00000)	(0.00000)
Change in income	-0.00899***	-0.00899***
	(0.00182)	(0.00182)
Divorce	-0.01029	-0.01030
	(0.01037)	(0.01037)
Total assets	-0.00000***	-0.00000***
	(0.00000)	(0.00000)
Prior defaults	-0.09117***	-0.09170***
	(0.00957)	(0.00957)
Constant	9.30134***	9.33133***
	(0.10560)	(0.10328)
Observations	386,770	386,770
R-squared	0.09467	0.09490

Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table B4. The interaction of rural indicators and trustee-debtor distance. See complete description in Table 3-5.

Independent Variables	Interactions	Interactions
MIZ scale	-0.04635***	
	(0.00336)	
Closest* MIZ scale	0.00018***	
	(0.00004)	
Closest trustee-debtor distance	-0.00020	-0.00369***
	(0.00022)	(0.00126)
Rural indicator		-0.20299***
		(0.01069)
Closest*Rural indicator		0.00439***
		(0.00126)
Car	0.14692***	0.15268***
	(0.00686)	(0.00685)
Self-employment	0.43676***	0.43252***
	(0.01428)	(0.01428)
Numerical literacy	0.00138***	0.00128***
	(0.00041)	(0.00040)
Age	0.00377***	0.00367***
	(0.00025)	(0.00025)
Bankruptcy stigma	-0.10150***	-0.10438***
	(0.00790)	(0.00789)
Household size	0.09285***	0.09336***
	(0.00294)	(0.00293)
Average income	0.00001***	0.00001***
	(0.00000)	(0.00000)
Change in income	-0.00901***	-0.00906***
	(0.00182)	(0.00182)
Divorce	-0.01018	-0.01015
	(0.01037)	(0.01037)
Total assets	-0.00000***	-0.00000***
	(0.00000)	(0.00000)
Prior defaults	-0.09709***	-0.09627***
	(0.00957)	(0.00957)
Constant	9.22701***	9.22594***
	(0.10666)	(0.10538)
Observations	386,768	386,768
R-squared	0.09381	0.09434

Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table B5. The impact of bankruptcy trustees' proximity on financial benefits of bankruptcy before the crisis (January 2005 – August 2008). See complete description in Table 3-6.

Independent Variables	Whole sample	Rural	Urban
Number of trustees within 10 km	0.00479***	-0.02016***	0.00428***
	(0.00026)	(0.00519)	(0.00032)
Car	0.17597***	0.18021***	0.18159***
	(0.01016)	(0.01932)	(0.01193)
Self-employment	0.43674***	0.40180***	0.43784***
	(0.02268)	(0.05169)	(0.02487)
Numerical literacy	0.00017	0.00118	0.00134**
	(0.00058)	(0.00125)	(0.00067)
Age	0.00302***	0.00230***	0.00341***
	(0.00036)	(0.00069)	(0.00042)
Bankruptcy stigma	-0.12589***	-0.06853***	-0.14232***
	(0.01149)	(0.02228)	(0.01335)
Household size	0.10632***	0.09944***	0.11084***
	(0.00428)	(0.00807)	(0.00504)
Average income	0.00001***	0.00001***	0.00000***
	(0.00000)	(0.00000)	(0.00000)
Change in income	0.00203	-0.00009	0.00470
	(0.00205)	(0.00266)	(0.00316)
Divorce	-0.03403**	-0.09485***	-0.01133
	(0.01562)	(0.03165)	(0.01778)
Total assets	-0.00000***	-0.00000***	-0.00000***
	(0.00000)	(0.00000)	(0.00000)
Prior defaults	-0.09092***	-0.02663	-0.12044***
	(0.01465)	(0.02753)	(0.01724)
Constant	9.56569***	9.30226***	9.24831***
	(0.14196)	(0.29677)	(0.17532)
Observations	194,201	62,421	131,780
R-squared	0.09647	0.19063	0.03225

We estimate these regressions on samples of bankruptcy filers described in the first row of this Table. Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table B6. The impact of bankruptcy trustees' proximity on financial benefits of bankruptcy before the crisis (January 2005 – August 2008). See complete description in Table 3-6.

Independent Variables	Whole sample	Rural	Urban
Closest trustee-debtor distance	0.00001	0.00068***	-0.00293
	(0.00014)	(0.00018)	(0.00191)
Car	0.15008***	0.18290***	0.16076***
	(0.01005)	(0.01933)	(0.01181)
Self-employment	0.44599***	0.40580***	0.44489***
	(0.02268)	(0.05161)	(0.02486)
Numerical literacy	0.00091	0.00121	0.00135**
	(0.00058)	(0.00124)	(0.00066)
Age	0.00342***	0.00232***	0.00377***
	(0.00036)	(0.00069)	(0.00042)
Bankruptcy stigma	-0.12228***	-0.05901***	-0.13056***
	(0.01155)	(0.02210)	(0.01329)
Household size	0.10435***	0.09949***	0.10932***
	(0.00429)	(0.00807)	(0.00505)
Average income	0.00001***	0.00001***	0.00000***
	(0.00000)	(0.00000)	(0.00000)
Change in income	0.00225	0.00005	0.00493
	(0.00205)	(0.00265)	(0.00316)
Divorce	-0.03172**	-0.09497***	-0.01135
	(0.01563)	(0.03164)	(0.01779)
Total assets	-0.00000***	-0.00000***	-0.00000***
	(0.00000)	(0.00000)	(0.00000)
Prior defaults	-0.09656***	-0.02633	-0.12861***
	(0.01464)	(0.02755)	(0.01723)
Constant	9.38592***	9.13111***	9.27189***
	(0.14801)	(0.30463)	(0.17548)
Observations	194,200	62,421	131,779
R-squared	0.09520	0.19069	0.03109

We estimate these regressions on samples of bankruptcy filers described in the first row of this Table. Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table B7. The impact of bankruptcy trustees' proximity on financial benefits of bankruptcy after the crisis (September 2008 – December 2010). See complete description in Table 3-7.

Independent Variables	Whole sample	Rural	Urban
Number of trustees within 10 km	0.00472***	-0.01829***	0.00394***
	(0.00023)	(0.00515)	(0.00027)
Car	0.16648***	0.16291***	0.17477***
	(0.00938)	(0.01934)	(0.01055)
Self-employment	0.42434***	0.45463***	0.40514***
	(0.01811)	(0.04011)	(0.01993)
Numerical literacy	0.00263***	0.00531***	0.00242***
	(0.00049)	(0.00113)	(0.00055)
Age	0.00418***	0.00274***	0.00483***
	(0.00035)	(0.00065)	(0.00041)
Bankruptcy stigma	-0.11233***	-0.05792***	-0.12694***
	(0.01030)	(0.02044)	(0.01170)
Household size	0.08699***	0.08256***	0.08913***
	(0.00396)	(0.00750)	(0.00464)
Average income	0.00001***	0.00001***	0.00001***
	(0.00000)	(0.00000)	(0.00000)
Change in income	-0.00186	-0.00957	0.00415
	(0.00459)	(0.00765)	(0.00549)
Divorce	0.01595	0.02062	0.01241
	(0.01349)	(0.02863)	(0.01496)
Total assets	-0.00000***	-0.00000***	-0.00000***
	(0.00000)	(0.00000)	(0.00000)
Prior defaults	-0.08745***	-0.02444	-0.11673***
	(0.01247)	(0.02440)	(0.01433)
Constant	9.02748***	8.35229***	9.00530***
	(0.12594)	(0.27367)	(0.15097)
Observations	192,569	61,489	131,080
R-squared	0.09372	0.17787	0.02477

We estimate these regressions on samples of bankruptcy filers described in the first row of this Table. Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table B8. The impact of bankruptcy trustees' proximity on financial benefits of bankruptcy after the crisis (September 2008 – December 2010). See complete description in Table 3-7.

Independent Variables	Whole sample	Rural	Urban
Closest trustee-debtor distance	-0.00023*	0.00056***	-0.00320**
	(0.00014)	(0.00019)	(0.00144)
Car	0.13754***	0.16456***	0.15305***
	(0.00927)	(0.01934)	(0.01042)
Self-employment	0.43259***	0.45804***	0.41165***
	(0.01813)	(0.04012)	(0.01992)
Numerical literacy	0.00321***	0.00525***	0.00241***
	(0.00049)	(0.00113)	(0.00055)
Age	0.00450***	0.00273***	0.00511***
	(0.00035)	(0.00065)	(0.00041)
Bankruptcy stigma	-0.10464***	-0.04906**	-0.11339***
	(0.01029)	(0.02021)	(0.01166)
Household size	0.08545***	0.08282***	0.08796***
	(0.00398)	(0.00750)	(0.00464)
Average income	0.00001***	0.00001***	0.00001***
	(0.00000)	(0.00000)	(0.00000)
Change in income	-0.00178	-0.00971	0.00402
	(0.00458)	(0.00765)	(0.00549)
Divorce	0.01808	0.01977	0.01314
	(0.01351)	(0.02861)	(0.01497)
Total assets	-0.00000***	-0.00000***	-0.00000***
	(0.00000)	(0.00000)	(0.00000)
Prior defaults	-0.09366***	-0.02492	-0.12429***
	(0.01245)	(0.02443)	(0.01431)
Constant	8.88066***	8.23179***	9.09966***
	(0.13211)	(0.27785)	(0.14925)
Observations	192,568	61,489	131,079
R-squared	0.09218	0.17786	0.02345

We estimate these regressions on samples of bankruptcy filers described in the first row of this Table. Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table B9. The interaction of rural indicators and trustee-debtor distance (01/2005-08/2008). See complete description in Table 3-8.

Independent Variables	Interactions	Interactions
MIZ scale	-0.01140**	
	(0.00465)	
Trustees* MIZ scale	-0.01083***	
	(0.00280)	
Trustees within 10 km	0.01487***	0.00367***
	(0.00275)	(0.00030)
Rural indicator		-0.07269***
		(0.01674)
Trustees*Rural indicator		-0.02070***
		(0.00503)
Car	0.17681***	0.17754***
	(0.01015)	(0.01015)
Self-employment	0.43581***	0.43397***
	(0.02268)	(0.02268)
Numerical literacy	-0.00013	-0.00020
	(0.00059)	(0.00058)
Age	0.00304***	0.00303***
	(0.00036)	(0.00036)
Bankruptcy stigma	-0.12383***	-0.12346***
	(0.01151)	(0.01147)
Household size	0.10639***	0.10655***
	(0.00427)	(0.00428)
Average income	0.00001***	0.00001***
	(0.00000)	(0.00000)
Change in income	0.00205	0.00208
	(0.00205)	(0.00205)
Divorce	-0.03515**	-0.03519**
	(0.01562)	(0.01561)
Total assets	-0.00000***	-0.00000***
	(0.00000)	(0.00000)
Prior defaults	-0.09243***	-0.09296***
	(0.01463)	(0.01463)
Constant	9.68338***	9.70883***
	(0.14843)	(0.14475)
Observations	194,201	194,201
R-squared	0.09667	0.09685

Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table B10. The interaction of rural indicators and trustee-debtor distance (01/2005-08/2008). See complete description in Table 3-8.

Independent Variables	Interactions	Interactions
MIZ scale	-0.04656***	
	(0.00456)	
Closest* MIZ scale	0.00016***	
	(0.00005)	
Closest trustee-debtor distance	-0.00002	-0.00367**
	(0.00028)	(0.00186)
Rural indicator		-0.20139***
		(0.01483)
Closest*Rural indicator		0.00445**
		(0.00186)
Car	0.16094***	0.16602***
	(0.01006)	(0.01008)
Self-employment	0.44203***	0.43832***
	(0.02267)	(0.02267)
Numerical literacy	0.00025	0.00017
	(0.00058)	(0.00058)
Age	0.00336***	0.00327***
	(0.00036)	(0.00036)
Bankruptcy stigma	-0.11110***	-0.11418***
	(0.01145)	(0.01143)
Household size	0.10545***	0.10599***
	(0.00428)	(0.00428)
Average income	0.00001***	0.00001***
	(0.00000)	(0.00000)
Change in income	0.00229	0.00224
	(0.00205)	(0.00205)
Divorce	-0.03507**	-0.03515**
	(0.01563)	(0.01562)
Total assets	-0.00000***	-0.00000***
	(0.00000)	(0.00000)
Prior defaults	-0.09837***	-0.09746***
	(0.01462)	(0.01463)
Constant	9.59718***	9.58994***
	(0.15021)	(0.14827)
Observations	194,200	194,200
R-squared	0.09591	0.09636

Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table B11. The interaction of rural indicators and trustee-debtor distance (09/2008-12/2010). See complete description in Table 3-8.

Independent Variables	Interactions	Interactions
MIZ scale	-0.01488***	
	(0.00409)	
Trustees* MIZ scale	-0.01031***	
	(0.00313)	
Trustees within 10 km	0.01420***	0.00344***
	(0.00308)	(0.00025)
Rural indicator		-0.09237***
		(0.01497)
Trustees*Rural indicator		-0.01892***
		(0.00495)
Car	0.16804***	0.16939***
	(0.00936)	(0.00936)
Self-employment	0.42313***	0.42115***
	(0.01811)	(0.01810)
Numerical literacy	0.00234***	0.00225***
	(0.00049)	(0.00049)
Age	0.00421***	0.00419***
	(0.00035)	(0.00035)
Bankruptcy stigma	-0.10811***	-0.10702***
	(0.01032)	(0.01025)
Household size	0.08723***	0.08737***
	(0.00397)	(0.00397)
Average income	0.00001***	0.00001***
	(0.00000)	(0.00000)
Change in income	-0.00192	-0.00192
	(0.00459)	(0.00459)
Divorce	0.01490	0.01496
	(0.01349)	(0.01349)
Total assets	-0.00000***	-0.00000***
	(0.00000)	(0.00000)
Prior defaults	-0.08872***	-0.08923***
	(0.01245)	(0.01245)
Constant	9.16511***	9.19847***
	(0.13145)	(0.12957)
Observations	192,569	192,569
R-squared	0.09400	0.09432

Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Table B12. The interaction of rural indicators and trustee-debtor distance (09/2008-12/2010). See complete description in Table 3-8.

Independent Variables	Interactions	Interactions
MIZ scale	-0.04711***	
	(0.00413)	
Closest* MIZ scale	0.00020***	
	(0.00005)	
Closest trustee-debtor distance	-0.00038	-0.00380***
	(0.00027)	(0.00142)
Rural indicator		-0.20838***
		(0.01338)
Closest*Rural indicator		0.00442***
		(0.00142)
Car	0.14938***	0.15602***
	(0.00927)	(0.00926)
Self-employment	0.42990***	0.42511***
	(0.01811)	(0.01810)
Numerical literacy	0.00269***	0.00257***
	(0.00049)	(0.00048)
Age	0.00446***	0.00436***
	(0.00035)	(0.00035)
Bankruptcy stigma	-0.09322***	-0.09598***
	(0.01019)	(0.01019)
Household size	0.08660***	0.08708***
	(0.00398)	(0.00397)
Average income	0.00001***	0.00001***
	(0.00000)	(0.00000)
Change in income	-0.00210	-0.00220
	(0.00459)	(0.00459)
Divorce	0.01500	0.01518
	(0.01350)	(0.01349)
Total assets	-0.00000***	-0.00000***
	(0.00000)	(0.00000)
Prior defaults	-0.09444***	-0.09368***
	(0.01245)	(0.01245)
Constant	9.10926***	9.11288***
	(0.13353)	(0.13261)
Observations	192,568	192,568
R-squared	0.09304	0.09370

Standard errors are clustered at the DA level. *** denotes significance at 1%, ** - significance at 5%, * - significance at 10%.

Chapter 4: Do Mandated Credit Expansions Cause Foreclosure? Regression Discontinuity Evidence from the Community Reinvestment Act

4.1 Introduction

The performance of loans extended due to a government mandate is of prominent economic importance as governments around the world attempt to increase credit supply to specially designated groups (such as low income borrowers in the USA). Government-induced loans may have a higher rate of default if the credit quality and ability to carry a mortgage of loan recipients is lower than the quality and ability of those borrowers that would receive credit without the influence of the government.

Despite the importance of the performance of government-mandated loans, existing literature on this question provides rather mixed evidence. Several recent studies have argued that government interventions into the mortgage credit market such as the Community Reinvestment Act of 1977 (CRA) might have contributed to the subprime mortgage meltdown and the 2008 financial crisis (e.g. Rajan, 2010; Demyanyk and Van Hemert, 2011; Agarwal et al., 2012). Rajan (2010) suggests that inexpensive mortgage credit encouraged by the US government, but provided to low income and minority borrowers by private lenders was a primary method of attenuating income disparities in the USA. However, this method later elevated mortgage foreclosures and aggravated the financial crisis. Another strand of the literature on this topic, however, argues that government interventions had no effect on mortgage defaults and the financial crisis (e.g. Ding et al., 2008; Laderman and Reid, 2008; Canner and Bhutta, 2008; Avery and Brevoort, 2011).

Our paper contributes to this literature by arguing that lender's expected

cost of a foreclosure determines whether a mandated increase in credit supply leads to more mortgage defaults. We hypothesize that lenders subject to the CRA are less inclined to lend to high-risk borrowers in states with high foreclosure cost than lenders facing the same risky borrowers in low foreclosure cost states. This hypothesis is new to the literature on the impact of the CRA on mortgage defaults.

Different US states have dissimilar legal foreclosure costs as some states require a judicial foreclosure process, while other jurisdictions allow lenders to repossess a property based on a mortgage contract only and without a court hearing (Pence, 2006; Clauretie and Herzog, 1990; Mian, Sufi, and Trebbi, 2011). These differences in foreclosure procedures affect the duration and costs of a foreclosure process. We use this variation in foreclosure costs to test the hypothesis that lenders bound by the CRA to lend to high risk borrowers will be more careful in their lending in states with higher foreclosure costs compared to lenders in states with low foreclosure costs. This hypothesis is based on the idea that in case of a borrower's default, lenders will expect to receive more revenue from a foreclosure in low foreclosure cost states compared to high foreclosure cost states.

This hypothesis implies that the CRA will affect foreclosure rates in states with low costs of foreclosure to a greater degree than it will affect foreclosure rates in high foreclosure cost states. Intuitively, financial institutions in high foreclosure cost states may be more reluctant to lend to high risk borrowers, despite being required to do so by the CRA, because they would expect to recover less in a possible future foreclosure. Effectively, these lenders would be more careful in allocating credit in order to minimize their losses in case of a possible future mortgage default. On the other hand, financial institu-

tions may be more willing to extend credit to CRA designated borrowers in low foreclosure cost states because they would expect to retrieve more from a possible mortgage default. We test this new hypothesis.

The effect of state foreclosure laws and procedures on the supply of mortgage credit has been examined by Pence (2006), who argues that borrowers in states with high costs of foreclosure (i.e., those that require a judicial foreclosure) receive smaller loans compared to borrowers in low costs of foreclosure states. However, Pence (2006) does not address the question of whether mandated mortgage credit supply or the performance of mandated loans are affected by foreclosure laws and procedures. Therefore, our paper is the first in the literature to apply the insights of Pence (2006) on foreclosure costs to the analysis of the mandated increases in mortgage credit supply from the CRA.

In order to isolate the causal effect of the CRA on foreclosures, we apply a regression discontinuity (RD) design to US ZIP code level data on foreclosure metrics from Zillow.com and income and neighborhood characteristics from the US Census Bureau and the Bureau of Labor Statistics. Our RD approach uses a rule embedded in the CRA that define Low and Moderate Income (LMI) regions as those with median family incomes below 80 % of the median family income of the Metropolitan Statistical Area (MSA) in which they are located. In addition, for a region to receive the influence from the CRA, it has to be in the assessment area of regulated lenders.[32] A quasi-experimental design of the RD relies on the notion that regions just above and just below the arbitrary and exogenously imposed threshold are the same in their observable

[32]Not all mortgage lenders are regulated by the CRA. Only performance of national and state chartered banks, savings banks and associations are evaluated by CRA overseers in the areas where these lending institutions have deposit-taking offices or ATMs. These areas are called CRA assessment areas. CRA assessment areas are defined as low and moderate income regions where regulated mortgage lenders have physical deposit-taking operations through their offices or automated teller machines (ATMs) (JCHS, 2002).

and unobservable attributes, but due to the rule only those below the cutoff are directly affected by the CRA. Therefore, any differences in the mortgage loan outcomes between these two groups of areas could be attributed to the influence of the CRA.

We use three standard measures to capture legal foreclosure costs across states. These measures are based on Mian, Sufi, and Trebbi (2011), Pence (2006), and the National Consumer Law Center (NCLC). Our results are robust to the usage of these different measures. In addition to simple binary measures of whether a state has judicial or non-judicial foreclosure process, we examine how the average duration of foreclosure impacts foreclosure rates. This approach to measuring foreclosure costs is new to the literature. In essence, states with a longer foreclosure process impose more costs on the lender due to property depreciation in foreclosure, higher legal fees, and possible house price declines. Using this method of measuring foreclosure costs, we find that states with the shortest duration of foreclosure (low cost states) have the largest impact of the CRA on foreclosure rates, which is consistent with our main hypothesis.

Our main contribution is that we combine the issues related to the impact of the Community Reinvestment Act (CRA) on mortgage defaults with legal costs of foreclosure as measured by state foreclosure laws. We also use a new database on mortgage foreclosures from Zillow.com, not used in previous studies of the effect of the CRA on mortgage credit or loan defaults.

This paper contributes to the growing literature on the impact of government mandated credit expansions, and in particular, the Community Reinvestment Act, on mortgage defaults. Demyanyk and Van Hemert (2011) use LoanPerformance data on subprime mortgages made in 2001-2007 to examine

reasons behind the subprime mortgage crisis. These authors conclude that the status of a neighborhood as low or moderate income, which is the primary criterion for the protection under the CRA, increased the probability of subprime mortgage delinquency. Similarly, Agarwal et al. (2012) examine the performance of loans extended by lenders undergoing CRA exams, and thus subject to more CRA related scrutiny, and find that these loans are more likely to default.

Other studies, however, have argued that the CRA does not increase loan defaults. Ding et al. (2008) compare the performance of loans granted by CRA-regulated banks and independent lenders to subprime borrowers in 2003-2006, and conclude that, on average, default risks were higher for subprime loans than for CRA loans. Laderman and Reid (2008) examine if mortgages originated by CRA regulated institutions were more likely to foreclose compared to loans extended by unregulated lenders and find that CRA mandated loans were less likely to default. Avery and Brevoort (2011) use a regression discontinuity design based on the CRA provision that only regions with incomes below the 80 % of MSA income were influenced by the CRA, and argue that regions subject to the CRA had experienced less mortgage delinquencies.

This paper is the first in the literature to examine whether the interaction of two separate legal institutions affect mortgage foreclosure. The two exogenous legal institutions we use are: (1) the Community Reinvestment Act, and (2) state level differences in foreclosure costs. Our regression discontinuity analysis of the impact of the CRA on mortgage foreclosures reveals that the costs of foreclosure are important in determining this impact. When we ignore these costs and estimate the effect of the CRA on foreclosures, we find no impact of the CRA on foreclosure rates, a result that is similar to Ding et

al. (2008), Laderman and Reid (2008), Canner and Bhutta (2008), and Avery and Brevoort (2011). However, when we account for state level differences in the costs of foreclosure, we find that the CRA increases mortgage foreclosures in states with low legal cost of foreclosure.

This finding is similar to the results of Agarwal et al. (2012) in sense that both this study and Agarwal et al. (2012) concentrate on a specific group of lenders more likely to be influenced by the CRA, and find that the CRA increases loan default rates for these specific lenders. Agarwal et al. (2012) use CRA exam dates to separate and compare loans extended by lenders subject to CRA exams, and thus more likely to be compliant with the CRA, and loans of financial institutions not subject to CRA tests. Thus, these authors identify the impact of the CRA using the timing of CRA exams. We take a different approach. We separate lenders based on legal foreclosure costs, and use the variation in these costs and the 80 % income discontinuity to identify the effect of the CRA on mortgage performance.

4.2 The Community Reinvestment Act and Mortgage Credit

The CRA was one of the major components of the homeownership stimulation efforts of the US government in the 1990s. The CRA encourages regulated mortgage lenders to invest back into the communities where they take deposits. This policy may be especially relevant for low and moderate income neighborhoods which were often avoided by mortgage lenders before the promulgation of the law. Deposit-taking operations are generally used by regulators to define lenders' assessment areas; however, regulated lenders have some choice in deciding which regions belong to their assessment areas (Karikari, 2009). Moreover, banks may add activities of their affiliates into

a CRA test (Karikari, 2009). CRA regulated lenders consist of national and state chartered banks, savings banks and savings associations (Berry and Lee, 2007).

CRA regulated lenders may be stimulated to extend credit to high-risk mortgage applicants whom they would reject without the influence of the legislation.[33] However, the CRA encourages banks to follow 'safety and soundness limitations' in their lending to LMI communities (12 U.S.C. § 228.21). Also, CRA evaluations take into account local lending conditions and banks' specific circumstances in determining CRA ratings (Canner and Bhutta, 2008). Even though CRA regulated lenders may charge higher rates on riskier mortgages, these rates might not be high enough to offset default risks; especially, if they are not correctly modelled.

The question raised in this essay is important because the CRA influences activities of many US lenders and borrowers. The importance of CRA impact may be illustrated by the fact that in 1990-2000, CRA regulated lending institutions operating inside their assessment areas originated between 30 and 36 % of home purchase mortgages (JCHS, 2002) and 45 % of these loans in 2008 (Avery et al., 2010).[34] In addition, Bhutta (2011) argues that non-CRA-regulated lenders provide more credit to communities with CRA-regulated banks because those banks' mortgages reveal information on house values. This information can be used to make better appraisals and reduce loan risks of all lenders.

[33]Schill and Wachter (1994) find that in the early 1990s low income and minority mortgage applicants were more likely to get a loan in low income regions. Hence, the authors conclude that the CRA may indirectly keep these people in poor neighborhoods.

[34]Even though the share of CRA influenced lending as of total lending declined during the 1990s, it may rise again after the crisis because many non-regulated lenders such as independent mortgage companies and brokers reduced their operations or went out of business.

4.3 Data

4.3.1 Dependent Variable

Our primary variable of interest is the number of foreclosures on residential mortgages as a proportion of homes where homes are all single-family houses and condominiums. These foreclosure data come from Zillow.com and are available monthly for the 5512 US ZIP codes for our study period of January 2003 to September 2010.[35] To facilitate interpretation and analysis, these data are transformed from the percent of homes in foreclosure into the number of foreclosures per 100,000 homes. For instance, one foreclosure in a month in a region with 1000 homes was shown as 0.1 percent in the original data and it is presented as 100 foreclosures per 100,000 homes in the transformed data.

Figure 4-1 depicts a histogram of the number of foreclosures per 100,000 homes per ZIP code per month with the income ratio constrained to be between 0.75 and 0.85.[36] The density function of this variable appears to be similar to that of a log-normal distribution. Therefore, we take the logarithm of the foreclosure rate and use it as the dependent variable. Figure 4-2 displays the distribution of our dependent variable for the income ratio between 0.75 and 0.85.

4.3.2 Foreclosure Costs

In order to identify the impact the CRA on foreclosure rates more precisely, we introduce foreclosure costs into our analysis. One issue important

[35]Some ZIP codes have missing values for this variable in some months. Because of these missing observations, we cannot aggregate these monthly data to annual data or to simple cross-sectional data for the entire period.

[36]We use only those ZIP codes for which the ratio of their median family income to MSA income is between 0.75 and 0.85.

for these costs is judicial foreclosure state status. In judicial foreclosure states a foreclosure requires a court hearing with associated delays in disposing foreclosed property, legal expenses and possible additional lender's losses due to house price declines during the foreclosure process. A more expensive foreclosure process in judicial states should have made financial institutions more careful in their lending in these states. For instance, Pence (2006) argues that the judicial status reduces mortgage loan amount. This may imply higher down payments, less credit to riskier borrowers, less negative home equity when house prices decline, and a lower number of foreclosures later. Lenders in non-judicial foreclosure states are not subject to those incentives and hence they could be more willing to provide mortgage credit to riskier borrowers. Therefore, the CRA may have a larger impact on foreclosures in non-judicial foreclosure states than in judicial states.

The definition of judicial and non-judicial state foreclosure procedures used in this paper is based on Mian, Sufi and Trebbi (2011) which is derived from RealtyTrac.com.[37] Alternative definitions of the judicial foreclosure state status are from Pence (2006) and the National Consumer Law Center (NCLC).[38] While we follow the definition of the Mian, Sufi, and Trebbi (2011) to derive most results of this paper, we also use alternative definitions for a robustness check. All three definitions are coded as dummy variables which are equal to 1 for ZIP codes in judicial foreclosure states and 0 for postal codes in non-judicial foreclosure states.

The judicial status dummy variables based on the alternative classifications seem to be highly correlated.[39] Indeed, they coincide in defining most

[37]It can be found at the following page: http://www.realtytrac.com/foreclosure-laws/foreclosure-laws-comparison.asp

[38]NCLC classification is available at http://www.nclc.org/images/pdf/foreclosure_mortgage/state_laws/survey-foreclosure-card.pdf

[39]The correlation between the first and the second measure is 0.8248, and the first and

of the states as judicial or non-judicial. However, some distinctions emerge among them due to technical details. For instance, Massachusetts is a judicial state according to RealtyTrac and a non-judicial state according to Pence (2006) and NCLC. This discrepancy can be due to the fact that this state requires establishing in court that the defaulted borrower is not on active military duty; if this fact is established then no formal hearing about the mortgage foreclosure itself is required (Gerardi and Willen, 2011).[40] Without assessing the merits of each classifications, we decide to use all three to check the robustness of our results. We have no preferences over these definitions and results are the same no matter which definition we use.

We also use the average duration of foreclosure process to assess the impact of the CRA on different lenders. Data on the duration of foreclosure are derived from RealtyTrac.com (see, footnote 38). We use data on days in foreclosure to define five groups of regions starting with ZIP codes with less than 100 days in foreclosure and ending with postal codes with more than 210 days in foreclosure. A shorter foreclosure process is similar to low cost of foreclosure as the lender can repossess property swiftly before it depreciates due to neglect and vandalism. Less days in foreclosure also imply lower legal fees and less costs related to maintenance. Finally, in the environment of falling house prices, as was the case in 2007-2010, shorter foreclosures imply less losses to the lender because of price declines. Using these data on foreclosure duration, we test whether states with lower costs of foreclosure experienced more mortgage defaults due to the CRA.

---

the third is 0.7702. See Table C7 for correlations between the three dummy variables corresponding to the judicial/non-judicial classifications as well as the dummy variable indicating whether the foreclosure process is longer than 120 days.

[40]See Kris Gerardi and Paul Willen blogs at http://realestateresearch.frbatlanta.org/rer/foreclosure-laws/ and http://realestateresearch.frbatlanta.org/files/021411_mass.html

4.3.3 Control Variables

We use a large set of variables to control for local house market conditions and factors influencing mortgage foreclosures. Full details on all our control variables are provided in the Appendix C. Below we briefly outline the sources of data on the most important control variables.

We control for local house prices using data from Zillow.com on house values in January 2003 - September 2010. Since this series is likely to have a unit root during this period, real house value growth is used in the estimation. This growth is calculated as a difference in the logarithms of the current and previous value of the statistics minus monthly inflation.

We use median family incomes for ZIP codes and MSAs from the 2000 US census. This source provides data for ZIP code tabulation areas (ZCTA), which are made of census blocks and correspond to US ZIP codes. The 2000 census also supplies data on regional demographic and house market characteristics. The number of mortgages and owner-occupied housing units in 1990 are from the 1990 US census. Annual county-level data on housing units in 2002-2009 are from US Census Population Estimates Program. Annual growth rate of housing is defined as a logarithmic difference between the number of houses in the present and previous year. These growth rates are computed for all years in 2003-2009 and included as separate control variables into the regression. Since future realized growth of housing should not be used for past periods, these variables are multiplied by dummy variables which are equal to 1 for all periods after the growth is recorded. In this form, the growth rates of housing units are used in the regressions.

The Bureau of Labor Statistics provides data on not seasonally adjusted monthly unemployment rate for Metropolitan Statistical Areas in 2003-2010

and average weekly wages for US counties from the fourth quarter of 2002 and until the second quarter of 2010. Monthly and quarterly not seasonally adjusted U.S. city average Consumer Price Index (CPI) for all items comes from the same source and it is used to compute monthly and quarterly inflation. The growth rate of the real average weekly wage is the difference in the logarithms of the current and past value adjusted for inflation. Quarterly values of the growth in real wages are matched to the appropriate months in order to merge this variable with the rest of the dataset.

ZIP code level data are matched to county and census tract data using crosswalk files provided by the US Department of Housing and Urban Development (HUD). After all matching is done, the primary database, which is constrained to ZIP codes with the ratio of the median family income to MSA income of between 0.75 and 0.85 (more details on this constraint are provided in the next section), has 24529 observations, covers 90 months (January 2003 - June 2010) and 437 ZIP codes. These monthly data allow us to control for variation in the dependent variable using time-varying control variables. We cluster our standard errors in all regressions by ZIP codes and months, to take into account possible correlations among postal codes and months. A description of all variables used as well as summary statistics are in Tables C1 and C2.

4.4 Econometric Methodology

The empirical analysis of this paper uses the Regression Discontinuity design (RD).[41] According to Angrist and Pischke (2009), a sharp regression

[41]The literature on the Regression Discontinuity uses word 'design' instead of 'method' and we stick to their terminology.

discontinuity design is suited to problems where treatment depends on a certain rule splitting the entire sample into two or more parts.[42] Ideally, this rule should be deterministic, discrete, and arbitrary. It is crucial for this identification that entities analyzed could not perfectly manipulate the attributes used in their assignment to treatment and control groups, otherwise they could self-select into either of these groups (Lee and Lemieux, 2010).[43] Also, it is important that the covariates influencing the outcome variable do not change discontinuously at the threshold for the assignment variable or the RD design will not be valid. Below we provide a typical description of the RD based on Lee and Lemieux (2010).

In our case, the assignment variable is the ratio of the ZIP code median family income to MSA median family income, denoted *inc*.[44] A dummy variable defining the treatment group is therefore:

$$D_i = \begin{cases} 1 & \text{if } inc_i < 0.8 \\ 0 & \text{if } inc_i \geq 0.8 \end{cases} \quad (1)$$

This rule is based on the provisions of the Community Reinvestment Act which stipulate that regions with median family incomes below 80 % of their MSA incomes are defined as Low and Moderate Income (LMI) and they are subject to protection under the CRA.

In addition to define our treatment group, we use 2005-2009 lists of US census tracts included into CRA assessment areas as provided by the Federal

[42]Sharp RD is different from fuzzy RD. While sharp RD uses a deterministic and discontinuous treatment, fuzzy RD relies on the probability of treatment (Angrist and Pischke, 2009).

[43]Perfect manipulation of the assignment not just manipulation is required. In most cases, subjects can manipulate the assignment. For instance, persons may work harder, increase their incomes, and, eventually, move their neighborhood from the treatment to control group. However, they should not be able to control the assignment perfectly so that it is entirely their decision whether to be treated or not (Lee and Lemieux, 2010).

[44]The CRA uses census tracts as regions for assigning the LMI status and it does not deal with ZIP codes. However, since almost all of our data are for ZIPs, we assigned them as LMI using the rule described in the CRA. It is also possible to match postal codes and census tracts.

Financial Institutions Examination Council. While the 2005-2009 period does not cover the whole period used in our study (2003-2010),[45] assessment areas definitions do not change much over time. Also, the CRA mandated lending may influence foreclosures many years later, as typical mortgage contracts can last up to 30 years. Hence, even if a ZIP code is no longer inside a CRA assessment area, but it was in a CRA area in the last decade, it should be included into our analysis.

We check that there is no sorting near the threshold by looking at the density of the assignment variable, *inc*. Abrupt discontinuities or jumps in this density hint that regions could be engaged in sorting near the LMI cutoff.[46] Another way to test whether this assumption is satisfied is by comparing observed characteristics of ZIP codes just above and below the threshold and making sure that these two groups are similar in these characteristics. We check this by plotting additional covariates against median family income and examining whether these graphs have discontinuities at the income cutoff (see, Appendix D for more details).

The simplest way to estimate the treatment effect is with the so-called zero order polynomial. It is as follows:

$$Y_{it} = \alpha_r + \tau \cdot D_i + \epsilon_{it} \tag{2}$$

where, Y is the percentage of home mortgage foreclosures in a region, $\tau = \alpha_l - \alpha_r$ is the effect of the CRA, α_l and α_r are regression intercepts to the left and right of the median family income ratio cutoff, and ϵ is an error term.

In simple words, this specification allows us to compare average foreclosure rates in the treatment group, regions with family incomes below 80

[45] We were not able to access data for 2003, 2004, and 2010.

[46] It is difficult to imagine a situation in which the inhabitants of a ZIP code will be interested in changing the median family income of the ZIP and median family income of their MSA just to sort their region around the LMI cutoff.

% of MSA income, and in the control group, regions with incomes above 80 % of MSA income. Specification in equation (2) can be extended to allow for linear relationships between the outcome variable (foreclosure rates) and region's family income. If these relationships are allowed to be different in the treatment and control groups, then the first order polynomial will look as follows:

$$Y_{it} = \alpha_r + \tau \cdot D_i + \beta_{r1} \cdot (inc_i - c) + (\beta_{l1} - \beta_{r1}) \cdot D_i \cdot (inc_i - c) + \epsilon_{it} \quad (3)$$

where c is the value of this income ratio at the cutoff (0.8), β_{l1} and β_{r1} are slope coefficients of the function of the median family income to the left and right of the threshold. If this regression is run on pooled data from both sides of the cutoff, it will directly generate the estimate of the treatment effect τ and its standard error.

If we allow for the regression function to be non-linear, higher order polynomial terms will be included in equation (3). For instance, the J-th order polynomial is as follows:

$$Y_{it} = \alpha_r + \tau \cdot D_i + \sum_{j=1}^{J} \beta_{rj} \cdot (inc_i - c)^j + \sum_{j=1}^{J} (\beta_{lj} - \beta_{rj}) \cdot D_i \cdot (inc_i - c)^j + \epsilon_{it} (4)$$

We also include state and month fixed effects in order to control for unobservable characteristics of states which do not change over time and for period-specific business cycle fluctuations.[47] ZIP code fixed effects cannot be added because the income ratio used to define CRA eligibility is fixed during the period under study. Hence, the treatment indicator will be perfectly correlated with a group of ZIP fixed effects. While not allowing ZIP fixed

[47]State fixed effects may capture differences in state laws regarding foreclosure, bankruptcy or lending. While we control for the judicial foreclosure state process, other differences in state legal provisions such as the size of homestead exemption in bankruptcy, availability of deficiency judgements and a statutory right of redemption are not explicitly taken into account. We decide to use state fixed effects instead of MSA effects because many MSAs span multiple states and the District of Columbia.

effects, the stability of the CRA eligibility helps to define the CRA effect better because there are no regions which move in and out of the treatment group during the study period (2003-2010).[48]

Our final econometric specification has the following form:

$$\begin{aligned} Y_{it} &= \alpha_r + \tau \cdot D_i + \sum_{j=1}^{J} \beta_{rj} \cdot (inc_i - c)^j + \sum_{j=1}^{J} (\beta_{lj} - \beta_{rj}) \cdot D_i \cdot (inc_i - c)^j + \\ &+ \sum_{k=1}^{T} \gamma_k x_t^k + \sum_{l=1}^{L} \delta_l \rho_i^l + \epsilon_{it} \end{aligned} \tag{5}$$

where T is the total number of months, x^k is a dummy variable equal to 1 in period k, L is total number of states, and ρ^l is a dummy variable equal to 1 for the state l.[49]

We estimate equation (5) in a neighborhood h of the cutoff value restricting the sample to only those observations for which $c - h \leq inc_i \leq c + h$ holds. This restriction excludes observations with income ratio far above or far below the cutoff from the analysis. The Regression Discontinuity (RD) method uses this approach to compare regions just above and just below the threshold. This comparison forms the identification strategy of the RD.

A larger bandwidth h may allow us to use more data in estimation and may increase the precision of the estimates. However, it may also generate a bias in the estimates if the functional forms of the relationship between income and foreclosures are not linear. With a large bandwidth, one may also capture the effects of other factors with threshold assignment rules. For instance, Bhutta (2011) argues that the Government Sponsored Enterprises (GSE) act of 1992 may induce a discontinuity at 90 % of median income due to its loan

[48] We do not have data about foreclosures on mortgages of particular vintages, just total foreclosures in a month. Hence, postal codes which switched from the treatment to control group or vice versa would "contaminate" our estimates of the CRA effect. However, there were no switchers in 2003-2010 because CRA eligibility was fixed.

[49] One month dummy variable and one state dummy are omitted to avoid perfect collinearity with the constant term.

purchase targets for Fannie Mae and Freddie Mac.[50] As argued by Lee and Lemieux (2010), the treatment effect, τ, is 'a *weighted* average treatment effect across *all* individuals' or regions and not only those near the threshold.

We include variables representing ZIP codes characteristics into equation (5) to control for other possible factors affecting foreclosures. In addition, we use these variables to check whether the model is correctly specified and the lack of the perfect control assumption is not violated. If the inclusion of the additional variables increases the standard errors or changes the estimated effect substantially then the functional form may be misspecified or the lack of control assumption is not satisfied.

We would like to explore if the CRA had different impacts in 2003-2006 and 2007-2010. As can be seen from Figure 4-3, foreclosures began to increase in 2006 and grew rapidly after that. Also, 2007 is usually considered as the beginning of the financial crisis. Figure 4-4 demonstrates that the number of new mortgages declined from 2006 onward and that it began to increase again only in 2009. Value of mortgage loans to household income and number of mortgages outstanding peaked in 2008.[51]

4.5 Results

Data suggest that the assumptions of the Regression Discontinuity design are likely to be satisfied in the present setup.[52] In particular, as Figures

[50]Hence, in his analysis, the author concentrates more on regions with median family income above 0.75 and below 0.85 of their MSA median family income.

[51]Total US personal incomes in 1999-2010 are from the Bureau of Economic Analysis. They are used to calculate the ratio of the value of mortgage balances to family incomes presented in Figure 4-4. The quarterly report on household debt and credit provides annual data on the value and the number of mortgage loans outstanding. These two series are quarterly and they are aggregated to the annual frequency by taking arithmetic averages over corresponding quarters. The number of mortgages extended each year in 2000-2009 as reported in Avery et al. (2010).

[52]See appendices C and D for a detailed discussion as well as Figures and Tables in those

in Appendices C and D suggest, the density of the income ratio is continuous at the cutoff, there are no jumps in the control variables at the threshold, and the observable covariates also seem to be balanced in the treatment and control groups. Therefore, we apply the RD method to our data and report results in Tables 4-1 through 4-4. These Tables show only the estimates of the treatment effect, its standard error, number of observations, and R^2. Full results from these regressions are reported in the appendices.

The first panel of Table 4-1 presents results for the 2003-2010 period.[53] The second panel of Table 4-1 reports the impact of the CRA on foreclosures before the crisis (2003-2006), while the last panel summarizes results during the crisis (2007-2010). Column 1 shows estimates of the CRA impact for the specification without the control variables as shown by equation (5), while Column 2 reports estimates of the impact with the control variables.[54] This Table shows that only in one case CRA coefficient is different from zero statistically. For most specifications, the CRA coefficient is not significant. This result is consistent with several studies finding that there is no impact of the CRA on mortgage defaults in the whole sample (Ding et al., 2008; Laderman and Reid, 2008; Canner and Bhutta, 2008; Avery and Brevoort, 2011).

In order to assess whether the duration of foreclosure process influenced the CRA impact, we split the sample into five subgroups based on days in

appendices.

[53]Standard errors reported in this paper are clustered by ZIP and month using a Stata program of Petersen (2009).

[54]Judicial foreclosure state status is one of these controls. It is perfectly collinear with a group of state fixed effects. Hence, one of the state dummy variables is omitted to avoid perfect collinearity.

foreclosure.[55] Results are outlined in Table 4-2. This Table also shows the impact of the CRA before and during the crisis. The results suggest that in the least foreclosure costly states (the shortest duration in foreclosure), panel 1 of Table 4-2, the CRA impact estimate is positive and significant for the entire sample as well as before and during the crisis. For the whole period (2003-2010), it is between 38 to 61 percent in various specifications. Before the crisis (2003-2006), the impact is around 49 percent. During the crisis (2007-2010), it is in the 56 to 66 percent range. Consequently, the positive sign and statistical significance of the effect is similar both before and during the crisis.

The impact of the CRA is not different from zero in the states with a longer and costlier foreclosure process (as reported in days in foreclosure), as panels 2-5 of Table 4-2 confirm. For instance, when foreclosure process takes between 100 and 120 days, panel 2, the effect of the CRA is not significant in all specifications and both before and during the crisis. This tendency implies that the CRA impact on foreclosures declines when days in foreclosure increase. This finding confirms the theoretical prediction that lenders provide mortgages considering costs of a possible future foreclosure. A shorter duration of the foreclosure process translates into lower foreclosure costs. This could have made financial institutions more willing to extend mortgages to riskier borrowers and be less careful in lending. This pattern holds for the whole period as well as for the periods before and during the crisis.

Table 4-3 reports results for high and low cost of foreclosure subsamples as measured by the judicial or non-judicial foreclosure dummy variables. In

[55]It is difficult to make finer distinctions of the time in foreclosure process because there is little variation among states in this variable. So, selecting smaller groups based on the duration of the foreclosure process amounts to choosing one state only or having no observations. When one state is selected it is impossible to distinguish whether the CRA effect is due to a particular time in foreclosure process or a particular state.

the first panel, we use the RealtyTrac classification of judicial vs. non-judicial states to define high and low cost subsamples.[56] Column 1 shows estimates for the judicial subsample without the controls, while column 2 presents CRA impact coefficients for the same sample with all controls. In this specification and elsewhere, we omit the judicial dummy variable from the controls when the sample is split into high and low foreclosure cost subsamples. None of the estimates for the judicial sample is different from zero statistically. Columns 3 and 4 outline similar estimates but for the non-judicial subsample. They are positive and statistically significant at the 10 and 5 percent levels. The estimates imply that the CRA increased residential mortgage foreclosures in non-judicial states by 28 to 33 percent over 2003-2010.

As can be seen from panels 2 and 3 of Table 4-3, the impact of the CRA is robust to the alternative classifications of the judicial foreclosure states. For instance, Pence (2006) definition of judicial foreclosure states, panel 2, yields statistically significant estimates of the CRA impact in the non-judicial sample of around 29 to 31 percent and insignificant coefficients for the judicial sample. Panel 3, which depicts results obtained using the NCLC judicial state classification, shows significant CRA impact estimate of around 30 to 35 percent for the non-judicial group, and estimates which are not different from zero statistically for the judicial group.

As an additional robustness check, we use days in foreclosure process to split the sample into two groups, panel 4 of Table 4-3. The status of having more than 120 days in foreclosure is equivalent to high foreclosure costs (judicial) and less than 120 days - to low foreclosure costs (non-judicial).[57] However, the duration of the foreclosure process may capture a different as-

[56]Mian, Sufi, and Trebbi (2011) also use this classification.

[57]The mean of the duration in foreclosure is 145 days.

pect of foreclosure costs because some judicial states have a short duration of foreclosure and vice versa. The correlation between the RealtyTrac judicial status variable and the dummy variable indicating more than 120 days in foreclosure process is 0.3281.[58] As in the previous panels, all coefficients but one in the more foreclosure expensive sample are not different from zero statistically, while in the less expensive foreclosure sample they are positive and significant. The estimates suggest that the CRA increased foreclosures by 33 to 43 percent in the less costly (shorter duration of foreclosure) group of states. These estimates are a bit larger but very similar to the estimates in the previous panels of Table 4-3.

Table 4-4 reports coefficients when we split the study period into the before-crisis and during-crisis periods and judicial and non-judicial subsamples.[59] The first panel of Table 4-4 shows results for the subsample of judicial only states before and during the crisis. The CRA impact is not statistically significant either before or during the crisis. The second panel of Table 4-4 displays estimates for the non-judicial foreclosure subsample. The coefficients are positive and significant only before the crisis. The estimate implies that the CRA increased the number of foreclosures by around 44-47 percent before the crisis in non-judicial foreclosure states, but it had no significant impact during the crisis. The difference between the results before the crisis and during the crisis may occur because many CRA loans defaulted before the start of the crisis. So, the effect of the CRA on foreclosures after 2007 is not distinguishable from zero statistically. We also repeated regressions reported in Tables 4-1 to 4-4 for different bandwidths around the income ratio threshold

[58]The correlation matrix of the three judicial foreclosure dummy variables and the dummy variable indicating more than 120 days in foreclosure is shown in Table C7.

[59]Judicial or non-judicial as defined using the RealtyTrac classification.

(e.g., 0.77-0.83 and 0.78-0.82) and obtained qualitatively similar results.[60]

Overall, our data suggest that the credit expansion due to the Community Reinvestment Act increased the number of residential mortgage foreclosures in 2003-2010 in the states with lower costs of foreclosure such as non-judicial and short foreclosure process states. The precise effect of the CRA varies across different groups of regions and regression equation specifications, with a more conservative estimate of this effect being around 30 percent. While this effect may seem to be large, in practice, it implies around 3.3 additional foreclosures per average ZIP code per month due to the CRA. The average number of foreclosures per postal code in our sample is 111 per 100,000 homes. Since the average number of homes in a ZIP code in 2000 is 10000, there are about 11 foreclosures per code, so 30 percent of 11 is 3.3.

4.6 Conclusion

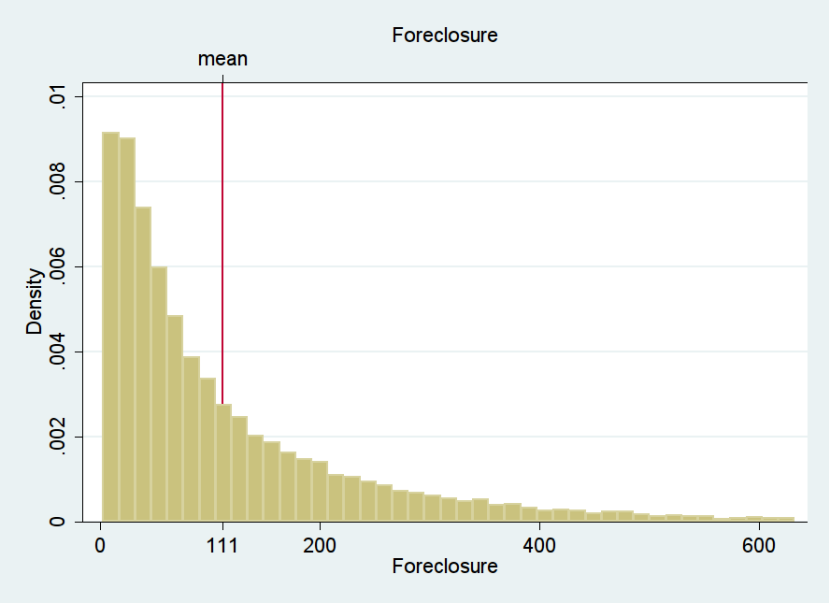
This paper proposes an estimation strategy to assess whether and how the Community Reinvestment Act of 1977 influenced US housing market and home mortgage foreclosures in 2003-2010. The proposed strategy relies on a regression discontinuity design which uses a CRA rule. This rule makes CRA eligible only those regions which have median family incomes below 80 percent of their MSA median family incomes and which are located inside assessment areas of CRA regulated institutions. Therefore, ZIP codes just above and just below the CRA eligibility threshold should be similar in their observable and unobservable attributes except for the CRA influence. This design allows us to make casual interpretations of any estimated difference between the number of mortgage foreclosures in these two groups.

[60] These results are not reported to preserve space.

We contribute to the literature by combining two separate and exogenous legal issues in mortgage credit. These issues are the Community Reinvestment Act and legal cost of foreclosure process. Using these two issues we are able to identify specific groups of lenders influenced by the CRA. This distinction allows us to explain a seeming contradiction in the literature, one branch of which argues that the CRA had no effect on mortgage defaults, while the other branch suggests that the CRA increased mortgage defaults. We find very weak evidence of the impact of the CRA on foreclosure for the whole sample, but a positive effect for particular groups of lenders.

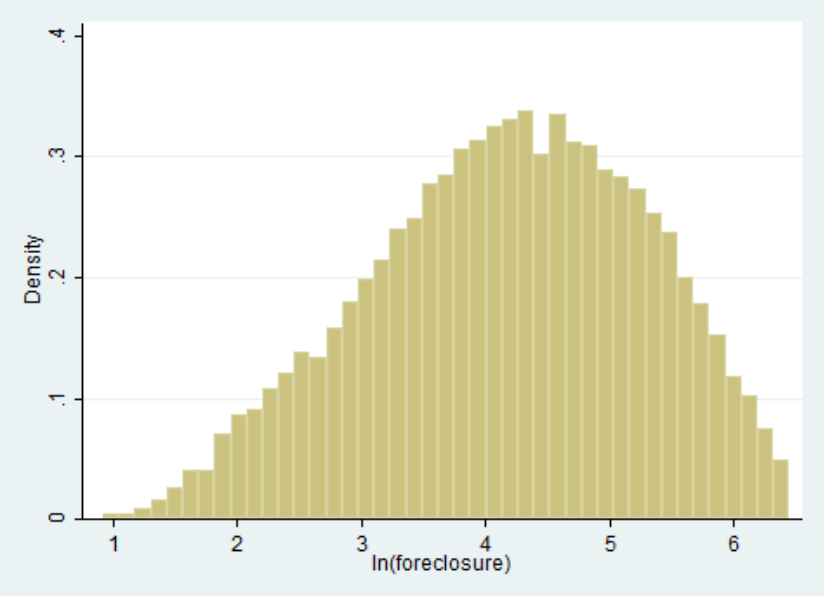
The data on the number of residential mortgage foreclosures per month across ZIP codes allow us to conclude that the CRA had different impacts in judicial and non-judicial foreclosure states. While the CRA effect in judicial states was not different from zero statistically, it was positive and significant in non-judicial states in 2003-2010. This discrepancy in the effects may be explained by less costly foreclosure process in non-judicial states which allowed financial institutions to have more incentives to lend to riskier borrowers. The estimated coefficients imply that the CRA on average increased the number of foreclosures per 100,000 homes by about 28 to 45 percent in 2003-2010.

Figure 4-1: A Histogram of the Number of Foreclosures per 100,000 Homes



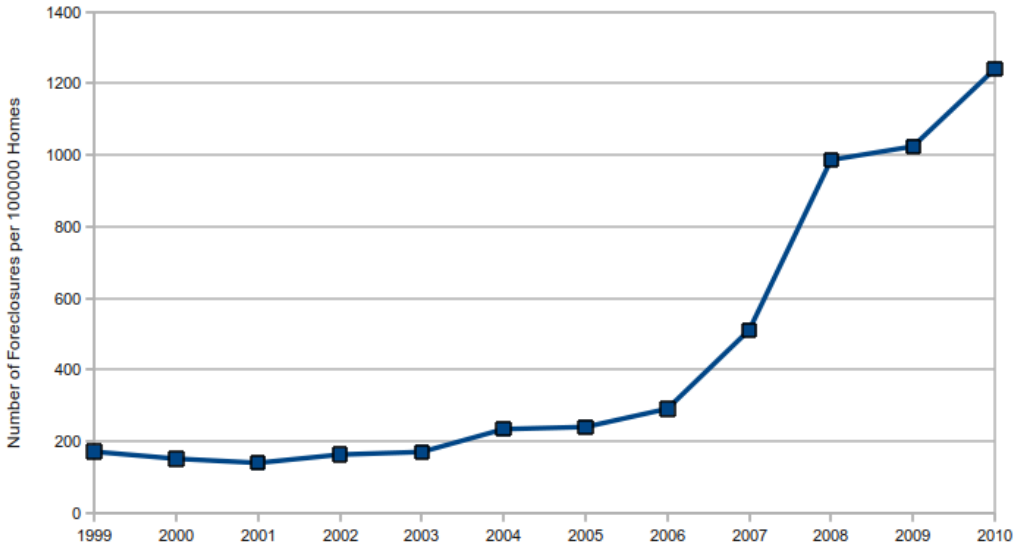
The income ratio is constrained to be between 0.75 and 0.85.

Figure 4-2: A Histogram of the Logarithm of the Number of Foreclosures per 100,000 Homes



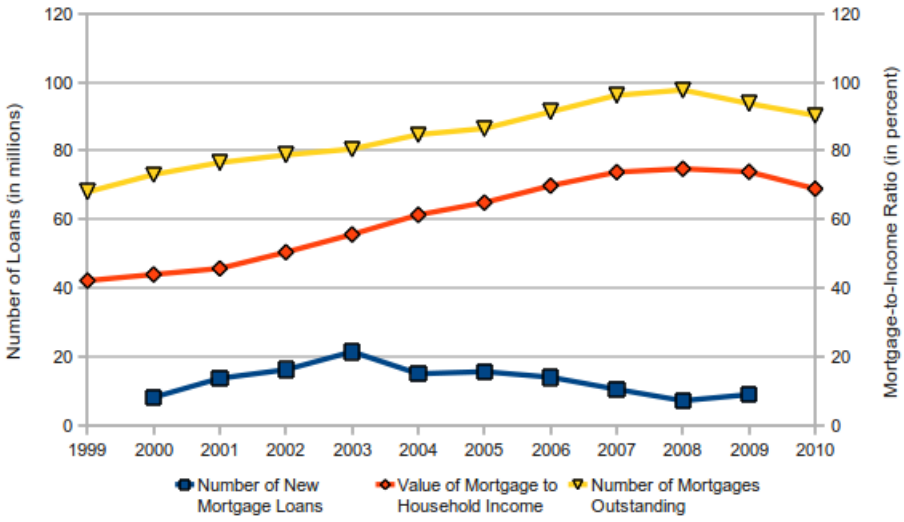
The income ratio is constrained to be between 0.75 and 0.85.

Figure 4-3: US National Foreclosure Trends in 2003-2010



Source: Zillow.com.

Figure 4-4: US National Mortgage Debt Trends in 2003-2010



Sources: New York Fed, Avery et al. (2010), and Bureau of Economic Analysis.

Table 4-1: CRA Impact on Foreclosures in 2003-2010

CRA impact in 2003-2010		
N	24529	24529
R^2	0.3956	0.5144
τ	0.215	0.2293*
s.e. of τ	0.1495	0.1341

CRA impact in 2003-2006		
N	10469	10469
R^2	0.371	0.4843
τ	0.2904	0.2335
s.e. of τ	0.177	0.147

CRA impact in 2007-2010		
N	14060	14060
R^2	0.4512	0.5315
τ	0.1502	0.147
s.e. of τ	0.1626	0.143
Polynom. order	2	2
Controls	N	Y

The dependent variable is the logarithm of the number of foreclosures per 100,000 homes at the ZIP code level. Specifications are as shown in equation (5). The treatment effect of the CRA is denoted τ and it is defined in equation (2). Controls as outlined in Table C1. Month and postal code clustered standard errors are reported. Observations with the number of foreclosures per 100,000 homes $>$ the mean plus 3 standard deviations are omitted as outliers. Only observations with the income ratio between 0.75 and 0.85 are included into the sample. Polynomial order with statistically significant results is reported, it may be dissimilar to the optimal polynomial order reported in Table C5. However, the order in Table C5 is only suggestive, not binding. Also, it applies to the full sample only and does not apply to any subsamples such as judicial or non-judicial, and before crisis or during crisis. ***, **, * denote statistical significance at 1, 5, and 10 %.

Table 4-2: CRA Impact on Foreclosures in States with Various Foreclosure Process before and during the crisis

	2003-2010		2003-2006		2007-2010	
Foreclosure Process < 100 days						
N	4917	4917	2102	2102	2815	2815
R^2	0.5146	0.6193	0.5611	0.6777	0.53	0.6264
τ	0.3822**	0.6108**	0.485**	0.3559	0.6609***	0.5559**
s.e. of τ	0.1854	0.2701	0.2298	0.2722	0.2547	0.2819
100 days ≤ Foreclosure Process < 120 days						
N	8065	8065	2899	2899	5166	5166
R^2	0.6572	0.751	0.1937	0.455	0.3466	0.5831
τ	0.0266	0.191	-0.0239	0.0471	0.1035	0.1013
s.e. of τ	0.1634	0.1591	0.2512	0.1552	0.2777	0.1887
120 days ≤ Foreclosure Process < 150 days						
N	3196	3196	1554	1554	1642	1642
R^2	0.3609	0.6153	0.2786	0.6153	0.3308	0.5388
τ	0.1564	0.0636	0.3494	-0.0169	0.0686	-0.3284
s.e. of τ	0.2687	0.2782	0.3027	0.2134	0.4581	0.3277
150 days ≤ Foreclosure Process < 210 days						
N	3445	3445	1705	1705	1740	1740
R^2	0.1572	0.4575	0.0914	0.3831	0.1978	0.5061
τ	-0.0801	-0.0838	0.0285	-0.1952	0.03	-0.2563
s.e. of τ	0.2595	0.329	0.2869	0.2266	0.383	0.2972
210 days ≤ Foreclosure Process						
N	4906	4906	2209	2209	2697	2697
R^2	0.3937	0.5419	0.3842	0.5766	0.4052	0.5572
τ	-0.1753	-0.1246	-0.2466	-0.5693**	0.0462	-0.0767
s.e. of τ	0.2992	0.3446	0.3137	0.2841	0.4536	0.3755
p-value	0.558	0.718	0.432	0.045	0.919	0.838
Pol. order	1	2	1	1	2	2
Controls	N	Y	N	Y	N	Y

The dependent variable is the logarithm of the number of foreclosures per 100,000 homes at the ZIP code level. Specifications are as shown in equation (5). Controls as outlined in Table C1. Judicial foreclosure state status dummy variable is not included into the controls. Month and postal code clustered standard errors are reported. Observations with the number of foreclosures per 100,000 homes > the mean plus 3 standard deviations are omitted as outliers. Only observations with the income ratio between 0.75 and 0.85 are included into the sample. ***, **, * denote statistical significance at 1, 5, and 10 %.

Table 4-3: CRA Impact on Foreclosures by different measures of foreclosure costs in 2003-2010

	High-cost		Low-cost	
	RealtyTrac judicial measure			
	Judicial		Non-Judicial	
N	7598	7598	16931	16931
R^2	0.3918	0.4567	0.443	0.5964
τ	0.0899	0.049	0.2781*	0.3219**
s.e. of τ	0.2672	0.2496	0.163	0.1383
	Pence (2006) measure			
	Judicial		Non-Judicial	
N	8577	8577	15952	15952
R^2	0.3678	0.4361	0.4699	0.6131
τ	0.0897	0.0897	0.2892*	0.309**
s.e. of τ	0.2661	0.2381	0.1647	0.1388
	NCLC measure			
	Judicial		Non-Judicial	
N	6453	6453	18076	18076
R^2	0.4047	0.499	0.4465	0.5871
τ	0.0576	-0.0967	0.2953*	0.3543***
s.e. of τ	0.2997	0.2707	0.1529	0.1276
	Days in foreclosure process			
	over 120		less than 120	
N	11547	11547	12982	12982
R^2	0.3039	0.4392	0.5647	0.6643
τ	0.1696	0.0626	0.3259*	0.4261***
s.e. of τ	0.2242	0.1955	0.1832	0.1414
Polynom. order	2	2	2	2
Controls	N	Y	N	Y

The dependent variable is the logarithm of the number of foreclosures per 100,000 homes at the ZIP code level. Specifications are as shown in equation (5). Controls as outlined in Table C1. Judicial foreclosure state status dummy variable is not included into the controls. Month and postal code clustered standard errors are reported. Observations with the number of foreclosures per 100,000 homes > the mean plus 3 standard deviations are omitted as outliers. Only observations with the income ratio between 0.75 and 0.85 are included into the sample. ***, **, * denote statistical significance at 1, 5, and 10 %.

Table 4-4: CRA Impact on Foreclosures before and during the crisis and according to the judicial state status in 2003-2010

	CRA impact in 2003-2006		CRA impact in 2007-2010	
	Judicial Only (RealtyTrac)			
N	3326	3326	4272	4272
R^2	0.4016	0.5045	0.3922	0.4616
τ	0.011	-0.1845	0.1478	0.205
s.e. of τ	0.3045	0.2546	0.2808	0.2779
	Non-Judicial Only (RealtyTrac)			
N	7143	7143	9788	9788
R^2	0.3607	0.5014	0.4053	0.5469
τ	0.466**	0.4418***	0.1395	0.1588
s.e. of τ	0.2053	0.1702	0.1879	0.1445
Polynom. order	2	2	2	2
Controls	N	Y	N	Y

The dependent variable is the logarithm of the number of foreclosures per 100,000 homes at the ZIP code level. Specifications are as shown in equation (5). Controls as outlined in Table C1. Judicial foreclosure state status dummy variable is not included into the controls. Month and postal code clustered standard errors are reported. Observations with the number of foreclosures per 100,000 homes > the mean plus 3 standard deviations are omitted as outliers. Only observations with the income ratio between 0.75 and 0.85 are included into the sample. ***, **, * denote statistical significance at 1, 5, and 10 %.

Bibliography

12 US Code. Sec. 2901.

Agarwal, S., Benmelech, E., Bergman, N., and A. Seru (2012). "Did the Community Reinvestment Act (CRA) Lead to Risky Lending?" National Bureau of Economic Research Working Paper No. 18609.

Angrist, J., and J. Pischke (2009). *Mostly harmless econometrics: An empiricist's companion.* Princeton: Princeton University Press.

Avery, R., Bhutta, N., Brevoort, K., Canner, G., and C. Gibbs (2010). "The 2008 HMDA Data: The Mortgage Market during a Turbulent Year," *Federal Reserve Bulletin, 96*, A169-A211.

Avery, R., and K. Brevoort (2011). "The Subprime Crisis: Is Government Housing Policy to Blame?" Social Science Research Network Working Paper No. 1726192.

Berry, C., and S. Lee (2007). "The Community Reinvestment Act: A regression discontinuity analysis," Harris School Working Paper 07.04.

Bhutta, B. (2011). "The Community Reinvestment Act and Mortgage Lending to Lower Income Borrowers and Neighborhoods," *Journal of Law and Economics, 54*(4), 953-983.

Canner, G., and N. Bhutta (2008). "Staff Analysis of the Relationship between the CRA and the Subprime Crisis," Board of Governors of the Federal Reserve System.

Clauretie, T., and T. Herzog (1990). "The Effect of State Foreclosure Laws on Loan Losses: Evidence from the Mortgage Insurance Industry," *Journal*

of Money, Credit and Banking, 22 (2), 221-233.

Demyanyk, Y., and O. Van Hemert (2011). "Understanding the subprime mortgage crisis," *The Review of Financial Studies, 24*(6): 1848-1880.

Ding, L., Quercia, R., Ratcliffe, J., and W. Li (2008). "Risky borrowers or risky mortgages: Disaggregating effects using propensity score models," Working Paper, Center for Community Capital, University of North Carolina.

Joint Center for Housing Studies (2002). "The 25th Anniversary of the Community Reinvestment Act: Access to Capital in an Evolving Financial Services System." Ford Foundation Sponsored Report.

Karikari, J. (2009). "The CRA and Subprime Lending in Neighborhoods: Myth or Reality?" Social Science Research Network Working Paper No. 1448831.

Laderman, E., and C. Reid (2008). "Lending in low- and moderate-income neighborhoods in California: The performance of CRA lending during the subprime meltdown," Federal Reserve Bank of San Francisco Working Paper No. 2008-05.

Lee, D., and T. Lemieux (2010). "Regression discontinuity designs in economics," *Journal of Economic Literature, 48*(2), 281-355.

Lin, Z., Rosenblatt, E., and V. Yao (2009). "Spillover effects of foreclosures on neighborhood property values," *Journal of Real Estate Finance and Economics, 38*, 387-407.

Mian, A., Sufi, A., and F. Trebbi (2011). "Foreclosures, house prices, and the

real economy," National Bureau of Economic Research Working Paper No. 16685.

Pence K. (2006). "Foreclosing on Opportunity: State Laws and Mortgage Credit," *The Review of Economics and Statistics, 88*(1), 177-182.

Petersen, M. (2009). "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches," *The Review of Financial Studies, 22*(1), 435-480.

Quarterly Report on Household Debt and Credit (2011). Federal Reserve Bank of New York.

Rajan, R. (2010). *Fault Lines: How Hidden Fractures Still Threaten the World Economy.* Princeton: Princeton University Press.

Schill, M., and S. Wachter (1994). "Borrower and neighborhood racial and income characteristics and financial institution mortgage application screening," *Journal of Real Estate Finance and Economics, 9*, 223-239.

Shlay, A. (2006). "Low-income homeownership: American dream or delusion?" *Urban Studies, 43*(3), 511-531.

Appendix C: Detailed Data Description, Results, Figures and Tables

The Dependent Variable

Figure C1 portrays the number of foreclosures per 100,000 homes per month across postal codes against the income ratio. This ratio is defined as the ZIP code median family income divided by Metropolitan Statistical Area (MSA) median family income. As required by the CRA, estimates of these incomes from the 2000 US Census are used to define the ratio and CRA eligibility status. Even though the CRA defines low and moderate income census tracts based on this income ratio, the same methodology is used in this paper to define low and moderate income ZIP codes. These ZIP codes are uniquely matched to census tracts using crosswalk files provided by the HUD. Each dot from Figure C1 represents one observation of a ZIP code in a month. This Figure shows that the number of foreclosures is a declining function of the income ratio.

Figure C1 also highlights that some ZIP codes tend to have extraordinary foreclosure values in certain months. This can be due to the factors not influential for the majority of the ZIP codes. Therefore, all observations with foreclosures higher than this series mean plus three standard deviations are omitted from further analysis. We delete only one observation where the number of foreclosures for a certain postal code in a month is outside of the range, but keep all other observations for this ZIP code in the other months unless they are over the limit as well. This operation results in 3069 out of 208378 (or about 1.47 percent) observations being dropped.

One might expect that the number of foreclosures per 1000 mortgages is a better dependent variable than the number of foreclosures per 1000 houses.

This is because foreclosure is a loan outcome not a house outcome and number of mortgages per house may differ across ZIP codes and over time. In an effort to account for the varying prevalence of mortgages across observations and extraneous exogenous variables, we use a set of control variables. Those are discussed later in this Appendix.

Legitimacy of the Regression Discontinuity Model

Figure C2, which depicts the density of the income ratio between 0.4 and 1.2, indicates that the Regression Discontinuity design is appropriate to use on our data. As argued by Lee and Lemieux (2010), no discontinuity in the density of the assignment variable, which is the income ratio, should be present at the threshold to make sure that subjects cannot perfectly manipulate the assignment to the treatment and control groups. Dots on Figure C2 represent individual densities, the sum of heights of all dots is equal to 1. In order to compute them, we first partition the income ratio range (0.4 to 1.2) into 200 equal intervals or bins with the width of 0.004. After that, the number of observations from the original data in every interval is counted and the result is divided by the total number of observations in the entire sample. In this way, each dot is obtained. Figure C2 shows that there is no discontinuity in the density of the income ratio at 0.8.

Figure C3 and C4 outline the relation between the ZIP-to-MSA income ratio and the logarithm of the number of foreclosures per 100,000 homes per month across ZIP codes. Instead of showing all data points, this and subsequent graphs show as dots local averages computed using non-overlapping intervals or bins of the original data. These bins are used to reduce the number of dots only so that it is possible to get a better sense of the distribution

of the observations and not to get confused by a glob of points. These bins are subsamples of the data in between two values of the income ratio. For instance, Figure C3 has 40 bins at each side of the cutoff, this means that the first dot is the average of the logarithm of foreclosures in regions where the income ratio is between 0.4 and 0.41, the second dot is the average logarithm of foreclosures for all observations with the income ratio between 0.41 and 0.42 and so on.

Figure C3 presents a large picture of the relation between the income ratio and the logarithm of foreclosures. Looking at the dots, no observable shift in the function at the threshold could be noticed. Figure C4, which portrays a closer picture of the relation, appears to suggest that the function is non-linear. Local averages (dots) do not seem to imply any discontinuities. Of course, a formal statistical test of the discontinuity such as regression analysis is required before making any conclusions. It should be noted that discontinuities for various subsamples considered may be dissimilar to the pictures presented here, but they are not included to preserve space.

Another way to show that a Regression Discontinuity design can be used to analyze the foreclosure data is to compare observable covariates in the control and treatment groups. Such comparison could be done by looking at Tables C3 and C4. It can be noted from these Tables that for most variables the mean and the standard deviation are of very similar magnitude in both groups. Figures D1-D24 also show no discontinuities in the control variables at the income ratio of 0.8.

Bandwidth and Polynomial Order

As suggested by Lee and Lemieux (2010), the sensitivity of results to a range of bandwidth and polynomial order is explored. The results of this

exercise are summarized in Table C5. Bin dummies, where the width of the bin is equal to 0.01, are included into equation (4) to help choose the polynomial order. This order should be increased as long as the bin dummies are jointly different from zero. The optimal order of the polynomial is then chosen among the specifications with statistically significant bin dummies using the Akaike's criterion. It can be noted from Table C5 that in most models the effect of the CRA on foreclosures is not different from zero statistically. When the effect is significant, it is positive. The effect of the CRA is the only thing which matters in this exercise. The optimal order of polynomial is used in this Table only and the bandwidth is chosen by the researcher. The results reported in this Table are obtained based on all observations available for a particular bandwidth. Their quantity is indicated in the row of the Table named 'N' which means number of observations.

Control Variables

The four sets of controls described in Table C1 are designed to capture the effect of mortgage growth and other factors influencing residential mortgage foreclosures. In particular, the first set of variables controls for the mortgage to house ratio. This ratio is not available directly, but it may influence foreclosures. The second set of covariates consists of demographic characteristics of the ZIP and may help explain mortgage lending and foreclosures. The third set of variables represent local economic conditions which influence borrower's ability to carry a mortgage and avoid foreclosure. Finally, the last set of controls includes some demographic and legislative factors that may affect foreclosures. In particular, minority share and proportion of population in

group quarters[61] might have an impact on foreclosures. The judicial foreclosure state status could also reduce foreclosures. This variable is only included into the controls when there are no splits based on the judicial variable, its alternative definitions or the duration in foreclosure process variables. Summary statistics for those variables may be found in Table C1. We decide to exclude from further analysis those ZIP codes where the number of owner-occupied houses in 2000 to owner-occupied houses in 1990 is greater than 8. It is very likely that large values of this variable are obtained due to redrawing postal code borders.

Table C6 presents estimation results with different bundles of control variables added progressively to the specification in equation (5). The first two columns of this Table show coefficients and standard errors of the treatment effect (LMI) and the second order polynomial of the income ratio.[62] The coefficients for the income ratio and the square of the ratio indicate that the logarithm of foreclosures declines with the income ratio. The same shape may be spotted in Figure C1. The interactive terms of the treatment indicator (LMI) and income ratio as well as income ratio squared suggest that slope coefficients are different to the left and right of the cutoff. In particular, the coefficient on the income ratio is higher by 5.64 to the right and the coefficient of the income ratio squared is higher by 522.9 to the right. Only the last result is statistically different from zero. These coefficients are jointly significant as indicated by the p-value of the F-statistic.

Results with several controls added are shown in column 3 of Table C6. The treatment indicator is still not significant and equal to 0.217, other results

[61]Those include armed forces barracks, educational institution residences, hospitals and nursing homes, prisons, and dormitories of various kinds.

[62]See equation (5) for the definition of the order of polynomial.

are similar to the previous columns. Most coefficients on the control variables have expected signs. For instance, the positive coefficient on the proportion of owner occupied houses means that regions with more homeowners receive more foreclosures. Also, more foreclosures happen in regions with more mortgages per owner-occupied house. The impact of the proportion of mortgages in 2000 to mortgages in 1990 is negative and significant indicating that loans made in the 1990s were probably less risky than those in the 2000s or, maybe, most risky of those loans defaulted or prepaid before 2003. The coefficient on the proportion of houses built in 1980-89 is positive and shows that this segment of the real estate market received more foreclosures in 2003-2010.

Column 5 of Table C6 shows that the effect of the median age of population is positive because more mortgages are obtained by persons in the middle of their life. Proportion of population of age over 65 has a negative coefficient implying that less mortgages and foreclosures occur in regions with this population group since it usually do not get regular mortgages. Post-secondary education seems to reduce foreclosures, maybe, because it captures the effect of financial literacy. The other variables have similar estimates as those in column 3 of this Table and their interpretation is unchanged. Coefficients of this bundle of controls are jointly different from zero as can be seen from the p-value of the F-statistic.

House values in 2000 and their growth in 2003-2010 have negative coefficients suggesting that higher house prices reduce foreclosures due to less negative home equity, column 7 of Table C6. Wage growth and unemployment generate negative and positive effects on foreclosures, respectively, as expected. The impact of the annual growth in housing units is difficult to predict in advance. However, it seems that regions with higher growth in housing

and therefore mortgages before 2007 experienced more foreclosures, while the growth after 2007 has a negative impact on foreclosures. This can be due to the tightening of lending standards during and after the crisis.

Column 9 of Table C6 reports results with final three controls. The coefficient on the judicial indicator is positive which is counterintuitive because foreclosures seem to be less likely in judicial states. However, when alternatives measures of judicial status from Pence (2006) or NCLC are used, the coefficient is negative and significant. So, those alternative measures are probably capturing judicial status better. Proportion of population Hispanic and share of population in group quarters have no significant impact on foreclosures.

Figure C1: Relation between Income Ratio and the Number of Foreclosures per 100,000 Homes per Month across ZIP codes

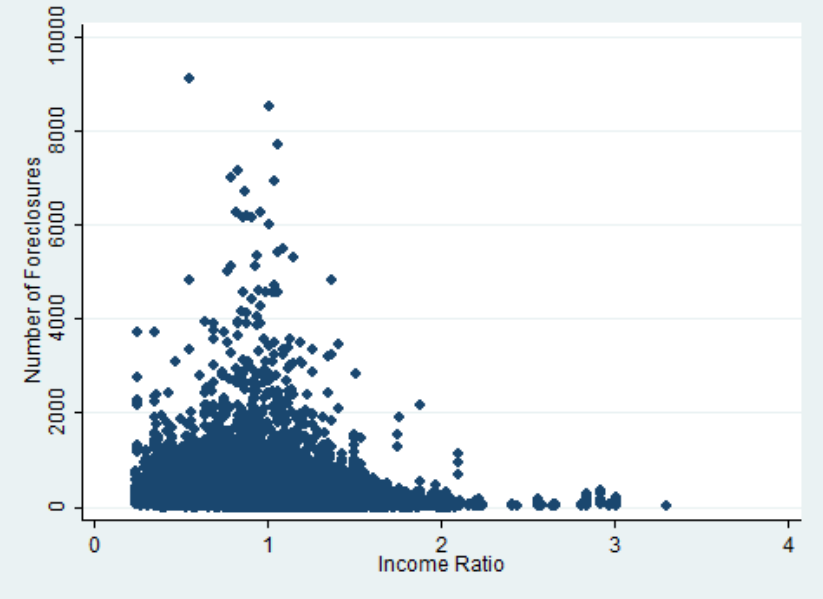
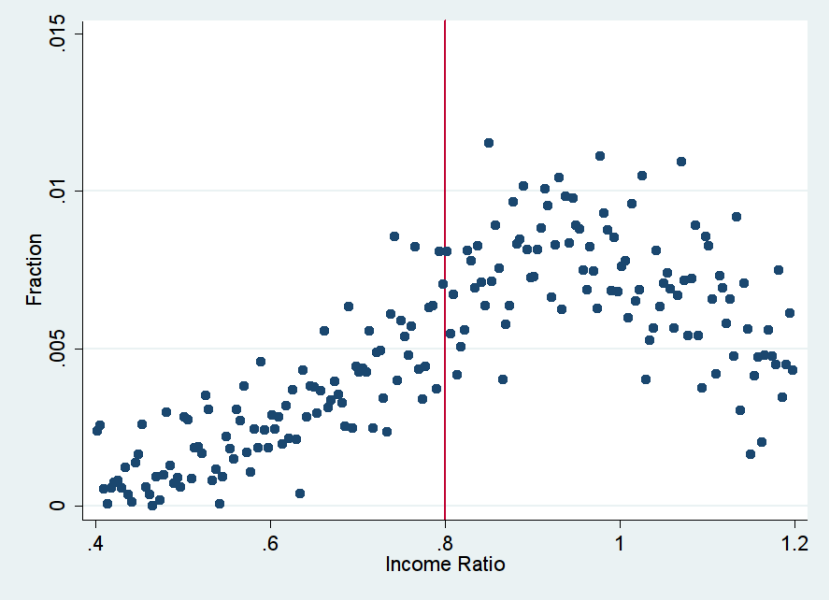
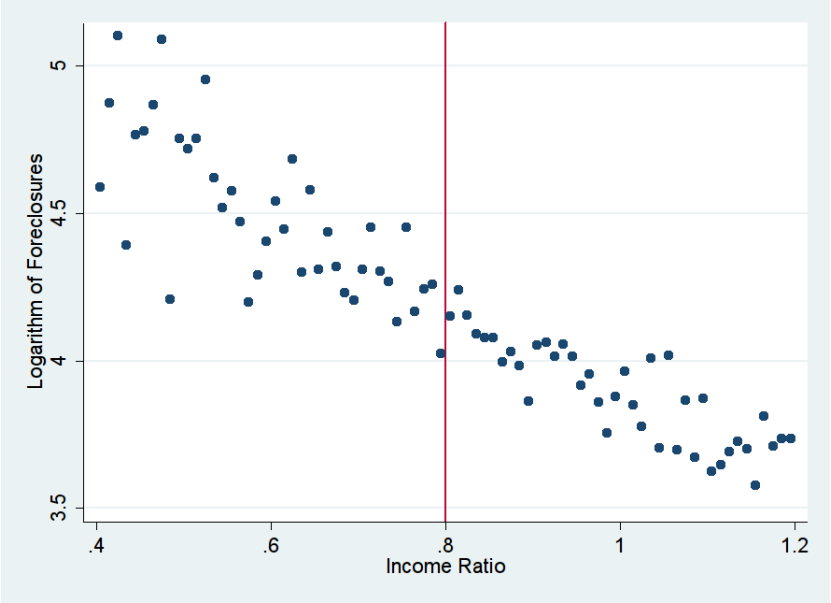


Figure C2: Density of the Income Ratio



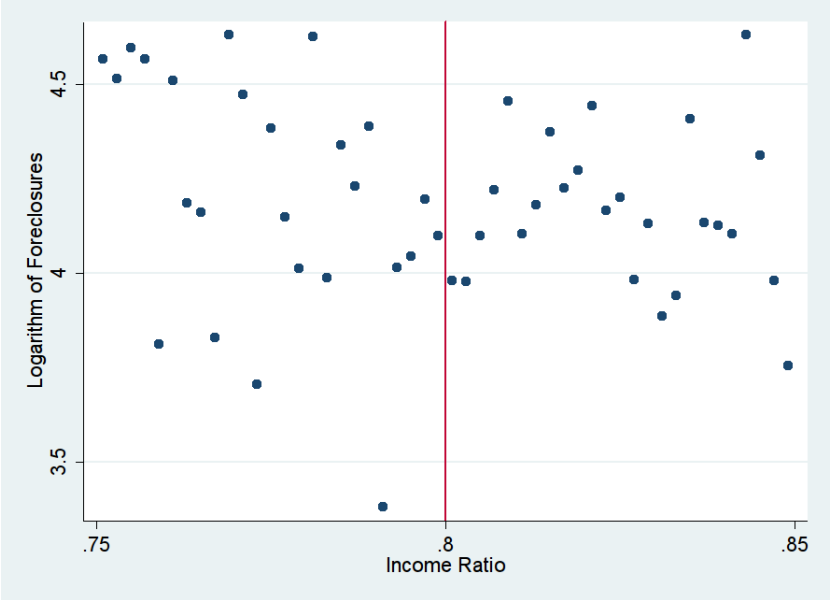
The income ratio is constrained to be between 0.4 and 1.2.

Figure C3: Relation between Income Ratio and the Percent of Homes Foreclosed per Month across ZIP codes



The income ratio is constrained to be between 0.4 and 1.2.

Figure C4: Relation between Income Ratio and the Percent of Homes Foreclosed per Month across ZIP codes



The income ratio is constrained to be between 0.75 and 0.85.

Table C1: Summary Statistics

Variable	N	Mean	Std. Dev.	Min	Max	Units of measurement
Dependent Variable						
Logarithm of foreclosures per 100,000 homes	24529	4.1714	1.1159	0.9163	6.4497	number
1-st Set of Controls						
Proportion of owner-occupied houses	24529	0.6077	0.1405	0.0694	0.9417	fraction of 1
Proportion of mortgaged owner-occupied houses	24529	0.5780	0.1307	0.0992	0.8764	fraction of 1
Mortgages in 2000 to mortgages in 1990	24529	1.2850	0.6390	0.4085	16.5	fraction
Number of owner-occupied houses in 2000 to owner-occupied houses in 1990	24529	1.5412	0.8956	0.4964	7.6429	fraction
Proportion of houses built in 1980-89	24529	0.1367	0.0945	0.0065	0.6122	fraction of 1
Proportion of houses built in 1940-69	24529	0.4113	0.1708	0.006	0.8548	fraction of 1
Proportion of houses detached	24529	0.5715	0.1831	0.0214	0.9793	fraction of 1
2-nd Set of Controls						
Median age of population	24529	34.6289	5.6597	23.3	74.5	years
Proportion of population of age over 65	24529	0.1268	0.0666	0.0338	0.7920	fraction of 1
% of population with post-secondary education	24529	45.7291	11.2909	14.2857	79.7575	percent
3-rd Set of Controls						
Growth of real average weekly wages	24529	-0.0012	0.0698	-0.4228	0.2985	fraction
Unemployment rate	24529	6.4873	2.6652	2.5	22.1	percent
Median value of owner-occupied houses	24529	122072	62814	38800	506700	dollars
Real house value growth	24529	-0.0028	0.0142	-0.0755	0.0667	fraction
Growth of housing units in 2003	24529	0.0126	0.0105	-0.0032	0.0633	fraction
Growth of housing units in 2004	24529	0.0120	0.0117	-0.0030	0.0575	fraction
Growth of housing units in 2005	24529	0.0109	0.0121	-0.0029	0.0687	fraction
Growth of housing units in 2006	24529	0.0089	0.0118	0.0000	0.0617	fraction
Growth of housing units in 2007	24529	0.0050	0.0082	-0.0003	0.0423	fraction
Growth of housing units in 2008	24529	0.0022	0.0050	-0.0017	0.0382	fraction
Growth of housing units in 2009	24529	0.0004	0.0017	-0.0011	0.0230	fraction
4-th Set of Controls						
Proportion of population Hispanic	24529	0.1751	0.2107	0	0.9202	fraction of 1
Proportion of population in group quarters	24529	0.0237	0.0461	0	0.6076	fraction of 1
Judicial foreclosure state status	24529	0.3098	0.4624	0	1	0 or 1

The income ratio is constrained to be between 0.75 and 0.85. Observations with the number of foreclosures per 100,000 homes > the mean plus 3 standard deviations are omitted as outliers.

Table C2: Description of Variables

Variable	Frequency	Time period available	Aggregation	Data Source
Logarithm of foreclosures per 100,000 homes	monthly	Jan2003-Jun2010	ZIP	Zillow.com
Median age of population	one observation	2000	ZIP	US Census
Proportion of population of age over 65	one observation	2000	ZIP	US Census
% of population with post-secondary education	one observation	2000	ZIP	US Census
Growth of housing units in 2003	one observation	2003	county	US Census
Growth of housing units in 2004	one observation	2004	county	US Census
Growth of housing units in 2005	one observation	2005	county	US Census
Growth of housing units in 2006	one observation	2006	county	US Census
Growth of housing units in 2007	one observation	2007	county	US Census
Growth of housing units in 2008	one observation	2008	county	US Census
Growth of housing units in 2009	one observation	2009	county	US Census
Growth of real average weekly wages	quarterly	q1:2003-q2:2010	county	BLS
Unemployment rate	monthly	Jan2003-Jun2010	MSA	BLS
Judicial foreclosure state status	one observation	2010	ZIP	RealtyTrac.com
Proportion of owner-occupied houses	one observation	2000	ZIP	US Census
Proportion of mortgaged owner-occupied houses	one observation	2000	ZIP	US Census
Mortgages in 2000 to mortgages in 1990	one observation	2000	ZIP	US Census
Number of owner-occupied houses in 2000 to owner-occupied houses in 1990	one observation	2000	ZIP	US Census
Median value of owner-occupied houses	one observation	2000	ZIP	US Census
Real house value growth	monthly	Jan2003-Jun2010	ZIP	Zillow.com
Proportion of houses detached	one observation	2000	ZIP	US Census
Proportion of houses built in 1980-89	one observation	2000	ZIP	US Census
Proportion of houses built in 1940-69	one observation	2000	ZIP	US Census
Proportion of population in group quarters	one observation	2000	ZIP	US Census
Proportion of population Hispanic	one observation	2000	ZIP	US Census

Table C3: Summary Statistics for the Treatment Group (0.75≤Income Ratio <0.80)

Variable	N	Mean	Std. Dev.	Min	Max
Logarithm of foreclosures per 100,000 homes	11149	4.2156	1.0994	0.9163	6.450
Median age of population	11149	34.7499	6.2161	23.3000	74.5
Proportion of population of age over 65	11149	0.1328	0.0779	0.0362	0.7920
% of population with post-secondary education	11149	44.6383	11.1208	17.4670	72.7197
Growth of housing units in 2003	11149	0.0118	0.0101	-0.0032	0.0565
Growth of housing units in 2004	11149	0.0114	0.0114	-0.0030	0.0531
Growth of housing units in 2005	11149	0.0104	0.0117	-0.0029	0.0501
Growth of housing units in 2006	11149	0.0084	0.0115	0.0000	0.0536
Growth of housing units in 2007	11149	0.0047	0.0078	-0.0003	0.0405
Growth of housing units in 2008	11149	0.0020	0.0045	-0.0017	0.0275
Growth of housing units in 2009	11149	0.0003	0.0015	-0.0011	0.0140
Growth of real average weekly wages	11149	-0.0009	0.0689	-0.4228	0.2985
Unemployment rate	11149	6.4312	2.5569	2.7000	22.1
Judicial foreclosure state status	11149	0.3374	0.4729	0.0000	1
Proportion of owner-occupied houses	11149	0.5857	0.1382	0.1206	0.9281
Proportion of mortgaged owner-occupied houses	11149	0.5645	0.1369	0.1214	0.8589
Mortgages in 2000 to mortgages in 1990	11149	1.2308	0.3763	0.4085	3.5333
Number of owner-occupied houses in 2000 to owner-occupied houses in 1990	11149	1.4932	0.9490	0.4964	7.5589
Median value of owner-occupied houses	11149	113726	55317.94	38800	466700
Real house value growth	11149	-0.0025	0.0143	-0.0717	0.0667
Proportion of houses detached	11149	0.5463	0.1858	0.0235	0.9641
Proportion of houses built in 1980-89	11149	0.1265	0.0875	0.0065	0.3950
Proportion of houses built in 1940-69	11149	0.4316	0.1646	0.0583	0.8253
Proportion of population in group quarters	11149	0.0270	0.0571	0	0.6076
Proportion of population Hispanic	11149	0.1797	0.2163	0	0.9202

Observations with the number of foreclosures per 100,000 homes > the mean plus 3 standard deviations are omitted as outliers.

Table C4: Summary Statistics for the Control Group (0.80≤Income Ratio ≤0.85)

Variable	Obs	Mean	Std. Dev.	Min	Max
Logarithm of foreclosures per 100,000 homes	13380	4.134569	1.128123	1.029619	6.44968
Median age of population	13380	34.52802	5.148518	23.6	69.3
Proportion of population of age over 65	13380	0.1217971	0.0550037	0.0337587	0.5833831
% of population with post-secondary education	13380	46.63795	11.35127	14.28571	79.75745
Growth of housing units in 2003	13380	0.0131949	0.0108022	-0.0017638	0.0632853
Growth of housing units in 2004	13380	0.0124756	0.0119529	0	0.0575032
Growth of housing units in 2005	13380	0.0113302	0.0123286	0	0.0687127
Growth of housing units in 2006	13380	0.0093401	0.0119835	0	0.0616785
Growth of housing units in 2007	13380	0.0052934	0.0085295	-0.0000161	0.0422692
Growth of housing units in 2008	13380	0.0023723	0.0053444	-0.000697	0.0381685
Growth of housing units in 2009	13380	0.0004091	0.0018029	-0.0011408	0.0229983
Growth of real average weekly wages	13380	-0.0013528	0.0704471	-0.4227614	0.2508155
Unemployment rate	13380	6.534066	2.751386	2.5	22.1
Judicial foreclosure state status	13380	0.2866966	0.4522355	0	1
Proportion of owner-occupied houses	13380	0.6260677	0.1397232	0.069435	0.9417346
Proportion of mortgaged owner-occupied houses	13380	0.5893504	0.1241949	0.0991652	0.8763717
Mortgages in 2000 to mortgages in 1990	13380	1.330133	0.7913106	0.5822342	16.5
Number of owner-occupied houses in 2000 to owner-occupied houses in 1990	13380	1.581276	0.8465105	0.5381816	7.642857
Median value of owner-occupied houses	13380	129026.4	67655.51	48700	506700
Real house value growth	13380	-0.0030192	0.0142036	-0.0754671	0.0619221
Proportion of houses detached	13380	0.5923769	0.1781417	0.0214256	0.9792627
Proportion of houses built in 1980-89	13380	0.1452654	0.0991302	0.0082645	0.6121988
Proportion of houses built in 1940-69	13380	0.3944597	0.1740758	0.0060241	0.8547816
Proportion of population in group quarters	13380	0.020923	0.0341773	0	0.252495
Proportion of population Hispanic	13380	0.1713797	0.2059449	0.002267	0.860988

Observations with the number of foreclosures per 100,000 homes > the mean plus 3 standard deviations are omitted as outliers.

Table C5: Sensitivity of the CRA Impact to the Bandwidth and Order of Polynomial Choice

	Bandwidth	0.10	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01
Pol. Order											
	N	49449	44394	39425	34454	30694	24529	19647	14247	9679	5433
Zero	τ	0.2286**	0.372***	0.0207	0.0018	0.3338***	0.1152	0.0272	-0.08	0.1667	0.0591
	s.e. of τ	0.1094	0.1313	0.1231	0.1208	0.1201	0.1146	0.1199	0.1321	0.1132	0.1386
	p-value	0.1578	0.5386	0.6179	0.4782	0.3588	0.4304	0.7812	0.5132	0.3618	
1-st	τ	-0.2229	-0.4334	-0.3470	-0.3406	-0.2169	-0.1361	0.0625	0.3723	0.6479**	0.4541**
	s.e. of τ	0.7139	0.9364	0.2932	0.312	0.5474	0.346	0.293	0.4482	0.3053	0.2097
	p-value	0.5394	0.663	0.5243	0.3736	0.2621	0.6165	0.6924	0.3823	0.2292	
2-nd	τ	-0.0696	-0.1534	0.1367	0.3121	0.2851	0.5945	0.7369*	0.6313	0.8301**	0.3022
	s.e. of τ	0.7996	0.969	0.431	0.451	0.557	0.4001	0.4116	0.4887	0.3538	0.2976
	p-value	0.5451	0.5866	0.3112	0.1668	0.1043	0.3548	0.2249	0.5934	0.2332	
3-rd	τ	-0.2951	-0.4176	0.5075	0.5373	-0.0912	0.5551	0.6877	0.6053	0.8076**	-0.2637
	s.e. of τ	0.8528	0.9524	0.4705	0.4733	0.5981	0.4037	0.4246	0.4955	0.3516	0.3481
	p-value	0.2042	0.0554	0.0678	0.0498	0.08	0.4091	0.4174	0.6008	0.0354	
Optimal Order		0	0	0	0	0	0	0	0	0	4

The dependent variable is the logarithm of the number of foreclosures per 100,000 homes at the ZIP code level. Specifications are as shown in equation (5). Month and postal code clustered standard errors are reported. Observations with the number of foreclosures per 100,000 homes > the mean plus 3 standard deviations are omitted as outliers. Bandwidth of 0.10 means observations for which 0.70≤ income ratio≤0.90. The optimal polynomial order is chosen using Akaike's criterion, which is defined as $AIC = (-2ln(L) + 2p)/N$, where L is the likelihood function, p is the number of parameters in the model, and N is the number of observations. Reported p-value is for the hypothesis of the joint significance of the bin dummy variables, where bin width is 0.01. ***, **, * denote statistical significance at 1, 5, and 10 %.

Table C6: Effects of the controls in 2003-2010

Variables	(1) coefficient	(2) s.e.	(3) coefficient	(4) s.e.	(5) coefficient	(6) s.e.
LMI status	0.215	0.150	0.217	0.150	0.184	0.146
Income ratio	7.978	8.520	5.571	8.441	4.186	8.302
LMI*Inc. Ratio	5.642	12.57	5.684	12.51	5.931	12.50
Inc. Ratio^2	-149.9	159.7	-117.8	157.9	-95.92	154.2
LMI*Inc. Ratio^2	522.9**	235.1	436.3*	228.3	404.1*	223.0
Proportion of owner-occupied houses			0.0976	0.311	0.145	0.341
Proportion of mortgaged owner-occupied houses			0.173	0.311	-0.348	0.324
Mortgages in 2000 to mortgages in 1990			-0.0976***	0.0362	-0.0563	0.0342
Owner-occupied houses in 2000 to these houses in 1990			0.136**	0.0580	0.0432	0.0494
Proportion of houses built in 1980-89			0.832*	0.500	1.070***	0.408
Proportion of houses built in 1940-69			0.0147	0.236	0.144	0.226
Proportion of houses detached			0.629**	0.252	0.503**	0.243
Median age of population					0.0193	0.0133
Proportion of population of age over 65					-3.268***	0.931
% of population with post-secondary education					-0.00733**	0.00324
Constant	3.710***	0.130	3.006***	0.326	3.533***	0.398
R-squared	0.396		0.410		0.423	
p-value of F-stat	0.0047		0.0002		0	

The dependent variable is the logarithm of the number of foreclosures per 100,000 homes at the ZIP code level. Specifications are as shown in equation (5). Month and postal code clustered standard errors are reported. The income ratio is constrained to be between 0.75 and 0.85. Observations with the number of foreclosures per 100,000 homes > the mean plus 3 standard deviations are omitted as outliers. F-statistic is for the hypothesis of joint significance of the added set of variables (all reported variables in the first column). ***, **, * denote statistical significance at 1, 5, and 10 %.

Table C6: continued

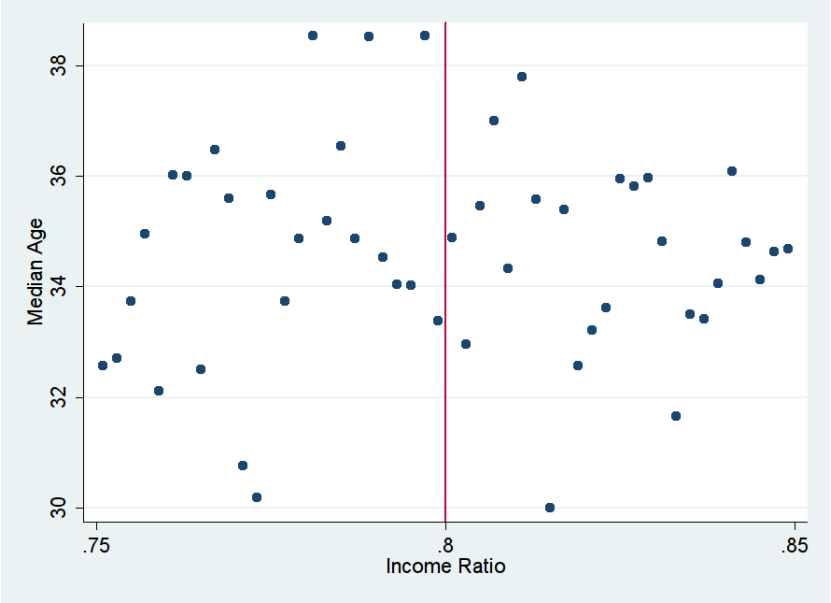
Variables	(7) coefficient	(8) s.e.	(9) coefficient	(10) s.e.
LMI status	0.231*	0.134	0.229*	0.134
Income ratio	7.167	7.542	7.006	7.549
LMI*Inc. Ratio	3.993	11.35	4.114	11.40
Inc. Ratio^2	-127.1	143.0	-125.0	142.9
LMI*Inc. Ratio^2	492.3**	203.2	488.0**	202.2
Proportion of owner-occupied houses	0.218	0.339	0.154	0.355
Proportion of mortgaged owner-occupied houses	-0.112	0.289	-0.128	0.290
Mortgages in 2000 to mortgages in 1990	-0.0529*	0.0321	-0.0543*	0.0329
Number of owner-occupied houses in 2000 to these houses in 1990	0.0464	0.0466	0.0486	0.0468
Proportion of houses built in 1980-89	0.796**	0.365	0.751**	0.373
Proportion of houses built in 1940-69	0.105	0.214	0.0932	0.217
Proportion of houses detached	0.204	0.250	0.219	0.249
Median age of population	0.0450***	0.0124	0.0439***	0.0131
Proportion of population of age over 65	-5.205***	0.918	-5.171***	0.930
% of population with post-secondary education	0.000551	0.00312	-0.000150	0.00387
Median value of owner-occupied houses	-1.48e-06***	5.28e-07	-1.47e-06***	5.32e-07
Real house value growth	-19.40***	1.708	-19.40***	1.710
Growth of real average weekly wages	-0.277	0.186	-0.275	0.187
Unemployment rate	0.162***	0.0192	0.162***	0.0192
Growth of housing units in 2003	0.453	4.578	0.450	4.583
Growth of housing units in 2004	-7.169	5.002	-6.811	5.016
Growth of housing units in 2005	9.613**	4.341	9.326**	4.385
Growth of housing units in 2006	6.763	5.201	6.857	5.183
Growth of housing units in 2007	-9.012**	4.216	-8.939**	4.228
Growth of housing units in 2008	-12.74***	4.365	-12.84***	4.379
Growth of housing units in 2009	-4.276	5.114	-4.424	5.140
Proportion of population Hispanic			-0.0818	0.238
Proportion of population in group quarters			-0.427	0.566
Judicial foreclosure state status			0.420***	0.137
Constant	1.816***	0.435	1.937***	0.517
R-squared	0.514		0.514	
p-value of F-stat	0		0.0226	

Table C7: Correlation matrix of the three classifications of judicial foreclosure states and foreclosure process longer than 120 days

	RealtyTrac	Pence (2006)	NCLC	Process > 120 days
RealtyTrac Judicial	1			
Pence (2006) Judicial	0.8248	1		
NCLC Judicial	0.7702	0.7477		
Process > 120 days	0.3281	0.4826	0.5502	1

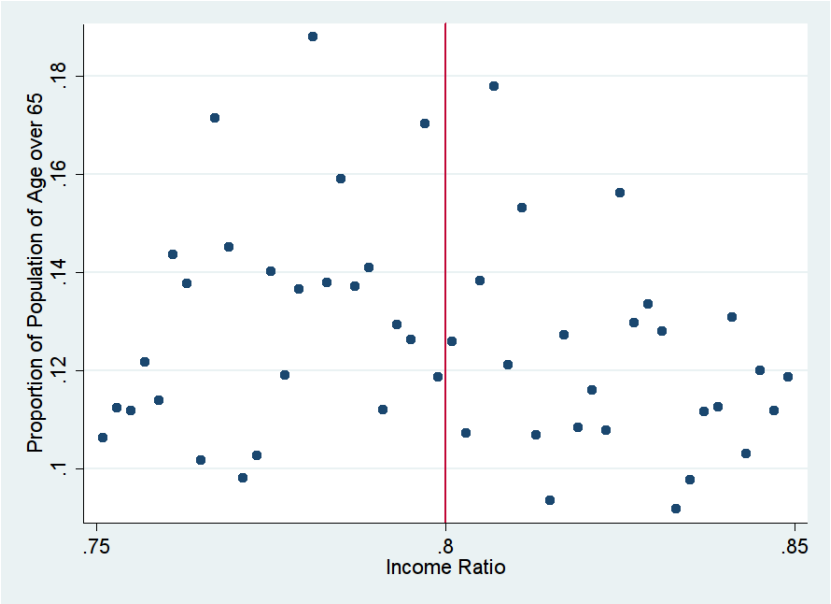
Appendix D: Discontinuities in the Control Variables

Figure D1: Relation between Median Age and Income Ratio



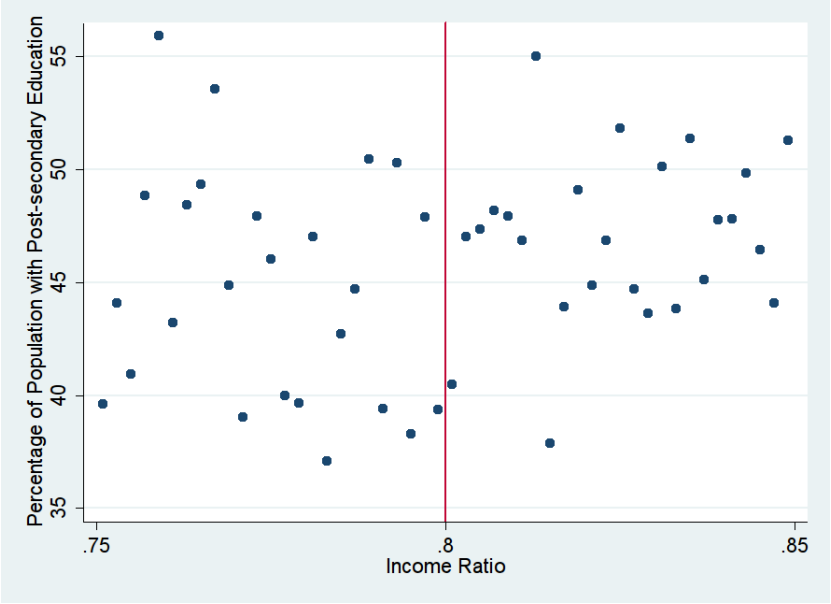
The income ratio is constrained to be between 0.75 and 0.85.

Figure D2: Relation between Proportion of Population of Age over 65 and Income Ratio



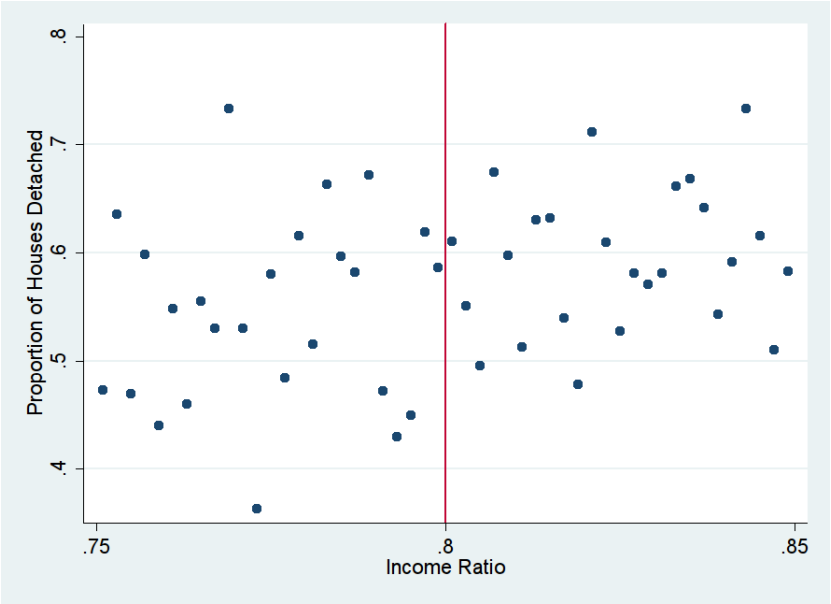
The income ratio is constrained to be between 0.75 and 0.85.

Figure D3: Relation between Percentage of Population with Post-secondary Education and Income Ratio



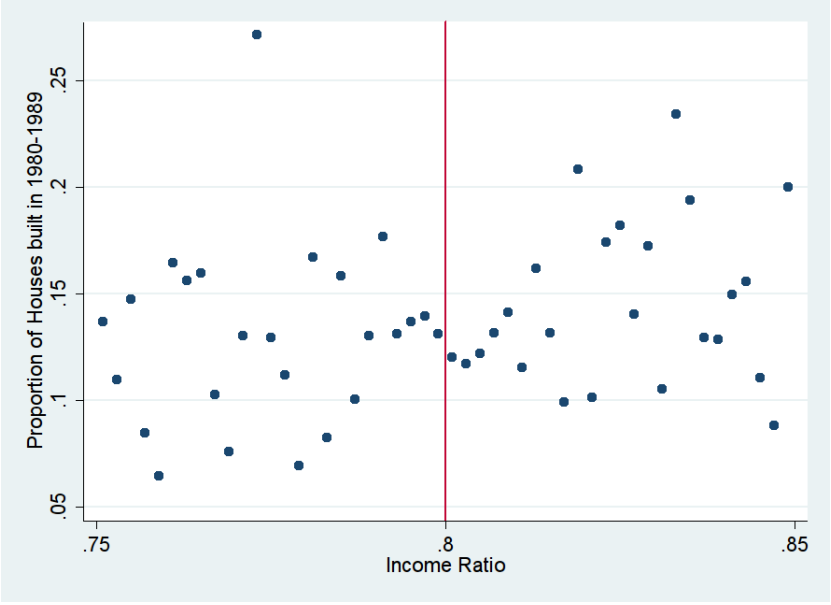
The income ratio is constrained to be between 0.75 and 0.85.

Figure D4: Relation between Proportion of Houses Detached and Income Ratio



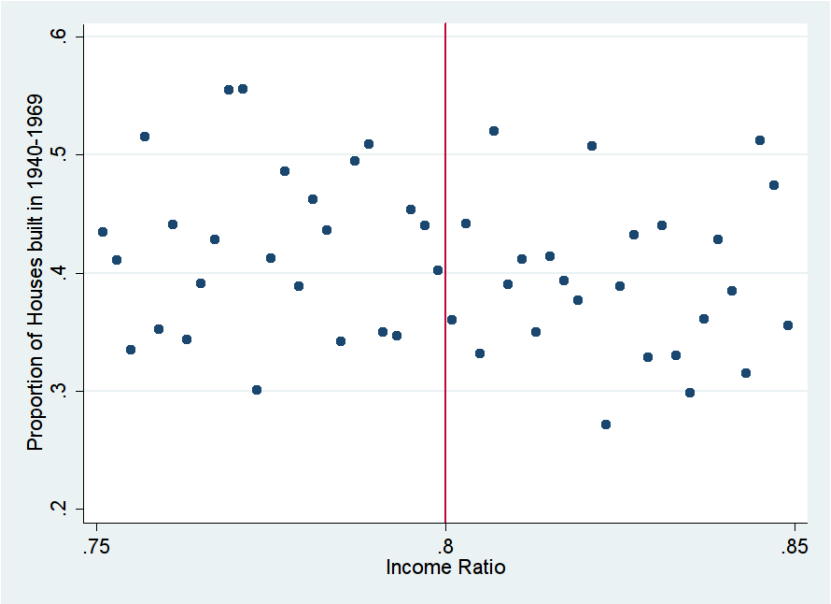
The income ratio is constrained to be between 0.75 and 0.85.

Figure D5: Relation between Proportion of Houses built in 1980-1989 and Income Ratio



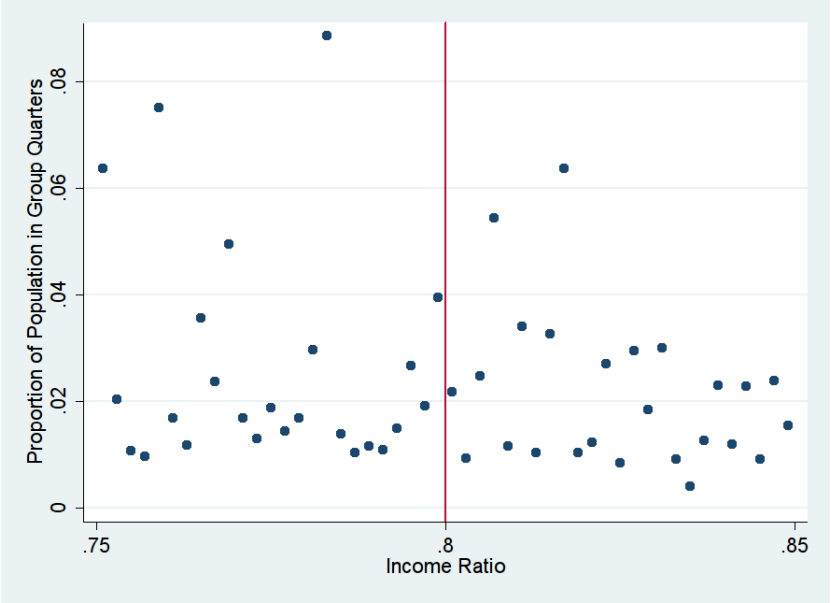
The income ratio is constrained to be between 0.75 and 0.85.

Figure D6: Relation between Proportion of Houses built in 1940-1969 and Income Ratio



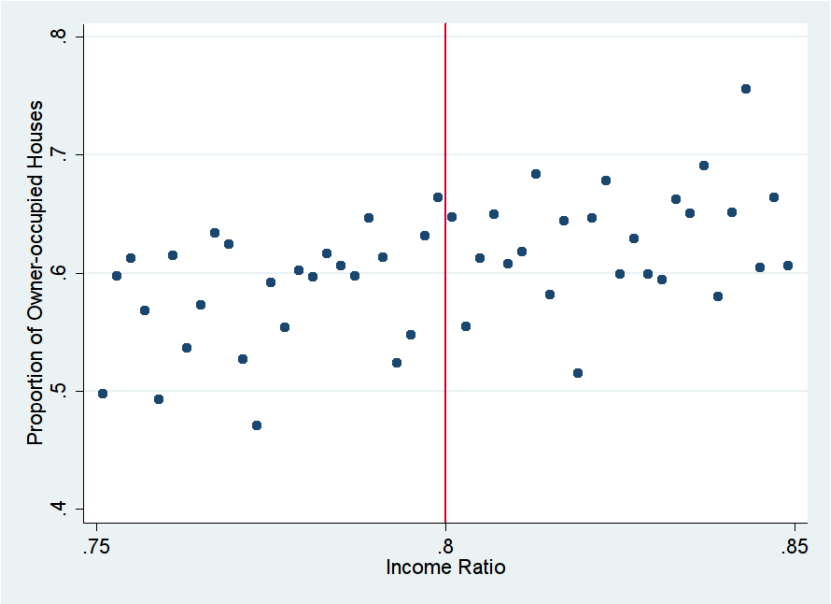
The income ratio is constrained to be between 0.75 and 0.85.

Figure D7: Relation between Proportion of Population in Group Quarters and Income Ratio



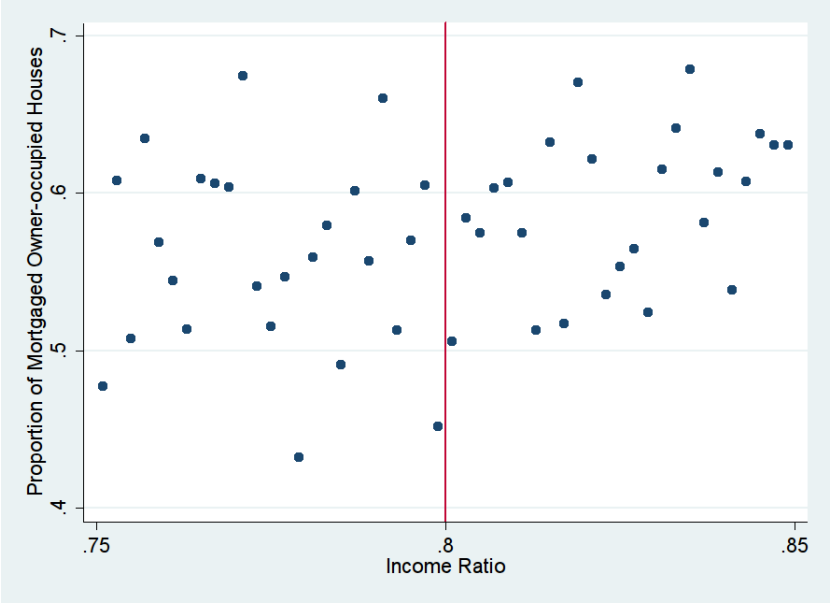
The income ratio is constrained to be between 0.75 and 0.85.

Figure D8: Relation between Proportion of Owner-occupied Houses and Income Ratio



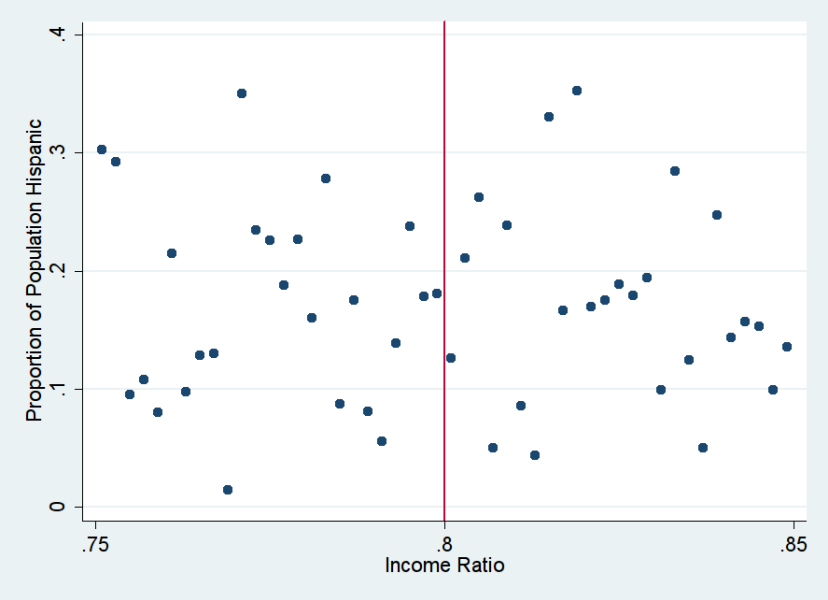
The income ratio is constrained to be between 0.75 and 0.85.

Figure D9: Relation between Proportion of Mortgaged Owner-occupied Houses and Income Ratio



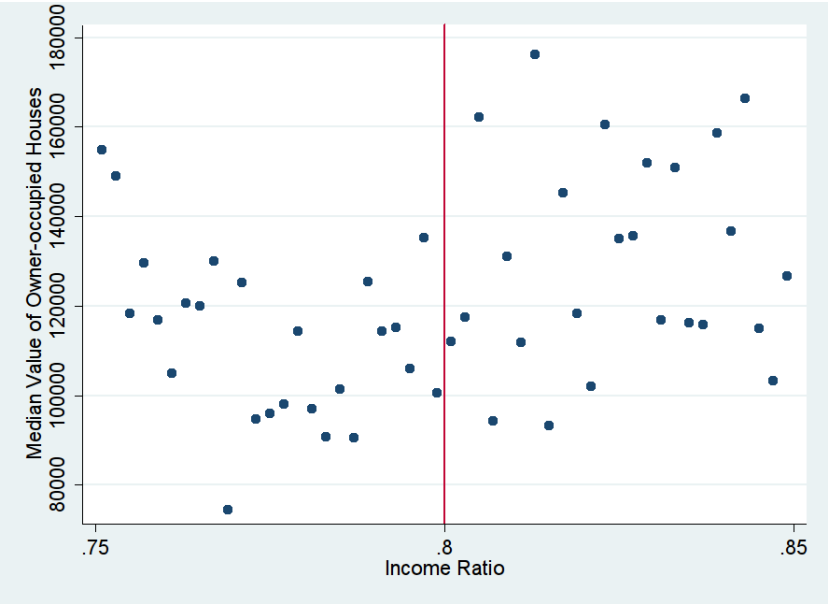
The income ratio is constrained to be between 0.75 and 0.85.

Figure D10: Relation between Proportion of Population Hispanic and Income Ratio



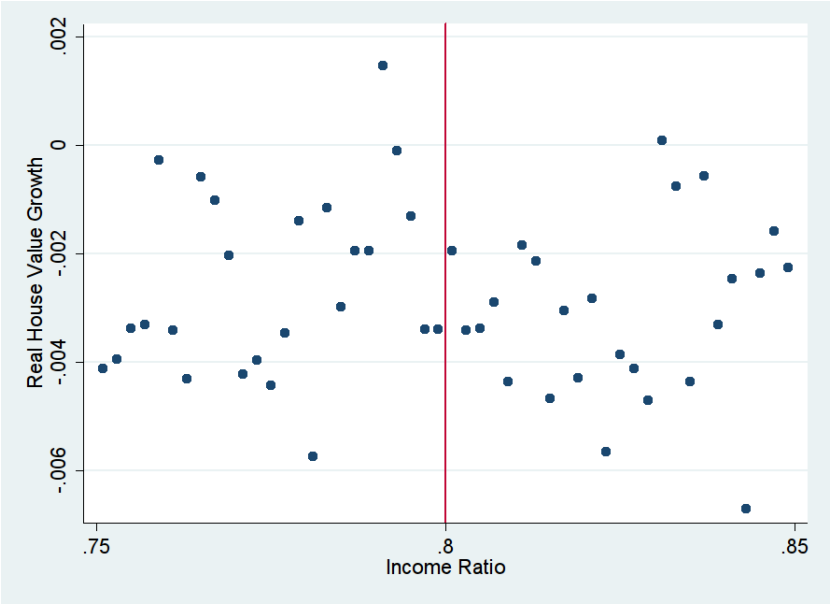
The income ratio is constrained to be between 0.75 and 0.85.

Figure D11: Relation between Median Value of Owner-occupied Houses and Income Ratio



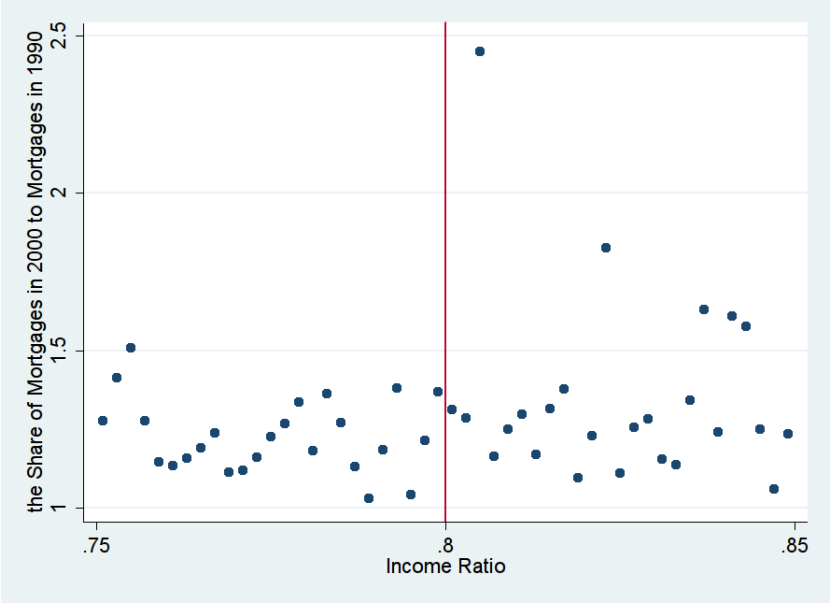
The income ratio is constrained to be between 0.75 and 0.85.

Figure D12: Relation between Real House Value Growth and Income Ratio



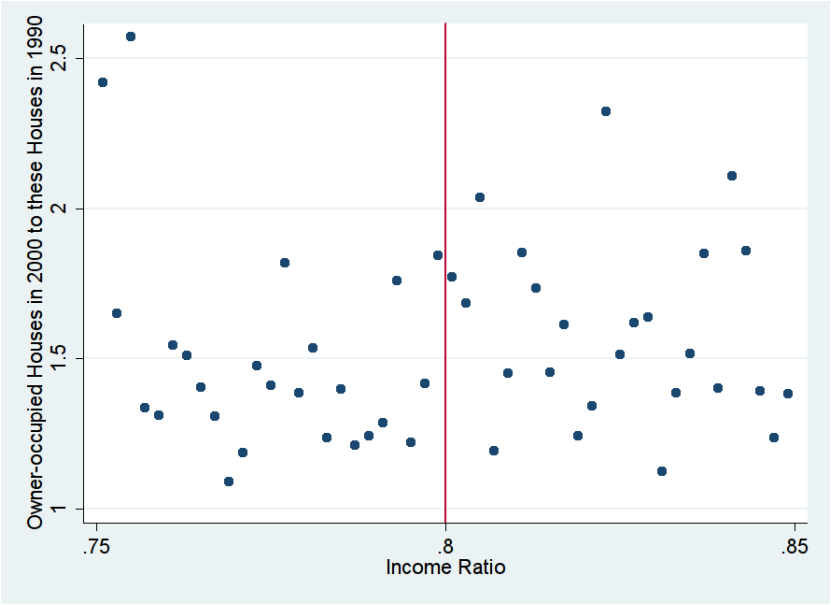
The income ratio is constrained to be between 0.75 and 0.85.

Figure D13: Relation between the Share of Mortgages in 2000 to Mortgages in 1990 and Income Ratio



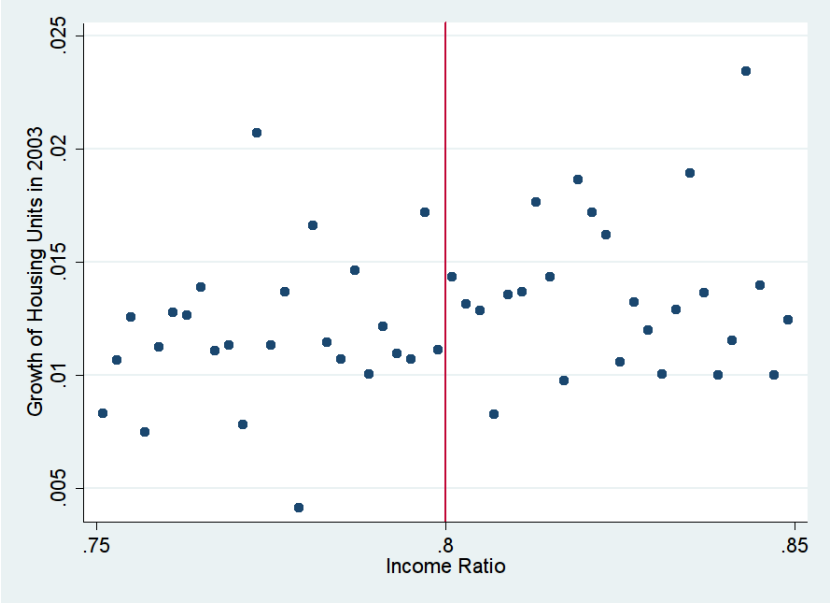
The income ratio is constrained to be between 0.75 and 0.85.

Figure D14: Relation between Number of Owner-occupied Houses in 2000 to Owner-occupied Houses in 1990 and Income Ratio



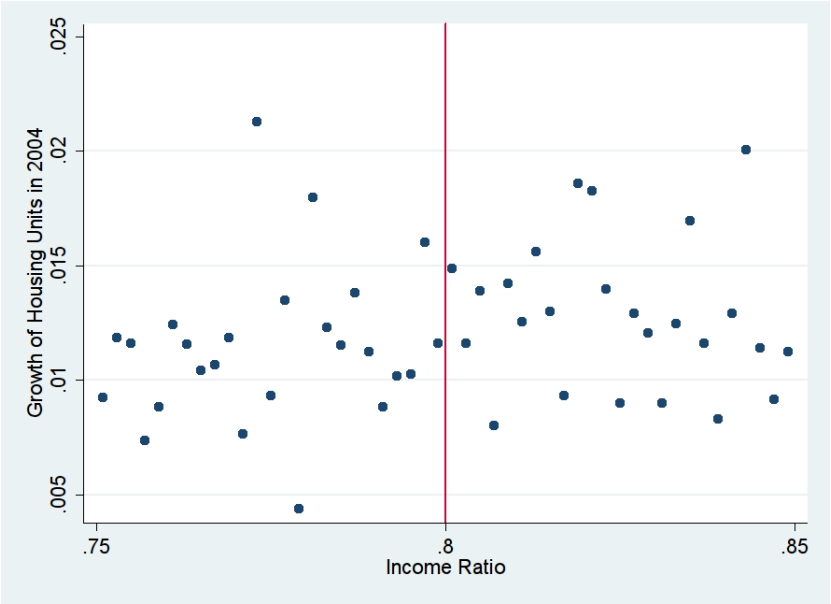
The income ratio is constrained to be between 0.75 and 0.85.

Figure D15: Relation between Growth of Housing Units in 2003 and Income Ratio



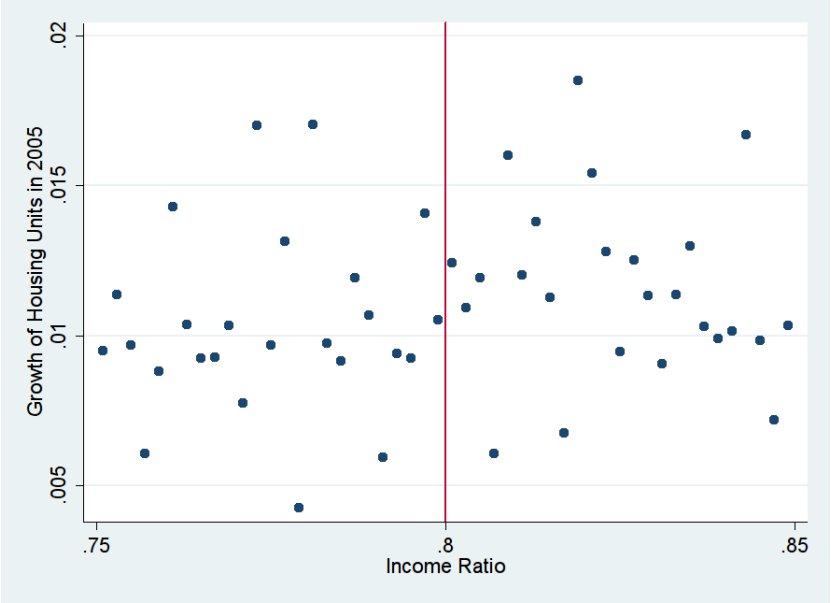
The income ratio is constrained to be between 0.75 and 0.85.

Figure D16: Relation between the Growth of Housing Units in 2004 and Income Ratio



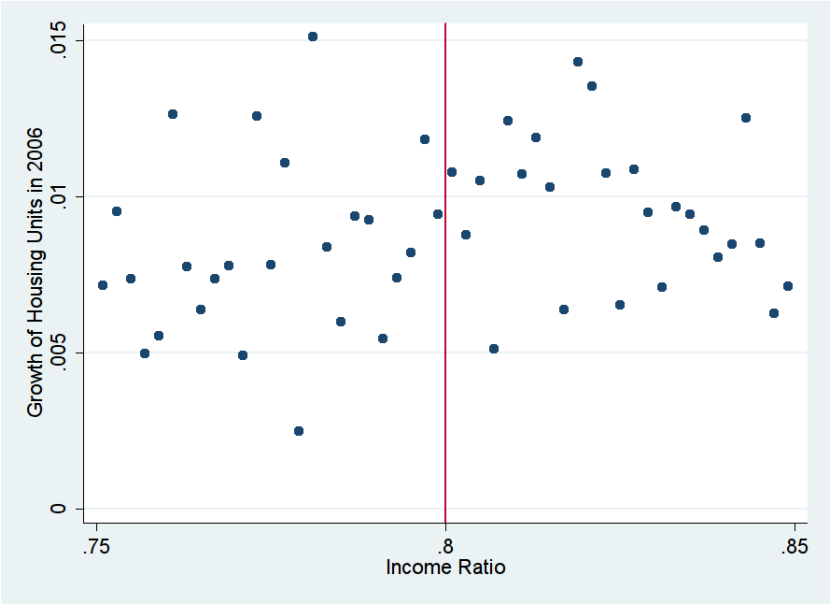
The income ratio is constrained to be between 0.75 and 0.85.

Figure D17: Relation between the Growth of Housing Units in 2005 and Income Ratio



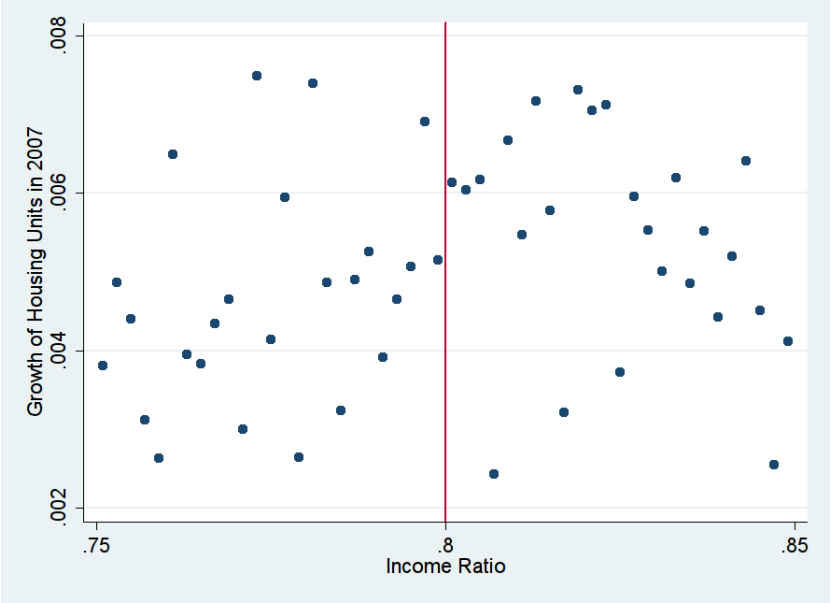
The income ratio is constrained to be between 0.75 and 0.85.

Figure D18: Relation between the Growth of Housing Units in 2006 and Income Ratio



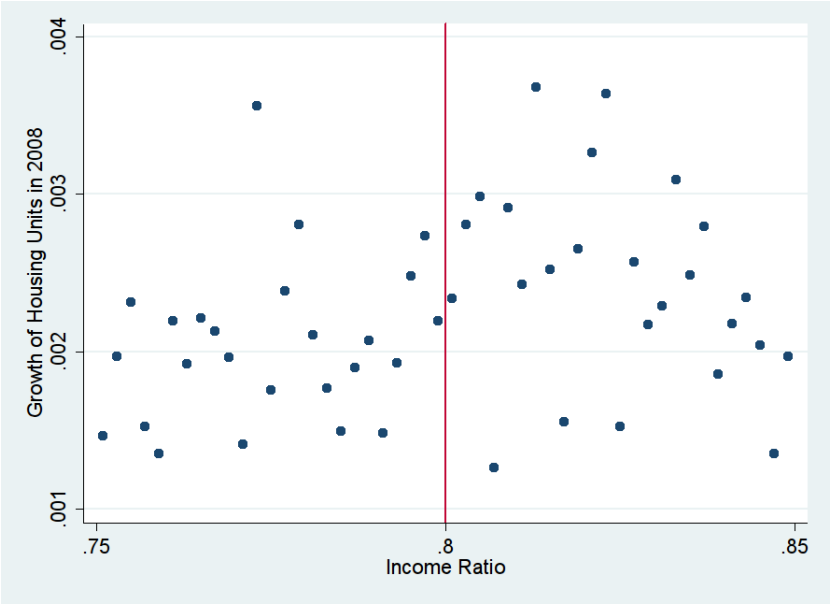
The income ratio is constrained to be between 0.75 and 0.85.

Figure D19: Relation between the Growth of Housing Units in 2007 and Income Ratio



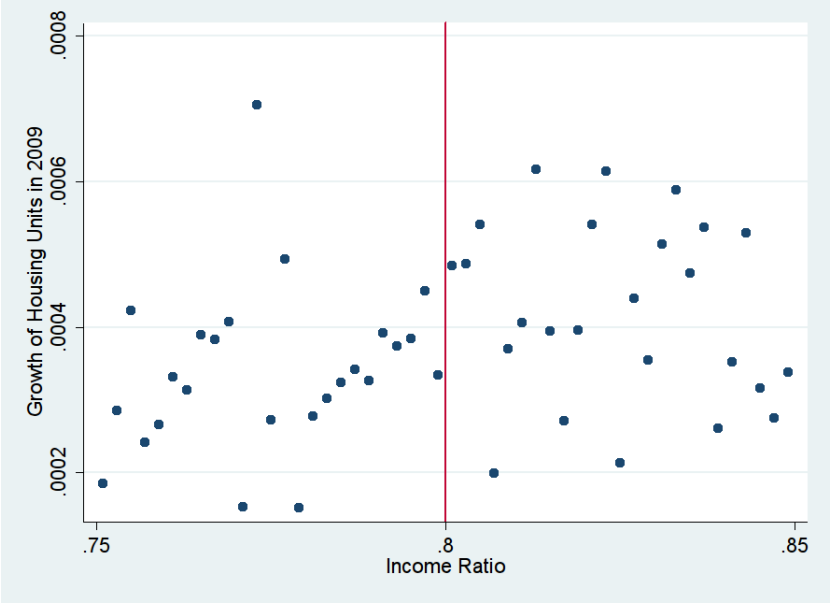
The income ratio is constrained to be between 0.75 and 0.85.

Figure D20: Relation between the Growth of Housing Units in 2008 and Income Ratio



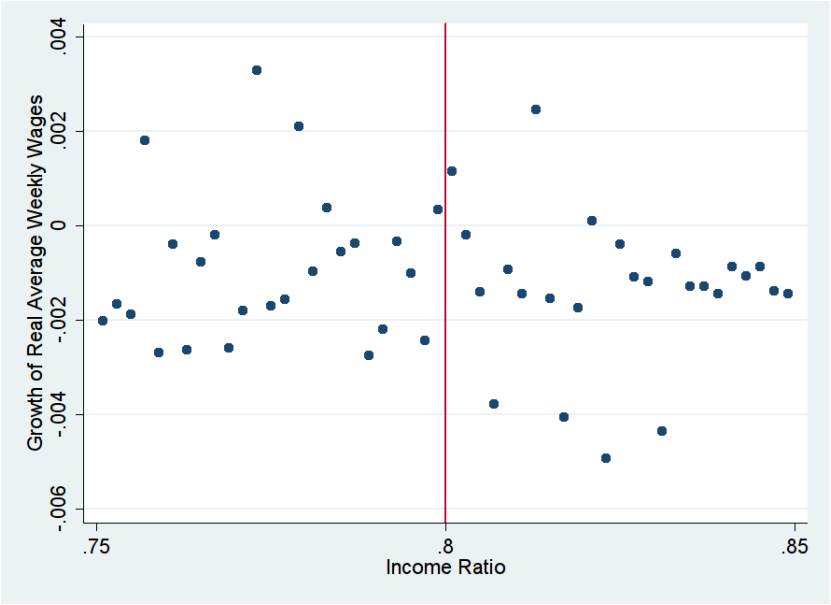
The income ratio is constrained to be between 0.75 and 0.85.

Figure D21: Relation between the Growth of Housing Units in 2009 and Income Ratio



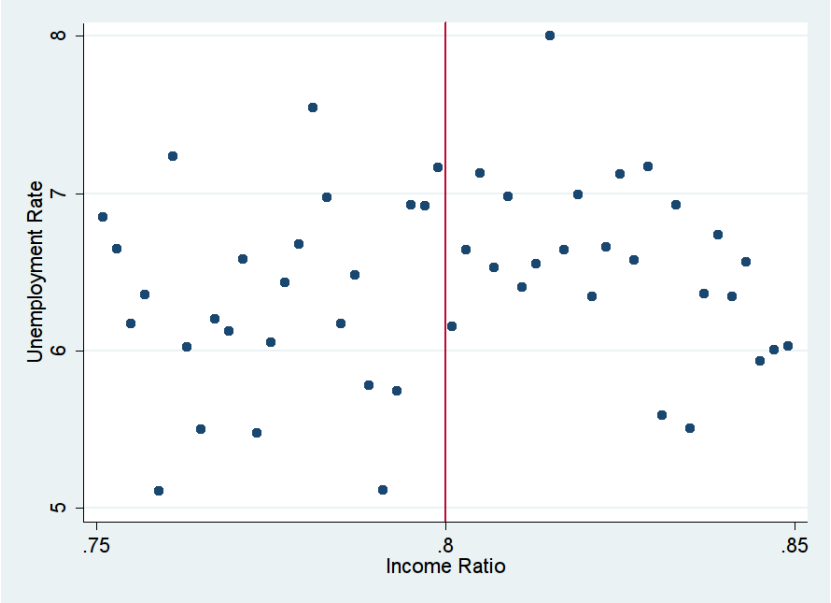
The income ratio is constrained to be between 0.75 and 0.85.

Figure D22: Relation between the Growth of Real Average Weekly Wages and Income Ratio



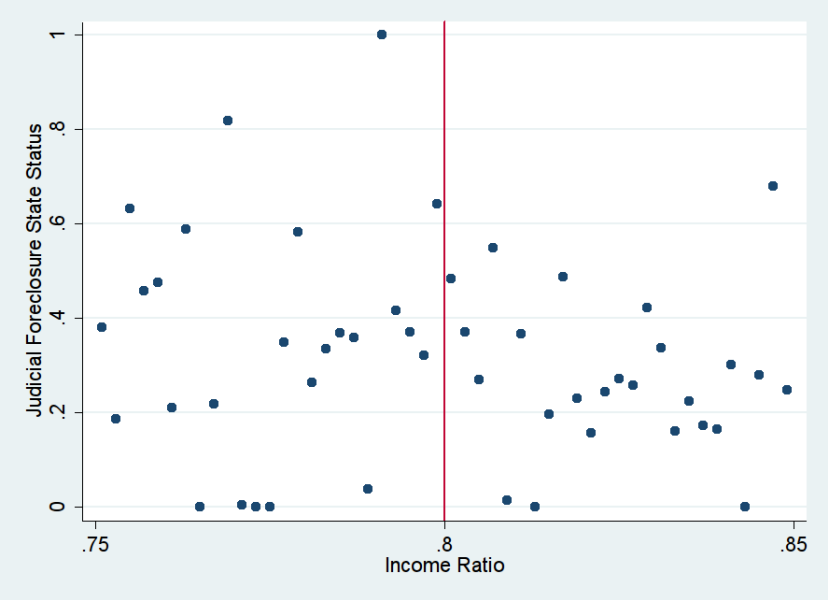
The income ratio is constrained to be between 0.75 and 0.85.

Figure D23: Relation between the Unemployment Rate and Income Ratio



The income ratio is constrained to be between 0.75 and 0.85.

Figure D24: Relation between the Judicial Foreclosure State Status and Income Ratio



The income ratio is constrained to be between 0.75 and 0.85.