Three Essays in Household Finance

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Hoonsuk Park, B.A.

Graduate Program in Business Administration

The Ohio State University

2017

Dissertation Committee:

Itzhak Ben-David, Advisor

Berk A. Sensoy

Michael S. Weisbach

Abstract

This dissertation examines three questions in the household finance literature. The three essays are broadly connected to each other in the sense that they examine household behavior, in particular consumption, in reaction to different events. There are theoretical predictions on how households change consumption in response to mortgage default, changes in income, and the implementation of sales taxes. I explore each question in the following essays in an attempt to further our understanding of households. The three essays also share the same dataset, which is transaction-level data on banking and credit card transactions. Research in household finance is limited by the data that is available, this dissertation also contributes to the literature by documenting detailed behavior in the new data.

In the first essay, I examine the effect of the household's lack of liquid assets on mortgage default. Using administrative data from banking and credit card transactions, I find that a significant number of households lack liquid assets, and that these households are more likely to default on their mortgages. The effect of the lack of liquid assets on mortgage default is amplified during unemployment. When comparing liquid assets with income, I find that high income households that lack liquid assets are more likely to default on their mortgages compared to low-income households

that have more savings. Finally, households that lack liquid assets reduce consumption dramatically during the period in which they default on their mortgages. These findings have implications for mortgage default theory and consumption theory.

In the second essay, I analyze household consumption surrounding mortgage payment changes following adjustable-rate mortgage rate adjustments. I find evidence of household consumption behavior that is consistent with loss aversion. Household consumption is sensitive to increases in adjustable-rate mortgage payments, while being statistically insignificant for decreases in adjustable-rate mortgage payments. This result remains for adjustments in mortgage payments over and under \$100. I find that the consumption sensitivity between households with low savings are not statistically different from households with high savings. These results are contrary to the predictions of other consumption theories that incorporate liquidity constraints, buffer stock savings, and myopia.

In the third essay, coauthored with Brian Baugh and Itzhak Ben-David, we study the effects of sales taxes on household consumption. For years, online retailers have maintained a price advantage over brick-and-mortar retailers by not collecting sales tax at the time of sale. Recently, several states have required that the online retailer Amazon collect sales tax during checkout. Using transaction-level data, we document that households living in these states reduce Amazon purchases by 9.4% after sales tax laws were implemented, implying elasticities ranging from 1.2 to 1.4. The effect is more pronounced for large purchases, for which we estimate a reduction of 29.1% in purchases, corresponding to an elasticity of 3.9. Studying competitors in the electronics field, we detect some evidence of substitution toward competing retailers.

Consistent with an income effect, we find a reduction in spending in other categories that is concentrated among the heaviest Amazon shoppers.

This is dedicated to my wife.

Acknowledgments

I would like to express my thanks and appreciation to my advisor Itzhak (Zahi) Ben-David. I am indebted to his support, encouragement, and patience throughout the six years of the Ph.D. program. I would also like to thank my dissertation committee Michael Weisbach and Berk Sensoy for their help and guidance.

I thank the friends that I met in the Ph.D. program, Brian Baugh, Sehoon Kim and Yoonkang Lee. I thank Youngsoo Jang, my neighbor and close friend. I thank my wife for going through this difficult period together. I thank my parents and sister for their support.

Vita

2009	B.A. Business Administration, Korea University
2011-2016	Graduate Research Associate, The Ohio State University.

Fields of Study

Major Field: Business Administration

Table of Contents

List of Tables

List of Figures

Chapter 1: Household Savings and Mortgage Default

1.1 Introduction

A fundamental issue in household finance concerns the factors that lead individuals to default on their loans. The largest loan for most households is their mortgage, and mortgage defaults have had a huge impact on the economy, illustrated most vividly through the 2008 financial crisis. Yet, little is known about the factors that lead households to default. Are mortgage defaults merely a reflection of households income shocks, or do they reflect something more fundamental about households financial acumen and habits? How does a households financial fragility, reflected by low savings and being prone to overdraft fees, affect the likelihood of defaulting on its mortgage? This paper addresses this question using a unique database on household financial behavior.

A large number of households lack liquid assets. For example, [70] surveyed 1,931 households and found that nearly half of the respondents could not come up with $2,000 in a month. Using data from the Panel Study of Income Dynamics (PSID), [30] find that over 46 percent of households had less than $5,000 in liquid assets. Another survey by the Federal Reserve Bank found that 47 percent of respondents could not come up with $400 without selling possessions or borrowing money. More surprising is the degree to which even high-income households also lack liquid savings. [70] find that 23% of households with incomes between $100,000 and $150,000 reported that they could not come up with $2,000 in a month, with similar results found in the survey by the Federal Reserve.

Assuming that these households with low savings are financially constrained, they are vulnerable to financial disruption such as income shocks, divorce, or emergency expenses. When these households face such shocks, they may have no other option than to forcibly adjust their consumption, in part by reducing housing expenditures.

In this paper, I empirically examine the importance of the lack of liquid assets on mortgage default, with an emphasis on exploiting unemployment shocks to highlight the effects of ex-ante financial fragility. To my knowledge, this study is the first to examine the household's savings in liquid assets and unemployment simultaneously. To do this, I employ a unique administrative dataset that includes detailed bank and credit card information at the transaction level for 2.7 million households from July 2010 to May 2015. Using this dataset, I find that a substantial number of households lack liquid assets. I consider households to lack liquid assets if they are in the low tercile in interest earned and if they have incurred overdraft fees. I measure unemployment shocks by the receipt of unemployment benefits.

I find that the households that incurred overdraft fees were 43 percent more likely to default on their mortgages, compared to households that did not incur overdraft fees. Similarly, I find that households in the lowest tercile of savings were also 43 percent more likely to default on their mortgages. On the other hand, households that participated in the financial markets through transacting on their brokerage accounts were 22 percent less likely to default. Compared to these measures of liquid assets, other household characteristics had smaller effects on mortgage default. Households in the lowest income tercile were only 22 percent more likely to default on their mortgages compared to households in the middle and high income terciles, and households that had a high ratio of spending relative to their income were only 8 percent more likely to default. Households that had a high mortgage-payment-to-income (MTI) ratio were 14 percent more likely to default. The economic magnitude of the effect of the lack of liquid assets on mortgage default, 43 percent, is much larger that of low income, 22 percent.

Financially fragile households that become unemployed are much more likely to default on their mortgages compared to non-fragile households that become unemployed. Households that had incurred overdraft fees were 160 percent more likely to default when they subsequently became unemployed. In contrast, households that had not incurred overdraft fees that subsequently became unemployed were only 50 percent more likely to default on their mortgages in the months in which they were unemployed. Households who were in the lowest tercile of interest earned that subsequently became unemployed were 97 percent more likely to default compared to the months in which households continued to be employed. However, households that were in the high tercile of interest earned who subsequently became unemployed were only 23 percent more likely to default in the months in which they became unemployed. The large contrast between fragile households (i.e. incurred overdraft fees or were in the lowest interest earned tercile) and households had more in liquid assets (i.e. did not incur overdraft fees or were in the highest interest earned tercile) highlight the importance of liquid assets as a significant determinant of mortgage default.

Next, I compare liquid assets and income as determinants of mortgage default. It is not just the poor who lack liquid assets. I find that 15 percent of households in the highest income tercile incurred overdraft fees and that 18 percent of households in the highest income tercile belong in the lowest interest earned tercile, consistent with the findings in the surveys above.

Households that incurred overdraft fees were more likely to default on their mortgages, regardless of whether they belonged in the low, mid, or high income tercile. High income households that incurred overdraft fees were 20 percent more likely to default on their mortgage, compared to households in the middle income tercile that did not incur overdraft fees. Similarly, high income households that were in the low interest earned tercile were 17 percent more likely to default on their mortgages compared to households in the middle income and middle interest earned tercile.

Finally, I examine consumption and debt repayment surrounding mortgage default. I find that households that lacked liquid assets reduced their consumption dramatically leading up to the time of default. Households that had more liquid assets also reduced their consumption, but not to the same extent. In addition, I find evidence of households paying back their credit card debt as they default on their mortgages This is consistent with the finding of [35] and [32], who find that consumers decide to preserve access to credit card borrowing as they default on their mortgages.

Theory provides two main spectrums of predictions on the effect of fragility on mortgage default. In a world without financial constraints, mortgage default does not depend on household fragility. Instead, households make default decisions based on an option-theoretical framework, where the household exercises its default option when the cost of the mortgage exceeds the value of the default option. In this setting, households default only when they have negative home equity ([64], [90], [37]).

In an alternative class of models that incorporate financial constraints, mortgage default is determined by a double trigger mechanism. In these models, negative home equity is a necessary, but not sufficient condition for default. Default is determined by borrowing constraints as well as income shocks such as unemployment and medical emergencies ([48], [13], [27], [82]). Households that are fragile are much more likely to encounter these borrowing constraints, and hence these models predict that financially fragile households will be more likely to default on their mortgages.

Other theories on household behavior explain why households are fragile. A particular strand of consumption theory suggests that fragility may even be preferred. The lack of liquid savings could be a self-commitment device [69] or simply an optimal portfolio allocation ([62] and [63]). Regardless of the rationale, financial fragility exposes the household to financial disruption.

Despite the emphasis on fragility in the theoretical literature, the empirical literature on mortgage default has mostly been focused on negative equity, primarily due to data constraints.[1] Data on household-level fragility or unemployment had been largely unavailable to researchers until recently. Notable exceptions are [41] and [**?**], who employ new datasets on household level credit card utilization and unemployment, respectively.

Continuing this trend of using novel datasets, this paper contributes to the empirical literature on household behavior by testing the theoretical predictions about

[1]The large empirical literature on mortgage default include [89], [29], [80], [90], [37], [34], [48], [11], [44], [72], [73], [41], [50], [13], [36], [46], [75], [42], and [**?**] among others.

household savings and mortgage default. More importantly, by identifying an important link between liquid assets and mortgage default, this paper sheds light on what has been at the center of economic policy debates over the recent decade.

1.2 Data and Empirical Methodology

1.2.1 Data

A key contribution of this paper is the data. Prior researchers have mostly relied on loan-level data to examine why households default on their mortgages. This loan-level data, often collected from mortgage servicers, is rich in information about loan performance and loan origination. For example, in the Freddie Mac Single Family Loan-Level Dataset, data on loan-to-value (LTV), debt-to-income (DTI), credit scores, and interest rates are provided at origination and unpaid principal balance (UPB) and delinquencies are kept track of as part of the monthly performance dataset.

If households make default decisions solely based on the option value of the mortgage, as in the option theoretic models, then the loan-level data would be sufficient in explaining why households default on their mortgages. However, if other factors such as household savings, unemployment, health shocks are also considered by households when they are making the decision on whether to default on their mortgage, we need to know how much savings the household has, whether they are unemployed, and if they face large medical bills. Unfortunately, information on these household events are unavailable in loan-level datasets.

In this paper, I attempt to advance the literature by using data that contains bank and credit card transactions from 2.7 million households to look at how savings, income, unemployment, and other characteristics affect the households default decision. The data is from an online website that aggregates bank and credit card accounts for households. Households use this aggregation service as a convenient way to keep track of their savings and spending, by providing the website of usernames and passwords of different financial institutions so that the website can gather this information and present the information in a single page.

This transaction-level data has become available to researchers in recent years. Examples of recent studies using this type of banking and credit card transaction data include [14], [17], [18], [19], [22], and [68]. These papers, with the exception of [22], are focused on testing consumption theory, since this type of data is naturally rich in detail on consumer spending. [22] looks at the effect of negative equity of mortgages on labor supply.

The data includes the date, description, amount, and other variables for each transaction. I use keyword searches to identify mortgage, income, unemployment benefits, overdrafts, interest earned, brokerage, and consumption transaction. After each transaction is classified, it is summed up at the monthly level. For credit card repayment, I add all the credit transactions for the month as repayment, and add all the debit transactions for the month as borrowing. Then I lag the sum of debit transactions so that it is matched with the sum of credit transactions one month later. This is done to match the payment cycle of credit cards. The difference between the two amounts is the amount that the household has repaid or borrowed from the credit card company.

Since unemployment benefits are administered at the state level, benefits for each state are identified separately. For example, unemployment benefits for New York often come with the description nys dol ui and unemployment benefits for Texas often come with twc benefit ui. nys dol stands for New York State Department of Labor and twc stands for the Texas Workforce Commission. Using this method, I identify 27 states, which are Alaska, Arkansas, Colorado, Connecticut, Delaware, Florida, Georgia, Iowa, Idaho, Louisiana, Massachusetts, Maine, Michigan, Minnesota, Missouri, North Carolina, New Mexico, New York, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Washington, West Virginia, and Wyoming. States such as California are missing because they issue unemployment benefits to a separate debit card that is not captured in the data. In other states such as Ohio, unemployment benefits are hard to identify using transaction descriptions since it is paid by the the Department of Job and Family Services, and do not distinguish between unemployment benefits and other welfare programs.

Mortgages are selected based on mortgage-related keywords such as “mtg”, “mortgage”, “home loan” or debit transactions to major mortgage servicers such as “ocwen” and “dovenmuehle”. Income is identified by searching for keywords such as “payroll”, “salary”, or “direct deposit”. Overdraft fees are identified by keywords such as “overdraft fee” or “nsf returned item fee”, retail consumption by keywords for major retailers such as “wal mart”, “target”, or “costco”, grocery consumption by keywords such as “kroger” and “safeway”, restaurant consumption by keywords such as “mcdonalds”, “burger”, “pizza”, brokerage transactions by keywords for large brokerages and mutual fund companies and so forth.

While these variables are directly observed, some other variables must be inferred. The most important of these is the indicator for whether the household has defaulted on its mortgage. I consider a household to have defaulted on its mortgage if the household has missed more than three mortgage payments. However, while I can observe mortgage payments, the termination or suspension of these payments does not always mean that the household has defaulted on its mortgage. It is possible that the household has fully paid off its mortgage, sold off the house, or started paying the mortgage from a different account.

In order to correctly identify mortgage default, I impose a number of filters in the data. First, I require that all the households in the sample to have not received social security payments at any time during the sample period. This is because households that receive social security payments are more likely to be old enough to have fully paid off their mortgages. Households that received social security payments were dropped from the sample.

Next, I identify households that have larger average mortgage payments in the last 3 months than compared to the previous 12 months. I consider these households to have fully paid off its mortgage or to have sold the house when they stop paying their mortgages. Similarly, I identify households that have larger average monthly sums of transactions larger than $5,000 in the final 6 months than compared to the previous 12 months. I also treat these households to have fully paid off its mortgage or to have sold the house when they stop paying their mortgages.

Finally, I look at additional ways of identifying default instead of solely relying on whether the household has stopped paying its mortgage. The first of these additional methods is to look at cured mortgages. I find mortgages that households have started paying off again, after more than 3 months of suspending the payment of mortgages. I also look at households that have reduced their consumption by more than 30 percent and households that experience income shocks. The results for these alternative definitions are not reported in this paper, but the results are consistent regardless of the definition of mortgage default.

Another drawback is that data lacks information on loan balances. Therefore, instead of using the loan-to-value ratios directly, I control for the changes in loan-to-value by using the changes in the real estate prices in the households city of residence. The city of residence is the most common city within a state in which a household makes its transactions. Then I take the Zillow Home Value Index from Zillow.com and match it to each household. Since I do not have loan balance data, I cannot see if the household has negative equity on its home. I also employ region-time fixed effects in some regressions to control for the changes in the house prices in the region as well as other economic factors such as regional unemployment.

The data in this paper complements prior research on the determinants of mortgage default. This paper can reliably observe household income, unemployment, savings, and consumption, but must estimate whether the household has defaulted on its mortgage, and tries to control for other variables such as LTV. In contrast, prior papers in general can observe delinquency, LTV, and credit scores but did not have access to household savings, unemployment, and consumption. Since theory predicts that unemployment is an important predictor of default, many studies use regional unemployment as a proxy for individual unemployment. This attenuates the actual effect of unemployment, leading to conclusions that unemployment is not as important as other variables that were directly observable such as negative equity ([41], [50], [55]). The data on savings is even less accessible.

Despite the restrictions in the data, there are two studies that analyze the effect of liquidity constraints and unemployment on mortgage default. [**?**] use a supplement to the PSID, which combines the households mortgage data with information on employment, income and other characteristics and find that unemployment is an important predictor of default. [41] combine mortgage servicer data with household credit card information from Equifax using balance, date, and zip code to measure household liquidity and find that liquidity as defined by credit card utilization was as important as negative equity in predicting default.

There are differences between the data used in this paper and the PSID or Equifax data used in [?] and [41]. First, the PSID surveys used in [?] are cross-sectional in nature. For example, it has information on savings but it is recorded at the time of the survey along with other variables such as mortgage default. Therefore it is not possible to find the effect of savings prior to a shock that would lead the household to default on its mortgage. Additionally, there is a difference in sample size. The final sample used by [?] has 5,281 households, compared to the 265,144 households in the final sample of this paper. The Equifax data used by [41] does not have information on unemployment and savings. However, the PSID and Equifax also have their advantages. The PSID is rich in demographic information such as race and age, and also has better data regarding LTV and default. The Equifax data has information on credit card limits and it is possible to link it to loan-level data which has very accurate information on default.

In the discussion on the importance of middle-income, high-income, and prime borrowers on the financial crisis of 2008, [74] point out that income may have been overstated during the housing boom of the mid-2000s due to mortgage fraud. Though not covering the same period, I can directly observe income receipt and provide additional evidence on the determinants of mortgage default for the middle class.

The raw data contains transactions for 2.7 million households from July 2010 to April 2015. Of the 2.7 million, I restrict the data to the 792,786 households that have both mortgages and income transactions. Next, I restrict the data to the households that live in the 27 states in which unemployment benefits can be identified. Households are considered to be living in a particular state if more than two-thirds of the transactions that have location information are in that one state. After this restriction, the sample is reduced to 430,024 households. Finally, I require that households receive no social security payments at any time in the data and that there be six consecutive months of both income receipt and mortgage payment. This is done so that the households in the final sample are those that are likely to have their main bank accounts registered with the online aggregator. This six month period is where I observe the households financial conditions prior to the default period. This leaves a final sample of 265,144 households.

The final sample of 265,144 households is not a representative sample of the US population. The households in the data select into the sample by choosing the services of the online aggregator. This may mean that the data includes households that are more technologically advanced, higher educated, younger, and also have more interest in managing their finances compared to the rest of the population. This is reflected to some degree in the geographic distribution of households in the sample. Figure 1.1 shows the distribution of households in each state, compared to the 2010 United States Census. Panel A shows the proportion of households residing each state and Panel B shows the percentage point difference between the final sample of households used in this data and the 2010 Census. Households in this sample are particularly heavily concentrated in New York and to a lesser extent in Texas. Households are also over-represented in Florida, Georgia, and Connecticut. In contrast, households are heavily under-represented in Pennsylvania and Michigan.

Table 1.1 presents the summary statistics. Out of the 265,144 households in the sample period, 28,082 households experienced default. Fifty percent of households received \$0.36 or less in monthly bank interest during the 6 month pre-period. Assuming a five basis-point interest rate, this equates to savings of around \$8,640 in the bank savings account, which is higher than the survey results from the Federal Reserve Bank that showed 47 percent of the respondents could not come up with \$400 without selling possessions or borrowing money. Eighty three percent of households did not have any transactions with brokerages in the pre-period, though when averaged, had \$547 in brokerage transactions. Nineteen percent of households had incurred overdraft fees, which average to \$3 for the whole sample. Retail consumption was\$413, restaurant consumption \$148, and grocery consumption \$115 on average for the sample.

In Table 1.2, I look at the household's liquid assets in more detail. In Panel A, I find that a third of the households had less than \$0.103 in monthly interest earned per month, which tranlates into \$2,472 in bank balance when assuming a five basis-point interest rate. Even for households in the high income tercile, 18 percent of houseoholds had less than \$2,472 of cash in the bank. In Panel B, I find similar results. 19 percent of households had incurred overdraft fees in the 6 month pre-period and around 15 percent of households in the high income tercile had incurred overdraft fees. These results are consistent with the surveys that find that over 20 percent high income households have trouble in coming up with \$2,000 in 30 days and not having \$400 to cover emergency expenses.

1.2.2 Empirical Methodology

The main analysis consists of a dynamic logit model for default that is equivalent to discrete hazard models ([85]). The dependent variable is mortgage default which is defined to be households that have missed more than three mortgage payments. Observations before the end of the pre-period and observations that follow delinquency are dropped from the sample. For example, if a household had a pre-period from Jan 2011 to June 2011 and became delinquent in May 2012, then the main sample would start from June 2011 and end in May 2012. The dependent variable would equal zero for all months except May 2012, which would equal one.

In the pre-period, I divide the sample into terciles based on savings, income, spending, and mortgage-payment-to-income (MTI) ratios. Savings consists of three dummies based on how much interest the household had earned in the pre-period. If the household belongs in the tercile with the lowest interest earned, it would equal one for the low savings dummy and zero for the mid savings and high savings dummies. The similar procedure is applied to income and spending. Spending is defined as the monthly average of the ratio between consumption and income in the pre-period.

Stock ownership and overdrafts are dummies that equal one if the household had brokerage transactions and overdraft fees in the pre-period, respectively. The Zillow Index represents that change in real estate prices in the city in which the household resides. I also include state and year-month fixed effects and cluster the standard errors by household. Following [54] and [41], I include a fifth order polynomial in account age to allow the hazard function to vary nonparametrically. The logit regression results are presented as relative risk unless specified otherwise.

When comparing the effect of income and savings on the likelihood of mortgage default, I use categorical indicator variables instead of interaction terms. This is due to the difference in the interpretation of the interaction in logit regressions ([9]). For example, when I estimate the effect of high-income of mortgage default, the logit regressions estimate the log-odds of mortgage default for the high-income households. The logit regressions estimates an interaction between income and savings as the interaction between the log-odds of each variable, and thus the coefficient on the interaction is multiplicative instead of additive, as it is for linear regressions. For example, if the likelihood of default for high-income households goes from 1 percent to 10 percent when becoming unemployed and the likelihood of default for low-income households goes from 20 percent to 80 percent, the logit regression would consider the increase in default likelihood for high-income households to be more important since it is a 10-fold increase, whereas the default likelihood for low-income households is less important since it is a 8-fold increase. While the multiplicative interpretation can be useful in some applications, I use categorical dummies to avoid these problems.

1.3 Results

In this paper, I present three primary results. First, I find that the lack liquid assets is an important determinant of mortgage default. The effect of liquid assets on mortgage default is stronger when the household becomes unemployed, as households that lack liquid assets are more fragile, thus vulnerable to shocks. These results are consistent with the double trigger models of mortgage default.

Second, I find that the lack of liquid assets is a more important predictor of mortgage default than income. Households that lack liquid savings that have high income are more likely to default on their mortgages than households with more liquid assets that have low income. This finding is broadly consistent with the existence of the wealthy hand-to-mouth documented by [62] and [63].

Third, I find that the decline in consumption surrounding mortgage default is large for households that lack liquid assets. These households dramatically decrease their consumption leading up to mortgage default, but maintains consumption after delinquency. Compared to households that lack savings in liquid assets, the decline in consumption for households with higher savings in liquid assets is more gradual. These results are consistent with the financial constraints literature in consumption theory ([56], [91]).

I also look at credit card debt repayment surrounding mortgage default. I find that households that lacked liquid assets tended to repay their credit card debt, perhaps in an attempt to preserve their access to financing. This repayment of credit card debt is not as clearly prevalent for households with savings in liquid assets. [32] show similar results, where they find that households defaulted on their mortgages whiles prioritizing the repayment of credit card debt and auto loans. The results for consumption and credit card debt also shines some light into the pecking order of coping methods proposed by [70].

1.3.1 Household Savings and Mortgage Default

In Table 1.3, I look at the measures of household savings in liquid assets and examine how they are related to mortgage default. In column (1), I find that households than incurred overdraft fees in the pre-period were 43 percent more likely to default on their mortgages compared to households that did not incur overdraft fees.

If overdraft fees are generally related to the poor, financial market participation is about the rich. For most households, the largest asset class is usually their house and not the stocks ([26]). In column (2), I look at households that had participated in the financial markets through brokerage transactions and find that these households were 22 percent less likely to default on their mortgages.

In columns (3) and (4), I look at the interest earned terciles and their relationship to mortgage default. In column (3), I find that households in the low interest earned tercile were 43 percent more likely to default on their mortgages. In column (4), I find that households in the middle interest earned tercile were 26 percent less likely to default on their mortgages than households in the low interest earned tercile. Households in the high interest earned tercile were 34 percent less likely to default on their mortgages, compared to the low savings tercile.

In columns (5), I put the overdraft, brokerage, and low savings dummies in single regressions for comparison. I find that the magnitude of the effect of overdrafts and low savings on mortgage default does not decrease by much, and the effect for financial market participation through brokerages only decreased by around 2 percentage points.

In Table 1.11, I exclude defaults that happen in the 3 months immediately following the pre-period to control for possible confounding factors. I find very similar results to that in Table 1.3.

In Table 1.4, I look at other household characteristics such as income, spending, and mortgage-to-income and examine how they are related to mortgage default. In columns (1) and (2), I look at the income terciles and their relationship to mortgage default. In column (1), I find that households in the low income tercile were 24 percent more likely to default on their mortgages. In column (2), I find that households in the middle income tercile were 17 percent less likely to default compared to households in the low income tercile. Households in the high income tercile were 22 percent less likely to default compared to the low income tercile.

In column (3) I find that households in the high spending tercile in the pre-period were 3 percent more likely to default compared to households that were not in the high spending tercile. In column (4), I find that households in the high mortgage-to-income ratio tercile were 12 percent more likely to default compared to households that were not in the high mortgage-to-income ratio tercile.

In columns (5), the variables in Table 1.2 and Table 1.3 are put into a single regression. The variables related to fragility are not much different from their coefficients in Table 1.2. However, the magnitudes for other household characteristics were decreased significantly in the multivariate regression. For example, the the low income tercile are now only 14 percent more likely to become delinquent when controlling for other variables, and other variables such as mortgage-to-income drops to 3 percent. Table 1.2 and Table 1.3 show that the measures for household savings in liquid assets tend to be more important as determinants of mortgage default than household characteristics such as income, spending, or leverage.

In Table 1.5, I look at the vulnerability of financially fragile households to shocks by explicitly looking at unemployment shocks and look that how it affects mortgage delinquency. In column (1), I find that households that incurred overdraft fees in the pre-period were 160 percent more likely to become delinquent on their mortgages when they became unemployed, compared to the months in which households in the sample were employed. Households that did not incur overdraft fees in the pre-period were only 50 percent more likely to become delinquent on their mortgages when unemployed, compared to the months in which households were employed.

In column (2), I find that households that had brokerage transaction in the pre-period were only 26 percent more likely to default in the months in which those households become unemployed compared to the months where the households remained employed. the 26 percent increase in likelihood of default is also insignificant at the 5 percent level. Households that did not have brokerage transactions and became unemployed were 82 percent more likely to default in the months in which those households were unemployed.

In column (3), I find that households in the low tercile of interest earned in the pre-period that became unemployed were 97 percent more likely to default on their mortgages compared to the months in which households were still employed. For households that were in the middle tercile of interest earned, the likelihood of defaulting on its mortgage went up by 96 percent when they became unemployed. Households that were in the high savings tercile were 23 percent more likely to default in the months where the households became unemployed.

In Table 1.6, I perform a similar type of analysis as with Table 1.5, but using income, spending, and mortgage-to-income terciles. In contrast to Table 1.5, I find that spread between the extreme terciles to be smaller. High income households that become unemployed are 59 percent more likely to default in those months, compared to the months in which households stay employed. Low income households that become unemployed are 81 percent more likely to be delinquent than employed households. The difference between the two terciles are only 22 percentage points. Similarly, the difference between high spending households and low spending households are only

42 percentage points and the difference between high MTI households and low MTI households is only 1 percentage point. These are small compared to the 110 percentage point difference between overdraft households and non-overdraft households and the 74 percentage point difference between low savings households and high savings households.

In this section, I find that overdraft fees and savings, as measures of household savings in liquid assets, were significantly associated with mortgage default. When these variables were interacted with unemployment shocks, it implied a significantly higher risk of default. These results strongly support the double-trigger models of mortgage default.

In untabulated results, I find that the effect of unemployment alone on mortgage delinquency to be large. Previous research typically found small effects for unemployment ([41], [50]), due to measurement issues ([55]). Households that become unemployed were 82 percent more likely to default than households that remained employed. These large results are consistent with the findings of [**?**], who also utilize individual-level data.

1.3.2 Household Savings versus High Income

Next, in Table 1.7, I look at how the effect of the household financial fragility on mortgage delinquency varies with income level. In Column (1), I find that households that incurred overdraft expenses, regardless of whether the household is in the high, middle, or low income tercile, had higher likelihoods of being delinquent than households that did not incur overdraft fees.

In columns (2) and (3), I interact spending and mortgage-to-income instead of income and find results that emphasize the importance of liquid assets. However since income levels are generally more meaningful in describing household outcomes than spending habits or leverage, I focus the analysis on income.

In Table 1.8, column (1), I look at how the effect of savings on mortgage delinquency varies with income level. The results are similar to the ones in Table 1.6. When households were in the low savings tercile, they were more likely to be delinquent whether or not they were in the high income tercile. The high income but low savings households were 17 percent more likely to be delinquent compared to middle income and middle savings households.

In table 1.9, I control for employment and find similar results. Compared to an omitted category of middle income, non-overdraft, and employed households, the low income, non-overdraft, and employed households were 16 percent more likely to default. When looking at the employed households that are high income but have overdraft fees, the increased likelihood of default is 20 percent.

These findings also support the work by [62] on the wealthy hand-to-mouth, households that are wealthy on paper but have their investment in illiquid assets and have high consumption sensitivity to income. I find many households in my sample are wealthy in the sense that they are homeowners, and who are hand-to-mouth in the sense that they are liquidity constrained (i.e. incurring overdraft fees) These wealthy hand-to-mouth households have a high propensity to become delinquent on their mortgages.

These results are also relevant to the discussion on the causes of the financial crsis. When we think of the causes of the financial crisis of 2008, we often visualize

Wall Street bankers and mortgage brokers that lent money to subprime borrowers who bought houses that they could not afford. In particular, [73] provide evidence that the expansion of mortgage credit to subprime borrowers was closely linked to the mortgage defaults during the crisis.

However, recent research shows that high-income, middle-income, and prime borrowers may have played a larger role in the 2008 financial crisis than previously realized. [2] and [3] find that mortgage originations increased for all income levels and FICO scores in the pre-2007 period. They also find that middle-income, high-income, and prime borrowers sharply increased their share of delinquencies during the crisis. Also, [42] find that about twice as many prime borrowers lost their homes compared to subprime borrowers during 2006 to 2012. Figure 1.1 uses data from HOPE NOW, an organization that helps homeowners pay their mortgages, to show the trend in foreclosures for prime and subprime mortgages during the crisis. The number of prime borrowers that lost their homes is much larger than the number of subprime borrowers that lost their homes. This is consistent with the findings of [2], [3] and [42]. The importance of liquid assets as a determinant of mortgage default for high income households, combiend with the finding that many high income household do not have much in liquid assets, show why high income households may have defaulted in large numbers during the financial crsis.

1.3.3 Consumption and Credit Card Debt Surrounding Mortgage Delinquency

Finally, I look at consumption and debt repayment surrounding mortgage default. Adverse shocks that accompany mortgage delinquency are not easy for households to deal with. If the household has sufficient savings then it can try to smooth the impact of the shock. However, if the household lacks savings, then they must respond to shocks in other ways, one of which is to reduce consumption.

In Figure 1.3, I show the trend in retail consumption surrounding mortgage default. Households that incurred overdraft expenses decreased their consumption by $98 at delinquency, which is a significant amount (27%) compared to the mean spending on retail of $369 by defaulting households. After default, the decline in consumption tended to stablize, though the household eventually decreases their consumption again later on. For non-overdraft households, the effect is not as large, but these households still reduce their consumption. The reduction is more gradual compared to overdraft households, but their consumption also decreases as time goes by.

In Figures 1.5 and 1.6, similar reduction of 20 to 30 percent can be found for restaurant and grocery spending. The difference in the reduction in consumption between fragile and non-fragile households are narrower for grocery consumption, due to it being less of a discretionary good, and also due to the opportunities of home production ([8]).

In Table 1.10, I look at credit card debt repayment surrounding mortgage delinquency. I find that households tend to repay their credit card debt during the same time that they dramatically reduce consumption and default on their mortgages. This result is also confirmed by [32]. This gives implications on the importance that households put on the availability of consumer credit.

These results shed some light in to the pecking order of coping methods proposed by [70]. I find that households initially reduce their consumption when hit by an adverse shock, which is followed by mortgage default. The suspension of mortgage payment gives some relief to the households in terms of stablized consumption. Despite the willingness to use credit to cope with shocks in [70]'s survey, I fine that households, on the contrary, pay back their credit card debt.

1.4 Conclusion

In this paper, I use administrative data from banking and credit card transactions to obtain data on ex-ante household savings in liquid assets and unemployment. Using this data, I find that households that lack liquid assets are more likely to default on their mortgages. They are also more sensitive to unemployment shocks compared to households with more savings in liquid assets. These finding are in support of the double-trigger hypothesis in mortgage default theory.

I also compare household savings in liquid assets and income and find that household savings is a more significant predictor of mortgage default than income. High income households that lack liquid savings are more likely to become delinquent on their mortgages than low income households that have more in liquid assets. These findings are broadly consistent with the research on the wealthy hand-to-mouth, where high income households may find it optimal to put themselves into a liquidity constrained position to gain higher returns on illiquid assets such as housing.

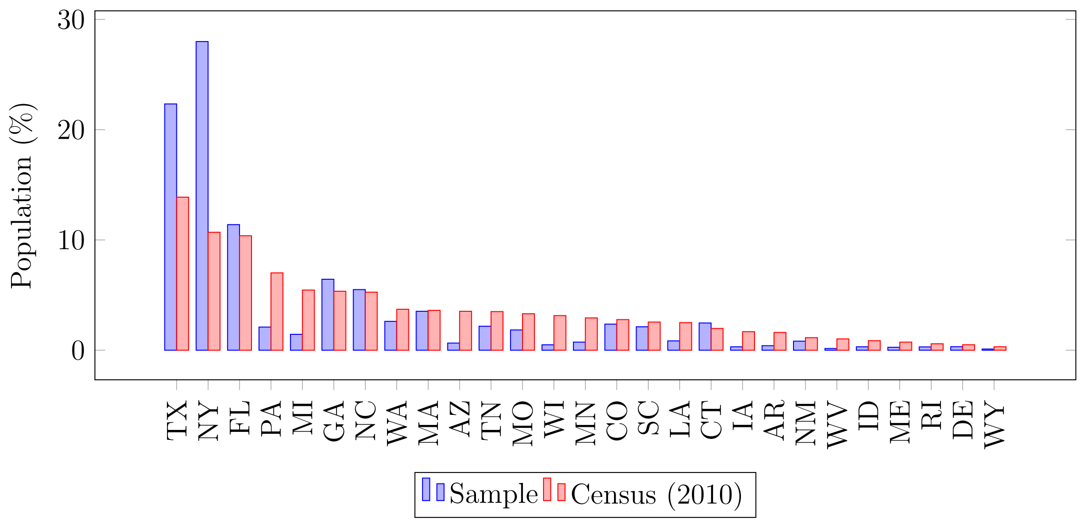
Finally, I look at the consumption and credit card debt repayment surrounding mortgage delinquency. I find that households decrease their consumption at default, but at the same time repay their credit card debt. Instead of using consumer credit as a means to smooth consumption, households seem to more concerned about preserving their access to credit.

These results emphasize the important role of households savings in liquid assets on mortgage default. Policies that encourage more liquid savings, such as escrow accounts may be useful in providing households with the buffer to withstand income shocks.

Figure 1.1: Geographic Distribution of Sample Households

This figure shows the distribution of households in each state . Panel A shows the proportion of the total households that reside in each state. The dark bars indicate the percentages for the sample of households used in this paper. The light bars indicate the percentages for the 2010 United States Census. Panel B shows the percentage point differences between the two percentages for each state.

Panel A



Panel B

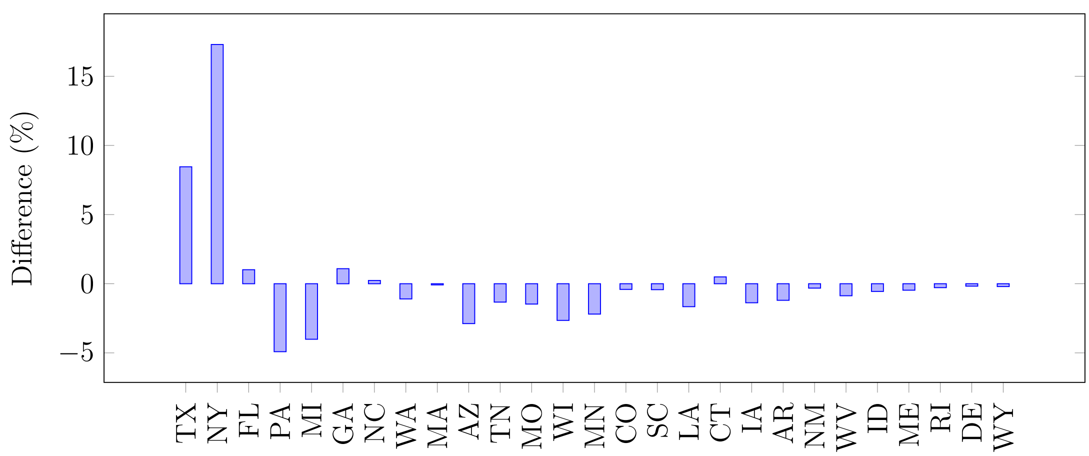


Figure 1.2: Distribution of Income

This figure shows the distribution of income of the sample, compared to the distribution of income in the 2013 American Housing Survey (AHS) and the 2012 Current Population Survey (CPS). The annual income for the sample are the average monthly income in the pre-period that is annualized. The sample income is after withholdings, so it is biased downward compared to the survey distribution.

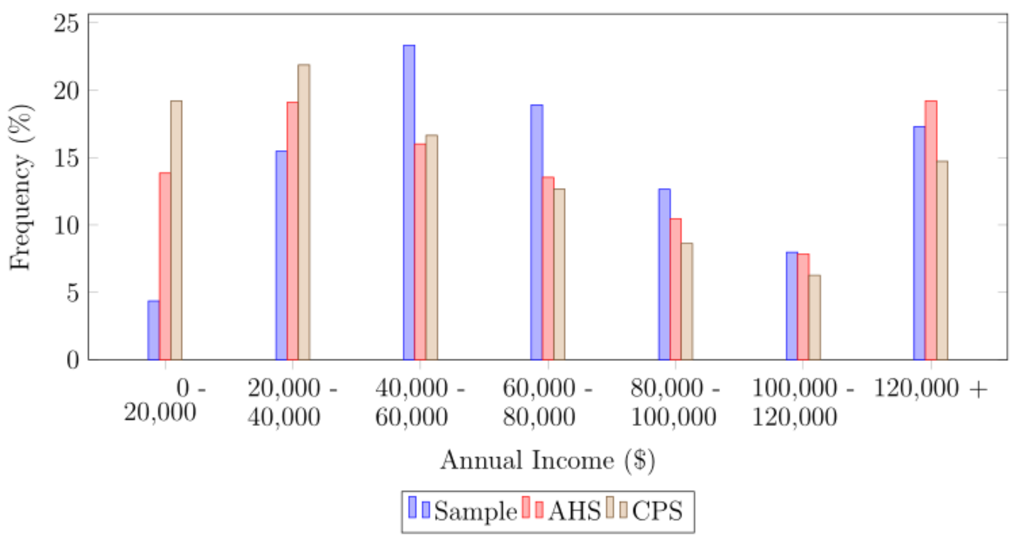
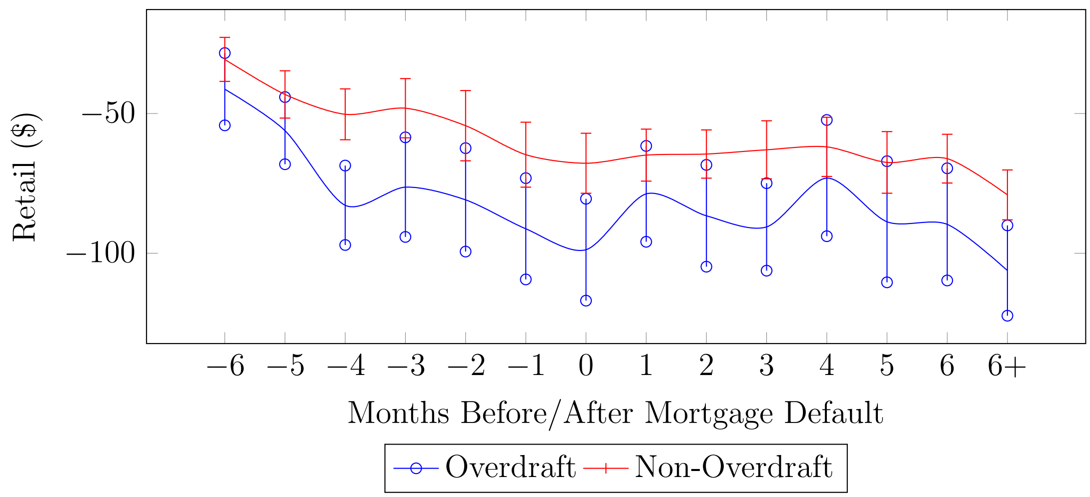


Figure 1.3: Retail Expenditures Surrounding Mortgage Default

This figure plots the OLS regression coefficients for retail consumption surrounding mortgage default. The regressions use a difference-in-differences design where the consumption for a particular month for households that default are compared to the consumption for households that do not default for the same month. The omitted variable is the all months prior to 6 months before default. Standard errors are clustered by household and year-month. The vertical lines represent 95% confidence intervals. The confidence intervals for the mid savings in Panel B are dropped for visual clarity.

Panel A



Panel B

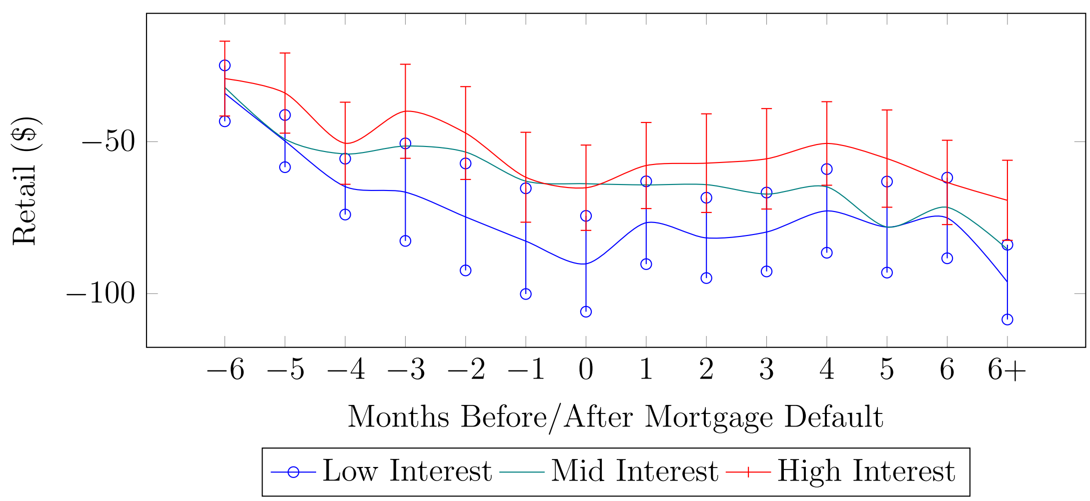


Figure 1.4: Restaurant Expenditures Surrounding Mortgage Default

This figure plots the OLS regression coefficients for restaurant consumption surrounding mortgage default. The regressions use a difference-in-differences design where the consumption for a particular month for households that default are compared to the consumption for households that do not default for the same month. The omitted variable is the all months prior to 6 months before default. Standard errors are clustered by household and year-month. The vertical lines represent 95% confidence intervals. The confidence intervals for the mid savings in Panel B are dropped for visual clarity.

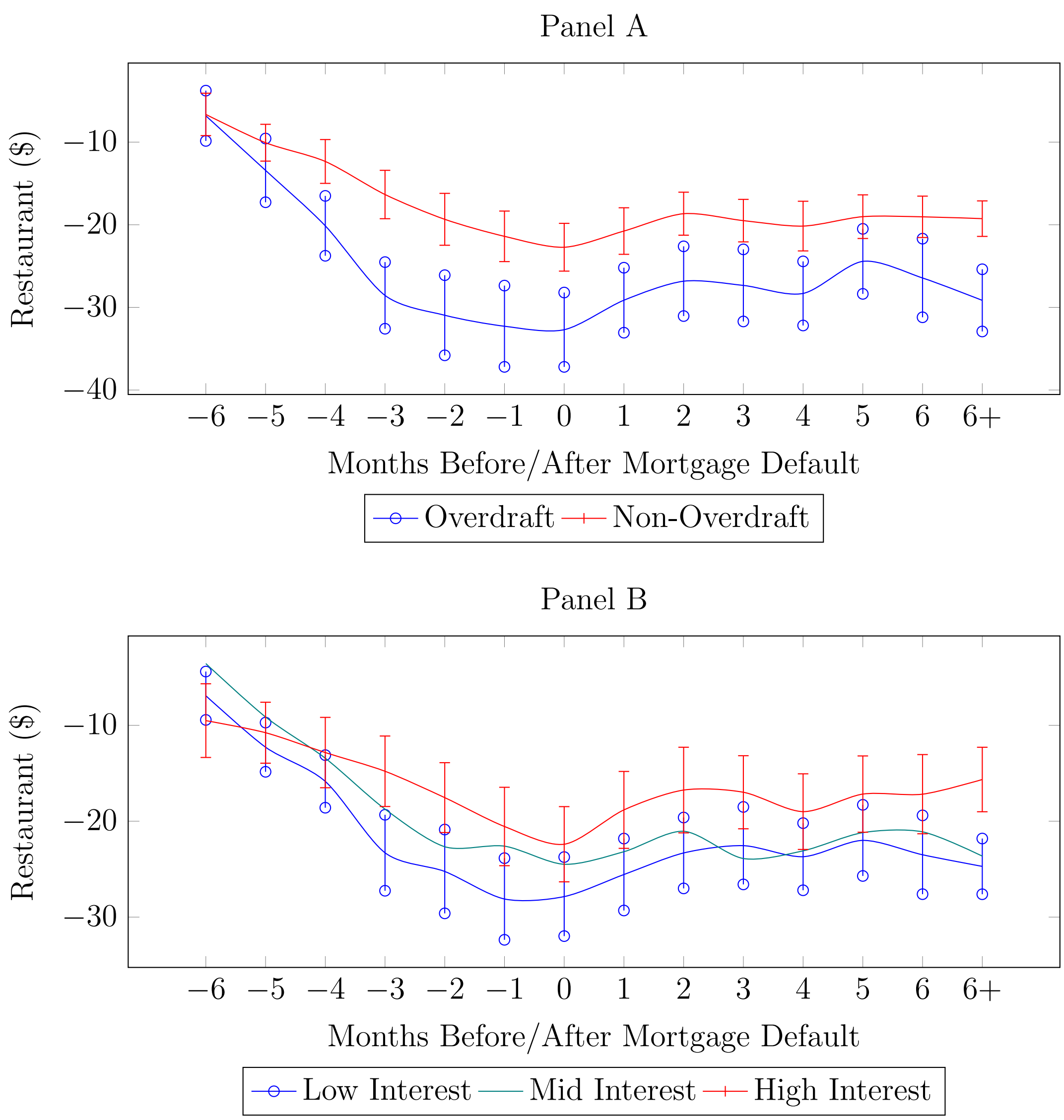


Figure 1.5: Grocery Expenditures Surrounding Mortgage Default

This figure plots the OLS regression coefficients for grocery consumption surrounding mortgage default. The regressions use a difference-in-differences design where the consumption for a particular month for households that default are compared to the consumption for households that do not default for the same month. The omitted variable is the all months prior to 6 months before default. Standard errors are clustered by household and year-month. The vertical lines represent 95% confidence intervals. The confidence intervals for the mid savings in Panel B are dropped for visual clarity.

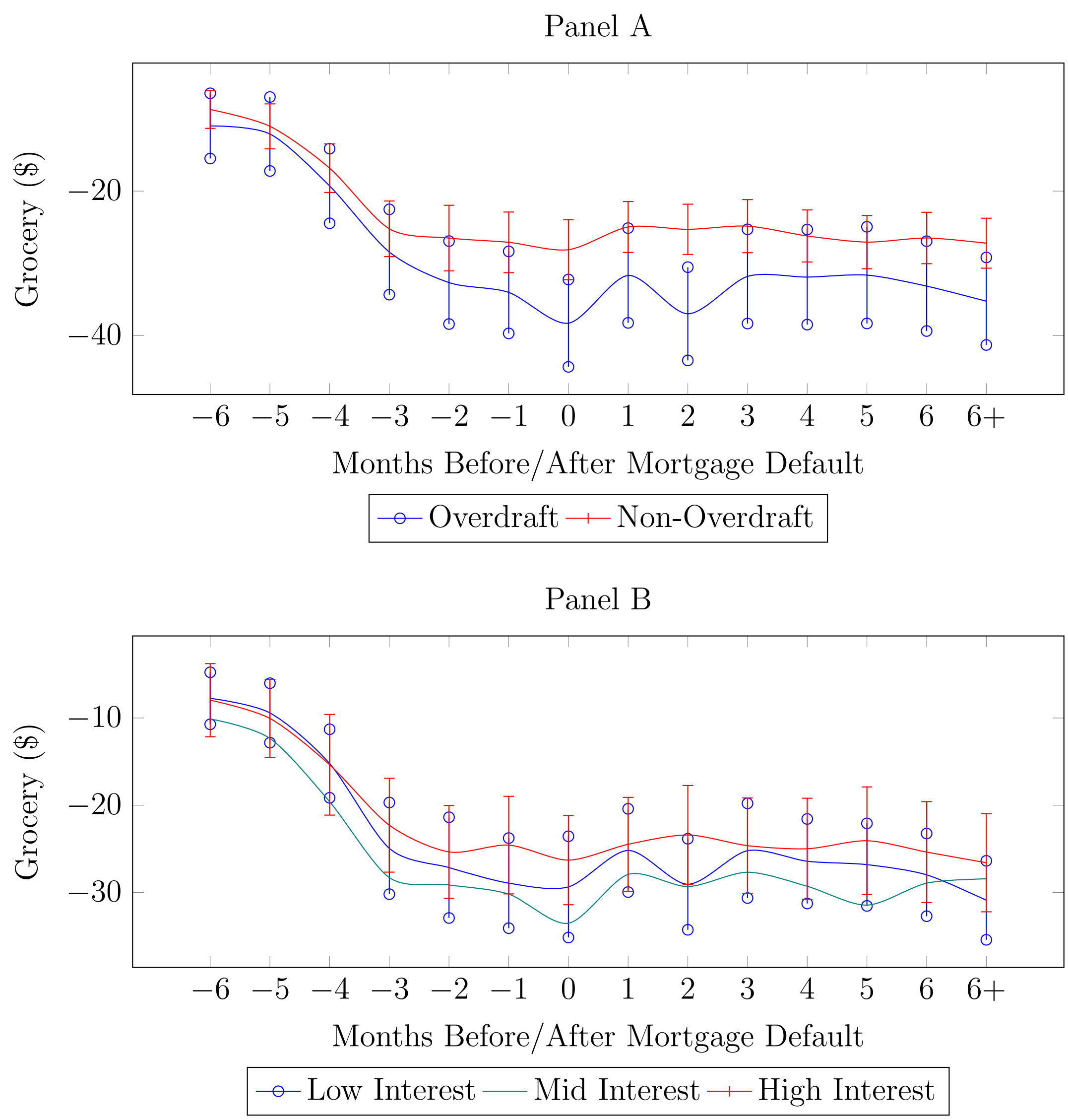


Table 1.1: Summary Statistics

This table presents the summary statistics for the households in the sample. Panel A shows the summary statistics for all households in the sample. Panel B shows the summary statistics for households that default on their mortgages during the sample period. Panel C shows the summary statistics for households that do not default on their mortgages during the sample period. All variables, except for the number of households, are measured as the monthly average amount in the pre-period for each household. All variables, except for the number of households, are rounded to the nearest dollar unless the amount is less than a dollar, in which case it is rounded to the nearest cent. The pre-period is the 6 month period in which households have both income and mortgage payment transactions.

Panel A

	Mean	StDev	Min	10%	50%	90%	Max
Mortgage Payment	2,035	2,525	1	800	1,612	3,537	839,568
Income	7,660	12,561	0.02	2,483	5,539	12,828	1,324,175
Interest Earned	8	112	0	0	0.36	13	30,281
Brokerage	547	7,617	0	0	0	70	2,501,883
Overdraft Fees	3	26	0	0	0	9	9,309
Consumption							
Retail	413	706	0	0	232	974	66,327
Restaurant	148	484	0	0	94	328	84,172
Grocery	115	284	0	0	16	367	84,522
Households (#)	265,144						

Continued

Table 1.1 Continued

Panel B

	Mean	StDev	Min	10%	50%	90%	Max
Mortgage Payment	2,049	2,158	1	764	1,578	3,572	83,247
Income	7,757	14,458	2	2,339	5,298	13,026	820,491
Interest Earned	10	221	0	0	0.29	13	30,281
Brokerage	536	9,284	0	0	0	50	1,108,582
Overdraft Fees	5	16	0	0	0	12	635
Consumption							
Retail	369	725	0	0	192	868	50,003
Restaurant	143	362	0	0	90	318	27,410
Grocery	99	199	0	0	10	319	5,129
Households (#)	28,082						

Panel C

	Mean	StDev	Min	10%	50%	90%	Max
Mortgage Payment	2,033	2,565	1	805	1,617	3,533	839,568
Income	7,648	12,317	0.02	2,500	5,565	12,805	1,324,175
Interest Earned	8	91	0	0	0.37	13	22,085
Brokerage	448	7,394	0	0	0	75	2,501,883
Overdraft Fees	3	27	0	0	0	8	9,309
Consumption							
Retail	418	703	0	0	237	986	66,327
Restaurant	149	496	0	0	94	329	84,173
Grocery	117	293	0	0	16	373	84,521
Households (#)	237,062						

Table 1.2: Households with Low Liquid Assets

This table presents the proportion of households with low liquid assets. Panel A shows the tercile breakpoints that separates the low-middle interest tecile and middle-high interest tercile, and also the proportion of households in each income tercile that belong in each interest tercile. The dollar breakpoints are the capitalized dollar amount that correspondes to the interest earned. Panel B shows that proportion of households that have incurred overdraft fees, for the entire sample and for each income tercile.

Panel A

	Low-interest	Mid-interest	High-interest
Tercile breakpoints ($)	0.103	1.285	
Dollar breakpoints ($)	2,472	30,840	
Low-income (%)	48.5	31.5	20.0
Mid-income (%)	33.6	37.0	29.4
High-income (%)	18.1	31.4	50.5

Panel B

	Overdrafts	Non-overdrafts
All (%)	19.4	80.6
Low-income (%)	24.1	75.9
Mid-income (%)	19.4	80.6
High-income (%)	14.7	85.3

Table 1.3: Household Savings and Mortgage Default

This table explores the relationship between household savings and mortgage default. The regressions in this table are dynamic logit models. All regressions include state and year-month fixed effects, changes in regional real estate prices, and a fifth-order polynomial in account age as controls. The regression coefficients are expressed in terms of relative risk. Mortgage delinquency equals one if a household is more than three months delinquent on its mortgage. Overdraft is a dummy that equals one if the household incurred overdraft fees during the pre-period. Brokerage is a dummy that equals one if the household had transactions with a financial brokerage or mutual fund companies during the pre-period. Low interest, mid interest, and high interest are dummies that equal one if the interest that the household received during the pre-period belongs in the low, middle, high tercile of households, respectively. Standard errors are clustered at the household level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Mortgage Default				
	(1)	(2)	(3)	(4)	(5)
Overdraft	1.43***				1.31***
	(0.02)				(0.02)
Brokerage		0.78***			0.80***
		(0.01)			(0.01)
Low Interest			1.43***		1.34***
			(0.02)		(0.02)
Mid Interest				0.74***	
				(0.01)	
High Interest				0.66***	
				(0.01)	
State FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
No. Obs.	7,660,189	7,660,189	7,660,189	7,660,189	7,660,189
Log Pseudolikelihood	-173,189	-173,331	-173,078	-173,051	-172,835
Pseudo-R^2	0.01	0.01	0.02	0.02	0.02

Table 1.4: Income, Spending, Mortgage-to-Income, and Mortgage Default

This table explores the relationship between other household characteristics and mortgage default. The regressions in this table are dynamic logit models. All regressions include state and year-month fixed effects, changes in regional real estate prices, and a fifth-order polynomial in account age as controls. The regression coefficients are expressed in terms of relative risk. Mortgage default equals one if a household is more than three months delinquent on its mortgage. Low income, mid income, and high income are dummies that equal one if the average household income in the pre-period belong in the low, middle, high tercile of households, respectively. High spending is a dummy that equals one if the ratio of consumption to income is in the high tercile of households in the pre-period. High Mortgage-to-Income is a dummy that equals one if the mortgage payment to income ratio for the household is in the high tercile of households in the pre-period. Standard errors are clustered at the household level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Mortgage Default				
	(1)	(2)	(3)	(4)	(5)
Low Income	1.24***				1.14***
	(0.02)				(0.02)
Mid Income		0.83***			
		(0.01)			
High Income		0.78***			
		(0.01)			
High Spending			1.03***		0.95
			(0.02)		(0.02)
High Mortgage/Income				1.12***	1.03***
				(0.02)	(0.02)
Overdraft					1.32***
					(0.02)
Brokerage					0.81***
					(0.02)
Low Interest					1.32***
					(0.02)
State FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
No. Obs.	7,660,189	7,660,189	6,989,505	7,660,189	6,989,505
Log Pseudolikelihood	-173,305	-173,299	-158,526	-173,404	-157,862
Pseudo-R^2	0.01	0.01	0.01	0.02	

Table 1.5: Unemployment, Household Savings, and Mortgage Default

This table explores the relationship between unemployment, household savings, and mortgage default. The regressions in this table are dynamic logit models. All regressions include state and year-month fixed effects, changes in regional real estate prices, and a fifth-order polynomial in account age as controls. The regression coefficients are expressed in terms of relative risk. Mortgage default equals one if a household is more than three months delinquent on its mortgage. Unemployment & Overdraft is a dummy variable that equals one if the household had incurred overdraft fees in the pre-period and if the household is unemployed in the current month. The definitions are similar for other variables. Standard errors are clustered at the household level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Mortgage Default		
	(1)	(2)	(3)
Unemployment & Overdraft	2.60***		
	(0.23)		
Unemployment & Non-Overdraft	1.50***		
	(0.09)		
Unemployment & Brokerage		1.26*	
		(0.17)	
Unemployment & Non-Brokerage		1.82***	
		(0.10)	
Unemployment & Low Interest			1.97***
			(0.15)
Unemployment & Mid Interest			1.96***
			(0.16)
Unemployment & High Interest			1.23**
			(0.12)
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
No. Obs.	7,660,189	7,660,189	7,660,189
Log Pseudolikelihood	-173,374	-173,383	-173,377
Pseudo-R^2	0.01	0.01	0.01

Table 1.6: Unemployment, Income, Spending, MTI, and Mortgage Default

This table explores the relationship between unemployment, other household characteristics, and mortgage default. The regressions in this table are dynamic logit models. All regressions include state and year-month fixed effects, changes in regional real estate prices, and a fifth-order polynomial in account age as controls. The regression coefficients are expressed in terms of relative risk. Mortgage default equals one if a household is more than three months delinquent on its mortgage. Unemployment & Low Income is a dummy variable that equals one if the household belonged in the low income tercile in the pre-period and if the household is unemployed in the current month. The definitions are similar for other variables. Standard errors are clustered at the household level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Mortgage Default		
	(1)	(2)	(3)
Unemployment & Low Income	1.81***		
	(0.13)		
Unemployment & Mid Income	1.68***		
	(0.15)		
Unemployment & High income	1.59***		
	(0.16)		
Unemployment & Low Spending		1.79***	
		(0.15)	
Unemployment & Mid Spending		1.80***	
		(0.16)	
Unemployment & High Spending		1.80***	
		(0.17)	
Unemployment & Low Mortgage/Income			1.54***
			(0.15)
Unemployment & Mid Mortgage/Income			1.88***
			(0.16)
Unemployment & High Mortgage/Income			1.72***
			(0.13)
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
No. Obs.	7,660,189	7,660,189	7,660,189
Log Pseudolikelihood	-173,386	-173,383	-173,386
Pseudo-R^2	0.01	0.01	0.01

Table 1.7: Overdrafts vs. Income, Spending, and Mortgage-to-Income

This table explores the effect of overdrafts against income, spending, and mortgage-to-income on mortgage default. The regressions in this table are dynamic logit models. All regressions include state and year-month fixed effects, a fifth-order polynomial in account age, and the changes in regional real estate prices as controls. The regression coefficients are expressed in terms of relative risk. Mortgage default equals one if a household is more than three months delinquent on its mortgage. Standard errors are clustered at the household level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Mortgage Default		
	Category:		
	Income	Spending	MTI
	(1)	(2)	(3)
Overdraft & Low Category	1.77***	1.42***	1.34***
	(0.04)	(0.04)	(0.04)
Overdraft & Mid Category	1.43***	1.41***	1.44***
	(0.04)	(0.03)	(0.04)
Overdraft & High Category	1.20***	1.47***	1.64***
	(0.04)	(0.04)	(0.04)
Non-Overdraft & Low Category	1.17***	1.00	1.02
	(0.02)	(0.02)	(0.02)
Non-Overdraft & Mid Category			
Non-Overdraft & High Category	0.98	0.98	1.09***
	(0.02)	(0.02)	(0.02)
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
No. Obs.	7,660,189	7,660,189	7,660,189
Log Pseudolikelihood	-173,027	-173,151	-173,124
Pseudo-R^2	0.02	0.01	0.02

Table 1.8: Savings vs. Income, Spending, and Mortgage-to-Income

This table explores the effect of savings against income, spending, and mortgage-to-income on mortgage default. The regressions in this table are dynamic logit models. All regressions include state and year-month fixed effects, a fifth-order polynomial in account age, and the changes in regional real estate prices as controls. The regression coefficients are expressed in terms of relative risk. Mortgage default equals one if a household is more than three months delinquent on its mortgage. Standard errors are clustered at the household level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Mortgage Default		
	Category:		
	Income	Spending	MTI
	(1)	(2)	(3)
Low Interest & Low Category	1.53***	1.34***	1.31***
	(0.04)	(0.03)	(0.04)
Low Interest & Mid Category	1.23***	1.35***	1.33***
	(0.03)	(0.03)	(0.04)
Low Interest & High Category	1.17***	1.38***	1.47***
	(0.04)	(0.03)	(0.04)
Mid Interest & Low Category	1.08***	1.04	0.98
	(0.03)	(0.02)	(0.03)
Mid Interest & Mid Category			
Mid Interest & High Category	0.97	0.95*	1.06**
	(0.03)	(0.02)	(0.03)
High Interest & Low Category	0.97	0.89***	0.94***
	(0.03)	(0.02)	(0.02)
High Interest & Mid Category	0.84***	0.85***	0.84***
	(0.02)	(0.02)	(0.02)
High Interest & High Category	0.91***	0.85***	0.93***
	(0.02)	(0.03)	(0.03)
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
No. Obs.	7,660,189	7,660,189	7,660,189
Log Pseudolikelihood	-172,189	-173,032	-173,024
Pseudo-R^2	0.02	0.02	0.02

Table 1.9: Unemployment vs. Overdrafts vs. Income

This table explores the effect of savings and unemployment against income on mortgage default. The regression in this table are dynamic logit models. The regression include state and year-month fixed effects, a fifth-order polynomial in account age, and the changes in regional real estate prices as controls. The regression coefficients are expressed in terms of relative risk. Mortgage default equals one if a household is more than three months delinquent on its mortgage. All numbers in this table are from a single regression with one omitted variable. Standard errors are clustered at the household level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Mortgage Default		
	Low Income	Mid Income	High Income
Employed & Overdraft	1.76***	1.42***	1.20***
	(0.04)	(0.04)	(0.04)
Employed & Non-Overdraft	1.16***		0.98
	(0.02)		(0.02)
Unemployed & Overdraft	2.90***	3.11***	2.69***
	(0.35)	(0.52)	(0.59)
Unemployed & Non-Overdraft	1.74***	1.60***	1.59***
	(0.15)	(0.18)	(0.18)
State FE		Yes	
Time FE		Yes	
Controls		Yes	
No. Obs.		7,660,189	
Log Pseudolikelihood		-172,980	
Pseudo-R^2		0.02	

Table 1.10: Credit Card Debt Repayment Surrounding Mortgage Default

This table explores credit card debt repayment surrounding mortgage default. The regressions in this table are OLS regressions, that use a difference-in-differences design where the debt repayment for a particular month for households that become delinquent are compared to the debt repayment for households that do not become delinquent for the same month. I(t $\geq Q_{-2}$) is a dummy variable for all months in and after the second quarter before default. I(t $= Q_t$) are dummy variable for the months in quarter t. I(t $> Q_2$) is a dummy variable for all months following the second month after default. Credit card debt repayment is the sum of all credit transactions for the current month less the sum of all debit transactions for the prior month, for the household's credit card accounts. The omitted variable is the all months prior to 6 months before default. Standard errors are clustered at the household and year-month level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A

	Credit Card Debt Repayment ($)			
	Subsample:			
	Overdraft		Non-Overdraft	
	(1)	(2)	(3)	(4)
I(t $\geq Q_{-2}$)	12.11***		7.74***	
	(4.42)		(2.65)	
I(t $= Q_{-2}$)		3.82		2.44
		(5.85)		(4.86)
I(t $= Q_{-1}$)		24.90***		19.21***
		(6.09)		(4.38)
I(t $= Q_0$)		12.16*		−3.68
		(6.61)		(5.30)
I(t $= Q_1$)		17.81**		2.75
		(7.34)		(6.11)
I(t $= Q_2$)		10.09*		14.55***
		(5.22)		(5.81)
I(t $> Q_2$)		8.39		9.53***
		(6.77)		(3.30)
State FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
No. Obs.	2,649,160	2,649,160	10,338,134	10,338,134
R^2	0.02	0.02	0.02	0.02

Continued

Table 1.10 Continued

Panel B

	Credit Card Debt Repayment ($)					
	Subsample:					
	Low Savings		Mid Savings		High Savings	
	(1)	(2)	(3)	(4)	(5)	(6)
I(t $\geq Q_{-2}$)	15.97***		6.21*		3.76	
	(4.02)		(3.55)		(3.47)	
I(t $= Q_{-2}$)		6.49		-5.68		8.19
		(5.67)		(6.53)		(8.02)
I(t $= Q_{-1}$)		26.64***		24.67***		9.39
		(4.52)		(7.26)		(6.96)
I(t $= Q_0$)		11.21**		2.59		-15.74**
		(5.07)		(8.37)		(6.50)
I(t $= Q_1$)		18.64***		5.16		-5.56
		(6.57)		(8.28)		(8.37)
I(t $= Q_2$)		20.64***		0.60		19.90**
		(5.19)		(6.85)		(9.03)
I(t $> Q_2$)		15.42**		6.96*		5.43
		(6.76)		(3.82)		(4.30)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
No. Obs.	4,351,982	4,351,982	4,473,961	4,473,961	4,161,351	4,161,351
R^2	0.02	0.02	0.02	0.02	0.02	0.02

Table 1.11: Household Savings and Mortgage Default, with Gaps

This table explores the relationship between household savings and mortgage default. These regressions are similar to those in Table 1.3, but includes a period of 3 months between the pre-period where default is not included. The regressions in this table are dynamic logit models. All regressions include state and year-month fixed effects, changes in regional real estate prices, and a fifth-order polynomial in account age as controls. The regression coefficients are expressed in terms of relative risk. Mortgage delinquency equals one if a household is more than three months delinquent on its mortgage. Overdraft is a dummy that equals one if the household incurred overdraft fees during the pre-period. Brokerage is a dummy that equals one if the household had transactions with a financial brokerage or mutual fund companies during the pre-period. Low interest, mid interest, and high interest are dummies that equal one if the interest that the household received during the pre-period belongs in the low, middle, high tercile of households, respectively. Standard errors are clustered at the household level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Mortgage Default				
	(1)	(2)	(3)	(4)	(5)
Overdraft	1.43***				1.31***
	(0.02)				(0.02)
Brokerage		0.78***			0.80***
		(0.01)			(0.01)
Low Interest			1.44***		1.34***
			(0.02)		(0.02)
Mid Interest				0.74***	
				(0.01)	
High Interest				0.66***	
				(0.01)	
State FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
No. Obs.	6,915,045	6,915,045	6,915,045	6,915,045	6,915,045
Log Pseudolikelihood	-171,919	-172,093	-171,837	-171,809	-171,594
Pseudo-R^2	0.01	0.01	0.01	0.01	0.01

Chapter 2: Loss Aversion and the Consumption Response to Adjustable-Rate Mortgage Payments

2.1 Introduction

The empirical literature on the life-cycle / permanent income hypothesis (LCPIH) has repeatedly demonstrated that household consumption is sensitive to changes in cash flows [91], [78], [84], [76], [86], [49], [87], [88], [6], [1], [92], and others). This is in contrast with the predictions of the LCPIH, which is that households should not show consumption sensitivity to expected changes in income. Researchers have developed two main methods of explaining why the empirical results reject the classical LCPIH ([60], [77]).

The first keeps the basic underlying assumptions of the LCPIH intact, but adds financial frictions such as liquidity constraints or income uncertainty. Households that are liquidity constrained would like to smooth income, but are unable to do so since they cannot borrow from future increases in income ([56], [91], [59]). Similarly, households that lack wealth create a buffer stock of savings to insure against future shocks, and therefore show high propensities to consume out of increases in income that exceed that buffer ([31]).

The other strand of theory posits that households are not optimizing for the future consumption. Instead, households are subject to behavioral biases that follow rules of thumb that are generally based on current, rather than permanent, income ([65], [43], [28], [69]). Under these assumptions, households either overvalue current consumption or follow a simple rule of consuming a fraction of current income, leading them to consume more when there is an increase in income.

Unfortunately, it is not easy to test which theory better fits the data. They both predict changes in consumption following increases in income. Even when looking at households that are liquidity constrained, the behavioral theories can explain that households that are subject to more behavioral biases become more liquidity constrained, leading to the same prediction that liquidity constrained households will show a higher sensitivity to income shocks than non-constrained households.

[84] found a way to empirically differentiate these theories, which was to distinguish household consumption after income increases and decreases. Under liquidity constraints, although the household is constrained against future income increases, if the household expects income to decrease, it is not constrained from savings current income to smooth the decline in future income. Therefore, we can expect households to show consumption sensitivity to future expected income increases while now showing consumption sensitivity to future expected income decreases. On the other hand, under the current income models of consumption, the households consumption changes based on current changes in income, whether it is positive or negative. If the households faces a positive income shock, then it will increase consumption when the income increases, and if the household faces a negative income shock, it will decrease its consumption then the income decreases.

Finally, [24] develop a theory based on [61]s prospect theory that predicts that households will resist consuming below their reference points even when expected income is below the reference point and change consumption when the income shock is realized. This reluctance to realize losses is due to loss aversion. When expected income increases consumption adjustment is immediate, so that there may be no changes following the realization of the income changes. What is required in testing [24]s model is that there be changes in expected income accompanied by uncertainty and that the current or previous income is a reference point for the household in making consumption decisions.

To summarize, if households increase consumption following expected income increases but not income decreases, then this is consistent with LCPIH under frictions such as liquidity constraints. If households increase consumption following expected income increases and decrease consumption following expected income decreases, this is consistent with the behavioral models based on current income. If households decrease consumption following expected income decreases but not income increases, then this is consistent with consumption based on loss aversion.

Consumption surrounding changes in adjustable-rate mortgage (ARM) payments is a setting in which this can be tested. Due to the variability of interest rates, there are both cases of positive and negative income changes for households, as shown in Figure 1. There are also expected changes in future income that follow from changes in interest rates, but remain uncertainty since interest rates can continue to change in the future until the next adjustment.

I find that households reduce their consumption when their mortgage payments increase, but do not increase their consumption when their mortgage payments decrease. This result is consistent with [24]s model of consumption based on loss aversion. Households decrease their consumption by 13 dollars when mortgage payments increase following ARM rate adjustments. However, there is no significant increase in consumption following mortgage payment decreases.

Next, I examine large mortgage payment changes of over $100 and compare it with smaller mortgage payment changes under $100 and find consistent results that show households are sensitive to mortgage payment increases but not mortgage payment decreases. I also split households into subsamples based on how likely the household is to be liquidity constrained. I also find no consistent result in these subsamples that show more liquidity constrained households show more consumption sensitivity to income changes than less constrained households. What is more important in this subsample analysis was whether the income change was an increase (mortgage payment decrease) or a decrease (mortgage payment increase).

This paper contributes to the empirical literature that has attempted to differentiate between rational agent models and behavioral models. [84] look at long-term union contracts and finds that consumption is more sensitive to predicted income declines than to predicted income increases, which is consistent with the results of this paper. Other papers look at the expected income decline at retirement. [16] and) [21] found that households dramatically reduced consumption after retirement, which is contrary to the predictions of the rational agent models. However, [8] point out that the drop in consumption could be due to home production, leisure, and an increased amount of time available for shopping. In contrast to the results of this paper, [20] look at tax refund receipt and tax refund payment and find that households smooth consumption for tax payments but increase consumption after receiving tax refunds.

This paper also contributes to the recent literature on changes to mortgage payments and its effect on household outcomes. One of the policy responses to the Great Recession of 2008 was to create programs such as the Home Affordable Modification Program (HAMP) and the Home Affordable Refinance Program (HARP) where households could modify or refinance to lower debt burdens. [4] find that the HAMP had effects for lowering foreclosures and housing price declines but did not observe effects for consumption. [47] also look at HAMP and find payment reductions are more effects than principal reductions for default and consumption.

2.2 Data and Empirical Methodology

The data used in this paper comes from an internet aggregator of financial accounts that make it easier for households to take care of their finances. Research using this type of data is becoming more common. For example, [14], [17], [18], [20], [22], [47], and [68] use similar data. The data differs from prior surveys in that they do not suffer from self-reporting bias, and that they capture a wider range of household consumption. Also, in contrast to research using credit card data, researchers are able to observe household income, debt, and consumption simultaneously.

The data contains banking and credit card transactions for more than 2.7 million households from July 2010 to May 2015. I first select households that have more than 12 mortgage payments in a period that spans more than 12 months. This is done as an initial screen to reduce the size of the dataset and results in 823,332 households. I look at households that generally keep only one mortgage at a time, and eliminate households with three or more continuous months of overlapping payments to different mortgage lenders. This further reduces the sample to 677,602 households.

To identify then changes in mortgage payments due to ARM rate adjustments, I first find households that pay the identical mortgage for at least three months. In the sample of these household-payment pairs, I require that households show a change to a different payment amount, from the same mortgage lender, and that there be no gap in payment between the household-payment pairs. The reason for not allowing a gap in payment is that households sometimes skip payments when modifying mortgages from the same lender and therefore a skipped payment may indicate that it was a modification of the loan instead of an ARM readjustment. This period in which the mortgage payment changes is identified as the rate adjustment period and the household consumption in the periods before and after the rate adjustment is the focus of analysis. These requirements lead to 350,388 households.

One concern in the data is that I do not directly know if the change in mortgage payment was due to an ARM rate adjustment. Mortgage payments can change for multiple reasons, which include mortgage modifications, refinancing, changes in property taxes, changes in insurance premiums, or any other problems with the mortgage escrow account. For the purposes of this paper, changes in mortgage payment due to property tax changes are not an issue, since it is not the household that is making the decision in conjunction with consumption. However, payments due to modification can be a problem since that is initiated by the household. To address these concerns, I require that households have changes in their mortgage payments in the same month for at least three years. This finally gets the sample to 60,612 households that face adjustments in their mortgage payments 154,181 times. Each month in the sample is assigned to an ARM rate adjustment that is the closest to the month. I distinguish whether the month is before or after the adjustment and the entire period that surround the rate adjustment is referred to as the adjustment period. Another concern is that changes in interest rates can affect both permanent income and mortgage payments. However, this does not impact the results as long as the timing of interest rates and the timing of mortgage rate adjustments are uncorrelated.

Mortgages, income, brokerage transactions, overdraft fees, and interest earned are identified using a keyword search method. I identify commonly used phrases to categorize the financial transactions into each category. For example, if the transaction description includes the term wells fargo home loan, then I categorize that into a mortgage payment. For the consumption variables such as retail spending, I use the categorization provided by the internet aggregator. For each transaction description, I use the most commonly assigned category provided by the internet aggregator as the category for all uses of that transaction description. Then I combine the retail, restaurant, grocery, travel, and entertainment categories to form a single consumption variable.

I report the summary statistics in Table 2.1. In Panel A, I look at the mortgage payment, mortgage payment adjustments, income, and interest earned for all households in the sample. All values are in monthly averages and are measured in the months that precede ARM rate adjustment for each household. The median after-tax monthly income for the sample is $5,681, which is higher than the median income for the general population as measured by the Census.

In Panel B and Panel C, I report the summary statistics for each adjustment period, with each variable measured in the months preceding the rate adjustment. The mortgage payment adjustments are in general quite small. The median change in mortgage payments for increases is $30, and the median change in mortgage payments for decreases is $23. For the overall sample median change in payment is $8. In Figure 2, I show the distribution of increases and decreases in mortgage payments due to ARM rate adjustments. In Panel A, I show the distribution for increased mortgage payments, and show that 23 percent of the households that had increased mortgage payments experienced changes of less than $10. In Panel B, I show the distribution for decreased mortgage payments and show that 28 percent of households that had decreased mortgage payments experienced changes of less than $10.

The regressions in the paper compare the difference in consumption of households that faced ARM rate adjustment to households that did not face such adjustment, compared to the period in which both groups of households did not face ARM rate adjustments. I measure the total period after the rate adjustment until the month where it is closer to the next ARM rate adjustment compared to the previous rate adjustment. I also measure the consumption for the individual months following ARM rate adjustment. Variables are winsorized at the 99th percentile and standard errors are clustered for each adjustment period and for the year-month.

2.3 Results

2.3.1 Consumption after ARM Rate Adjustments

Table 2.2 presents the main results. Columns (1) and (2) report the regression results that measure changes in consumption after both increases and decreases in mortgage payments following ARM rate adjustments. Columns (3) and (4) report the regression results that measure changes in consumption after increases in mortgage payments following ARM rate adjustments. Finally, Columns (5) and (6) report the regression results that measure changes in consumption after decreases in mortgage payments following ARM adjustments. Columns (1), (3), and (5) report the average monthly changes in consumption following ARM rate adjustments, until the periods before the next ARM rate adjustment. In Columns (2), (4), and (6) I report the regression results for changes in consumption each month following ARM rate adjustments.

In Columns (1) and (2), $I(t = Mt)$ and $I(t \geq Mt)$ do not represent dummy variables. Instead, they are variables that equal 1 if households face mortgage payment increases following ARM rate adjustment and if the month of observation is the month t after the rate adjustment, in the case of I(t = Mt). If households face decreased mortgage payments following ARM rate adjustments, the variable equals -1. This means that the expected sign of the coefficient for this variable is negative if households decrease consumption after the size of the mortgage payments increase and increase consumption after the size of the mortgage payment decrease. The reason for including this variable is to analyze the average changes in household consumption that include household-adjustments that both increase and decrease in the same regression.

On the other hand, the variables $I(t = Mt)$ and $I(t \geq Mt)$ for Columns (3) to (6) do represent dummy variables. If households decrease consumption after increased mortgage payments, the coefficients shown in Columns (3) and (4) should be negative.

Similarly, if households increase consumption after decreased mortgage payments, the coefficients shown in Columns (5) and (6) should be positive.

In Column (1), I find that household consumption is not sensitive to ARM rate adjustments in general. The coefficient is negative, which is expected, but not close to being significant. In Column (3), I look at cases of ARM rate adjustments where the mortgage payments are increased, which leads to increased debt burdens for the household. In this regression, I find that consumption is sensitive to increased mortgage payments. Households reduce consumption by $14, which can be compared to the average increase in mortgage payments of $60, and a median increase in mortgage payments of $30. In Column (5), I look at cases of ARM rate adjustments where the mortgage payments are decreased, leading to a reduction in the debt burden for households. I find that household consumption is not sensitive to this reduction in mortgage payments. The coefficient for Column (5) is negative, if significant, would mean that households reduced consumption despite having more room in the household budget due to the decrease in mortgage payments.

In Columns (2), (4), and (6), I divide the period after the ARM rate adjustment into variables that represent each month up to the seventh month after adjustment, where all months after the seventh month after adjustment is represented by a single variable. In these regressions, I find similar results to that in Columns (1), (3), and (5). In Column (2), I find that households do not exhibit consumption sensitivity when looking at both positive and negative ARM rate adjustments, which is consistent with the results in Column (1). In Column (4), I find that households reduce consumption due to increased mortgage payment after ARM rate adjustments. Though the coefficients are negative from the first month after ARM rate adjustment to the seventh month after ARM rate adjustment, only the second month up to the sixth month after the rate adjustment is significant. The dollar reduction in consumption peaks at the fourth month after rate adjustment, at $22. This is a large reduction compared to the mean and median mortgage payment increases of $60 and $30, respectively. In Column (6), I analyze consumption following rate adjustments that lead to a smaller mortgage payment. I find that the consumption sensitivity to a smaller mortgage payment is generally insignificant, except for the counterintuitive results in the fourth month after rate adjustment where households consume less even though their mortgage payments had decreased.

In Panel B, I run similar regressions as in Panel A, but instead of including variables that indicate the timing of rate adjustment, I replace it with the amount of the mortgage payment changes due to the ARM rate adjustment. This specification allows me to directly estimate the amount of changes in consumption relative to the dollar change in mortgage payments. The overall results are consistent with the results in Panel A.

In Column (1), I find that household consumption is not sensitive to the amount of change in mortgage payments due to ARM rate adjustments. In Column (3), I find that household consumption declines when mortgage payments increase. For every dollar increase in mortgage payment, I find that households decrease their consumption by seventeen cents. In Column (5), I find that household consumption does not significantly change after rate adjustments. This is consistent with the results in Panel A. Unlike Column (5) of Panel A, the coefficient for Column (5) of Panel B is positive, though both are far from being statistically significant.

Columns (2), (4), and (6) show changes in consumption for each month after ARM rate adjustment, relative to the dollar amount of the mortgage payment changes. The results are similar to Panel A, where consumption after both increases and decreases in mortgage payment, and consumption after decreases in mortgage payment are insignificant. Consumption after increases in mortgage payment due to rate adjustment declines, though it is statistically significant for the third to sixth month after rate adjustment. The month where consumption reduction is highest if in the sixth month, where for every dollar increase in mortgage payments, consumption decreases by 35 cents. In Column (2), I find that consumption in the seventh month after rate adjustment is negative, but this seems to be driven by the counterintuitive result in Column (6) where household consumption declines despite have to pay smaller mortgages. Also, in Column (6) the coefficients for the regression are mostly positive though insignificant, except for the sixth month after rate adjustment.

The regressions in Table 2.2 provide evidence that is consistent with the predictions of Bowman, Minehart, and Rabin (1999)s model of consumption based on loss aversion. Household consumption is sensitive to decreases but not increases in income. Theories that add friction to the LCPIH such as liquidity constraints and buffer-stock predict that consumption should only be sensitive to increases in income, which is the opposite of what is found in Table 2.2. In addition, the family of behavioral models based on current income, present bias, or rule of thumb predict that consumption should be sensitive to both increases and decreases income.

2.3.2 Consumption after ARM Rate Adjustments, by Subsamples

ARM Rate Adjustments over $100

In the following tables, I divide the sample into subsamples for further analysis. In Table 2.3 and Table 2.4, I divide the sample into adjustment periods that had ARM rate adjustments which resulted in mortgage payment changes over $100 and under $100, respectively. One benefit from dividing the sample into large mortgage payment changes and small mortgage payment changes is to test the magnitude hypothesis, which is also tested by [83] who looks at consumption smoothing after the last mortgage payment.

[67],) [25], [58], [83], and [45] claim that due to bounded rationality or other mental costs households are less likely to be sensitive to small changes in income and be sensitive to large changes.

In my data, most changes in mortgage payments due to ARM rate adjustments are small. The median change in payments for increases and decreases are $30 and $23 respectively. In order for the change in mortgage payment to be over $100, it has to be near the 90th percentile for both mortgage payment increases and decreases. I compared results in Table 2.3 and Table 2.4 to examine if the magnitude hypothesis is relevant for households facing changing mortgage payments.

In Table 2.3, I perform similar regression to Table 2, except that the sample only includes adjustment periods that experienced changes in mortgage payments over $100. In Panel A, I find results that are similar to Panel A of Table 2.2. Household consumption is sensitive to increases in mortgage payments, while not being sensitive to both increases and decrease or only decreases in mortgage payments. The dollar change compared to Table 2.2, Panel A is larger, reflecting the larger change in mortgage payment. In Column (4) I find that the reduction in consumption is the largest in the sixth month after rate adjustment, where consumption declines by $67 on that month.

In Panel B, the results are again similar to Table 2.2. In Column (3), I find that for every dollar increase in mortgage payments, consumption declined by 13 cents for adjustment periods with changes in mortgage payments over $100. The monthly results in Column (4) are not as consistently significant as was in Table 2.2. Only the consumption in the third and sixth month after ARM rate adjustment are statistically significant, though this could be due to the lack of power from the smaller sample. The size of the coefficients are actually smaller than compared to Panel B of Table 2.2. In Table 2.2, consumption declined 17 cents for every dollar change in mortgage payments, while in Table 3 the number is 13 cents. The largest reduction in consumption is 33 cents in Table 2.3, whereas it is 35 cents in Table 2.2. This is inconsistent with the magnitude hypothesis since households would care more about the larger changes in income.

ARM Rate Adjustments under $100

In Table 2.4, I perform similar regressions on a sample that only includes adjustment periods that experienced changes in mortgage payments under $100. In Panel A, I again find similar results to Table 2.2 and Table 2.3. The coefficients for Panel A are smaller than that of Table 2.2 and Table 2.3, and this reflects that fact that the size of change in mortgage payments were also smaller. In Column (3), I find that households that experienced increased mortgage payments due to ARM rate adjustments reduced consumption by eleven dollars. In contrast, consumption sensitivity for regression (1) and (3) which look at both increases and decreases in mortgage payments and only decreases in mortgage payments were not significant. In Column (4) I find that the consumption was lower from the second month after rate adjustment to the sixth month, and the reduction was largest on the fourth month at $20.

In Column (3) of Panel B, I find that consumption declines by 18 cents for every dollar increase in mortgage payments. This is larger than the regressions in Table 2.2 and Table 2.3, which report decreases in consumption of 17 cents and 13 cents. The same result is found in Column (4). For every dollar increase in mortgage payment, the largest monthly reduction in consumption after rate adjustment for the sample in Table 2.4 was 42 cents. In Table 2.2 and Table 2.3 this was 35 cents and 33 cents respectively. The overall results of Table 2.3 and Table 2.4 are not consistent with the magnitude hypothesis. I find consumption sensitivity for small increases in mortgage payments and when I compare it with the size of the mortgage payments, it is larger for the small increases. I also find that consumption sensitivity to increased mortgage payments and non-significance of sensitivity to decreased mortgage payments remain regardless of the size of the change in the mortgage payments.

Income Subsamples

In Tables 2.5 through 2.8, I split the sample into groups based on household characteristics that reflect liquidity constraints. This method of testing for the effect of liquidity constraints by looking at subsamples has been used since Zeldes (1989). Results that show stronger sensitivity to income changes for more liquidity constrained households are typically interpreted as evidence in support of the liquidity constraints view of consumption, though it is also consistent with the view that households that have present-bias become liquidity constrained due to their myopia. In contrast, the loss-aversion models do not have an explicit prediction with respect to liquidity constraints.

In Table 2.5, I split the sample in to income quintiles for households that have income transactions in the data. Income is not a direct measure of household liquidity since low income households can still have high savings and thus high liquidity. Regardless of these concerns, I report results for the bottom, middle, and top income quintile. In Panel A, I examine the change in consumption after ARM rate adjustment for low, middle, and high income quintiles and find that consumption does not significantly change in response to the change in mortgage payment, even for households in the low income quintile. In Panel B, I examine households that experience an increase in mortgage payments and find that consumption is reduced for the low and middle income quintile. Households in the high income quintile also show a similar magnitude decline in consumption as measured by the coefficient, but it is not statistically significant. When looking at the change in consumption for each dollar increase in mortgage payment, the high and middle income quintiles show a statistically significant decrease in consumption, while households in the low income quintile only show a consumption decline that is significant at the 10% level. In Panel C, I find that consumption does not significantly change in response to the decrease in mortgage payment.

Under liquidity constraints, consumption is predicted to increase for decreases in mortgage payment and not for increases in mortgage payment. Also, the in measuring the size of the coefficients, it should be larger for the low income quintile compared to the high income quintile. In Table 2.5, I find that consumption is sensitive to increases in mortgage payment but not decreases, which is the opposite of the prediction under liquidity constraints. The size of the coefficients between the three quintiles, when they are statistically significant are not much different from one another, which is another contradiction to the predictions.

Interest Earned Subsamples

In Table 2.6, I perform a similar analysis to Table 2.5, but I split the sample in to interest earned quintiles. For these splits, I assume that households that do not report interest earned in their bank accounts do not have sufficient savings to earn interest instead of dropping them from the subsample analysis, as was done in Table 5. In Panel A, I find statistically significant results for consumption changes for each dollar change in mortgage payment for the low income quintile in Column (2). I find that households decrease or increase their consumption by 12 cents for each dollar in mortgage payment that was increased or decreased. Other than the single regression in Column (2), the results remain insignificant. In Panel B, I find significance for all regressions except for the change in consumption for low income quintiles following ARM rate adjustment in Column (1). The coefficient for consumption changes for the high interest earned quintile was $21, which was much larger than the coefficient for the low income quintile. The consumption change relative to the change in mortgage payments in Columns (2), (4), and (6) were also not much different from each other and the high income quintile actually had a larger coefficient than the low income quintile, which is contrary to the predictions of the liquidity constraints models. In Panel C, I find no significant results in consumption changes following decreases in mortgage payments.

Brokerage Transaction Subsamples

In Table 2.7, I split the sample into households that had brokerage transactions in the periods before the rate adjustment for each adjustment period. The typical household does not hold financial assets, and usually the largest asset is the house ([26]). Therefore, if the household has had transactions with financial brokerages indicating ownership of financial assets, it could be interpreted as indication of high liquidity. In Panel A and Panel B, I find that households that did not have any transactions with brokerages show significant consumption sensitivity to changes in mortgage payments. Unlike Table 2.5 and Table 2.6, this could be interpreted as evidence in support of liquidity constraints models. However the difference between the coefficients are not significant.

Overdraft Fee Subsamples

In Table 2.8, I split the sample into households that had incurred overdraft fees in the periods before the rate adjustment for each adjustment period. Households incur overdraft fees because of the lack of savings in the bank, and therefore can be regarded as households that lack liquidity. I do not consider household that have overdraft protection to have incurred overdraft fees when the protection is triggered. In Panel B, I find that households that did not incur overdraft fees to show reduction in consumption after increases in mortgage payments. Households that incurred overdraft fees either showed insignificant results, as in Column (3), or showed significance but had smaller coefficient than compared to the sample that did not incur overdraft fees, as in Column (4). These results are also inconsistent with the predictions of the liquidity constraints models.

Mortgage to Income Subsamples

In Table 2.9, I split the sample into households depending on the the ratio of mortgage payment to income in the period before the mortgage rate adjustment as another measure for liquidity constraints. Comparing households. The results for the mortgage to income splits are similar to the tables above, in the sense that household consumption is sensitive to increases in mortgage payments but not to decreases in mortgage payments.

2.3.3 Savings after ARM Rate Adjustments

In Table 2.10, I examine changes in measures of household savings instead of household consumption. I look at the amount of interest that the household incurred, the amount of brokerage transactions, and the amount of overdraft fees that the household had to pay. These measures of savings are compared to the period before the ARM rate adjustment, just as was the case for consumption. While there are significant results in some of the regressions, there is no general pattern of results that allow for an interpretation that support or reject consumption theories.

In Column (4) of Panel A, I find that households showed a 0.09 cent decrease in interest earned for every dollar increase in mortgage payment. However, there are no significant coefficients for other regressions in the panel. In Column (1) of Panel B, I find that brokerage transactions were sensitive to changes in mortgage payment changes. In Column (3), I find that households reduced brokerage transactions by $18 after mortgage payments increased. Finally in Column (6) of Panel C, I find that households paid less in overdraft fees after mortgage payments decreased.

2.4 Conclusion

In this paper, I test the predictions of consumption theories that explain the empirical inconsistencies with the LCPIH. The key to differentiating between these theories is to look at both increases and decreases in expected income. I find that households reduce their consumption when their mortgage payments increase due to ARM rate adjustments. However, I do not find evidence of households increasing their consumption when mortgage payments decrease. These results are in support of the consumption model of [24], which is based on loss aversion.

The finding that consumption does not increase after a reduction in mortgage payment is rather peculiar, since there are many papers that document increased consumption after increases in income. [38] and [66] also look at ARM rate adjustments and find that household consumption increases when rates are lowered. One difference between [38] and [66]s results and this paper is the difference in the magnitude of the rate adjustment. They find results for larger rate adjustments, while this paper focuses on smaller rate adjustments. The measure of consumption is also different.

Figure 2.1: 1-Year LIBOR Rates

This figure presents trends in the 1-year LIBOR rate during the sample period. The data was downloaded from the Federal Reserve Bank of St. Louis.

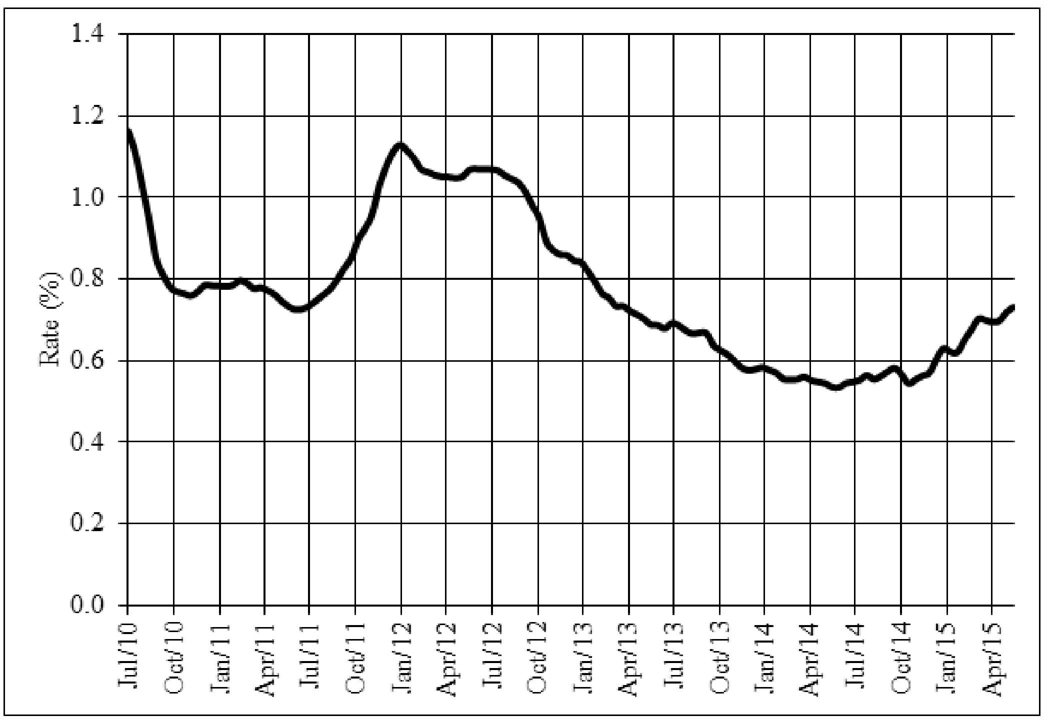
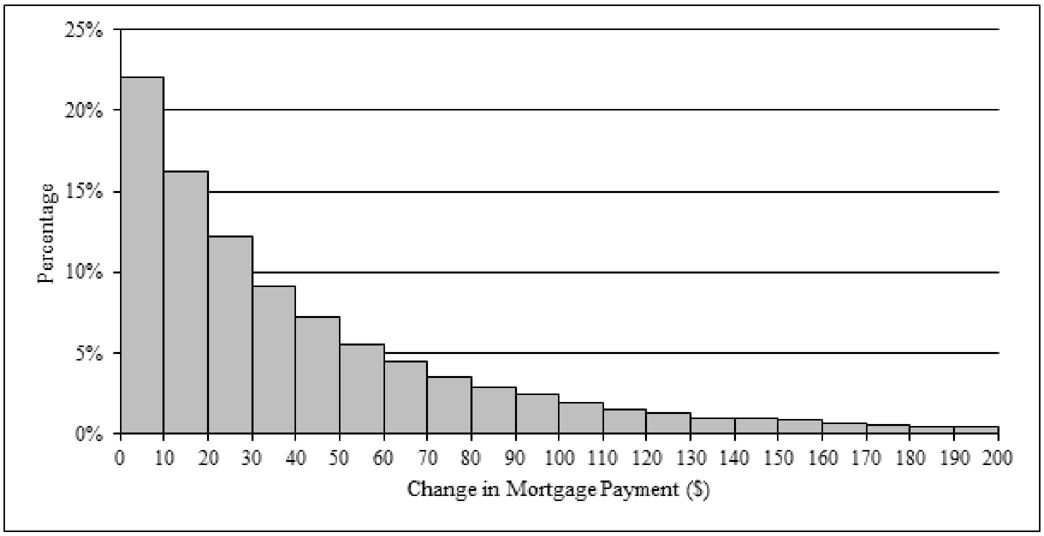


Figure 2.2: Distribution of Annual Income

This figure presents the distribution of the changes in mortgage payments after ARM rate adjustments. Panel A shows the distribution of the mortgage payment increases and Panel B shows the distribution of mortgage payment decreases. Changes over $200 are omitted from the figures.

Panel A



Panel B

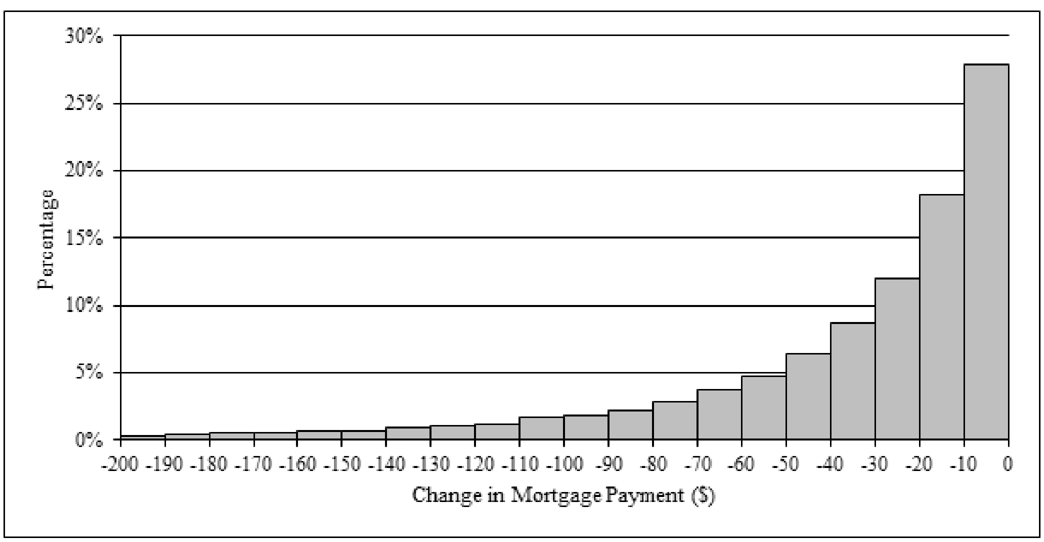


Table 2.1: Summary Statistics

This table presents the summary statistics for the households in the sample. Panel A shows the summary statistics for all households in the sample. Panel B shows the household summary statistics for the periods in which households pay more in mortgages due to adjustments in ARM payments. Panel C shows the household summary statistics for the periods in which households pay less in mortgages due to adjustments in ARM payments. All variables, except for the number of households, are measured as the monthly average amount in the 3 to 6 month period before the mortgage adjustment for each household. All variables, except for the number of households and the number of adjustment in mortgage payments, are rounded to the nearest dollar.

Panel A: All households

	Mean	StDev	Min	10%	50%	90%	Max
Mortgage	1,840	1,108	6	822	1,572	3,118	16,977
Adjustment	15	168	-17,662	-31	8	64	10,118
Income	7,030	10,375	0	2,242	5,681	12,070	1,494,000
Interest	6	53	0	0	0	10	4,871
Consumption	1,666	1,443	0	344	1,376	3,230	70,776
Retail	892	792	0	159	714	1,801	21,718
Groceries	300	541	0	20	196	686	70,139
Restaurant	234	289	0	16	169	512	17,895
Households (#)	60,612						

Continued

Table 2.1 Continued

Panel B: Increased Mortgage Payments

	Mean	StDev	Min	10%	50%	90%	Max
Mortgage	1,922	1,135	9	863	1,650	3,246	17,034
Adjustment	60	139	0	4	30	127	10,118
Income	7,125	10,541	0	2,209	5,748	12,424	1,494,000
Interest	6	68	0	0	0	10	12,502
Consumption	1,685	1,565	0	290	1,350	3,395	78,977
Retail	901	903	0	123	689	1,895	66,492
Groceries	309	552	0	9	183	748	78,349
Restaurant	233	303	0	8	160	529	13,557
Adjustment (#)	90,041						

Panel C: Decreased Mortgage Payments

	Mean	StDev	Min	10%	50%	90%	Max
Mortgage	1,759	1,082	2	776	1,497	2,989	17,225
Adjustment	-52	162	-17,662	-112	-23	-3	0
Income	7,000	9,478	0	2,107	5,564	12,024	656,370
Interest	6	89	0	0	0	10	16,227
Consumption	1,646	1,548	0	284	1,301	3,323	62,574
Retail	882	879	0	122	667	1,858	26,507
Groceries	289	546	0	7	167	705	61,929
Restaurant	233	335	0	8	157	529	28,249
Adjustment (#)	64,140						

Table 2.2: Consumption after ARM Rate Adjustments

This table presents the changes household consumption in the months after the mortgage payments are changed due to ARM adjustments. The regressions in this table are OLS regressions that use a difference-in-differences design where the consumption for a particular month for households that face adjustments in their mortgage payments are compared to the consumption for households that do not refinance their mortgages in the same month. I(t=Mt) are dummy variables for the month t, where t=1 is the month in which households face adjustments in mortgage payments. I(t=Mt) for column (1) and (2) are not dummy variables, instead these are variables that equal 1 if mortgage payments decrease and equal -1 if mortgage payments increase. The omitted variable is the all months prior to the month in which the household faces adjustments in mortgage payments. Consumption is winsorized at the 99th percentile. Standard errors are clustered at the adjustment period, which are the months surrounding the adjustment that are closest to it, and clustered at the year-month, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Change in Consumption

Dependent Variable	Consumption ($) (Overall)		Consumption ($) (Increased Payment)		Consumption ($) (Decreased Payment)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$	-1.963		-13.78***		-5.039	
	(1.796)		(3.615)		(4.078)	
$I(t = M_1)$		0.490		-3.873		1.170
		(3.056)		(5.169)		(4.799)
$I(t = M_2)$		-1.444		-14.12**		-7.936
		(3.533)		(5.944)		(5.530)
$I(t = M_3)$		-4.882*		-18.39***		-5.523
		(2.522)		(4.970)		(5.551)
$I(t = M_4)$		-2.431		-21.78***		-16.01**
		(3.407)		(5.983)		(6.598)
$I(t = M_5)$		-3.793		-19.49***		-9.660
		(3.404)		(5.962)		(6.644)
$I(t = M_6)$		-5.629*		-18.94***		-4.218
		(3.038)		(4.691)		(7.245)
$I(t = M_7)$		-2.753		-11.04		3.068
		(5.219)		(7.281)		(10.59)
$I(t > M_7)$		4.181		7.312		13.18
		(4.461)		(7.112)		(8.671)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1,850,637	1,850,637	1,092,818	1,092,818	757,454	757,454
R^2	0.616	0.616	0.637	0.637	0.646	0.646

Continued

Table 2.2 Continued

Panel B: Change in Consumption Relative to Adjustment in Mortgage Payment

Dependent Variable	Consumption ($) (Overall)		Consumption ($) (Increased Payment)		Consumption ($) (Decreased Payment)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$ x Adj	-0.0300		-0.169***		0.0195	
	(0.0186)		(0.0409)		(0.0221)	
$I(t = M_1)$ x Adj		-0.0388		-0.113*		-0.0125
		(0.0262)		(0.0630)		(0.0308)
$I(t = M_2)$ x Adj		-0.0268		-0.108*		0.00193
		(0.0348)		(0.0622)		(0.0381)
$I(t = M_3)$ x Adj		-0.0300		-0.227***		0.0374
		(0.0230)		(0.0499)		(0.0262)
$I(t = M_4)$ x Adj		-0.0155		-0.213***		0.0477
		(0.0393)		(0.0704)		(0.0521)
$I(t = M_5)$ x Adj		-0.0227		-0.213***		0.0413
		(0.0328)		(0.0652)		(0.0438)
$I(t = M_6)$ x Adj		-0.0374		-0.353***		0.0922**
		(0.0462)		(0.0549)		(0.0449)
$I(t = M_7)$ x Adj		-0.119***		-0.167*		-0.0876**
		(0.0411)		(0.0857)		(0.0422)
$I(t > M_7)$ x Adj		-0.0161		0.0662		-0.0128
		(0.0385)		(0.0930)		(0.0303)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1,850,637	1,850,637	1,092,818	1,092,818	757,454	757,454
R^2	0.616	0.616	0.637	0.637	0.646	0.646

Table 2.3: Consumption after ARM Rate Adjustments over $100

This table presents the changes household consumption in the months after the mortgage payments are changed due to ARM adjustments, for adjustments over $100. The regressions in this table are OLS regressions that use a difference-in-differences design where the consumption for a particular month for households that face adjustments in their mortgage payments are compared to the consumption for households that do not refinance their mortgages in the same month. I(t=Mt) are dummy variables for the month t, where t=1 is the month in which households face adjustments in mortgage payments. I(t=Mt) for column (1) and (2) are not dummy variables, instead these are variables that equal 1 if mortgage payments decrease and equal -1 if mortgage payments increase. The omitted variable is the all months prior to the month in which the household faces adjustments in mortgage payments. Standard errors are clustered at the period before and after adjustment in mortgage payment where in only change in mortgage payment is at adjustment for each household and year-month level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Change in Consumption

Dependent Variable	Consumption ($) (Overall)		Consumption ($) (Increased Payment)		Consumption ($) (Decreased Payment)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$	-7.945		-28.68***		13.04	
	(6.069)		(10.56)		(14.93)	
$I(t = M_1)$		-2.952		-8.115		23.96
		(9.718)		(13.72)		(15.81)
$I(t = M_2)$		-0.625		-16.38		6.598
		(8.691)		(13.25)		(17.46)
$I(t = M_3)$		-9.822		-34.40**		2.811
		(8.310)		(13.11)		(19.72)
$I(t = M_4)$		-11.74		-35.43**		15.94
		(9.461)		(15.79)		(19.75)
$I(t = M_5)$		-9.285		-34.99**		10.09
		(11.17)		(16.19)		(22.60)
$I(t = M_6)$		-31.75***		-66.93***		15.17
		(9.369)		(15.56)		(22.57)
$I(t = M_7)$		-17.35		-51.58**		22.86
		(15.46)		(23.69)		(39.72)
$I(t > M_7)$		13.37		-11.79		24.15
		(15.33)		(24.90)		(35.38)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	237,323	237,323	151,027	151,027	86,258	86,258
R^2	0.671	0.671	0.686	0.686	0.684	0.684

Continued

Table 2.3 Continued

Panel B: Change in Consumption Relative to Adjustment in Mortgage Payment

Dependent Variable	Consumption ($) (Overall)		Consumption ($) (Increased Payment)		Consumption ($) (Decreased Payment)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$ x Adj	-0.0324		-0.129**		-0.0127	
	(0.0212)		(0.0502)		(0.0241)	
$I(t = M_1)$ x Adj		-0.0471*		-0.0836		-0.0394
		(0.0265)		(0.0679)		(0.0322)
$I(t = M_2)$ x Adj		-0.0288		-0.0536		-0.0259
		(0.0358)		(0.0611)		(0.0369)
$I(t = M_3)$ x Adj		-0.0267		-0.177***		0.0123
		(0.0232)		(0.0596)		(0.0258)
$I(t = M_4)$ x Adj		-0.0135		-0.133*		0.00924
		(0.0434)		(0.0778)		(0.0551)
$I(t = M_5)$ x Adj		-0.0133		-0.127		0.00848
		(0.0357)		(0.0789)		(0.0468)
$I(t = M_6)$ x Adj		-0.0327		-0.333***		0.0600
		(0.0501)		(0.0714)		(0.0514)
$I(t = M_7)$ x Adj		-0.134***		-0.133		-0.137***
		(0.0351)		(0.109)		(0.0311)
$I(t > M_7)$ x Adj		-0.0374		0.0279		-0.0519
		(0.0382)		(0.124)		(0.0395)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	237,323	237,323	151,027	151,027	86,258	86,258
R^2	0.671	0.671	0.686	0.686	0.684	0.684

Table 2.4: Consumption after ARM Rate Adjustments under $100

This table presents the changes household consumption in the months after the mortgage payments are changed due to ARM adjustments, for adjustments at or under $100. The regressions in this table are OLS regressions that use a difference-in-differences design where the consumption for a particular month for households that face adjustments in their mortgage payments are compared to the consumption for households that do not refinance their mortgages in the same month. I(t=Mt) are dummy variables for the month t, where t=1 is the month in which households face adjustments in mortgage payments. I(t=Mt) for column (1) and (2) are not dummy variables, instead these are variables that equal 1 if mortgage payments decrease and equal -1 if mortgage payments increase. The omitted variable is the all months prior to the month in which the household faces adjustments in mortgage payments. Standard errors are clustered at the period before and after adjustment in mortgage payment where in only change in mortgage payment is at adjustment for each household and year-month level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Change in Consumption

Dependent Variable	Consumption ($) (Overall)		Consumption ($) (Increased Payment)		Consumption ($) (Decreased Payment)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$	-0.545		-11.22***		-5.863	
	(1.896)		(3.496)		(4.242)	
$I(t = M_1)$		2.004		-3.007		-0.857
		(3.107)		(4.939)		(4.880)
$I(t = M_2)$		-0.777		-13.63**		-8.667
		(3.667)		(5.688)		(5.869)
$I(t = M_3)$		-3.428		-15.86***		-5.304
		(2.698)		(5.126)		(5.944)
$I(t = M_4)$		-0.292		-19.89***		-18.53***
		(3.797)		(5.958)		(6.835)
$I(t = M_5)$		-1.998		-16.99***		-10.39
		(3.337)		(6.118)		(7.218)
$I(t = M_6)$		-1.021		-10.97**		-4.605
		(3.131)		(4.905)		(7.366)
$I(t = M_7)$		0.230		-3.144		4.061
		(5.310)		(7.213)		(10.97)
$I(t > M_7)$		1.520		9.770		15.87*
		(4.506)		(7.439)		(9.379)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1,613,155	1,613,155	941,717	941,717	671,158	671,158
R^2	0.617	0.617	0.636	0.636	0.645	0.645

Continued

Table 2.4 Continued

Panel B: Change in Consumption Relative to Adjustment in Mortgage Payment

Dependent Variable	Consumption ($) (Overall)		Consumption ($) (Increased Payment)		Consumption ($) (Decreased Payment)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$ x Adj	-0.00862		-0.183**		0.150	
	(0.0531)		(0.0783)		(0.0981)	
$I(t = M_1)$ x Adj		0.0967		-0.0348		0.170
		(0.0733)		(0.109)		(0.121)
$I(t = M_2)$ x Adj		0.00267		-0.169		0.155
		(0.0847)		(0.130)		(0.140)
$I(t = M_3)$ x Adj		-0.0963		-0.298**		0.124
		(0.0922)		(0.121)		(0.151)
$I(t = M_4)$ x Adj		-0.0592		-0.423***		0.422***
		(0.0884)		(0.124)		(0.157)
$I(t = M_5)$ x Adj		-0.0829		-0.368**		0.281
		(0.100)		(0.146)		(0.174)
$I(t = M_6)$ x Adj		-0.0340		-0.194		0.130
		(0.0858)		(0.131)		(0.168)
$I(t = M_7)$ x Adj		0.107		-0.0349		0.185
		(0.141)		(0.188)		(0.286)
$I(t > M_7)$ x Adj		0.0778		0.292*		-0.433*
		(0.125)		(0.171)		(0.244)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1,613,155	1,613,155	941,717	941,717	671,158	671,158
R^2	0.617	0.617	0.636	0.636	0.645	0.645

Table 2.5: Change in Consumption after ARM Rate Adjustments, by Income

This table presents the changes household consumption in the months after the mortgage payments are changed due to ARM adjustments, for three income quintiles. Low income represents the subsample for households that belong in the first income quintile. Mid income represents the subsample for households that belong in the third income quintile. High income represents the subsample for households that belong in the high income quintile. The regressions in this table are OLS regressions that use a difference-in-differences design where the consumption for a particular month for households that face adjustments in their mortgage payments are compared to the consumption for households that do not refinance their mortgages in the same month. I(t=Mt) are dummy variables for the month t, where t=1 is the month in which households face adjustments in mortgage payments. I(t=Mt) for Panel A are not dummy variables, instead these are variables that equal 1 if mortgage payments decrease and equal -1 if mortgage payments increase. The omitted variable is the all months prior to the month in which the household faces adjustments in mortgage payments. Standard errors are clustered at the period before and after adjustment in mortgage payment where in only change in mortgage payment is at adjustment and for each household and year-month level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: All Mortgage Payment Changes

Dependent Variable	Consumption ($) (Low Income)		Consumption ($) (Mid Income)		Consumption ($) (High Income)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$	2.812		-6.321		2.645	
	(2.908)		(3.905)		(5.290)	
$I(t \geq M_1)$ x Adj		-0.0153		-0.0371		-0.0241
		(0.0472)		(0.0457)		(0.0264)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	330,977	330,977	380,827	380,827	376,663	376,663
R^2	0.603	0.603	0.529	0.529	0.627	0.627

Continued

Table 2.5 Continued

Panel B: Increased Mortgage Payments

Dependent Variable	Consumption ($) (Low Income)		Consumption ($) (Mid Income)		Consumption ($) (High Income)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$	-11.14**		-18.65***		-15.53	
	(4.789)		(6.541)		(9.592)	
$I(t \geq M_1)$ x Adj		-0.153*		-0.228***		-0.172**
		(0.0804)		(0.0754)		(0.0800)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	188,550	188,550	223,879	223,879	229,998	229,998
R^2	0.635	0.635	0.554	0.554	0.647	0.647

Panel C: Decreased Mortgage Payment

Dependent Variable	Consumption ($) (Low Income)		Consumption ($) (Mid Income)		Consumption ($) (High Income)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$	-7.231		3.866		-11.22	
	(6.274)		(7.546)		(11.10)	
$I(t \geq M_1)$ x Adj		0.0406		0.0238		0.00879
		(0.0742)		(0.0725)		(0.0298)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	142,281	142,281	156,914	156,914	146,608	146,608
R^2	0.612	0.612	0.566	0.566	0.661	0.661

Table 2.6: Change in Consumption after ARM Rate Adjustments, by Interest Earned

This table presents the changes household consumption in the months after the mortgage payments are changed due to ARM adjustments, for three interest earned quintiles. Low interest represents the subsample for households that belong in the first interest earned quintile. Mid interest represents the subsample for households that belong in the third interest earned quintile. High interest represents the subsample for households that belong in the high interest earned quintile. The regressions in this table are OLS regressions that use a difference-in-differences design where the consumption for a particular month for households that face adjustments in their mortgage payments are compared to the consumption for households that do not refinance their mortgages in the same month. I(t=Mt) are dummy variables for the month t, where t=1 is the month in which households face adjustments in mortgage payments. I(t=Mt) for Panel A are not dummy variables, instead these are variables that equal 1 if mortgage payments decrease and equal -1 if mortgage payments increase. The omitted variable is the all months prior to the month in which the household faces adjustments in mortgage payments. Standard errors are clustered at the period before and after adjustment in mortgage payment where in only change in mortgage payment is at adjustment and for each household and year-month level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: All Mortgage Payment Changes

Dependent Variable	Consumption ($) (Low Interest)		Consumption ($) (Mid Interest)		Consumption ($) (High Interest)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$	-6.154*		-3.659		-4.654	
	(3.485)		(4.211)		(4.520)	
$I(t \geq M_1)$ x Adj		-0.119***		-0.0281		-0.0444*
		(0.0434)		(0.0426)		(0.0228)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	324,382	324,382	386,712	386,712	373,426	373,426
R^2	0.618	0.618	0.607	0.607	0.621	0.621

Continued

Table 2.6 Continued

Panel B: Increased Mortgage Payments

Dependent Variable	Consumption ($) (Low Interest)		Consumption ($) (Mid Interest)		Consumption ($) (High Interest)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$	-3.738		-18.11***		-21.52***	
	(5.531)		(6.167)		(7.371)	
$I(t \geq M_1)$ x Adj		-0.181**		-0.223***		-0.218**
		(0.0773)		(0.0818)		(0.0851)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	189,045	189,045	228,943	228,943	222,879	222,879
R^2	0.631	0.631	0.635	0.635	0.641	0.641

Panel C: Decreased Mortgage Payment

Dependent Variable	Consumption ($) (Low Interest)		Consumption ($) (Mid Interest)		Consumption ($) (High Interest)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$	3.097		-4.145		-6.594	
	(6.661)		(8.385)		(9.247)	
$I(t \geq M_1)$ x Adj		-0.0271		0.0896		-0.0152
		(0.0565)		(0.0587)		(0.0260)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	135,261	135,261	157,691	157,691	150,457	150,457
R^2	0.654	0.654	0.627	0.627	0.650	0.650

Table 2.7: Change in Consumption after ARM Rate Adjustments, by Brokerage Transaction

This table presents the changes household consumption in the months after the mortgage payments are changed due to ARM adjustments, for households that had or did not have brokerage transactions. sThe regressions in this table are OLS regressions that use a difference-in-differences design where the consumption for a particular month for households that face adjustments in their mortgage payments are compared to the consumption for households that do not refinance their mortgages in the same month. I(t=Mt) are dummy variables for the month t, where t=1 is the month in which households face adjustments in mortgage payments. I(t=Mt) for Panel A are not dummy variables, instead these are variables that equal 1 if mortgage payments decrease and equal -1 if mortgage payments increase. The omitted variable is the all months prior to the month in which the household faces adjustments in mortgage payments. Standard errors are clustered at the period before and after adjustment in mortgage payment where in only change in mortgage payment is at adjustment and for each household and year-month level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: All Mortgage Payment Changes

Dependent Variable	Consumption ($) (No Brokerage Txn)		Consumption ($) (Brokerage Txn)	
	(1)	(2)	(3)	(4)
$I(t \geq M_1)$	-0.780		-7.522*	
	(1.947)		(4.468)	
$I(t \geq M_1)$ x Adj		-0.0508**		0.00357
		(0.0226)		(0.0317)
Adjustment FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Obs	1,510,032	1,510,032	340,605	340,605
R^2	0.613	0.613	0.626	0.626

Continued

Table 2.7 Continued

Panel B: Increased Mortgage Payments

Dependent Variable	Consumption ($) (No Brokerage Txn)		Consumption ($) (Brokerage Txn)	
	(1)	(2)	(3)	(4)
$I(t \geq M_1)$	-14.23***		-12.15*	
	(3.811)		(7.221)	
$I(t \geq M_1)$ x Adj		-0.230***		0.0320
		(0.0439)		(0.0883)
Adjustment FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Obs	891,200	891,200	201,618	201,618
R^2	0.634	0.634	0.647	0.647

Panel C: Decreased Mortgage Payment

Dependent Variable	Consumption ($) (No Brokerage Txn)		Consumption ($) (Brokerage Txn)	
	(1)	(2)	(3)	(4)
$I(t \geq M_1)$	-3.864		-9.834	
	(3.895)		(9.665)	
$I(t \geq M_1)$ x Adj		0.0185		0.0159
		(0.0268)		(0.0371)
Adjustment FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Obs	618,547	618,547	138,907	138,907
R^2	0.642	0.642	0.656	0.656

Table 2.8: Change in Consumption after ARM Rate Adjustments, by Overdraft Fees

This table presents the changes household consumption in the months after the mortgage payments are changed due to ARM adjustments, for households that had or did not incur overdraft fees. The regressions in this table are OLS regressions that use a difference-in-differences design where the consumption for a particular month for households that face adjustments in their mortgage payments are compared to the consumption for households that do not refinance their mortgages in the same month. I(t=Mt) are dummy variables for the month t, where t=1 is the month in which households face adjustments in mortgage payments. I(t=Mt) for Panel A are not dummy variables, instead these are variables that equal 1 if mortgage payments decrease and equal -1 if mortgage payments increase. The omitted variable is the all months prior to the month in which the household faces adjustments in mortgage payments. Standard errors are clustered at the period before and after adjustment in mortgage payment where in only change in mortgage payment is at adjustment and for each household and year-month level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: All Mortgage Payment Changes

Dependent Variable	Consumption ($) (No Overdraft Txn)		Consumption ($) (Overdraft Txn)	
	(1)	(2)	(3)	(4)
$I(t \geq M_1)$	-2.610		-0.103	
	(2.196)		(3.699)	
$I(t \geq M_1)$ x Adj		-0.0359		-0.0123
		(0.0221)		(0.0342)
Adjustment FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Obs	1,358,708	1,358,708	491,929	491,929
R^2	0.617	0.617	0.608	0.608

Continued

Table 2.8 Continued

Panel B: Increased Mortgage Payments

Dependent Variable	Consumption ($) (No Overdraft Txn)		Consumption ($) (Overdraft Txn)	
	(1)	(2)	(3)	(4)
$I(t \geq M_1)$	-16.38***		-5.311	
	(4.391)		(5.997)	
$I(t \geq M_1)$ x Adj		-0.169***		-0.159**
		(0.0466)		(0.0762)
Adjustment FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Obs	805,151	805,151	287,667	287,667
R^2	0.638	0.638	0.630	0.630

Panel C: Decreased Mortgage Payment

Dependent Variable	Consumption ($) (No Overdraft Txn)		Consumption ($) (Overdraft Txn)	
	(1)	(2)	(3)	(4)
$I(t \geq M_1)$	-4.408		-6.112	
	(4.532)		(7.005)	
$I(t \geq M_1)$ x Adj		0.0114		0.0423
		(0.0241)		(0.0427)
Adjustment FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Obs	553,282	553,282	204,172	204,172
R^2	0.645	0.645	0.641	0.641

Table 2.9: Change in Consumption after ARM Rate Adjustments, by Mortgage to Income

This table presents the changes household consumption in the months after the mortgage payments are changed due to ARM adjustments, for three mortgage to income (MTI) quintiles. Low MTI represents the subsample for households that belong in the first MTI quintile. Mid MTI represents the subsample for households that belong in the third MTI quintile. High MTI represents the subsample for households that belong in the high MTI quintile. The regressions in this table are OLS regressions that use a difference-in-differences design where the consumption for a particular month for households that face adjustments in their mortgage payments are compared to the consumption for households that do not refinance their mortgages in the same month. I(t=Mt) are dummy variables for the month t, where t=1 is the month in which households face adjustments in mortgage payments. I(t=Mt) for Panel A are not dummy variables, instead these are variables that equal 1 if mortgage payments decrease and equal -1 if mortgage payments increase. The omitted variable is the all months prior to the month in which the household faces adjustments in mortgage payments. Standard errors are clustered at the period before and after adjustment in mortgage payment where in only change in mortgage payment is at adjustment and for each household and year-month level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: All Mortgage Payment Changes

Dependent Variable	Consumption ($) (Low MTI)		Consumption ($) (Mid MTI)		Consumption ($) (High MTI)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$	-5.729		1.424		-7.904**	
	(4.302)		(4.252)		(3.792)	
$I(t \geq M_1)$ x Adj		-0.0456		0.0046		-0.172***
		(0.0885)		(0.0584)		(0.0482)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	379,140	379,140	379,779	379,779	333,255	333,255
R^2	0.598	0.598	0.600	0.600	0.631	0.631

Continued

Table 2.9 Continued

Panel B: Increased Mortgage Payments

Dependent Variable	Consumption ($) (Low MTI)		Consumption ($) (Mid MTI)		Consumption ($) (High MTI)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$	-21.54***		-10.36		-12.36**	
	(7.287)		(6.677)		(5.720)	
$I(t \geq M_1)$ x Adj		-0.143		0.0161		-0.259***
		(0.120)		(0.0852)		(0.0598)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	214,335	214,335	224,354	224,354	203,262	203,262
R^2	0.618	0.618	0.622	0.622	0.653	0.653

Panel C: Decreased Mortgage Payment

Dependent Variable	Consumption ($) (Low MTI)		Consumption ($) (Mid MTI)		Consumption ($) (High MTI)	
	(1)	(2)	(3)	(4)	(5)	(6)
$I(t \geq M_1)$	7.820		-9.213		3.168	
	(8.798)		(7.497)		(7.769)	
$I(t \geq M_1)$ x Adj		-0.0619		0.0806		0.0316
		(0.133)		(0.114)		(0.0935)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	164,762	164,762	155,380	155,380	129,840	129,840
R^2	0.632	0.632	0.632	0.632	0.649	0.649

Table 2.10: Change in Savings after ARM Rate Adjustments

This table presents the changes household savings in the months after the mortgage payments are changed due to ARM adjustments. The regressions in this table are OLS regressions that use a difference-in-differences design where the consumption for a particular month for households that face adjustments in their mortgage payments are compared to the consumption for households that do not refinance their mortgages in the same month. I(t=Mt) are dummy variables for the month t, where t=1 is the month in which households face adjustments in mortgage payments. I(t=Mt) for Panel A are not dummy variables, instead these are variables that equal 1 if mortgage payments decrease and equal -1 if mortgage payments increase. The omitted variable is the all months prior to the month in which the household faces adjustments in mortgage payments. Standard errors are clustered at the period before and after adjustment in mortgage payment where in only change in mortgage payment is at adjustment and for each household and year-month level, and are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: All Mortgage Payment Changes

Dependent Variable	Interest Earned ($) (Overall)		Interest Earned ($) (Increased Payment)		Interest Earned ($) (Decreased Payment)	
	(1)	(2)	(3)	(4)	(5)	(6)
I(tM1)	-0.0106		-0.00296		-0.00742	
	(0.0142)		(0.0276)		(0.0236)	
I(tM1) x Adj		¿ 0.0001		-0.0009***		0.0003
		(0.0002)		(0.0003)		(0.0003)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1,850,637	1,850,637	1,092,818	1,092,818	757,454	757,454
R^2	0.797	0.797	0.821	0.821	0.844	0.844

Continued

Table 2.10 Continued

Panel B: Increased Mortgage Payments

Dependent Variable	Brokerage Txn ($) (Overall)		Brokerage Txn ($) (Increased Payment)		Brokerage Txn ($) (Decreased Payment)	
	(1)	(2)	(3)	(4)	(5)	(6)
I(tM1)	-10.31**		-18.17**		-1.416	
	(4.366)		(7.123)		(10.25)	
I(tM1) x Adj		-0.0797		-0.0562		-0.0870
		(0.114)		(0.139)		(0.194)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1,850,637	1,850,637	1,092,818	1,092,818	757,454	757,454
R^2	0.261	0.261	0.267	0.267	0.294	0.294

Panel C: Decreased Mortgage Payment

Dependent Variable	Overdraft Fees ($) (Overall)		Overdraft Fees ($) (Increased Payment)		Overdraft Fees ($) (Decreased Payment)	
	(1)	(2)	(3)	(4)	(5)	(6)
I(tM1)	0.008		0.0381		0.007	
	(0.017)		(0.0270)		(0.0313)	
I(tM1) x Adj		-0.0003*		0.0003		-0.0004**
		(0.0001)		(0.0003)		(0.0002)
Adjustment FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1,850,637	1,850,637	1,092,818	1,092,818	757,454	757,454
R^2	0.224	0.224	0.243	0.243	0.263	0.263

Chapter 3: Can Taxes Shape an Industry? Evidence from the Implementation of the "Amazon Tax"

3.1 Introduction

An important question in economics is the degree to which taxes distort decisions of firms and of individuals. Among other effects at the firm level, taxes may cause products and services to be more or less attractive to consumers, and thus potentially affect the growth and organization of businesses within an industry.

In this study, we focus on the role of sales tax imposed on consumers and collected by firms. Previous empirical work shows that consumers are indeed sensitive to sales tax. [5] provide evidence that consumers make cross-border trips to save on sales tax. In the online retail arena, [39] find that eBay customers avoid transactions in which they need to pay sales tax. A recent trend in state legislatures is to enforce the collection of sales tax on Internet retailers, particularly on Amazon, the largest online retailer. To date, there has been little evaluation of far-reaching and permanent sales tax policies on retail, competition, and consumers.

Between 2012 and 2015, 19 states began implementing laws requiring Amazon.com, the largest online retailer, to collect sales tax from its customers. These laws, commonly referred to as the Amazon Tax, provide an ideal setting for examining the effects of sales tax collection on consumer behavior. Enforcing the collection of sales tax has generated much discussion among legislators and the public as sales tax is a major source of income for many states. At the same time, taxes can shape the growth and organization of businesses because they affect the attractiveness of firms products and services to customers. Previous empirical work shows that consumers are sensitive to sales tax. [5] provide evidence that consumers make cross-border trips to save on sales tax. In the online retail arena, [39] find that eBay customers avoid transactions in which they need to pay sales tax.

Using a unique dataset containing transaction-level financial data for 2.7 million US households, we closely track consumers purchase behavior around the introduction of the tax. Our results shed light on the effects of the Amazon Tax on the demand for Amazon products. Because little prior empirical evidence has been gathered about the effects of wide implementation of such a tax on retail and as more and more states begin to implement Amazon Tax laws, this study contributes to our understanding its consequences.

Over the past decade, online retail transactions have increased dramatically in volume. According to the US Census, online sales constituted 2.5% of retail sales in 2006 and 7.7% of retail sales in 2016 (corresponding to $354 billion over the four quarters from 2015Q2 to 2016Q1).[2] Many factors have contributed to this growth in online sales, one of which is that out-of-state online retailers do not charge sales tax, which has generally given them a price advantage over retailers with a presence in the state. This sales tax collection loophole has not gone unnoticed by state governments or by competing retailers. State governments are concerned that these online sales

[2]www.census.gov/retail/index.html

depress local employment and erode tax revenues. From 2012 to 2015, many states responded by requiring that Amazon begin to collect sales tax.

State governments have begun paying increased attention to the issue of sales tax collection in light of the Great Recession and the recent growth in online retail volume. General sales taxes represent an important part of state revenue: For example, in 2011, general sales tax constituted 10.4% of revenues. Figure 3.1 shows that the importance of this tax varies considerably by state, ranging from 0% of revenues in states without sales tax (such as Oregon and Alaska) to as high as 21.0% of state revenues for Washington.[3] Recently, the issue has received federal attention. The Marketplace Fairness Acts of 2013 and 2015 were attempts by lawmakers to enable all states to force retailers to collect sales tax on purchases made by out-of-state customers, but neither act has been adopted into law.[4] Proponents of the online sales tax collection bill often tout the elimination of the Internet retailer sales tax advantage as leveling the playing field and helping to restore business and jobs to local economies.

Online retailers, including Amazon, that are not required to collect sales tax enjoy a price advantage. As a result, we hypothesize that the introduction of the Amazon Tax will lead to a decline in Amazons sales and substitution to alternative retailers. With effective sales tax rates as high as 10% in some jurisdictions (after accounting for state, county, and city taxes), this price advantage can be sizable. Gene DeFelice, vice president of Barnes and Noble, the largest book retailer in the United States, summarized the issue succinctly: "We are at a serious competitive disadvantage against

[3] 2011 US Census Annual Survey of State & Local Government Finance: www.census.gov/govs/local/

[4] The text and status of the bill are found here: www.govtrack.us/congress/bills/113/s743, www.govtrack.us/congress/bills/114/s698

out-of-state, online retailers who pay no taxes."[5] An additional factor that is likely to facilitate customer migration from Amazon to alternative outlets is the low search cost of online shopping.

Our analysis of the effects of the Amazon Tax on purchasing behavior uses data from an online financial account aggregator. This financial service enables subscribers to concentrate all of their accounts in one place for viewing and monitoring purposes. Our base dataset includes data on 2.7 million households and contains transaction-level information similar to what is found on bank and credit card statements.

We begin our analysis by using a traditional difference-in-differences (diff-in-diff) methodology to test whether households decreased their Amazon purchases following the introduction of the law. Each state that adopted the Amazon Tax during our sample period is considered treated following the adoption, and other states are considered controls. Our results show that the introduction of the Amazon Tax resulted in a persistent decline of 9.4% in the amount spent on products (net of sales tax, which we hereafter refer to as the tax-exclusive price) through Amazon, corresponding to an average elasticity of 1.2. In an alternative specification, we find that a one percentage point increase in sales tax leads to a $54.33 reduction in tax-exclusive Amazon spending, corresponding to an elasticity of 1.4. We also test whether these effects are more sensitive to households in high-tax jurisdictions and find that indeed these consumers have higher elasticity.

We next investigate whether consumers decreased their tax-inclusive spending on Amazon after the tax was introduced. Our results show a reduction in tax-inclusive spending on Amazon in the wake of the laws implementation.

[5]articles.latimes.com/2011/jan/20/business/la-fi-internet-tax20110120

We find that low-income households reduced their tax-exclusive spending on Amazon slightly more than high-income households (9.9% versus 7.0%, respectively). Further, the percent reduction in spending on Amazon was slightly higher among heavy Amazon customers. The highest tercile of Amazon spending in 2011 reduces Amazon purchases by $6.22, corresponding to a 9.4% reduction, whereas the lowest tercile of Amazon shoppers reduces expenditures by a statistically insignificant $1.65, corresponding to an 8.0% reduction.

Consistent with the idea that consumers trade off sales tax with search costs, we find that the decline in Amazon purchases is more pronounced for larger purchases, as consumers would garner the greatest savings by avoiding tax on such purchases. We document strong evidence that the effect of the Amazon Tax increases with the size of the purchase, suggesting that households are particularly likely to engage in Internet shopping to avoid sales tax for large purchases. Consumers decrease their spending by 29.1% on transactions of at least $250, implying an elasticity of 3.9. In a more refined analysis into smaller transaction amount bins, we show that the elasticity is increasing in the transaction amount.

Next, we study substitution effects. Because many of Amazons large competitors are companies with a larger scope of products than that of Amazon (e.g., groceries at Walmart, Costco), we focus on a particular industry: electronics retailers. We find that Newegg, one of Amazons direct competitors, experienced an increase in sales thanks to the implementation of the Amazon Tax. On average, Neweggs sales increased by 13.0%. We also observe that the share of retail purchases coming from Amazon decreases for treated households and that this effect is primarily driven by heavy Amazon shoppers.

Finally, we analyze the income effects induced by the Amazon Tax. We find that after implementation of the Amazon tax, heavy Amazon shoppers reduce spending in each of the categories we investigate: restaurants, groceries, and entertainment. The magnitude of this reduction is increasing in the households spending on Amazon during the pre-treatment year of 2011.

Our work relates to two recent strands of the literature. First, several empirical studies have documented that consumers are price and tax sensitive, and thus attempt to avoid sales taxes. [79] and [23] find that price levels in locations with high sales tax are lower than those in locations with lower sales tax. [5] find that consumers who live near state borders often shop in the neighboring state when there are positive sales tax differences. [7] show that consumers increase their purchases during sales tax holidays. [33] use an experimental setting to show that sales tax that is salient to consumers reduces the demand for the product.

Second, several studies explore the sensitivity to sales tax in the specific context of online retail. The closest study to ours in this strand of the literature is [39] (EKLS). These researchers document a strong preference among eBay customers for out-of-state sellers, for whom sales taxes do not apply. They observe eBay shoppers reactions when they discover that the seller is from the same state, which requires them to collect sales tax. They document that eBay shoppers are indeed sensitive to sales tax and thus less likely to buy from sellers who reside in the same state. In this setting, they estimate an elasticity 1.7.

Our research contributes to the literature beyond EKLS on multiple accounts. First, our paper directly studies the effects of a permanent change in sales tax for the largest internet retailer in the world. We rely on state-level implementation of laws;

consequently, our results directly measure the effect of these laws on Amazon and on Amazons competitors. While the results of EKLS indicate that online shoppers are sensitive to taxes, their evidence does not translate directly to the effect of the tax implementation and thus is less conducive to measuring the policy impact and less relevant to the debate. Second, we are able to study how the imposition of the Amazon Tax affects the sales of competitors as well as the effect on other, unrelated, consumption items, such as restaurants, groceries, and entertainment (income effect). Third, our empirical setting is different from that of EKLS, validating both sets of results. The EKLS study is based on a limited sample of transactions (about 270,000). Conversely, our analysis is based on millions of transactions made by a sample of over 460,000 households in our broadest regressions. In addition, Amazon is larger than eBay: As of 2014, Amazons revenue was five times larger. Finally, the time periods of the studies are distinct, although chronologically close. EKLS use a sample from 2010; our data are from 20112015. Given that speed that online commerce is evolving, it important to monitor the persistence of effects over time.

Several additional studies examine the intersection of online sales and sales tax. [51], [52] uses survey data to estimate that the number of online shoppers would drop by 24% if the tax-advantaged status of Internet retailers were removed. [10], [15], and [81] address this question as well, though they find smaller magnitudes for the effect. [53] ascertain that the penetration of the Internet is correlated with lower sensitivity of cigarette sales to local taxes, suggesting that smokers use the Internet to purchase tax-free cigarettes. [40] explore the price elasticity of memory modules sold by a particular retailer and determine that consumers are price sensitive both to tax-exclusive prices and to state taxes. [12] show that when retail chains open their first store in a new state, they experience a decline in their Internet sales shipped to that state because of the sales tax, but the researchers find no similar effect on catalog sales. Finally, [57] find that Internet retailers exhibit negative stock market returns following legislative proposals to collect sales tax from customers, such as the Marketplace Fairness Act of 2013.

3.2 Background and Empirical Setting

Sales tax is not collected on purchases from online retailers due to the Commerce Clause in the US Constitution. Current interpretation of the law, which has been consistently upheld by the US Supreme Court, is that online retailers must only collect sales tax on out-of-state purchases if the retailer has a nexus (or a substantial physical presence) in the state. Due to the nature of their business structure, online retailers have a physical presence in very few states. Ten years ago, Amazon was only required to collect sales taxes in states in which it had a nexus (for example, where it was headquartered or had fulfillment centers).

In recent years, states have attempted to collect sales taxes by broadening the definition of a nexus. Legislation by these states has defined the presence of affiliate programs or subsidiaries as constituting a nexus.[6] Even when this legislation has been ruled constitutional by state courts, the effectiveness of this method of tax collection has been mixed. Overstock.com, for example, has responded to these laws by simply dropping its affiliates in these states. Amazon has acted similarly in some states but

[6] Online retailers such as Amazon and Overstock will often advertise on websites such as blogs. If a website reader clicks on the advertisement and subsequently purchases the Amazon product, the website owner will receive a commission on the sale. These website owners who allow Amazon to advertise on their websites are referred to as affiliates.

in other states has chosen to accede to the Amazon Tax laws due to various political and operational issues.

As of February 2015, Amazon was collecting sales tax in 24 states, comprising more than half of the US population. Over our sample period, 19 states implemented Amazon Tax laws, resulting in the beginning of sales tax collection on well-defined dates for each of these states.[7] Our diff-in-diff study relies on this change in tax policy over time for these states, relative to a control group of other states that did not change their tax policy contemporaneously.

Our study investigates the impact of the Amazon Tax in 19 states in which Amazon started collecting sales taxes between 2012 and 2014. These states are Texas (7 / 1 / 2012), Pennsylvania (9 / 1 / 2012), California (9 / 16 / 2012), Arizona (2 / 1 / 2013), New Jersey (7 / 1 / 2013), Virginia (9 / 1 / 2013), Georgia (9 / 1 / 2013), West Virginia (10 / 1 / 2013), Connecticut (11 / 1 / 2013), Massachusetts (11 / 1 / 2013), Wisconsin (11 / 1 / 2013), Indiana (1 / 1 / 2014), Nevada (1 / 1 / 2014), Tennessee (1 / 1 / 2014), North Carolina (2 / 1 / 2014), Florida (5 / 1 / 2014), Maryland (10 / 1 / 2014), Minnesota (10 / 1 / 2014), and Illinois (2 / 1 / 2015).

A critical facet of the diff-in-diff methodology is the parallel trends assumption. One concern with our setting is that many states require that households pay sales taxes that are not collected at the time of purchase. These taxes are referred to as use taxes and are collected by states annually at the time of tax filing. However, compliance with this use tax has been abysmal. [71] finds that only 22 states have use tax provisions in their state income tax forms and that the vast majority of households

[7]Before our sample period begins, five states collected sales tax from Amazon, including Washington where Amazon is headquartered. After our sample period ends, more states already have or will shortly begin collecting sales tax on Amazon purchases.

residing in these states do not report any use tax liability. For example, only 0.2% of households in Rhode Island report any use taxes, and only 0.3% of households in California and New Jersey report use taxes. However, some states have higher participation rates, such as Vermont and Maine, with 7.9% and 9.8% of households in each state reporting use taxes, respectively. Unlike income tax reporting, systems for tracking and enforcing collection of these sales taxes are weak.[8] Note that these figures do not necessarily represent the percentage of compliance with the law. In particular, the quoted numbers do not account for underreporting of use taxes conditional on reporting a use tax liability.

3.3 Data

The data we use were provided by an online account aggregator. This service allows subscribers to view their various financial information in one place, e.g., view spending by category, monitor investments, etc. The service also provides alerts for upcoming bills and for approaching credit limits, and the like. Households join the service for free and provide their username and passwords to various financial institutions so that the service can extract relevant bank and credit card information.

The information we use consists of daily transactions for 2.7 million households from January 2011 to May 2015, and includes both banking (i.e., checking, savings, and debit card) and credit card transactions. We observe the date, amount, and description of each transaction. Thus, our dataset contains transaction-level data

[8]For example, Colorados version of the Amazon Tax legislation tried to force online retailers to report to both customers and the state tax authority summaries of use tax incurred, but it was later declared unconstitutional by the District Court. However, Amazon makes annual spending reports available to residents of South Carolina and Tennessee to aid households in tax filing, though this information is not reported to state tax authorities by Amazon.

similar to those typically found on monthly bank or credit card statements. Because each household is assigned a unique identifier, we are able to follow each household through time.

Identifying the state of residence of the household is integral to our analysis, because this allows us to determine whether the household lived in one of the 19 treatment states affected by an Amazon Tax. We identify the state of residence of households in our dataset by requiring that 75% of transactions occur within a given state. We then assign the most common city as the city of residence of the household, though our results are robust to alternative methods of identifying the city of residence of the household as described in Section A1 of the Appendix.

Because we are primarily interested in how Amazon customers respond after the implementation of the Amazon Tax, we focus our analysis on households who made some purchases on Amazon prior to implementation. We include households that spent more than $200 on Amazon during 2011, though the results are robust to using alternative spending thresholds, as demonstrated in Section A2 of the Appendix. After applying these two filters, our sample size is reduced to 275,437 households, 180,330 of which live in one of the 19 states that implemented the Amazon Tax during our sample period.

The unit of observation in our analyses is the household-month. For each household-month, we sum all Amazon expenditures. For all transactions in our database, we adjust by the households sales tax to determine the tax-exclusive amount spent on goods purchased. For Amazon purchases by households in the 19 states that implemented an Amazon Tax, we only adjust transactions after the law has been implemented.[9] All variables are winsorized at the 99th percentile.

Table 3.1 shows the geographic distribution of households in our sample relative to the 2010 US Census. Our sample is quite geographically diverse and maps fairly well to the US Census data. Our sample does contain more California and New York residents than the general population, potentially raising the concern that our results are attenuated to reflect the behavior of households in these states. However, New York implemented an Amazon Tax law prior to the study data period (2008), and thus is always in the control sample. California implemented the law during the study period. To ensure that the results are not driven by California-specific behavior, we rerun our main analyses excluding California, and find that the results remain virtually unchanged.

Figure 3.2 shows annual income of households in our sample and in US Census data. Our dataset maps fairly well to the US Census, but with a few caveats. The income we observe flows through to a households checking or savings account. Thus, it will be equal to gross income minus the sum of withholdings (payroll tax, state tax, federal tax, healthcare contributions, retirement contributions, etc.). Consequently, a households gross income will be higher than what we directly observe. Nonetheless, the data are well dispersed across income groups and seem to be reasonably representative of the US income distribution.

We provide the average tax-exclusive Amazon spending before and after the Amazon tax implementation of each state in Table 3.2. In this table, the tax-exclusive

[9] For two states (Pennsylvania and California), the implementation of the Amazon tax took place at the middle of the month. In these cases, we removed the household-month observations from the transition month.

spending for a particular state is reported along with that of the control states. As shown in this table, treated states reduce tax-exclusive spending at Amazon relative to control states. We analyze this formally in the subsequent sections.

3.4 States Implementing the Amazon Tax

States that decide to implement the Amazon Tax are, of course, not drawn randomly. This fact raises the concern that the decline in Amazon sales that we document occurs due to an unobservable confounding factor that pushes states to embrace the Amazon Tax and at the same time causes a decline in Amazon sales. Perhaps the most obvious potential latent factor is a state-level economic weakness that leads states to adopt the Amazon Tax in order to increase revenues, and at the same time causes a decline in consumption.

We address this concern in three ways. First, we explore whether states that implement the Amazon Tax during our sample period experience significantly different gross domestic product (GDP) growth around the implementation of the tax than states that did not implement the tax. We collect five-year GDP growth data around the implementation year. Then, we test whether the average GDP growth is different for state-quarters following the implementation of the Amazon Tax. Table 3.3, Columns (1) and (2) present the results. The regressions indicate no significant difference in state-level GDP following the Amazon Tax implementation.

Second, we test whether households income changed around the implementation of the tax using household-month data. We extract a households income from its cash flows. We regress household income on time dummies, surrounding the implementation of the Amazon Tax. In addition, we include month fixed effects and household fixed effects. The results, found in Columns (3) and (4) of Table 3.3, show that households did not experience a meaningful change in income around the implementation of the tax. Hence, it is not likely that our main results are due to changes in the purchasing power of households.

Third, because a state-level slowdown typically is accompanied by a general decline in consumption, we examine whether the pattern of purchasing at electronics retailers changed after the taxs implementation (Section 5.5). We find no such decline in consumption.

In sum, we conclude that our results are not likely to be driven by a state-level economic weakness that caused states to implement the Amazon Tax and at the same time caused a slowdown in consumption.

3.5 The Effect of the Amazon Tax on Amazon Sales

In this section, we examine how Amazons sales in the treated states changed after implementation of the tax and compare these results to Amazons sales in states that did not change their laws. We perform this analysis using both the tax-exclusive price and the tax-inclusive price. We also expect that different types of households might react to the new tax differently. Thus, we repeat our analysis but split our sample into terciles based first on household income and then Amazon historical spending intensity. Finally, we examine the taxs effect on large purchases exclusively.

We use a diff-in-diff methodology in which we measure the consumption effects after states started imposing sales tax on Amazon purchases. Our basic empirical specification is

$$Y_{h,t} = \beta_0 + \beta_1 \times TreatedState_h \times I(t \geq Q)_{h,t} + CostofLivingIndex_{c,t}$$
$$+ MonthFixedEffects_t + HouseholdFixedEffects_h + \varepsilon_{h,t}$$

where $Y_{h,c,t}$ is the dependent variable of interest and takes on the value of monthly Amazon expenditures (both tax-exclusive or tax-inclusive spending on Amazon). $TreatedState_h \times I(t \geq Q)_{h,t}$ is an indicator variable that takes a value of 1 for treated households after implementation of the Amazon Tax, and 0 otherwise. In a slightly modified empirical specification, we divide the $TreatedState_h \times I(t \geq Q)_{h,t}$ term into a more granular interactive term to investigate short- versus long-term responses to the treatment at a quarterly frequency. To account for regional differences in cost of living that vary over time, we introduce a time-varying cost of living index at the city-month level, denoted $CostofLivingIndex_{c,t}$. This index is computed by calculating the mean expenditures in the categories of gas, restaurants, groceries, and retail (excluding Amazon purchases) for each city-month.[10]

3.5.1 Average Value of Purchased Goods (Tax-Exclusive Price)

We begin our analysis by examining whether the average monthly amount that households spend on Amazon purchases changes as a result of the new sales tax. For each household in the sample, we aggregate the dollar amount spent on Amazon products within each month. Because we are interested in the impact of the sales tax on Amazons sales and the value to households, we create the tax-exclusive price by dividing by one plus the local tax rate.

[10]We thank the referee for the suggestion to control for time-varying differences in cost of living across locations.

Table 3.4 presents the results of this analysis. Column (1) shows the change in average monthly Amazon spending after the tax was implemented. The results indicate that consumers in affected states reduced their average monthly purchases on Amazon by \$3.65, a 9.4% (3.65/39.00) reduction in purchases relative to mean monthly spending among the treated states before the tax was implemented. This result is statistically and economically significant and corresponds to an elasticity of 1.2.[11] Because these values are tax-exclusive, the drop in spending reflects a drop in Amazons revenues in the affected states.

In Column (2), we examine the timing of the Amazon purchases in the quarter preceding and in the quarters following the tax implementation. $I(t = Q_{-1})_{h,t}$, $I(t = Q_0)_{h,t}$, and $I(t = Q_{+1})_{h,t}$ are indicator variables for the quarter(s) before, quarter after, and subsequent quarters following the tax implementation, respectively. We find some evidence of a buildup in purchases before the Amazon Tax was implemented, corresponding to an increase of 3.6% (1.42/39.00).

In the quarter immediately following the sales tax implementation, consumers in the affected states reduced their monthly Amazon purchases by \$3.29, corresponding to an 8.4% (3.29/39.00) reduction from the mean. In subsequent quarters, the reduction of Amazon purchases is \$3.21, corresponding to an 8.2% (3.21/39.00) reduction from the mean. The results are highly statistically significant.

In Column (3), we interact our $TreatedState_h \times I(tgeQ)_{h,t}$ term with the local tax rate of each household to examine whether the households that lived in localities with a high sales tax were more sensitive to the implementation of an Amazon Tax.

[11] $(\$3.65/\$39.00)/7.5\% = 1.24$.

Indeed, we find that a 1% increase in sales taxes leads to a \$54.32 reduction in monthly Amazon spending, corresponding to an elasticity of 1.4.[12]

3.5.2 Average Spending (Tax-Inclusive Price)

We also assess whether households changed their overall expenditure on Amazon (tax-inclusive price, which includes the effect of sales tax on price). We rerun our analysis from the previous section but use the tax-inclusive price. This analysis examines whether households spend less money overall on Amazon when the Amazon Tax is in effect. It is difficult to predict ex-ante the direction of the results in this analysis because households may increase their overall expenditure, keep it the same, or even decrease it in the wake of the new sales tax.

In Table 3.4, Columns (4) through (6), we repeat the previous tests using as the dependent variable the tax-inclusive Amazon expenditures. The coefficient in Column (4) is 1.21, corresponding to a 3.0% (1.21/40.73) reduction in tax-inclusive Amazon expenditures after implementation of the Amazon Tax. However, this coefficient is only marginally significant. Column (5) confirms a run-up in spending in the quarter prior to treatment but shows no significant change in tax-inclusive Amazon spending in subsequent quarters. Finally, Column (6) confirms that tax-inclusive Amazon spending is sensitive to the sales tax rate of the household. Treated households reduce Amazon spending, inclusive of tax, by \$21.21 per month for every 1% increase in sales tax, corresponding to an elasticity of 0.5.

[12] (\$54.32/\$39.00) = 1.39

3.5.3 The Cross-Section of Households

Different households may react to the inclusion of sales tax differently. In this section, we explore heterogeneity in household responses along two dimensions: income and historical purchases on Amazon. The analysis in Table 3.5 repeats the main specification (Column (1) of Table 3.4) but uses subsets of the population.

We first split the sample into terciles based on observable household income and perform our main specification for each tercile. Columns (1) to (3) of Table 3.5 indicate that low-income households are the most sensitive to the Amazon Tax, reducing Amazon purchases by \$3.04 per month, corresponding to a 9.9% reduction in spending relative to mean and an elasticity of 1.3. High-income households reduce their purchases by \$3.76, corresponding to a 7.0% reduction in spending and an elasticity of 1.0. These results are consistent with low-income households being more price sensitive than high-income households. Further, the results are also consistent with lower income households having lower opportunity costs and being willing to bear search costs to substitute to alternative retailers.

We also split households into terciles by the total amount of Amazon purchases in 2011 to explore how past Amazon shopping behavior might affect a households response to the new tax. Columns (4) through (6) of Table 3.5 present the results. We find that households with high Amazon spending in 2011 exhibited the biggest dollar decline in spending. Such households reduced Amazon purchases by \$6.22, corresponding to a 9.4% reduction in Amazon purchases and an elasticity of 1.3. This coefficient is highly statistically significant. In contrast, households with low Amazon spending in 2011 exhibited the lowest decline in spending. Such households reduced Amazon purchases by \$1.65, which corresponds to an 8.0% reduction and an elasticity of 1.1.

3.5.4 Large Purchases

Given that the amount of sales tax charged on an item is proportional to its price, we expect households to be more sensitive to sales taxes as the size of the purchase increases, especially when assuming some sort of fixed search costs. For example, assume a household has a sales tax rate of 10%. If the household were to purchase a \$10 (\$1,000) item at a local brick-and-mortar retailer, it would result in a \$1 (\$100) sales tax charge. When there is a fixed search cost associated with finding the tax savings, this household would be more likely to purchase the \$1,000 item online as opposed to the \$10 item. However, after implementation of the Amazon Tax, the tax avoidance incentive to make large purchases through Amazon is removed, and any observed change in behavior surrounding this event could be attributed to the Amazon Tax.

We test this prediction in Table 3.6, which repeats the base regressions (from Table 3.4) with a new dependent variable consisting of transactions of at least \$250. Specifically, for each household in the sample, we include only Amazon transactions of at least \$250 using tax-exclusive prices. Transactions below these amounts are set to zero. Then, we aggregate the large transactions at the household-month level.

The results show that the effects are substantially stronger for large purchases. Column (1) shows the average decline in Amazon sales is 29.1% (2.25/7.73), corresponding to an elasticity of 3.9. In the more granular specification, Column (2) shows that there is some buildup in purchases before the tax took effect and that the decline in purchases following the tax implementation is persistent at a rate of 27.2% (2.10/7.73),. Column (3) shows that the reduction in large purchases increases with the tax rate of the household: A 1% increase in sales tax results in a \$31.92 reduction in large purchases, corresponding to an elasticity of 4.1 for large purchases. Columns (4) through (6) repeat these regressions for tax-inclusive spending and find similar patterns.

To better understand the persistence of these effects, we plot the coefficients for the regression in Figure 3.3 using month dummies instead of quarter dummies. We see a buildup in purchases in the quarter prior to the Amazon Tax taking effect, after which there is a large and persistent reduction in Amazon purchases. This trend is true for both total Amazon purchases and Amazon purchases over \$250. Both Table 3.6 and Figure 3.3 highlight the fact that a large portion of the aggregate results are driven by large purchases.

In Table 3.7, we further examine the relation between large purchases and the tax increase, by subpopulations. As before, we split the sample by income and by historical Amazon purchases in 2011. We detect similar patterns to those we found in Table 3.5. Column (3) shows that low-income households reduce their large purchases at Amazon by 34.3% after implementation of the Amazon Tax, corresponding to an elasticity of 4.5. In contrast, Column (1) shows that high-income households reduce their large purchases only by 24.8%, corresponding to an elasticity of 3.3.

Column (4) shows that those with high past Amazon expenditures reduce tax-exclusive spending by 30.1% (implying an elasticity of 4.0), while Column (6) shows that those with low past Amazon expenditures reduce spending by 26.0% (implying an elasticity of 3.5).

In Appendix Section A4, we explore how the probability of purchasing through Amazon changes as a function of purchase size. We find that the probability of treated households making large purchases declines following implementation of the Amazon Tax but detect no change in the probability of making other purchases.

3.5.5 Substitution to Competing Retailers

We are interested in whether the forgone sales of Amazon went to competing firms and whether these firms are brick-and-mortar stores or other online retailers. Previous studies have found that the imposition of sales tax pushes consumers to look for alternative sellers who do not collect sales tax. For example, evidence of cross-border shopping (e.g., [15]; [5] indicates substitution in the physical sphere. In the online arena, [39] find that eBay customers back out of transactions once they find that they need to pay sales tax and that they are more likely to instead buy another item from an out-of-state seller who does not collect sales tax. [40] document that buyers of memory modules choose to purchase from sellers who do not collect sales tax. The substitution observed in these studies of online retailers is performed on the same platform (either eBay or Pricewatch, respectively), making it is easy for the consumer to substitute within the platform and for researchers to identify the effect. In the case of Amazon, substitution may be costlier for customers and is more difficult for researchers to detect.

In our tests of substitution, we face a data issue. While we observe transaction amounts at Amazon and the competing firms, we do not know what products were purchased. Furthermore, if there is substitution to other retailers, it is likely spread among several competitors rather than one retailer. Finally, it is empirically difficult to detect an increase in sales in giant competitors like Walmart, Costco, or Target that sell a wide array of products including some that are not usually offered by Amazon (e.g., groceries).

Nevertheless, we can provide some evidence about substitution in specific areas. In this section, we investigate electronics retailers as well as broad Internet merchants. We focus on electronics products for several reasons. First, these are often large purchases, making it worth the shoppers time to find a good deal. Second, these products are easily identifiable by brand and model; hence, shoppers can easily compare prices across outlets. Third, competing retailers in the electronics space specialize in electronics only, sharpening the empirical test. We, therefore, look at the largest competing electronics stores: Best Buy and Newegg. Best Buy is the largest electronics retailer in the United States, and Newegg is the second largest online-only retailer after Amazon. Best Buy has physical presence in most states and thus collects sales tax both for physical and online sales. Newegg, however, is headquartered in California and has limited operations in two other states, so it is only required to collect sales tax from purchases in three states.[13] To gain more insights into household behavior, we divide Best Buy transactions into brick-and-mortar and online purchases.

Next, we identify transactions through eBay, which is a viable competitor to Amazon, selling a wide variety of products in its online marketplace. Unfortunately, there is no easy way to identify eBay transactions in our dataset because the majority of

[13]www.newegg.com/HelpInfo/FAQDetail.aspx?Module=2

these transactions occur through PayPal payments directly to eBay sellers.[14] The portion of these transactions that contain the keyword eBay we unambiguously classify as eBay transactions. All other PayPal transactions we leave in their own PayPal category, with the understanding that this is an imperfect proxy for eBay transactions. Next, we identify all other Internet merchants by searching for the keyword .com for all retail transactions not previously classified into the other categories in an attempt to capture a wide breadth of online retailers.

To test for the possibility that competing electronics retailers benefited from some of Amazons forgone sales, we regress total spending of the competing retailers sales on the $TreatedState_h \times I(t \geq Q)_{h,t}$ variable introduced earlier. As with the previous regressions, we also include household and month fixed effects. The results of the substitution analysis are presented in Panel A of Table 3.8. We find no significant results for Best Buy in Columns (1) and (2). However, we find evidence of substitution toward Newegg in Column (3). On average, households increase their purchases at Newegg by \$0.25 per month, corresponding to a 13.0% increase in expenditures. The result is highly statistically significant and could be attributable to the fact that it retains its tax advantage over Amazon and Best Buy. In Columns (4) and (5), we find no significant results for eBay or PayPal, respectively. Likewise, Column (6) indicates no evidence of substitution toward other retailers captured with the .com retail query.

In Panel B of Table 3.8, we look at substitution using an alternative approach. In this panel, we explore whether the ratio of Amazon to total retail purchases (including Amazon) changes as a result of the Amazon Tax. In Column (1), we find that treated

[14]Paypal, owned by eBay, is the primary payment system on the eBay platform.

households reduce the share of Amazon purchases by 0.5 percentage points. When we investigate more carefully in Column (2), we find that this substitution to other retailers is driven primarily by heavy Amazon shoppers who reduce the share of Amazon purchases by 1.3 percentage points.

3.5.6 Substitution to Amazon Marketplace

We also analyze potential substitution of Amazon customers to Amazon Marketplace. Amazon Marketplace is a platform that allows third-party sellers to sell products directly on Amazons website. Many products on Amazon are sold by both Amazon.com and Amazon Marketplace within a single product page. Amazon handles the billing and often the shipping of these orders, so Amazon Marketplace sellers are an almost perfect substitute for Amazon. Because these third-party Amazon Marketplace sellers have limited geographical footprints and are not subject to the Amazon Tax laws, products sold by these sellers are not generally taxed. However, the sales tax advantage of these Marketplace sellers may not be immediately evident to the casual shopper who mistakenly assumes that the Amazon Tax laws apply to both Amazon and Amazon Marketplace transactions.

We test the effect of the Amazon Tax on Marketplace sales in Column (7) of Table 3.8 and find a marginally significant negative coefficient on the variable, corresponding to a 2.3% reduction in Amazon Marketplace expenditures among treated households. This surprising result could stem from treated Amazon shoppers not knowing that Amazon Marketplace transactions allow them to avoid paying sales tax. Thus, any positive effects from the more attractive treatment of sales tax of Marketplace transactions appear to be offset by the negative effects of the perceived increases in taxes by the casual Amazon shopper.

3.5.7 Income Effects Caused by the Amazon Tax

In this section, we explore the income effects resulting from implementation of the Amazon Tax. It is reasonable to assume that those who were the heaviest Amazon spenders would be most impacted by the implementation of the Amazon Tax. We formally test this in Table 3.9.

We divide households into terciles based on their Amazon spending in 2011, with tercile 3 being the highest spending group. We then interact these tercile indicators with the $TreatedState_h \times I(t \geq Q)_{h,t}$ term from previous tables to understand the differential income effects across these groups. We omit $TreatedState_h \times I(t \geq Q)_{h,t} \times AmazonTercile1$ from the regressions, which will serve as our baseline group.

The categories of consumption we analyze constitute a large share of a typical households spending: restaurants ($253/month), groceries ($315/month), and entertainment ($35/month).

Regression results are found in Columns (1) through (3) of Table 3.9. In each of these categories, there is a clear monotonic relationship between 2011 Amazon spending and the reduction in spending after treatment. These results suggest that implementation of the Amazon Tax produced a negative income effect concentrated among the heaviest Amazon shoppers.

3.6 Conclusion

Taxes affect not only business decisions by managers, but also purchasing decisions by customers. In the aggregate, purchasing decisions have significant effects on corporations. In this study, we analyze the effects of implementing the Amazon Tax law in various states. The law requires Amazon to collect sales tax, which in turn makes Amazons products less competitive.

Using transaction-level data of 275,437 households in our main specifications, we examine the effects of the Amazon Tax on the purchasing behavior of residents living in 19 states that adopted such laws during the 2012-2015. We find that Amazon sales fall by 9.4% after implementation of an Amazon Tax, corresponding to an elasticity of 1.2. We further find that a one percentage point increase in the tax rate of the household leads to a $54.33 reduction in tax-exclusive Amazon spending, corresponding to an elasticity of 1.4. We find the effect to be concentrated among large purchases of at least $250. For this subset of purchases, we find that Amazon sales fall by 29.1% after implementation of the Amazon Tax, corresponding to an elasticity of 3.9.

To understand whether Amazons competitors benefit from the law, we examine the sales of the online retailers competitors in the electronics industry. We find no evidence of substitution toward Best Buy, Amazons largest competitor in the electronics space, but our results indicate substitution to Newegg. Finally, we find evidence of an income effect spilling over into other categories of consumption such as restaurants, groceries, and entertainment. As expected, we find that the income effect is concentrated among the heaviest Amazon shoppers, who reduce spending in each of the observed categories by the largest amount.

Figure 3.1: Histogram of (Sales Tax Revenue / Total State Revenue) for the 50 States in 2011

This figure illustrates the importance of sales tax revenues as a percentage of total state revenues. The data come from 2011 US Census Annual Survey of State and Local Government Finance: www.census.gov/govs/local/. This figure shows that the importance of this tax varies considerably across states, ranging from 0% of state revenues in states without a sales tax (such as Oregon and Alaska) to as high as 21.0% of state revenues for Washington.

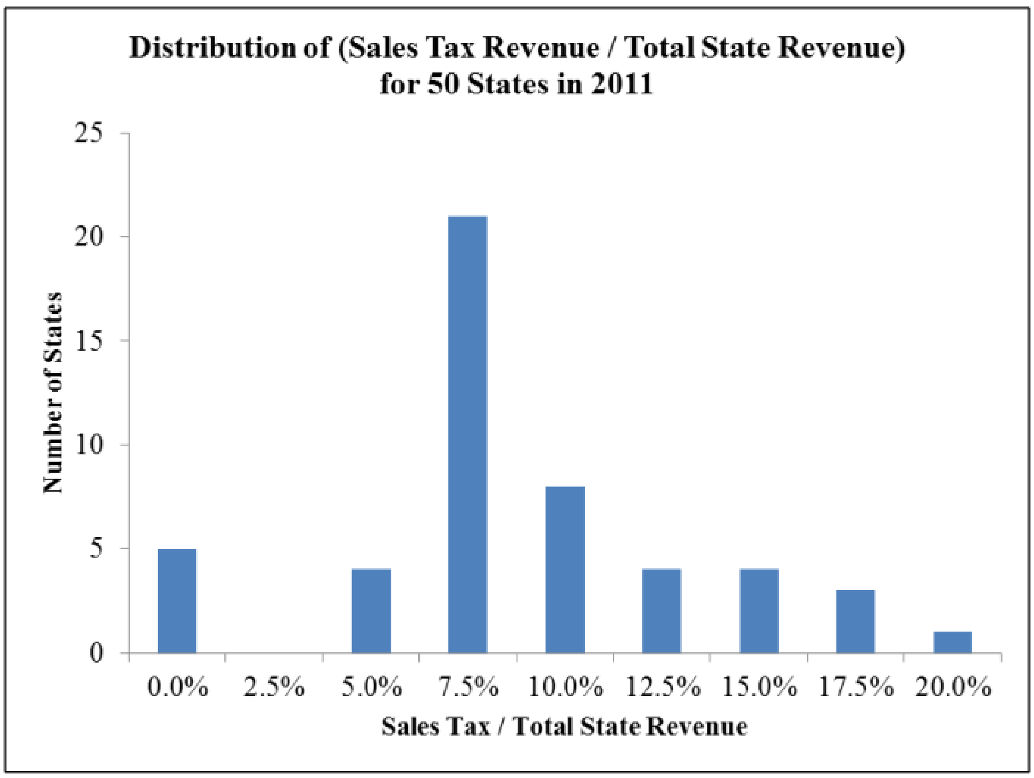


Figure 3.2: Distribution of Annual Income

This figure illustrates the differences in the distribution of annual income between our sample and the US Census. The income observed in our data is that which arrives in households checking and savings accounts. Therefore, it equals gross income minus the sum of withholdings (payroll tax, state tax, federal tax, healthcare contributions, retirement contributions, etc.). These omissions will result in a gross income that is higher than what we directly observe.

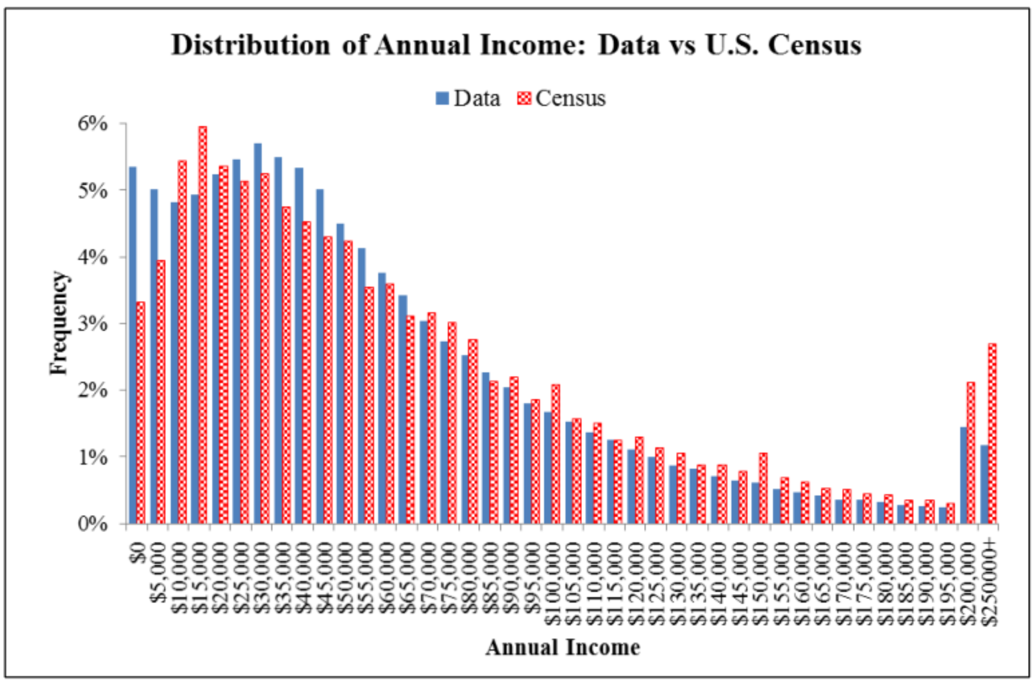


Figure 3.3: Amazon Spending Before and After the Amazon Tax

This figure illustrates the trend of the regression coefficients of monthly Amazon spending in the 6 to +32-month window surrounding implementation of Amazon Tax laws. The specification is similar to the base specification described previously but with a series of months-after-treatment indicator variables rather than quarters-after-treatment indicators. We run two different regressions. The dependent variable for the first regression is the sum of all tax-exclusive Amazon purchases. The dependent variable for the second regression is the sum of all tax-exclusive Amazon purchases that are at least $250 in size. Regression coefficients for the two regressions are plotted. To account for regional differences in cost of living that vary over time, we introduce a time-varying cost of living index at the city-month level, denoted Cost of Living Indexc,t. This index is computed by calculating the mean expenditures in the categories of gas, restaurants, groceries, and retail (excluding Amazon purchases) for each city-month.

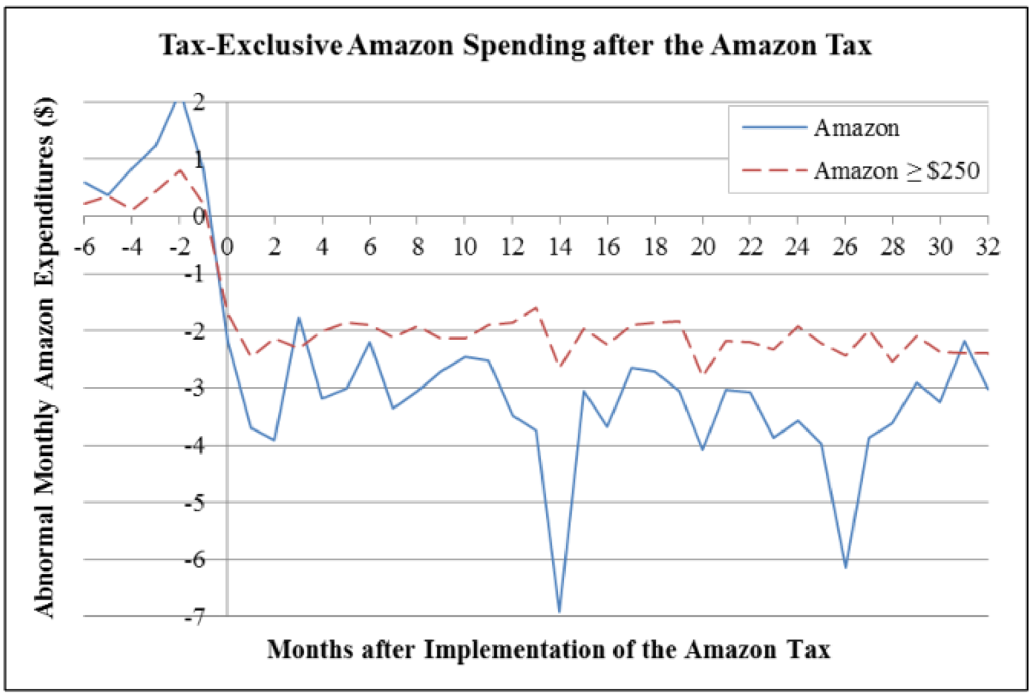


Table 3.1: Summary Statistics

This table shows the geographic distribution of the households in the sample relative to the 2010 US Census.

State	% Households Residing			State	% Households Residing		
	Data	US Census	Data - US Census		Data	US Census	Data - US Census
Alabama	0.6%	1.5%	-1.0%	Montana	0.1%	0.3%	-0.2%
Alaska	0.3%	0.2%	0.0%	Nebraska	0.3%	0.6%	-0.3%
Arizona	1.8%	2.1%	-0.2%	Nevada	0.9%	0.9%	0.0%
Arkansas	0.3%	0.9%	-0.6%	New Hampshire	0.2%	0.4%	-0.2%
California	21.5%	12.1%	9.5%	New Jersey	2.1%	2.8%	-0.8%
Colorado	1.1%	1.6%	-0.5%	New Mexico	0.4%	0.7%	-0.2%
Connecticut	1.2%	1.2%	0.1%	New York	19.2%	6.3%	13.0%
Delaware	0.1%	0.3%	-0.1%	North Carolina	2.5%	3.1%	-0.6%
District of Columbia	0.4%	0.2%	0.2%	North Dakota	0.1%	0.2%	-0.1%
Florida	6.2%	6.1%	0.1%	Ohio	0.7%	3.7%	-3.0%
Georgia	2.6%	3.1%	-0.5%	Oklahoma	0.6%	1.2%	-0.6%
Hawaii	0.4%	0.4%	-0.1%	Oregon	0.7%	1.2%	-0.5%
Idaho	0.2%	0.5%	-0.3%	Pennsylvania	1.2%	4.1%	-2.9%
Illinois	5.4%	4.2%	1.3%	Rhode Island	0.2%	0.3%	-0.2%
Indiana	0.4%	2.1%	-1.7%	South Carolina	0.9%	1.5%	-0.6%
Iowa	0.2%	1.0%	-0.8%	South Dakota	0.1%	0.3%	-0.2%
Kansas	0.4%	0.9%	-0.5%	Tennessee	1.0%	2.1%	-1.0%
Kentucky	0.3%	1.4%	-1.1%	Texas	10.9%	8.1%	2.8%
Louisiana	0.4%	1.5%	-1.0%	Utah	0.3%	0.9%	-0.6%
Maine	0.2%	0.4%	-0.3%	Vermont	0.1%	0.2%	-0.1%
Maryland	2.4%	1.9%	0.5%	Virginia	4.1%	2.6%	1.5%
Massachusetts	2.8%	2.1%	0.6%	Washington	1.7%	2.2%	-0.4%
Michigan	0.7%	3.2%	-2.5%	West Virginia	0.1%	0.6%	-0.5%
Minnesota	0.4%	1.7%	-1.3%	Wisconsin	0.3%	1.8%	-1.5%
Mississippi	0.2%	1.0%	-0.8%	Wyoming	0.1%	0.2%	-0.1%
Missouri	0.8%	1.9%	-1.1%				

Table 3.2: Average Monthly Tax-Exclusive Expenditures Before and After Sales Tax Change

This summary table presents average tax-exclusive spending at Amazon in the +/3-month window before and after implementation of Amazon Tax laws. We include only households that spent over $200 on Amazon during 2011. If an Amazon transaction occurs after the tax law changes and the household resides in one of the 19 affected states, we adjust the post-implementation transactions by dividing by one plus the local sales tax rate to create the tax-exclusive amount. Control states are the 31 states that do not change their Amazon tax status during our sample period.

	States (3-month window)									
	All	TX	PA	CA	AZ	NJ	VA	GA	WV	CT
Before implementation										
Treated state(s)	$40.51	$32.45	$37.56	$37.21	$51.00	$36.31	$44.05	$36.75	$38.30	$42.75
Control states	$35.71	$30.72	$31.09	$31.19	$46.83	$33.66	$34.38	$34.38	$34.45	$34.66
After implementation										
Treated state(s)	$39.93	$29.98	$37.64	$44.06	$31.90	$35.10	$45.63	$37.58	$58.65	$59.82
Control states	$39.68	$31.32	$35.27	$45.52	$32.06	$34.45	$37.74	$37.74	$51.13	$51.89
	MA	WI	IN	NV	TN	NC	FL	MD	MN	IL
Before implementation										
Treated state	$41.14	$44.75	$60.87	$54.06	$61.45	$58.11	$38.91	$42.83	$44.97	$49.18
Control states	$34.66	$34.66	$51.13	$51.13	$51.13	$51.88	$35.23	$36.68	$36.68	$46.66
After implementation										
Treated state	$56.07	$60.42	$39.43	$33.69	$35.95	$35.65	$36.02	$54.73	$52.59	$31.42
Control states	$51.89	$51.89	$35.50	$35.50	$35.50	$35.23	$36.98	$47.85	$47.89	$33.41

Table 3.3: State GDP Growth and household income around Amazon Tax Implementation

This table explores whether states that implemented the Amazon Tax experienced a different GDP growth (Columns (1) and (2)) or a change in household income (Columns (3) and (4)) than states that did not implement the tax. All regressions are ordinary least squares (OLS) regressions and include time and state fixed effects. The unit of observation in Columns (1) and (2) is the state quarter. The regression in Column (1) is weighted by the GDP of the each state. The regression in Column (2) is weighted by the relative number of households in each state in the sample. The unit of observation in Columns (3) and (4) is the household month. Column (3) looks at household income after the tax implementation in the treated states. Column (4) looks at the short-term and long-term changes in household income after the tax implementation in the treated states. Standard errors are clustered by state and time. Treated State is an indicator variable that takes a value of 1 for states that implemented an Amazon Tax during our sample period. $I(t \geq Q)_{h,t}$ is an indicator variable that takes a value of 1 for all months after implementation of the Amazon Tax. $I(t = Q_{-1})_{h,t}$, $I(t = Q_0)_{h,t}$, and $I(t \geq Q_{+1})_{h,t}$ are indicator variables for the quarter(s) before, quarter after, and subsequent quarters following the tax implementation, respectively. t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable:	State-level GDP growth (%)		Income	
	(1)	(2)	(3)	(4)
$TreatedState \times I(t \geq Q)$	0.184	-0.104	58.224	
	(0.42)	(-0.22)	(1.68)	
$TreatedState \times I(t = Q-1)$				-3.130
				(-0.09)
$TreatedState \times I(t = Q0)$				36.061
				(1.11)
$TreatedState \times I(t \geq Q+1)$				65.934
				(1.41)
State Fixed Effect	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes
Weighting	GDP	#Households		
Obs	757	757	10,436,160	10,436,160
R^2	48%	52%	73%	73%

Table 3.4: Effect of Amazon Tax on Monthly Amazon Expenditures

This table explores the effect of the Amazon Tax on Amazon expenditures. The unit of observation is the household month, and the dependent variable is the sum of monthly Amazon transactions per household. Columns (1) through (3) evaluate tax-exclusive expenditures, while Columns (4) through (6) evaluate tax-inclusive expenditures. Treated State is an indicator variable that takes a value of 1 for states that implemented an Amazon Tax during our sample period. $I(t \geq Q)_{h,t}$ is an indicator variable that takes a value of 1 for all months after implementation of the Amazon Tax. $I(t = Q_{-1})_{h,t}$, $I(t = Q_0)_{h,t}$, and $I(t \geq Q_{+1})_{h,t}$ are indicator variables for the quarter(s) before, quarter after, and subsequent quarters following the tax implementation, respectively. $TaxRate_h$ is the households sales tax rate. To account for regional differences in cost of living that vary over time, we introduce a time-varying cost of living index at the city-month level, denoted $CostofLivingIndex_{c,t}$. This index is computed by calculating the mean expenditures in the categories of gas, restaurants, groceries, and retail (excluding Amazon purchases) for each city-month. All regressions are ordinary least squares (OLS) regressions and include household and year-month fixed effects. Standard errors are clustered by state and time. t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable:	Amazon spending (tax-exclusive)			Amazon spending (tax-inclusive)		
	(1)	(2)	(3)	(4)	(5)	(6)
$TreatedState \times I(t \geq Q)$	-3.648***			-1.205*		
	(-5.07)			(-1.76)		
$TreatedState \times I(t = Q_{-1})$		1.421***			1.324***	
		(2.87)			(2.71)	
$TreatedState \times I(t = Q_0)$		-3.289***			-0.850	
		(-3.82)			(-1.08)	
$TreatedState \times I(t \geq Q_{+1})$		-3.208***			-0.803	
		(-4.47)			(-1.16)	
$TreatedState \times I(t \geq Q) \times Taxrate$			-54.328***			-21.210**
			(-7.05)			(-2.63)
Household Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
YYYYMM Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Cost of Living Index (City-Month)	Yes	Yes	Yes	Yes	Yes	Yes
Obs	10,436,160	10,436,160	10,436,160	10,436,160	10,436,160	10,436,160
R^2	28%	28%	28%	28%	28%	28%
Mean spending of treated	\$39.00	\$39.00	\$39.00	\$40.73	\$40.73	\$40.73
Mean tax rate of treated	7.5%	7.5%	7.5%	7.5%	7.5%	7.5%
Implied Elasticity ($TreatedState \times I(t \geq Q)$)	-1.24			-0.39		
Implied Elasticity ($TreatedState \times I(t = Q_{-1})$)		0.48			0.43	
Implied Elasticity ($TreatedState \times I(t = Q_0)$)		-1.12			-0.28	
Implied Elasticity ($TreatedState \times I(t = Q_{+1})$)		-1.09			-0.26	
Implied Elasticity ($TreatedState \times I(t \geq Q) \times Taxrate$)			-1.39			-0.52

Table 3.5: Effect of Amazon Tax on Different Types of Households

This table explores the effect of the Amazon Tax on different types of households. The unit of observation is the household month, and the dependent variable is the tax-exclusive sum of monthly Amazon transactions per household. Households are divided into three groups depending on their monthly income and total Amazon spending in 2011. . Treated State is an indicator variable that takes a value of 1 for states that implemented an Amazon Tax during our sample period. $I(t \geq Q)_{h,t}$ is an indicator variable that takes a value of 1 for all months after implementation of the Amazon Tax. To account for regional differences in cost of living that vary over time, we introduce a time-varying cost of living index at the city-month level, denoted $CostofLivingIndex_{c,t}$. This index is computed by calculating the mean expenditures in the categories of gas, restaurants, groceries, and retail (excluding Amazon purchases) for each city-month. All regressions are ordinary least squares (OLS) regressions and include household and year-month fixed effects. Standard errors are clustered by state and time. t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Amazon spending (tax-exlusive)					
	Income terciles			Amazon spending terciles		
	High	Mid	Low	High	Mid	Low
	(1)	(2)	(3)	(4)	(5)	(6)
$TreatedState \times I(t \geq Q)$	-3.755***	-3.675***	-3.038***	-6.224***	-2.830***	-1.649***
	(-3.53)	(-4.85)	(-5.81)	(-5.37)	(-4.82)	(-2.89)
Household Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
YYYYMM Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Cost of Living Index (City-Month)	Yes	Yes	Yes	Yes	Yes	Yes
Obs	2,501,723	2,501,759	2,501,788	3,478,700	3,478,723	3,478,737
R^2	30%	26%	24%	30%	20%	17%
Mean spending of treated	$53.34	$38.72	$30.57	$65.97	$29.86	$20.53
Mean tax rate of treated	7.4%	7.5%	7.6%	7.5%	7.5%	7.5%
Implied Elasticity	-0.95	-1.27	-1.31	-1.25	-1.26	-1.07
$TreatedState \times I(t = Q)$ / Mean spending	-7.0%	-9.5%	-9.9%	-9.4%	-9.5%	-8.0%

Table 3.6: Effect of Amazon Tax on Large Amazon Expenditures

This table explores the effect of the Amazon Tax on large Amazon expenditures. The unit of observation is the household month, and the dependent variable is the sum of monthly Amazon transactions per household that are at least $250. Columns (1) through (3) evaluate tax-exclusive expenditures, while Columns (4) through (6) evaluate tax-inclusive expenditures. Treated State is an indicator variable that takes a value of 1 for states that implemented an Amazon Tax during our sample period. $I(t \geq Q)_{h,t}$ is an indicator variable that takes a value of 1 for all months after implementation of the Amazon Tax. $I(t = Q_{-1})_{h,t}$, $I(t = Q_0)_{h,t}$, and $I(t \geq Q_{+1})_{h,t}$ are indicator variables for the quarter(s) before, quarter after, and subsequent quarters following the tax implementation, respectively. $TaxRate_h$ is the households sales tax rate. To account for regional differences in cost of living that vary over time, we introduce a time-varying cost of living index at the city-month level, denoted $CostofLivingIndex_{c,t}$. This index is computed by calculating the mean expenditures in the categories of gas, restaurants, groceries, and retail (excluding Amazon purchases) for each city-month. All regressions are ordinary least squares (OLS) regressions and include household and year-month fixed effects. Standard errors are clustered by state and time. t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable:	Amazon spending $\geq$ $250 (tax-exclusive)			Amazon spending $\geq$ $250 (tax-inclusive)		
	(1)	(2)	(3)	(4)	(5)	(6)
$TreatedState \times I(t \geq Q)$	-2.249***			-1.791***		
	(-7.92)			(-6.48)		
$TreatedState \times I(t = Q_{-1})$		0.471**			0.441*	
		(2.16)			(1.98)	
$TreatedState \times I(t = Q_0)$		-2.128***			-1.668***	
		(-6.78)			(-5.51)	
$TreatedState \times I(t \geq Q_{+1})$		-2.103***			-1.659***	
		(-7.04)			(-5.58)	
$TreatedState \times I(t \geq Q)$ Tax rate			-31.923***			-25.705***
			(-11.31)			(-8.93)
Household Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
YYYYMM Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Cost of Living Index (City-Month)	Yes	Yes	Yes	Yes	Yes	Yes
Obs	10,436,160	10,436,160	10,436,160	10,436,160	10,436,160	10,436,160
R^2	8%	8%	8%	8%	8%	8%
Mean spending of treated	$7.73	$7.73	$7.73	$8.05	$8.05	$8.05
Mean tax rate of treated	7.5%	7.5%	7.5%	7.5%	7.5%	7.5%
Implied Elasticity ($TreatedState \times I(t \geq Q)$)	-3.87			-2.96		
Implied Elasticity ($TreatedState \times I(t = Q_{-1})$)		0.81			0.73	
Implied Elasticity ($TreatedState \times I(t = Q_0)$)		-3.66			-2.75	
Implied Elasticity ($TreatedState \times I(t = Q_{+1})$)		-3.61			-2.74	
Implied Elasticity ($TreatedState \times I(t \geq Q) \times Taxrate$)			-4.13			-3.19

Table 3.7: Effect of Amazon Tax on Different Types of Households for Large Purchases

This table explores the effect of the Amazon Tax on different types of households for large purchases. The unit of observation is the household month, and the dependent variable is the tax-exclusive sum of monthly Amazon transactions per household that are at least $250. Households are divided into three groups depending on their monthly income and total Amazon spending in 2011. . Treated State is an indicator variable that takes a value of 1 for states that implemented an Amazon Tax during our sample period. $I(t \geq Q)_{h,t}$ is an indicator variable that takes a value of 1 for all months after implementation of the Amazon Tax. To account for regional differences in cost of living that vary over time, we introduce a time-varying cost of living index at the city-month level, denoted $CostofLivingIndex_{c,t}$. This index is computed by calculating the mean expenditures in the categories of gas, restaurants, groceries, and retail (excluding Amazon purchases) for each city-month. All regressions are ordinary least squares (OLS) regressions and include household and year-month fixed effects. Standard errors are clustered by state and time. t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Amazon spending $250 (tax-exclusive)					
	Income terciles			Amazon spending terciles		
	High	Mid	Low	High	Mid	Low
	(1)	(2)	(3)	(4)	(5)	(6)
$TreatedState \times I(t \geq Q)$	-2.601***	-2.157***	-1.879***	-4.250***	-1.467***	-0.935***
	(-5.56)	(-6.12)	(-9.45)	(-7.59)	(-7.53)	(-4.51)
Household Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
YYYYMM Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Cost of Living Index (City-Month)	Yes	Yes	Yes	Yes	Yes	Yes
Obs	2,501,723	2,501,759	2,501,788	3,478,700	3,478,723	3,478,737
R^2	7%	6%	6%	9%	5%	5%
Mean spending of treated	$10.50	$6.83	$5.48	$14.12	$5.33	$3.59
Mean tax rate of treated	7.4%	7.5%	7.6%	7.5%	7.5%	7.5%
Implied Elasticity	-3.34	-4.21	-4.53	-3.99	-3.66	-3.45
$TreatedState \times I(t = Q)$ / Mean spending	-24.8%	-31.6%	-34.3%	-30.1%	-27.5%	-26.0%

Table 3.8: Substitution Effects from the Amazon Tax

This table explores the effect of the Amazon Tax on other retailers. Panel A investigates the dollar value spent at Best Buy, Newegg, eBay, PayPal, generic online merchants, and Amazon Marketplace. Panel B investigates the percentage of retail spending occurring at Amazon. In both panels, the unit of observation is the household month. In Panel A, the dependent variable is the tax-inclusive sum of monthly retail transactions for a given retailer. Best Buy sales are categorized as either brick-and-mortar or online transactions. DotCom corresponds to a generic query intended to capture all other online merchants using the term .com in the description that are not otherwise classified in the other columns. We include households that spent at least $200 on Amazon during 2011. Treated State is an indicator variable that takes a value of 1 for states that implemented an Amazon Tax during our sample period. $I(t \geq Q)_{h,t}$ is an indicator variable that takes a value of 1 for all months after implementation of the Amazon Tax. To account for regional differences in cost of living that vary over time, we introduce a time-varying cost of living index at the city-month level, denoted $CostofLivingIndex_{c,t}$. This index is computed by calculating the mean expenditures in the categories of gas, restaurants, groceries, and retail (excluding Amazon purchases) for each city-month. All regressions are ordinary least squares (OLS) regressions and include household and year-month fixed effects. Standard errors are clustered by state and time. t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A

Dependent variable:	Best Buy (Brick)	Best Buy (Online)	Newegg	eBay	PayPal	DotCom	Amazon Marketplace
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$TreatedState \times I(t \geq Q)$	-0.018	-0.066	0.247***	0.030	-1.698	-0.271	-0.948*
	(-0.07)	(-0.53)	(2.99)	(1.22)	(-1.15)	(-0.19)	(-1.77)
Household Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YYYYMM Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cost of Living Index (City-Month)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	10,436,160	10,436,160	10,436,160	10,436,160	10,436,160	10,436,160	10,436,160
R^2	8%	5%	12%	27%	26%	21%	27%
Mean spending of treated	$11.63	$2.28	$1.89	$0.51	$36.31	$58.89	$41.51
Mean tax rate of treated	7.5%	7.5%	7.5%	7.5%	7.5%	7.5%	7.5%
$TreatedState \times I(t \geq Q)$ / Mean spending	-0.2%	-2.9%	13.0%	5.9%	-4.7%	-0.5%	-2.3%

Continued

Table 3.8 Continued

Panel B

Dependent variable:	Amazon / (Amazon + Other Retail)	Amazon / (Amazon + Other Retail)
	(1)	(2)
$TreatedState \times I(t \geq Q)$	-0.005***	0.000
	(-4.68)	(0.09)
$TreatedState \times I(t \geq Q)$ Amazon Tercile 2		-0.002***
		(-4.41)
$TreatedState \times I(t \geq Q)$ Amazon Tercile 3		-0.013***
		(-9.86)
Household Fixed Effect	Yes	Yes
YYYYMM Fixed Effect	Yes	Yes
Cost of Living Index (City-Month)	Yes	Yes
Obs	9,592,627	9,592,627
R^2	28%	28%

Table 3.9: Income Effects from the Amazon Tax

This table investigates the income effects following implementation of the Amazon Tax by exploring expenditures in the categories of restaurants, groceries, and entertainment. The unit of observation is the household month, and the dependent variable is the tax-inclusive expenditures for the given spending category. We include households that spent at least $200 on Amazon during 2011. Treated State is an indicator variable that takes a value of 1 for states that implemented an Amazon Tax during our sample period. $I(t \geq Q)_{h,t}$ is an indicator variable that takes a value of 1 for all months after implementation of the Amazon Tax. Households are divided into three groups depending on their monthly income and total Amazon spending in 2011. Amazon Tercile 3 as the group of Amazon shoppers with the highest Amazon expenditures in 2011. To account for regional differences in cost of living that vary over time, we introduce a time-varying cost of living index at the city-month level, denoted $CostofLivingIndex_{c,t}$. This index is computed by calculating the mean expenditures in the categories of gas, restaurants, groceries, and retail (excluding Amazon purchases) for each city-month. All regressions are ordinary least squares (OLS) regressions and include household and year-month fixed effects. Standard errors are clustered by state and time. t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable:	Restaurants	Groceries	Entertainment
	(1)	(2)	(3)
$TreatedState \times I(t \geq Q)$ Amazon Tercile 3	-9.281***	-5.454**	-4.739***
	(-4.77)	(-2.54)	(-8.30)
$TreatedState \times I(t \geq Q)$ Amazon Tercile 2	-4.775***	-1.926	-1.781***
	(-4.20)	(-1.27)	(-4.18)
$TreatedState \times I(t \geq Q)$	4.167***	1.379	2.659***
	(4.32)	(1.29)	(3.02)
Household Fixed Effect	Yes	Yes	Yes
YYYYMM Fixed Effect	Yes	Yes	Yes
Cost of Living Index (City-Month)	Yes	Yes	Yes
Obs	10,436,160	10,436,160	10,436,160
R^2	61%	68%	29%

Mean spending of treated	$252.61	$314.98	$65.17
Mean tax rate of treated	7.5%	7.5%	7.5%
$TreatedState \times I(t \geq Q)$ Amazon Tercile 3 / Mean spending	-3.7%	-1.7%	-7.3%
$TreatedState \times I(t \geq Q)$ Amazon Tercile 2 / Mean spending	-1.9%	-0.6%	-2.7%
$TreatedState \times I(t \geq Q)$ / Mean spending	1.6%	0.4%	4.1%

Bibliography

[1] Daniel Aaronson, Sumit Agarwal, and Eric French. The spending and debt response to minimum wage hikes. *The American Economic Review*, 102(7):3111–3139, 2012.

[2] Manuel Adelino, Antoinette Schoar, and Felipe Severino. Loan Originations and Defaults in the Mortgage Crisis: Further Evidence. 2015.

[3] Manuel Adelino, Antoinette Schoar, and Felipe Severino. Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class. 2016.

[4] Sumit Agarwal, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, Tomasz Piskorski, and Amit Seru. Policy Intervention in Debt Renegotiation: Evidence from the Home Affordable Modification Program. *Journal of Political Economy*, (conditionally accepted), 2016.

[5] Sumit Agarwal, Souphala Chomsisengphet, Teck-Hua Ho, and Wenlan Quian. Cross-Border Shopping: Do Consumers Respond to Taxes or Prices? 2013.

[6] Sumit Agarwal, Chunlin Liu, and Nicholas S Souleles. The reaction of consumer spending and debt to tax rebatesevidence from consumer credit data. *Journal of political Economy*, 115(6):986–1019, 2007.

[7] Sumit Agarwal, Nathan Marwell, and Leslie McGranahan. Consumption Responses to Temporary Tax Incentives: Evidence from State Sales Holidays. 2013.

[8] Mark Aguiar and Erik Hurst. Consumption versus Expenditure. *Journal of Political Economy*, 113(5):919–948, 2005.

[9] Chunrong Ai and Edward C Norton. Interaction terms in logit and probit models. *Economics letters*, 80(1):123–129, 2003.

[10] James Alm and Mikhail I Melnik. Sales taxes and the decision to purchase online. *Public Finance Review*, 33(2):184–212, 2005.

[11] Gene Amromin and Anna L. Paulson. Comparing patterns of default among prime and subprime mortgages. *Economic Perspectives*, 33(2):18–37, 2009.

[12] Eric T Anderson, Nathan M Fong, Duncan I Simester, and Catherine E Tucker. How sales taxes affect customer and firm behavior: The role of search on the internet. *Journal of Marketing Research*, 47(2):229–239, 2010.

[13] Patrick Bajari, Chenghuan Sean Chu, and Minjung Park. An Empirical Model of Subprime Mortgage Default From 2000 to 2007. *National Bureau of Economic Research Working Paper Series*, 14625:1–57, 2008.

[14] Scott R Baker. Debt and the Consumption Response to Household Income Shocks. 2015.

[15] Charles L Ballard and Jaimin Lee. Internet purchases, cross-border shopping, and sales taxes. *National Tax Journal*, pages 711–725, 2007.

[16] James Banks, Richard Blundell, and Sarah Tanner. Is There a Retirement-Savings Puzzle? *The American Economic Review*, 88(4):769–788, 1998.

[17] Brian Baugh. Payday borrowing and household outcomes; Evidence from a natural experiment. 2015.

[18] Brian Baugh, Itzhak Ben-David, and Hoonsuk Park. Can Taxes Shape an Industry? Evidence from the Implementation of the Amazon Tax. 2015.

[19] Brian Baugh, Itzhak Ben-David, Hoonsuk Park, and Jonathan A. Parker. Household Spend Tax Refunds and Smooth Tax Payments. 2017.

[20] Brian Baugh, Itzhak Ben-David, Hoonsuk Park, and Jonathan A. Parker. Household Spend Tax Refunds and Smooth Tax Payments. 2017.

[21] B. Douglas Bernheim, Jonathan Skinner, and Steven Weinberg. What Accounts for the Variation in Retirement Wealth among U.S. Households? *The American Economic Review*, 91(4):832–857, 2001.

[22] Asaf Bernstein. Household Debt Overhang and Labor Supply. 2016.

[23] Timothy J Besley and Harvey S Rosen. Vertical externalities in tax setting: evidence from gasoline and cigarettes. *Journal of Public Economics*, 70(3):383–398, 1998.

[24] David Bowman, Deborah Minehart, and Matthew Rabin. Loss aversion in a consumptionsavings model. *Journal of Economic Behavior & Organization*, 38:155–178, 1999.

[25] Martin Browning and Thomas F. Crossley. The Life-Cycle Model of Consumption and Saving. *Journal of Economic Perspectives*, 15:3–22, 2001.

[26] John Y Campbell and João F Cocco. Household Risk Management and Optimal Mortgage Choice. *The Quarterly Journal of Economics*, 118(4):1449–1494, 2003.

[27] John Y. Campbell and João F. Cocco. A Model of Mortgage Default. *Journal of Finance*, 70(4):1495–1554, 2015.

[28] John Y. Campbell and N. Gregory Mankiw. Permanent Income, Current Income, and Consumption. *Journal of Business & Economic Statistics*, 8(3):265–279, 1990.

[29] Tim S. Campbell and J. Kimball Dietrich. The Determinants of Default on Insured Conventional Residential Mortgage Loans. *The Journal of Finance*, 38(5):1569–1581, 1983.

[30] Asena Caner and Edward N Wolff. Asset Poverty in the United States, 1984-99: Evidence from the Panel Study of Income Dynamics. *Review of Income and Wealth*, 50(4):493–518, 2004.

[31] Christopher D Carroll. Buffer-stock saving and the life cycle/permanent income hypothesis. *The Quarterly journal of economics*, 112(1):1–55, 1997.

[32] Sewin Chan, Andrew Haughwout, Andrew Hayashi, and Wilbert Van Der Klaauw. Determinants of Mortgage Default and Consumer Credit Use: The Effects of Foreclosure Laws and Foreclosure Delays. *Journal of Money, Credit and Banking*, 48(2-3):393–413, 2016.

[33] Raj Chetty, Adam Looney, and Kory Kroft. Salience and taxation: Theory and evidence. *The American Economic Review*, 99(4):1145–1177, 2009.

[34] J.M. Clapp, Yongheng Deng, and Xudong An. Alternative models for competing risks of mortgage termination. *Real Estate Economics*, 34(2):243–273, 2006.

[35] Ethan Cohen-Cole and Jonathan Morse. Your house or your credit card, which would you choose? personal delinquency tradeoffs and precautionary liquidity motives. 2010.

[36] Yuliya Demyanyk and Otto Van Hemert. Understanding the subprime mortgage crisis. *Review of financial Studies*, 24(6):1848–1880, 2011.

[37] Yongheng Deng, John M Quigley, and Robert Order. Mortgage terminations, heterogeneity and the exercise of mortgage options. *Econometrica*, 68(2):275–307, 2000.

[38] Marco Di Maggio, Amir Kermani, and Rodney Ramcharan. Monetary Policy Pass-Through: Household Consumption and Voluntary Deleveraging. 2015.

[39] Liran Einav, Dan Knoepfle, Jonathan Levin, and Neel Sundaresan. Sales taxes and internet commerce. *The American Economic Review*, 104(1):1–26, 2014.

[40] Glenn Ellison and Sara Fisher Ellison. Tax sensitivity and home state preferences in internet purchasing. *American Economic Journal: Economic Policy*, 1(2):53–71, 2009.

[41] Ronel Elul, Nicholas S Souleles, Souphala Chomsisengphet, Dennis Glennon, and Robert Hunt. What "Triggers" Mortgage Default. *The American Economic Review Papers and Proceedings*, 100(2):490–494, 2010.

[42] Fernando Ferreira and Joseph Gyourko. A new look at the us foreclosure crisis: Panel data evidence of prime and subprime borrowers from 1997 to 2012. Technical report, National Bureau of Economic Research, 2015.

[43] Marjorie Flavin and Marjorie Flavin. Excess Sensitivity of Consumption to Current Income: Liquidity Constraints or Myopia? *Canadian Journal of Economics*, 18(1):117–136, 1985.

[44] Christopher Foote, Kristopher Gerardi, Lorenz Goette, and Paul Willen. Reducing foreclosures: No easy answers. *NBER Macroeconomics Annual*, 24(1):89–138, 2010.

[45] Nicola Fuchs-Schuendeln and Tarek Alexander Hassan. Natural Experiments in Macroeconomics. In *Handbook of Macroeconomics*, volume 2, chapter 12, pages 923–1012. Elsevier B.V., 2016.

[46] Andreas Fuster and Paul S Willen. Payment size, negative equity, and mortgage default. Technical report, National Bureau of Economic Research, 2013.

[47] Peter Ganong and Pascal Noel. The Effect of Debt on Default and Consumption : Evidence from Housing Policy in the Great Recession. 2016.

[48] Kristopher Gerardi, Adam H Shapiro, and Paul Willen. Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures. 2007.

[49] Paul Gertler and Jonathan Gruber. Insuring Consumption Against Illness. *The American Economic Review*, 92(1):51–70, 2002.

[50] Laurie S Goodman, Roger Ashworth, Brian Landy, and Ke Yin. Negative equity trumps unemployment in predicting defaults. *The Journal of Fixed Income*, 19(4):67–72, 2010.

[51] Austan Goolsbee. In a world without borders: The impact of taxes on internet commerce. *The Quarterly Journal of Economics*, 115(2):561–576, 2000.

[52] Austan Goolsbee. Internet commerce, tax sensitivity, and the generation gap. *Tax policy and the economy*, 14:45–65, 2000.

[53] Austan Goolsbee, Michael F Lovenheim, and Joel Slemrod. Playing with fire: Cigarettes, taxes, and competition from the internet. *American Economic Journal: Economic Policy*, 2(1):131–154, 2010.

[54] David B Gross and Nicholas S Souleles. An empirical analysis of personal bankruptcy and delinquency. *Review of financial studies*, 15(1):319–347, 2002.

[55] Joseph Gyourko and Joseph Tracy. Reconciling theory and empirics on the role of unemployment in mortgage default. *Journal of Urban Economics*, 80:87–96, 2014.

[56] Fumio Hayashi. The effect of liquidity constraints on consumption: a cross-sectional analysis. *The Quarterly Journal of Economics*, 100(1):183–206, 1985.

[57] Jeffrey L Hoopes, Jacob R Thornock, and Braden M Williams. Is Sales Tax Avoidance a Competitive Advantage? 2014.

[58] Chang-Tai Hsieh. Do Consumers React to Anticipated Income Changes? Evidence from the Alaska Permanent Fund. *The American Economic Review*, 93(1):397–405, 2003.

[59] Tullio Jappelli, Jörn-Steffen Pischke, and Nicholas S Souleles. Testing for liquidity constraints in euler equations with complementary data sources. *Review of Economics and statistics*, 80(2):251–262, 1998.

[60] Tullio Jappelli and Luigi Pistaferri. The Consumption Response to Income Changes. *Annual Review of Economics*, 2(1):479–506, 2010.

[61] Daniel Kahneman and Amos Tversky. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–292, 1979.

[62] Greg Kaplan and Giovanni L Violante. A model of the consumption response to fiscal stimulus payments. *Econometrica*, 82(4):1199–1239, 2014.

[63] Greg Kaplan, Giovanni L. Violante, and Justin Weidner. The Wealthy Hand-to-Mouth. *Brookings Papers on Economic Activity*, 48(1 (Spring)):77–153, 2014.

[64] James B. Kau, Donald C. Keenan, and Taewon Kim. Default Probabilities for Mortgages, 1994.

[65] John Maynard Keynes. *The General Theory of Employment, Interest and Money.* Macmillan, London, 1936.

[66] Benjamin J. Keys, Tomasz Piskorski, Amit Seru, and Vincent Yao. Mortgage Rates, Household Balance Sheets, and the Real Economy. 2014.

[67] Mordechai E. Kreinin. Windfall Income and Consumption: Additional Evidence. *The American Economic Review*, 51:388–390, 1961.

[68] Theresa Kuchler. Sticking to your plan: Empirical evidence on the role of present bias for credit card paydown. 2015.

[69] David Laibson. Golden Eggs and Hyperbolic Discounting. *The Quarterly Journal of Economics*, 112(2):443–477, 1997.

[70] Annamaria Lusardi, Daniel Schneider, and Peter Tufano. Financially Fragile Households : Evidence and Implications. *Brookings Papers on Economic Activity*, Spring:83–150, 2011.

[71] Nina Manzi. *Use tax collection on income tax returns in other states*. Research Department, Minnesota House of Representatives, 2004.

[72] Christopher Mayer, Karen Pence, and Shane M Sherlund. The rise in mortgage defaults. *The Journal of Economic Perspectives*, 23(1):27–50, 2009.

[73] Atif Mian and Amir Sufi. The consequences of mortgage credit expansion: Evidence from the us mortgage default crisis. *The Quarterly Journal of Economics*, 124(4):1449–1496, 2009.

[74] Atif Mian and Amir Sufi. Fraudulent income overstatement on mortgage applications during the credit expansion of 2002 to 2005. *Review of Financial Studies*, page hhw104, 2016.

[75] Christopher Palmer. Why Did So Many Subprime Borrowers Default During the Crisis: Loose Credit or Plummeting Prices? 2013.

[76] Jonathan A Parker. The reaction of household consumption to predictable changes in social security taxes. *The American Economic Review*, 89(4):959–973, 1999.

[77] Jonathan A. Parker. Why Don't Households Smooth Consumption? Evidence from a 25 Million Dollar Experiment. 2014.

[78] James M Poterba. Are consumers forward looking? evidence from fiscal experiments. *The American Economic Review*, 78(2):413–418, 1988.

[79] James M Poterba. Retail price reactions to changes in state and local sales taxes. *National Tax Journal*, pages 165–176, 1996.

[80] Roberto G Quercia and Michael A Stegman. Residential mortgage default: a review of the literature. *Journal of Housing Research*, 3(2):341, 1992.

[81] Mark A Scanlan. Tax sensitivity in electronic commerce. *Fiscal Studies*, pages 417–436, 2007.

[82] Thomas Schelkle. Mortgage Default during the U . S . Mortgage Crisis. 2012.

[83] Barry Scholnick. Consumption Smoothing after the Final Mortgage Payment: Testing the Magnitude Hypothesis. *Review of Economics and Statistics*, 95(4):1444–1449, 2013.

[84] John Shea. Union contracts and the life-cycle/permanent-income hypothesis. *The American Economic Review*, pages 186–200, 1995.

[85] Tyler Shumway. Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business*, 74(1):101–124, 2001.

[86] Nicholas S Souleles. The response of household consumption to income tax refunds. *The American Economic Review*, 89(4):947–958, 1999.

[87] Melvin Stephens. "3rd of tha month": Do social security recipients smooth consumption between checks? *The American Economic Review*, 93(1):406–422, 2003.

[88] Melvin Stephens. Paycheque receipt and the timing of consumption. *The Economic Journal*, 116(513):680–701, 2006.

[89] Kerry D Vandell. Default risk under alternative mortgage instruments. *The Journal of Finance*, 33(5):1279–1296, 1978.

[90] Kerry D Vandell. How ruthless is mortgage default? a review and synthesis of the evidence. *Journal of Housing Research*, 6(2):245, 1995.

[91] Stephen P Zeldes. Consumption and liquidity constraints: an empirical investigation. *Journal of political economy*, 97(2):305–346, 1989.

[92] C Yiwei Zhang. Monthly budgeting heuristics: Evidence from extrapaychecks. Technical report, Mimeo, University of Pennsylvania, Philadelphia, 2014.

Appendix A: Appendix to Chapter 3

A.1 Alternative Methods of Identifying City of Residence

Because the observed reduction in Amazon demand is dependent on the correct calculation of tax-exclusive Amazon expenditures, it is important that the correct sales tax value is used. If we use an incorrect sales tax rate, then the inferred tax-exclusive Amazon expenditures will be incorrect, leading to a potential overstatement (understatement) of the results in the event that the actual sales tax of the household is lower (higher) than the sales tax we assign to the household.

As mentioned in the Section 3, we identify the state of residence of households in our dataset by requiring that 75% of transactions occur within a given state. We then assign the most common city as the city of residence of the household.

In this section, we provide results for two alternative methods of identifying city of residence, both of which are straightforward. The first method simply takes the second most common city where transactions occur and assign the corresponding tax rate to this household. If, for example, an individual works in downtown Chicago and frequently gets coffee or lunch, we will mistakenly assign the city of residence of the household as Chicago, IL, rather than its actual hometown of Naperville, IL. Because

Chicago has a higher sales tax rate than Naperville (10.25% vs 7.25% at the time of this writing), this would lead to an overstatement of our results.

The second method is the most conservative. It takes the minimum sales tax of the first- and second-most common cities observed. Continuing from the example above, we would conservatively assume that the household resided in Naperville, IL and assign the more conservative 7.25% sales tax rate to the household. This lower sales tax rate would lead to a higher value for tax-exclusive Amazon purchases, and thus reduce the magnitude of our main results.

The results from these alternative methods are presented in Table A1. The main coefficient in Column (1) is 3.735 and is highly statistically significant, corresponding to an elasticity of 1.3. The coefficient is larger in magnitude than the coefficient of 3.648 found in Column (1) of Table 3.4. Similar results hold in Columns (2) through (3). When we repeat the activity using the second alternative method, the observed magnitude is 3.617 in Column (4) (corresponding to an elasticity of 1.2), which is only slightly lower in magnitude than the initial value. Similar results hold for Columns (5) and (6).

As a result, it does not appear that misclassification of city of residence is driving the results. Using the most conservative of the three methods, the results are still highly statistically and economically significant.

A.2 Removal of the $200 Amazon Spending Filter in 2011

As explained in Section 3, we are primarily interested in how Amazon customers respond after the implementation of the Amazon Tax. As a result, our main results focus on households that spent at least $200 on Amazon during 2011, prior to any of the tax law changes exploited in the paper. In this section, we relax this filter and instead include any household that had non-zero spending on Amazon during 2011. Doing so increases our sample size from 275,437 households (180,330 of which are in the treatment group) to 460,983 households (301,830 of which are in the sample group).

The results hold for this broader group as shown in Table A2. The main coefficient in Column (1) indicates a \$2.67 per month reduction in spending at Amazon and is highly significant. Note that this sample has lower mean monthly spending on Amazon of \$28.78 per month as opposed to \$39.00 in the main body of the paper. This reduction in mean spending is a natural result of the entry of households who spent less than \$200 in 2011 and thus are likely less frequent Amazon shoppers. When our main coefficient in Table A2 is normalized by the mean spending, it shows a reduction in spending of 9.3% (2.67/28.78), which is very close to that found in the main body of the paper of 9.4%. As a result, the implied elasticities are nearly equivalent with either method (1.23 using the whole sample vs. 1.24 using the restricted sample).

A.3 Alternative Calculation of Elasticity

In the main body of the text, we estimate elasticity in two straightforward ways. First, we use our difference-in-difference framework to estimate the change in the level of tax-exclusive Amazon spending among treated households. We then divide the level change by the mean Amazon spending of treated households to arrive at the percent reduction in tax-exclusive spending. We next divide by the sales tax rate to arrive at the elasticity.

The second way we estimate elasticities is by using the same difference-in-difference framework to estimate the dollar change in tax-exclusive Amazon spending for a one percentage point increase in the sales tax. We then normalize by the mean spending to arrive at the elasticity. For comparison purposes, our main estimations of elasticity shown in Table 3.4 are reproduced in Columns (1) and (2) of Table A3.

An alternative approach is to log the dependent variable and directly observe the elasticity from the regression coefficient. We do so in Column (3) of Table A3. In this regression, the dependent variable is the log of (1 + tax-exclusive Amazon spending). The regression coefficient, and elasticity, is 0.834 and highly statistically significant. However, the magnitude of the coefficient is smaller than that estimated in the main specifications as reproduced in Columns (1) and (2). To understand why the logged specification produces an elasticity that is significantly smaller than the others, we explore how the elasticity varies with purchase size in Table A4. Similar to the analysis performed in Table 3.5 with purchases over \$250, we create more refined bins of Amazon purchases in \$100 intervals. Column (1) in Table A4 corresponds to Amazon purchase sizes of \$0.01-\$99.99, Column (2) corresponds to Amazon purchase sizes of \$100.00-\$199.99, and so forth. Elasticities are computed in the bottom row and also plotted in Figure A1.

There is a clear negative trend between purchase size and elasticity. The largest observed elasticity is 6.8 in Column (8), corresponding to tax-exclusive purchases of \$700 and up. The smallest observed elasticity is 0.7 in Column (1), corresponding to tax-exclusive purchases of \$99.99 or less.

Taking into account the results in Table A4 and Figure A1, it is easy to understand why the estimated elasticity in Column (2) of Table A3 is smaller in magnitude than the elasticity estimates elsewhere in the paper, as the higher elasticities resulting from the bigger purchases are muted from the log transformation.

A.4 Probability of Amazon Purchasess

In the main body of the text, we estimate the dollar reduction in tax-exclusive Amazon spending following the introduction of the Amazon Tax. In this section, we explore how the probability of shopping at Amazon changes after implementation of the Amazon Tax. The results of our logit specification are found in Table A5. Columns (1) and (2) shows that households do not reduce the likelihood of shopping at Amazon during a given month after implementation of the Amazon Tax for all purchases and purchases under $250, respectively. Finally, Column (3) shows that households reduce the likelihood of making purchases over $250. The coefficient of 0.0023 corresponds to a 16% reduction in the probability of making a purchase of at least $250 after implementation of the Amazon Tax.

Figure A.1: Elasticities as a Function of Purchase Size

This figure presents the elasticities of Amazon shoppers as a function of tax-exclusive purchase size. The elasticities are coefficients of regressions of Amazon purchase amounts on an indicator of treatment state and post-tax. The dependent variable equals the purchase amount if it falls within the bracket being investigated (e.g., \$200 to \$299.99), and zero otherwise. The dashed lines indicate the 95% confidence interval, and standard errors are clustered by state and time. The last bucket (\$700+) includes all of the tax-exclusive transactions that are greater than \$700.

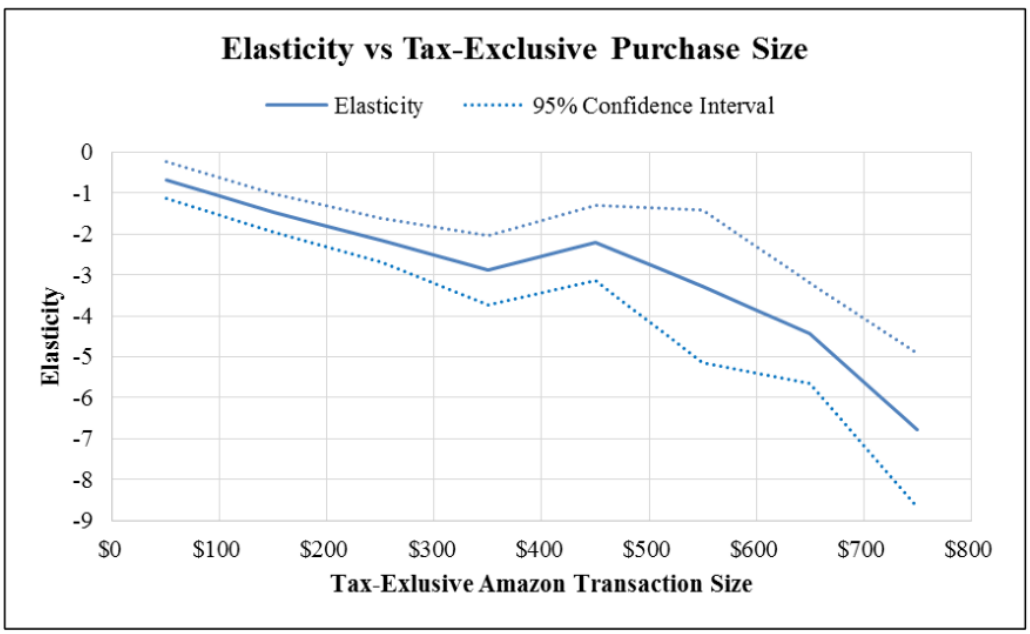


Table A.1: Effect of Amazon Tax on Monthly Amazon Expenditures Using Alternative Methods of Identifying City of Residence (Replication of Table 3.4, Columns (1) to (3))

This table explores the effect of the Amazon Tax on Amazon expenditures for any household that shopped at Amazon at any point in our sample. (The tables in the main body require that the household spent at least $200 at Amazon in 2011.) The unit of observation is the household month, and the dependent variable is the sum of monthly Amazon transactions per household. Columns (1) through (3) evaluate tax-exclusive expenditures, while Columns (4) through (6) evaluate tax-inclusive expenditures. Treated State is an indicator variable that takes a value of 1 for states that implemented an Amazon Tax during our sample period. $I(t \geq Q)_{h,t}$ is an indicator variable that takes a value of 1 for all months after implementation of the Amazon Tax. $I(t = Q_{-1})_{h,t}$, $I(t = Q_0)_{h,t}$, and $I(t \geq Q_{+1})_{h,t}$ are indicator variables for the quarter(s) before, quarter after, and subsequent quarters following the tax implementation, respectively. $TaxRate_h$ is the households sales tax rate. To account for regional differences in cost of living that vary over time, we introduce a time-varying cost of living index at the city-month level, denoted $CostofLivingIndex_{c,t}$. This index is computed by calculating the mean expenditures in the categories of gas, restaurants, groceries, and retail (excluding Amazon purchases) for each city-month. All regressions are ordinary least squares (OLS) regressions and include household and year-month fixed effects. Standard errors are clustered by state and time. t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable:	Amazon spending (tax-exclusive using second-most common city's tax rate)			Amazon spending (tax-exclusive using minimum of first- and second-most common city's tax rates)		
	(1)	(2)	(3)	(4)	(5)	(6)
$TreatedState \times I(t \geq Q)$	-3.735***			-3.617***		
	(-5.22)			(-5.03)		
$reatedState \times I(t = Q - 1)$		1.418***			1.416***	
		(2.85)			(2.85)	
$TreatedState \times I(t = Q0)$		-3.383***			-3.255***	
		(-3.93)			(-3.79)	
$TreatedState \times I(t \geq Q + 1)$		-3.293***			-3.180***	
		(-4.61)			(-4.44)	
$TreatedState \times I(t \geq Q) \times Taxrate$			-55.338***			-53.825***
			(-7.27)			(-6.96)
Household Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
YYYYMM Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Cost of Living Index (City-Month)	Yes	Yes	Yes	Yes	Yes	Yes
Obs	10,436,160	10,436,160	10,436,160	10,436,160	10,436,160	10,436,160
R2	28%	28%	28%	28%	28%	28%
Mean spending of treated	$38.95	$38.95	$38.95	$39.02	$39.02	$39.02
Mean tax rate of treated	7.5%	7.5%	7.5%	7.5%	7.5%	7.5%
Implied Elasticity ($TreatedState \times I(t \geq Q)$)	-1.27			-1.23		
Implied Elasticity ($TreatedState \times I(t \geq Q)$)		0.48			0.48	
Implied Elasticity ($TreatedState \times I(t = Q0)$)		-1.15			-1.11	
Implied Elasticity ($TreatedState \times I(t = Q + 1)$)		-1.12			-1.08	
Implied Elasticity ($TreatedState \times I(tgeQ) \times Taxrate$)			-1.42			-1.38
$TreatedState \times I(t \geq Q)$ / Mean spending	-9.6%			-9.3%		
$TreatedState \times I(t \geq Q)$ / Mean spending		3.6%			3.6%	
$TreatedState \times I(t = Q0)$ / Mean spending		-8.7%			-8.3%	
$TreatedState \times I(t \geq Q + 1)$ / Mean spending		-8.5%			-8.1%	

Table A.2: Effect of Amazon Tax on Monthly Amazon Expenditures after Removing \$200 Spending Requirement in 2011 (Replication Table 3.4)

This table explores the effect of the Amazon Tax on Amazon expenditures for any household that shopped at Amazon at any point in our sample. (The tables in the main body require that the household spent at least \$200 at Amazon in 2011.) The unit of observation is the household month, and the dependent variable is the sum of monthly Amazon transactions per household. Columns (1) through (3) evaluate tax-exclusive expenditures, while Columns (4) through (6) evaluate tax-inclusive expenditures. Treated State is an indicator variable that takes a value of 1 for states that implemented an Amazon Tax during our sample period. $I(t \geq Q)_{h,t}$ is an indicator variable that takes a value of 1 for all months after implementation of the Amazon Tax. $I(t = Q_{-1})_{h,t}$, $I(t = Q_0)_{h,t}$, and $I(t \geq Q_{+1})_{h,t}$ are indicator variables for the quarter(s) before, quarter after, and subsequent quarters following the tax implementation, respectively. $TaxRate_h$ is the households sales tax rate. To account for regional differences in cost of living that vary over time, we introduce a time-varying cost of living index at the city-month level, denoted $CostofLivingIndex_{c,t}$. This index is computed by calculating the mean expenditures in the categories of gas, restaurants, groceries, and retail (excluding Amazon purchases) for each city-month. All regressions are ordinary least squares (OLS) regressions and include household and year-month fixed effects. Standard errors are clustered by state and time. t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Amazon spending (tax-exclusive)			Amazon spending (tax-inclusive)		
(1)	(2)	(3)	(4)	(5)	(6)
-2.665***			-0.868		
(-4.83)			(-1.64)		
	1.101***			0.999**	
	(2.72)			(2.52)	
	-2.381***			-0.605	
	(-3.61)			(-0.99)	
	-2.326***			-0.562	
	(-4.18)			(-1.05)	
		-39.668***			-15.349**
		(-6.58)			(-2.46)
Yes	Yes	Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes	Yes	Yes
17,483,777	17,483,777	17,483,777	17,483,777	17,483,777	17,483,777
28%	28%	28%	28%	28%	28%
\$28.78	\$28.78	\$28.78	\$30.08	\$30.08	\$30.08
7.5%	7.5%	7.5%	7.5%	7.5%	7.5%
-1.23			-0.38		
	0.51			0.44	
	-1.10			-0.27	
	-1.07			-0.25	
		-1.38			-0.51

Table A.3: Alternative Methods of Calculating Elasticities

This table explores alternative methods of estimating the elasticities driven by the Amazon Tax. The unit of observation is the household month. The dependent variable in Column (1) is the tax-exclusive sum of monthly Amazon transactions per household. The dependent variable in Column (2) is the log of 1 plus the tax-exclusive sum of monthly Amazon transactions. Treated State is an indicator variable that takes a value of 1 for states that implemented an Amazon Tax during our sample period. $I(t \geq Q)_{h,t}$ is an indicator variable that takes a value of 1 for all months after implementation of the Amazon Tax. To account for regional differences in cost of living that vary over time, we introduce a time-varying cost of living index at the city-month level, denoted $CostofLivingIndex_{c,t}$. This index is computed by calculating the mean expenditures in the categories of gas, restaurants, groceries, and retail (excluding Amazon purchases) for each city-month. All regressions are ordinary least squares (OLS) regressions and include household and year-month fixed effects. Standard errors are clustered by state and time. t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable:	Amazon spending (tax-exclusive)		log (1+Amazon spending) (tax-exclusive)
	(1)	(2)	(3)
$TreatedState \times I(t \geq Q)$	-3.648***		
	(-5.07)		
$TreatedState \times I(t \geq Q) \times TaxRate$		-54.328***	-0.834***
		(-7.05)	(-3.49)
Household Fixed Effect	Yes	Yes	Yes
YYYYMM Fixed Effect	Yes	Yes	Yes
Cost of Living Index (City-Month)	Yes	Yes	Yes
Obs	10,436,160	10,436,160	10,436,160
R2	28%	28%	34%
Mean spending of treated	$39.00	$39.00	$39.00
Mean tax rate of treated	7.5%	7.5%	7.5%
Implied Elasticity	-1.24	-1.39	-0.83

Table A.4: Elasticities as a Function of Purchase Sizes

This table explores how elasticity varies with purchase size. The dependent variable in Column (1) is the tax-exclusive sum of monthly Amazon transactions per household. The dependent variable in Columns (1) through (8) is the sum of tax-exclusive monthly Amazon transactions for various sized bins. Column (1) corresponds to purchases with tax-exclusive prices of $0-$100, Column (2) corresponds to purchases with tax-exclusive prices of $200-$300, and so on. Treated State is an indicator variable that takes a value of 1 for states that implemented an Amazon Tax during our sample period. $I(t \geq Q)_{h,t}$ is an indicator variable that takes a value of 1 for all months after implementation of the Amazon Tax. To account for regional differences in cost of living that vary over time, we introduce a time-varying cost of living index at the city-month level, denoted Cost of Living Indexc,t. This index is computed by calculating the mean expenditures in the categories of gas, restaurants, groceries, and retail (excluding Amazon purchases) for each city-month. All regressions are ordinary least squares (OLS) regressions and include household and year-month fixed effects. Standard errors are clustered by state and time. t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable:	Tax-Exclusive Amazon Purchase Size in bracket							
	$0.01 -$99.99	$100.00 - $199.99	$200.00 - $299.99	$300.00 - $399.99	$400.00 - $499.99	$500.00 - $599.99	$600.00 - $699.99	$700 and up
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$TreatedState \times I(t \geq Q)$	-1.245***	-0.832***	-0.533***	-0.381***	-0.183***	-0.172***	-0.157***	-1.269***
	(-3.03)	(-6.13)	(-7.73)	(-6.66)	(-4.70)	(-3.44)	(-7.08)	(-7.07)
Household Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YYYYMM Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cost of Living Index	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	10,436,160	10,436,160	10,436,160	10,436,160	10,436,160	10,436,160	10,436,160	10,436,160
R2	38%	11%	6%	5%	5%	4%	3%	6%
Mean spending of treated	$24.22	$7.46	$3.30	$1.76	$1.10	$0.70	$0.47	$2.48
Mean tax rate of treated	7.5%	7.5%	7.5%	7.5%	7.5%	7.5%	7.5%	7.5%
Implied Elasticity	-0.68	-1.48	-2.15	-2.88	-2.22	-3.28	-4.43	-6.79
$TreatedState \times I(t \geq Q)$ / Mean	-5.1%	-11.2%	-16.2%	-21.7%	-16.7%	-24.7%	-33.3%	-51.1%

Table A.5: Effect of the Amazon Tax on the Probability of Amazon Expenditures

This table explores the effect of the Amazon Tax on the probability of Amazon expenditures. The unit of observation is the household month, and the dependent variable is an indicator variable that takes a value of 1 if the household has purchased from Amazon in a given month. Column (1) explores the probability of any Amazon expenditure. Column (2) explores the probability of any Amazon expenditure less than \$250. Column (3) explores the probability of any Amazon expenditure over \$250. Treated State is an indicator variable that takes a value of 1 for states that implemented an Amazon Tax during our sample period. $I(t \geq Q)_{h,t}$ is an indicator variable that takes a value of 1 for all months after implementation of the Amazon Tax. The regression is a logit specification, and standard errors are clustered by state and time. t-statistics are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable:	Tax-Exclusive Amazon $> \$0$	Tax-Exclusive Amazon $< \$250$	Tax-Exclusive Amazon $\geq \$250$
	(1)	(2)	(3)
Treated State $\times$ $I(t \geq Q)$	-0.0038	-0.0015	-0.0023***
	(-0.51)	(-0.22)	(-3.03)
Obs	10,436,160	10,436,160	10,436,160
Mean probability of treated	0.3437	0.3289	0.0148
Treated State $I(t \geq Q)$ / Mean probability	-1.1%	-0.5%	-15.5%