



Financial Literacy and Other Determinants of Consumer Complaint Behavior

An empirical study utilizing the CFPB's complaint database

Jostein Lund Halvorsen & Petter Johnsen Møkkelgård

Supervisor: Nataliya Gerasimova

Master Thesis in Finance

NORWEGIAN SCHOOL OF ECONOMICS

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible – through the approval of this thesis – for the theories and methods used, or results and conclusions drawn in this work.

Preface

This thesis is written as part of our Master of Science in Economics and Business Administration at Norwegian School of Economics (NHH). The thesis is written within the field of Finance.

Working with this thesis has been both instructive and exciting. There has been a number of challenges along the way, collection of relevant data was particularly demanding.

We would like to extend a special thanks to our supervisor, Nataliya Gerasimova. During our work with this thesis, she has been a valuable source of knowledge and constructive criticism. Her feedback has been very helpful, and her guidance has been both useful and pedagogical.

Bergen, December 2018



Jostein Lund Halvorsen



Petter Johnsen Møkkelgård

Abstract

Too many in the United States lack the necessary skills to make sound financial choices. In combination with a wide range of options in the consumer-finance market and the inherent conflict of interest that exists between profit-maximizing financial-services providers and their financially naïve customers, consumer-protection regulation is of great importance. The Consumer Financial Protection Bureau was established subsequent to the Great Recession, and has enabled consumers to submit complaints about unfair, deceptive, or abusive acts or practices by financial services companies.

Several empirical studies have attempted to explain consumer complaint behavior. However, there are few contemporary studies focusing on financial literacy and socio-economic characteristics regarding consumer complaint behavior. Accordingly, the research question of the thesis is:

How do financial literacy and other socio-economic characteristics relate to mortgage and student loan complaints?

The research question is answered using a truncated regression with a lower limit of zero, with and without the cluster command on the individual variable. The data consists of over 218,000 mortgage complaints and 34,000 student loan complaints originated from the Consumer Financial Protection Bureau over the time period 2012-2017. In addition, there are 22 relevant variables from other sources.

The results from the study suggest that areas containing (1) more upscale socio-economic consumers, (2) more highly educated consumers and (3) consumers with more time on their hands complain more frequently. The variable for financial literacy ended up insignificant. However, the insignificant result may owe to the fact that people with higher degree of financial knowledge are better at distinguishing good financial products and services from bad, which may neutralize the tendency to complain.

Contents

PREFACE	2
ABSTRACT	3
1. INTRODUCTION.....	6
1.1 PRACTICAL BACKGROUND AND MOTIVATION	6
1.2 STRUCTURE.....	7
2. LITERATURE REVIEW	8
2.1 STUDIES ON CONSUMER COMPLAINT BEHAVIOR	8
2.2 STUDIES ON DETERMINANTS OF CUSTOMER SATISFACTION	9
2.3 STUDIES ON BEHAVIORAL ECONOMICS, CONSUMER PROTECTION AND FINANCIAL LITERACY	10
2.4 STUDIES ON CONSUMER TYPES AND COMPLAINT BEHAVIOR	11
2.5 SUMMARY LITERATURE REVIEW.....	12
3. ABOUT THE CONSUMER FINANCIAL PROTECTION BUREAU	13
3.1 HISTORY OF THE CONSUMER FINANCIAL PROTECTION BUREAU.....	13
3.2 TODAY	15
4. DATA	16
4.1 MORTGAGE COMPLAINTS	16
4.2 STUDENT LOAN COMPLAINTS	21
4.2.1 <i>Modifications</i>	21
4.3 SOURCES	23
5. METHODOLOGY	25
5.1 PANEL DATA ESTIMATION METHODS.....	25
5.2 ESTIMATION METHOD	26
5.2.1 <i>Mortgage Complaints Analysis</i>	27
5.2.2 <i>Student Loan Complaints Analysis</i>	30

6.	ANALYSIS.....	33
6.1	ROBUSTNESS OF OUR ESTIMATION MODELS	34
6.1.1	<i>Mortgage Complaints</i>	<i>34</i>
6.1.2	<i>Student Loan Complaints.....</i>	<i>37</i>
6.2	HYPOTHESES	40
6.3	EMPIRICAL RESULTS - MORTGAGE COMPLAINTS	41
6.4	EMPIRICAL RESULTS - STUDENT LOAN COMPLAINTS	47
6.5	OTHER RESULTS.....	50
6.5.1	<i>Regression Without Cluster Command.....</i>	<i>50</i>
6.5.2	<i>Regression with Outlier</i>	<i>51</i>
6.5.3	<i>One-year Regression.....</i>	<i>51</i>
6.5.4	<i>First-difference Regression.....</i>	<i>51</i>
7.	CONCLUSION	53
7.1	SUMMARY.....	53
7.2	MAIN FINDINGS AND CONCLUDING REMARKS	54
7.3	LIMITATIONS	55
7.4	RECOMMENDATION FOR FUTURE RESEARCH.....	57
8.	REFERENCES	58
	APPENDIX	65

1. Introduction

A resilient consumer finance marketplace is a key feature for any economy in the world. A lot of governments safeguard its marketplace with a regulatory agency charged with overseeing financial products and services that are offered to consumers. The Consumer Financial Protection Bureau in the United States is an example of such an agency. CFPB has collected consumer complaints since 2011. This thesis analyses the publicly available consumer complaint database (Consumer Financial Protection Bureau, 2018). First, we present some background information, some existing literature and the motivation for examining the research question. Secondly, we present the structure of the thesis.

1.1 Practical Background and Motivation

Subsequent to the Great Recession, there has been implemented several changes in the regulatory system in the United States. One of these is the Consumer Financial Protection Bureau. The overall aim of the CFPB is to facilitate the development of the consumer finance marketplace. This includes ensuring consumer rights, preventing financial harm to consumers and educating and empowering consumers to live better lives (Consumer Financial Protection Bureau, 2018).

Since the Consumer Financial Protection Bureau was established, it has enabled consumers in the United States to submit complaints about “unfair, deceptive, or abusive acts or practices” by financial-services companies. The independent consumer protection agency’s creation was authorized by the Dodd-Frank Wall Street Reform and Consumer Protection Act, whose passage in 2010 was a legislative response to the financial crisis of 2007-08 (Consumer Financial Protection Bureau, 2018).

The CFPB is a regulatory agency responsible for consumer protection in the financial sector. Its main role is to protect and educate consumers about the various types of financial products and services (Investopedia, 2018). According to a study by Day & Bodur (1978), consumers tend to be quite dissatisfied with financial services. In combination with the fact that the consumer-finance market is characterized by a wide range of choices and the inherent conflict of interest that exists between profit-maximizing financial-services providers and their financially naïve customers, consumer-protection regulation is of great importance (Emmons, 2005).

According to a large-scale review of financial literacy in the United States by Brookings, the levels of financial literacy, especially concerning the younger generations, are persistently low (Kasman, Heuberger, & Hammond, 2018). These findings are supported by a report from Champlain College which says that too many adults in the United States lack the necessary skills to make good choices when it comes to saving for the future, buying a car or a home and planning for retirement (Pelletier, 2016). Thus, it is important to raise the average level of financial literacy in the United States. With all the complaints the CFPB receives, it enables them to improve legislation and regulations as well as to unveil practices perceived as consumer-unfriendly and potential regulatory violations that it can target both with new guidelines and public enlightenment.

The data is publicly available because the Bureau encourages the public and other interested parties to analyze, augment and build on the database. This thesis is in keeping with that encouragement. There are many studies on customer satisfaction and consumer complaint behavior. However, the amount of contemporary studies on financial literacy and socio-economic characteristics regarding consumer complaint behavior seem to be scarce. To our knowledge there has only been one contemporary study related to this particular topic. Since the mortgage complaint data are comprehensive and students are a particular focus of the CFPB, our research question are as follows:

How do financial literacy and other socio-economic characteristics relate to mortgage and student loan complaints?

1.2 Structure

The thesis has six sections. Section 1 introduce the topic as well as presenting the research question. Section 2 reviews the existing literature. Section 3 describes the Consumer Financial Protection Bureau's history, its responsibilities and some challenges ahead. Section 4 presents and discusses the collected data set. Section 5 describes our chosen empirical methodology. Section 6 presents test diagnostics of the estimation model and discuss the findings. The principal analysis concerns mortgage complaints, whereas the student loans analysis is secondary. Section 7 summarize the most important findings, discuss the limitations of the sample and makes recommendations of future research on the topic.

2. Literature Review

In this section we will discuss relevant research on similar topics, as well as an attempt to motivate how this thesis fits into the existing literature. There are many studies on customer satisfaction and consumer complaints behavior (CCB). In contrast, the amount of contemporary studies on socio-economic characteristics regarding consumer complaint behavior is limited. The section will examine the existing research and generate a baseline of results for which the findings of this study will be compared to. To our knowledge the study by Ayres, Lingwall, & Steinway (2013) is the only contemporary study related to this particular topic. Moreover, studies utilizing financial literacy data seem to be scarce. Hence, it would be interesting to further investigate and build on that study with a new approach accompanied by new insight and more comprehensive data.

2.1 Studies on Consumer Complaint Behavior

Albert O. Hirschman was a pioneer within the research regarding consumer complaint behavior. A vast majority of succeeding literature utilizes his findings and model as a starting point. Hirschman (1970) includes a basic distinction between alternative ways of consumer behaviors. *Exit* is for consumers who switches to a competing product, *loyalty* is for consumers who tolerates and/or forget and *voice* is for consumers who are agitated and wish to exert influence for change.

He also argues that “if the exit option was blocked or unavailable, complaining would be the only way the dissatisfied consumer could react”. Hence, economic theory predicts that there should be more complaints in competitive markets, such as in the financial markets, rather than in markets resembling the notion of monopoly. However, *voice* responses depend on several other factors. As Hirschman claims, the probability of *voice* depends on the value gained from a successful complaint times the probability of achieving a successful outcome Hirschman (1970). The relative power between a buyer and the seller is therefore a critical measure, which has changed after the implementation of the CFPB.

Although the study by Hirschman is quite old, it gained strong empirical support for its validity regarding classification of consumer complaint behavior responses in a study by Maute & Forrester Jr. (1993). They highlighted that it clearly distinguishes between active and passive dissatisfaction responses.

Furthermore, Day & Landon Jr. (1976) build on the findings introduced by Hirschman (1970). They propose the following structure of consumer behavioral responses: (1) do nothing at all, i.e., make no behavioral response; (2) take some “private action” by modifying one’s own behavior or seeking to influence the behavior of family and friends; and (3) take some “public action” such as contacting business firms, consumer organizations, or governmental agencies Day & Landon Jr. (1976).

Moreover, as people are very different and will react differently, Day & Landon introduce the term *propensity to complain*, to describe different patterns of complaining behavior. They suggest a function of four factors: (1) the individual's propensity to complain when dissatisfied; (2) the individual's opportunities to become dissatisfied with products or services; (3) the opportunities available to the individual to obtain redress and/or register complaints; and (4) disparity in consumer knowledge. Landon Jr. offers the following definition of a consumer complaint: “An expression of dissatisfaction on a consumer's behalf to a responsible party” (1980).

2.2 Studies on Determinants of Customer Satisfaction

There are many determinants of customer satisfaction and the causes are numerous. According to Oliver (1977), the two biggest determinants of consumer satisfaction are expectations and disconfirmation. The study by Oliver is supported by the findings of Bearden & Teel (1983), which suggest that both expectations and disconfirmation are related positively to satisfaction. And as for a definition of the two Moore & Shuptrine suggest:

Expectations are conceptualized as predictions of product performance when consumed. Disconfirmation is based on the process of comparing perceived product performance with expectations and if performance meets, goes beyond, or falls short of expectations (1984).

On the contrary, Churchill Jr. & Surprenant (1982) argue that since nondurable goods had been used in a majority of previous studies, it could perhaps be other determinants which applies to durable goods. They found that for durable products performance differences are the major determinant of satisfaction, and conversely that the disconfirmation of initial expectations has little impact.

Concerning the customer satisfaction of financial services, Day & Bodur (1978) discovered that consumers in general are quite dissatisfied. The reasons for dissatisfaction were mostly related directly to the quality of the supplier's performance, which could involve services provided in a careless or unprofessional manner, which serve as a strong linkage for overall customer satisfaction (Anderson & Sullivan, 1993). Furthermore, Levesque & McDougall (1996) reveal that satisfaction with problem recovery and complaint handling is also a key determinant for customer satisfaction.

Regarding customer complaint handling for financial services, the switching cost in the market can be quite high, especially in retail banking, and consequently a financial service provider should focus on its customer complaint handling. According to Hart, Heskett, & Sasser Jr. (1990), anecdotal evidence suggests that when service providers accept responsibility and resolve the problem, the customer becomes "bonded" to the organization. This can lead to increased customer loyalty and retention, which moreover can have substantial impact on profits.

2.3 Studies on Behavioral Economics, Consumer Protection and Financial Literacy

According to the research within the field of behavioral economics, people are impatient and partially irrational. The concept of self-control, presented by Thaler & Shefrin (1981), and the present-bias preference, are often dominant factors. These refer to the tendency of people to give stronger weight to payoffs that are closer to the present when considering trade-offs between times of consumption. Formally, present-biased preferences can be seen as the result of the interplay of two separate decision-making systems: the affective system, which values immediate gratification and sharply discounts all future periods; and the deliberative system, which makes long-run plans and displays higher discount factors (O'Donoghue & Rabin, 1999). Meier & Sprenger (2010) provide evidence that present-biased individuals are more likely to have significantly higher amounts of credit card debt, controlling for disposable income, other socio-demographics, and credit constraint.

In addition to low levels of financial literacy (Kasman, Heuberger, & Hammond, 2018), the financial markets around the world have become increasingly accessible, as new products and financial services grow widespread. Alternative services, including payday loans, pawn shops, tax refund loans, and rent-to-own shops, have increased a lot in size and usage. The

combination between the growing autonomy and financial complexity, comes with great challenges. According to Lusardi & Mitchell (2014), all of these services are difficult to master for individuals with a lower degree of financial knowledge. As an example, Moore D. L. (2003) reported that the least financially literate were more likely to have costly mortgages.

In view of the foregoing, Rutledge (2010) argues why consumer protection is key to any market. Consumer protection ensures that consumers receive information *ex ante* that will exert influence towards better decisions, warn against unfair and deceptive practices and provide recourse mechanisms to resolve disputes. Furthermore, she argues that financial literacy initiatives give consumers the knowledge, skills and confidence to understand the information they receive and evaluate the risks and rewards of financial services and product.

2.4 Studies on Consumer Types and Complaint Behavior

There have also been some studies concerning the disparity in complaint behavior with regards to consumer types. According to Stokes (1974), involving the analysis of complaint letters indicated that they were heavily weighted by two groups: (1) people with time on their hands; and (2) highly educated, articulate people. Warland, Herrmann, & Willits (1975) studied the differences in the magnitude of complaining on three consumer types with respect to socio-economic characteristics. They found that upscale socio-economic people tended to complain more frequently. They were better educated, earned higher incomes, were more active in formal organizations, politically committed and liberal. They were in general younger and owned more stocks and bonds. They also found that those who complained more frequently than others were more interested in consumerism and were in favor of more consumer protection. These findings are supported by the results from a Norwegian study by Gronhaug (1977), which additionally concludes that those with a high education level and living close to a consumer agency make more use of their consumer representatives to pursue their interests than others did.

Further, Liu & McClure found that when dissatisfied, consumers from collectivistic cultures, often found in Asian countries, are “less likely to engage in voice behavior ... than those in an individualistic culture” (2001).

To our knowledge the study by Ayres, Lingwall, & Steinway (2013) is the only study of the Consumer Financial Protection Bureau’s consumer complaints. They found that there were significant increases in mortgage complaints in populations with a higher proportion of

Blacks, Hispanics, as well as an increase in untimeliness of responses from financial institutions towards senior citizens and college students.

Finally, a study by Andreassen & Streukens (2013) has shown that the ability to complain online increases the likelihood that consumers will take action towards a provider.

2.5 Summary Literature Review

In general, the results from the existing literature suggest that the same characteristics tend to affect consumer complaint behavior. However, most of the research is from the 70's, and potentially outdated. To our knowledge there has only been one contemporary study related to our particular research question. Moreover, studies utilizing financial literacy data seem to be scarce. Hence, it would be interesting to extend the study by Ayres, Lingwall, & Steinway (2013) with a new approach accompanied by new insight and more comprehensive data.

3. About the Consumer Financial Protection Bureau

The Consumer Financial Protection Bureau (CFPB) is a regulatory agency responsible for consumer protection in the financial sector. Consequently, its main role is to protect and educate consumers about the various types of financial products and services. Specifically, the CFPB helps consumer finance markets work more efficiently by providing rules, enforcing those rules and empowering consumers to take control of their personal financial lives. The CFPB works to educate and inform consumers against abusive financial practices, and to study data to better understand consumers and the financial markets they participate in (Investopedia, 2018).

The mission of the federal CFPB is to make the financial markets for consumer products and services work for Americans. This concerns applying for mortgages, choosing student loans, or any other consumer financial products. These and other measures are in order to empower individuals to live better lives (Consumer Financial Protection Bureau, 2018).

3.1 History of the Consumer Financial Protection Bureau

The initiative behind CFPB started in 2007. The United States faced the most severe financial crisis since the Great Depression. Millions of Americans lost vast amounts of savings, saw their home values shrink, their jobs eliminated, their businesses lose financing and countless of consumer loans went into default (Consumer Financial Protection Bureau, 2018).

After the world wars, the American society had rising wages and growth in savings. American families tended to incorporate moderate amounts of debt in this era. However, with the emergence of stagnating wages in the 70's, combined with rising expenses for housing and public services, a lot of families were pushed into debt. Simultaneously, Americans saw a significant increase in financial liberalization with credit substantially more available. Following this development, a lot of the old rules regulating the credit market became antiquated or disappeared. In the 2000s, there were widespread failures in consumer protection resulting in a rapid growth in irresponsible lending practices. Many lenders exploited the liberalized financial market and developed financial products and sold mortgages and other products that were overly complicated (World Bank, 2005).

This left many Americans with loans they did not fully understand and most importantly could not afford. Millions of Americans who behaved responsibly were lured into expensive loans

by deceptive promises of low payments. Competing in the market for honest lenders, which resisted the pressure of selling complex financial products, became challenging (Consumer Financial Protection Bureau, 2018).

The worst consequence of these vastly complicated structures of the securitized credit market was that even for those who avoided the temptations of excessively risky credit were caught in the same interconnected system. Even those who never took out an unaffordable mortgage, saw the values of their homes plummet when homes in the same area foreclosed. Additionally, those who had saved regularly saw their retirement funds lose significant value and saw their states cut back on important public services to make up for their own revenue losses. The costs of the crisis, with its origin in the irresponsible lending, were carried by the American families (Cappeli, Barankay, & Lewin, 2018).

After his election, President Obama incorporated elements to his economic agenda to address failures of consumer protection by establishing a new financial agency to focus directly on the consumers, rather than on bank safety or on monetary policy. In June of 2009, Obama included a proposal for a new consumer financial agency in his comprehensive financial reform plan. In particular, his goal of the agency was to “protect consumers and investors from financial abuse” and “reduce gaps in federal supervision and enforcement; improve coordination with the states; set higher standards for financial intermediaries; and promote consistent regulation of similar products” (Ayres, Lingwall, & Steinway, 2013).

The President urged Congress to give the consumer agency the same accountability and independence that other banking agencies had, and adequate funding that it could ensure that influential financial companies would comply with consumer laws. Although the financial industry executives lobbied heavily against the creation of a new regulatory agency (Andrews, 2009), the idea got legislative interest. After some modifications to garner Republican support and enable passage, the Congress passed, and President Obama signed the Dodd-Frank Wall Street Reform and Consumer Protection Act in July 2010. The Act established the Consumer Financial Protection Bureau under the Title X of Dodd-Frank as an independent executive agency with responsibility for regulating the market for financial products or services according to federal consumer financial laws (Consumer Financial Protection Bureau, 2018).

3.2 Today

The CFPB operates approximately on a \$663 million budget (Consumer Financial Protection Bureau, 2017), where ten percent is allocated to consumer response operations, which mainly consist of formal complaints processes. In addition to assisting consumers with specific complaints and aggregating complaint data, the CFPB also aims to educate consumers, study consumer behavior, supervise large financial companies and enforce federal consumer financial protection laws (Consumer Financial Protection Bureau, 2018).

The purpose of the complaints process is three-fold: to assist individual consumers with specific complaints; to focus the Bureau's enforcement and regulatory efforts on specific companies and general "business practices that may pose risks to consumers" based on aggregate consumer concerns; and, by making the data publicly available, to provide the financial services industry with a high-level view of what matters to consumers and to provide customers with a view into how companies are meeting those needs (Ayres, Lingwall, & Steinway, 2013).

In July 2011, the Bureau began accepting consumer complaints on its first day of operations. The Consumer Response division has since expanded to a lot of financial products. Since it opened, the Bureau has received over 1,150,000 consumer complaints. After only two years, the Bureau had received over 100,000 consumer complaints, where approximately 46% were submitted through the website and 34% were referrals from other state and federal agencies. In 2018 over 90% originated from the same sources, and only 7% from referrals. This underlines how CFPB has solidified its position as the one inter-agency regulatory interface. In addition, it is believable that consumers would have been confused prior to the creation of CFPB, as there were at least twelve federal agencies responsible for consumer financial protection (Consumer Financial Protection Bureau, 2018).

Despite the positive trend in consumer financial protection, there are forces that are fighting for rollbacks of the financial industry regulations. Since February 3rd, 2017, the Trump administration has been advocating changes, and on May 22nd, 2018, Trump signed two executive orders which approved the first big rollback of the Dodd-Frank Act. Later this year, regulators are expected to release a plan to dilute the Volcker Rule, which bans bank from making risky bets with depositors' money (Amadeo, 2018).

4. Data

In this section, we will discuss the structure of the collected data, go through the sources of the data and present a descriptive analysis. As mentioned in the introduction, the principal analysis concerns the mortgage complaints analysis. Thus, student loan specific data will be presented and discussed secondary.

4.1 Mortgage Complaints

The structure of the data set is panel data with time series, where Year is the time variable and State is the individual variable. Since the first available data originates from 2012, we have collected data from 2012 to 2017. Hence, the data set has $T=6$ and $N=300$. Figure A.1 in Appendix A illustrates the data set.

Moreover, we have structured the data set in five groups: main variables, education variables, state program variables, demographic variables, and supplemental variables. The data collected are in accordance with our research questions, and the hypotheses introduced in the next section.

Dependent Variable

We analyze the relationship between mortgage complaints and financial literacy and other socio-economic characteristics, hence the dependent variable is *Mortgage complaints*. As the population size differ from state to state, we divide the total number of complaints in each state by the total number of mortgages in each state. We extracted the complaints data from the CFPB (Consumer Financial Protection Bureau, 2018), while population and total mortgages is collected from the *American FactFinder* (United States Census Bureau, 2018). The mortgage complaints consist of all kinds of mortgages; conventional fixed mortgages, FHA mortgages, conventional adjusted rate mortgages, reverse mortgages and other mortgages.

Main Variables

This group consists of the main explaining variables concerning mortgage complaints: *Final grade*, *Unemployment rate*, *Debt-to-income* and *Families below poverty level*. Final grade is a scaled grade, from A+ to F, and describe the level of financial literacy in the given state. The grades are given based on a score ranging from 0 to 100, where a score below 60 is an F. Final

grade is a grade based on five factors: financial knowledge, total credit, saving and spending, retirement readiness, and protect and insure. The financial literacy grades are measured using 59 state specific data points, extracted from 18 different organizations (Pelletier, 2016). *Final grade* is collected from a report written by John Pelletier at the Center of Financial Literacy at Champlain College and has only been conducted once. Hence, *Final grade* is time-constant, and implications will be discussed in the limitations of the analysis.

Further, *Unemployment rate* is the annual unemployment rate for the given state in that given year. We collected this variable from United States Census Bureau, using their application called *American FactFinder*. The data in *American FactFinder* is the result of several censuses and surveys. United States Census Bureau is the organization that produces and disseminates the official estimates. (United States Census Bureau, 2018)

Moreover, the variable *Debt-to-income* is the median debt-to-income ratio in a state. This data is collected from the BLS (Bureau of Labor Statistics, 2018), where the debt is the median household debt and the income is the median household income.

We have included the variable *Families below poverty level*. These are percentages of a state's population, which represent the number of households below a poverty threshold each year. Census uses several thresholds based on annual income to determine poverty, and where the thresholds varies by family size and composition. For example, a family of four with two kids has a lower poverty threshold than a family of five with three kids (United States Census Bureau, 2018). The thresholds do not vary geographically.

Education Variables

The group of education variables consists of the variables *Bachelor's degree*, *Business school*, and *Legal occupation*. *Bachelor's degree* and *Business school* are collected from the *American FactFinder* (United States Census Bureau, 2018) while *Legal occupation* is collected from the United States BLS (Bureau of Labor Statistics, 2018). *Legal occupation* and *Business school* are in total numbers, while *Bachelor's degree* is in percentages. *Bachelor's degree* is an education level variable, describing the percentage of the population in a state with a bachelor's degree. Moreover, *Business school* is the total number of people with a bachelor's degree in business. The last education variable, *Legal occupation*, represents how many people that work within a legal field. This variable is, like *Final grade*, time-constant.

State Program Variables

State program variables consist of the variables *10-year lagged economic education* and *10-year lagged personal finance education*. These two variables are programs initiated by the state to improve financial literacy. We have scaled these variables from one to five, based on which standards the different states have implemented. There are five standards:

1. States that include economics/personal finance in their K-12 standards
2. States that require standards to be implemented
3. States where a high school course is required to be offered
4. States where a high school course is required to be taken
5. States with standardized testing of economic/personal finance concepts.

This data is collected by the Council of Economic Education (CEE) and published semi-annually. The CEE's mission is to teach K-12 students about economics and personal finance, and have been doing so for nearly 70 years. The CEE's goal is to reach every child in every district and school so that they can make better decisions for themselves, their families and their communities (Council for Economic Education , 2016).

Demographic Variables

This group of variables consists of different ethnicities in each state. We have included the following in the analysis: *African American*, *Indian or Alaskan*, *Asian*, *Native Hawaiian* and *White*. These are percentages of the total population in each state. These variables are collected from the *American FactFinder* (United States Census Bureau, 2018).

Supplemental Variables

In the fifth group, we have included supplemental variables which consist of *Internet users above 15 years* and *Median income*. Because of the relatedness with *Families below poverty level*, which we will discuss later, we have included *Median income* in this group to control for the income differences within states. Median income data is collected from the *American FactFinder* (United States Census Bureau, 2018).

Further, the variable *Internet users above 15 years* represents the percentage of the population above 15 years in a state with accessibility to the internet. This data is collected from NTIA (National Telecommunications and Information Administration, 2018).

Descriptive Analysis

Table 4.1 presents number of observations, mean, standard deviation, minimum and maximum value for the variables ranging from 2012 to 2017.

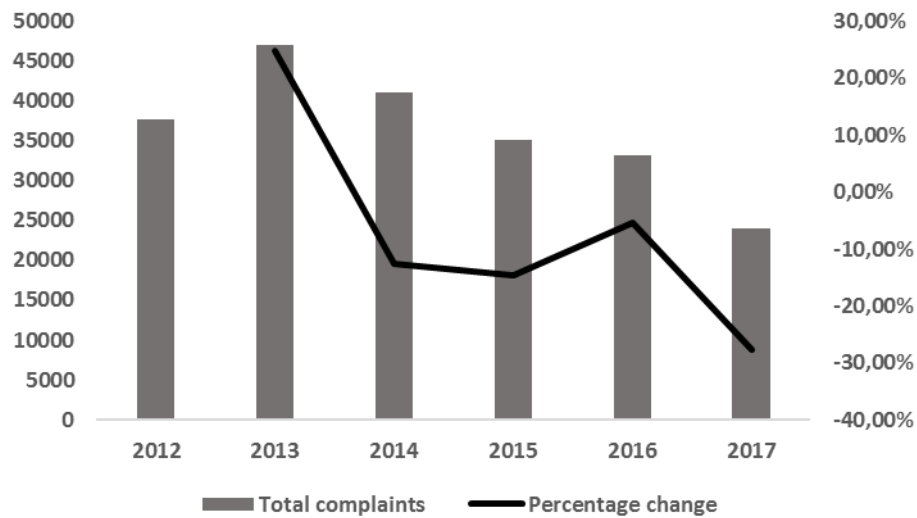
Table 4.1: Descriptive Statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
Total mortgage complaints	300	727.33	1139.932	10	8913
Mortgage complaints	300	0.000616	0.0003851	0.0000915	0.0030349
Final grade	300	77.37	8.288371	61	91
Unemployment rate	300	0.0634	0.0183709	0.024	0.112
Families below poverty level	300	0.10447	0.0272462	0.053	0.178
Debt-to-income	300	1.43075	0.3165555	0.74	43102
Business school	300	266699.9	305405	14173	1645595
Bachelor's degree	300	0.2975633	0.0528467	0.186	0.567
Legal occupation	300	6852.96	1666.457	4305	12066
10-year lagged economic education	300	31444	1.467735	0	5
10-year lagged personal finance	300	22647	1.415017	0	5
White	300	0.771227	0.1266915	0.2493244	0.9519818
African American	300	0.1050327	0.0950457	0.0031831	0.3801676
Indian or Alaskan	300	0.0163298	0.0284379	0.0000819	0.1596497
Asian	300	0.0479371	0.0710936	0.0008706	0.4183778
Native Hawaiian	300	0.0035662	0.0140369	0	0.1044019
Internet users above 15 years old	300	0.7709667	0.0486067	0.65	0.87
Median income	300	56300.67	9433.375	32338	81084

The gaps between the variables represent the different groups the variables are divided into: main variables, education variables, state program variables, demographic variables and supplemental variables. The big difference between the minimum and maximum value in *Mortgage complaints*, might owe to the fact that there are significantly more complaints in states where there are more issued mortgages. Hence, we divide the number of complaints by the total number of mortgages per state. If we look at the independent variables, we can see that there is a large difference in unemployment rate in the different states. The maximum value is 11,2% while the lowest is 2,4%. This could indicate two things; there has been drastic changes in the labor force between 2012 to 2017, and/or there are substantial differences in the different states. It is the same case with the variable *Families below poverty level*, with a minimum value of 5,3% and maximum value of 17,8%.

The graph below shows the development of the total mortgage complaints in the United States. There is a decreasing trend in the total number of complaints. The highest number of complaints was in 2013 with 47,041 complaints. 2017 had the largest decrease with a 27% fall in complaints.

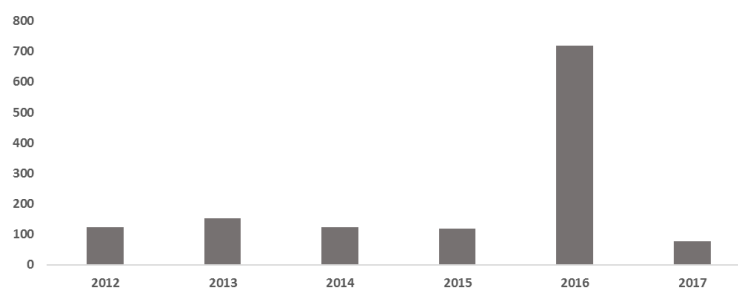
Figure 4.1: Development in Mortgage Complaints



Outliers in the Data Set

An outlier in the data set is an observation that is distant from the other observations (Grubbs, 1969). An outlier may be a result from variability in the measurement, an experimental error, or maybe just an extreme case. Most of the parameters used in statistics like means, standard deviations, correlations and other statistics based on these, are highly sensitive to outliers in the data set. Since the assumptions in linear regression are based on these statistics, outliers can affect and distort the analysis (Grace-Martin, 2018). Hence, we need to control for the outliers, and see if our results differ a lot from the untrimmed to the trimmed model. There are several ways to deal with outliers; remove the outliers, remove 1% from both sides of the extreme values, or replace the outlier with the average of the rest of the group the outlier belongs to. We choose the latter. There is one outlier in the data, specifically the number of mortgage complaints in Maine, 2016.

Table 4.2: Number of Mortgage Complaints in Maine



As illustrated in table 4.2, Maine had an average of 130 complaints per year before 2016, when suddenly the reported complaints amounted to 719. In 2017, the reported amount of complaints was 78. We believe there is some kind of measurement error that have led to this outlier and have therefore decided to replace the outlier with the average number of complaints before 2016. We will compare the results from the regressions with and without the outlier later.

4.2 Student Loan Complaints

In this subsection we will present the data regarding the secondary analysis. The structure of the data set is similar to the data set in the analysis of mortgage complaints. The data set is panel data with time series ranging from 2012 to 2017. Moreover, we will go through changes from the mortgage analysis data set and present a descriptive analysis of the new variables.

4.2.1 Modifications

In this data set, we choose to exclude the education variables *Bachelor's degree* and *Business school*, since students are working towards such accomplishments. Further, the variables are organized into four groups; main variables, state program variables, demographic variables, and supplemental variables. The main difference from the previous model, is that we add the variables *Public school* and *Private school* in the supplemental variables, and that we include the variable *Legal occupation* in the main variables group. The variables *Public school* and *Private school* is a percentage of the population between the ages of 3 to 17 which attend a public or a private school. The variables are extracted from the *American FactFinder* and are negatively correlated to the extreme since the proportion that does not go to public school, goes to private school and vice versa (United States Census Bureau, 2018).

Dependent Variable

We want to analyze the relationship between student loan complaints and financial literacy and other socio-economic characteristics, hence the dependent variable is *Student loan complaints*. As the population size differ between states, we control for this by dividing total number of complaints per state by the total number of student loans. Unfortunately, in contrast with *Mortgage complaints*, we could not find a total number of student loans. However, according to Hess (2017), approximately 70% of all students in the United States have student loans. Hence, we made a proxy for the number of student loans by multiplying the number of students in each state with 70%. We extracted the complaints data from the CFPB (Consumer

Financial Protection Bureau, 2018), while population and total students is collected from the *American FactFinder* (United States Census Bureau, 2018). The student loan complaints consist of both federal and private loans.

Descriptive Analysis

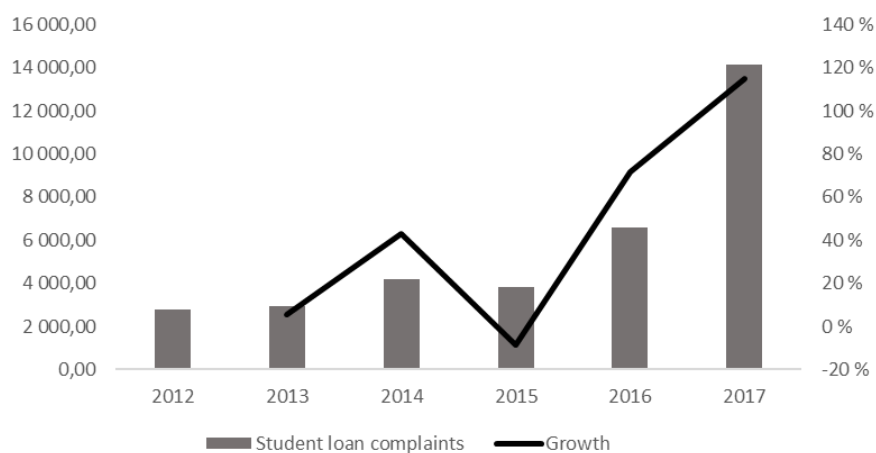
Table 4.3 presents number of observations, mean, standard deviation, minimum and maximum values for the variables ranging from 2012 to 2017. Compared to the analysis of mortgage complaints, there is a smaller amount of observations in student loan complaints. At maximum, there were 1,333 student complaints in 2017 in California. The mean is 115 complaints per year, which indicates that the maximum case was an extreme occurrence.

Table 4.3: Descriptive Statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
Student loan complaints	300	0.0002	0.0002	0.00002	0.0017
Total student loan complaints	300	115	172	1	1333
Public school	300	0.87	0.030	0.78	0.92
Private school	300	0.13	0.03	0.08	0.23

In contrast to mortgage complaints, number of complaints on student loans has, as illustrated in figure 4.2, skyrocketed the last five years. From 2012 to 2017, the annual number of student loan complaints has increased from 2,779 to 14,151, with an average growth rate of 45%.

Figure 4.2: Development in Student Loan Complaints



4.3 Sources

The data are collected from several sources: United States Census Bureau (CB), Consumer Financial Protection Bureau (CFPB), Champlain College, National Telecommunications and Information Administration in United States (NTIA), Bureau of Labor Statistics (BLS), Champlain's Center for Financial Literacy (CFL) and the Council for Economic Education (CEE).

The Consumer Financial Protection Bureau regulates the offering and provision of consumer financial products and services, under the federal consumer financial laws. In addition to assist consumers, they also assist with specific complaints. The CFPB also aims to educate consumers, research consumer behavior, supervise large financial companies and enforce federal consumer financial protection laws. As previously mentioned, CFPB has made the aggregated complaint data publicly accessible. The data concerning mortgage complaints and student loan complaints, are extracted from this database and spans complaint data from 2012 to 2017 in each state (Consumer Financial Protection Bureau, 2018).

Census Bureau is an American institute, providing facts and statistics about people, places and economy in America. Census Bureau is a part of United States Department of Commerce and employs about 4,285 staff members. Census Bureau collects data through surveys, in addition to directly conversation with businesses. However, their primary sources of data are federal, state, commercial entities and local governments. The entire information the bureau collects through surveys is confidential and protected by federal law. The majority of the collected data originates from Census Bureau. Census Bureau supplies data within education, demographics and basic count and estimates of population (United States Census Bureau, 2018).

Similar to Census, National Telecommunication and Information Administration is also a part of United States Department of Commerce. NTIA is the executive branch agency that is responsible by law for advising the President on telecommunications and information policy issues (National Telecommunications and Information Administration, 2018). From this source, we collected the variable *Internet users above 15*.

Moreover, The Bureau of Labor Statistics of the United States Department of Labor is the principal federal agency responsible for measuring labor market activity, working conditions, and other changes in the economy (Bureau of Labor Statistics, 2018). BLS provides products,

data and services that are accurate, objective, relevant, timely and accessible. We collected the variables *Legal occupation* and *Debt-to-income* from this source.

The Council for Economic Education is an organization, founded by business leaders in the United States in a collaboration with the President's council of economic advisors. The CEE's mission is to teach K-12 students about economics and personal finance, to make students able to undertake better decisions for themselves, their families and their communities. The institution reaches over 55,000 K-12 teachers a year, which again reach out to more than five million K-12 students throughout the whole country. In 2014, CEE started to provide information and data about the level of financial literacy and standards in high school. CEE obtain first-hand information through their business, hence we believe that CEE is a reliable source and that their data are suitable for analytical purposes (Council for Economic Education, 2016).

Finally, we used the report *Adult Financial Literacy* from Champlain College. The report is written by the Center of Financial Literacy at Champlain College. From this report, we collected the variable *Final grade* for each state in 2016. Champlain's Center for Financial Literacy was established in 2010 and was designed to promote and develop financial literacy skills in K-12 student, college students, K-12 teachers, and adults (Pelletier, 2016). The mission is to lead adults and youngsters to more sound decisions concerning spending, credit, debt, investments and other complex financial situations.

5. Methodology

In this section, we will present the methodical framework and discuss our hypotheses based on the literature presented in section 2. Additionally, we will discuss the choices made concerning our approach to answer our research question, which is important to improve the quality of our results.

5.1 Panel Data Estimation Methods

Traditionally, there are three estimation methods for panel data that combines cross-sectional and time-series data (Johnstone & DiNardo, 1997). We will go through these three common estimation methods, and discuss the chosen method. The three different methods are pooled OLS, fixed effects model, and random effects model.

Pooled OLS

Pooled OLS is, according to Johnstone and DiNardo, the simplest estimation method, which proceeds by essentially ignoring the panel structure of the data. The estimation of this model is straightforward and contains the same assumptions as the regular linear OLS. The assumptions are that for a given individual, observations are serially uncorrelated; across individuals and time, the errors are homoscedastic (Johnstone & DiNardo, 1997). The model for pooled OLS is as following:

$$Y = X\beta + \varepsilon \quad (1)$$

where we now assume that $\varepsilon_{it} \sim \text{iid}(0, \sigma^2)$ for all i and t .

The Random Effects Model

Random effects model (RE) has almost the same structure as the pooled OLS model, except that it takes the structure of the data set into account.

$$Y_{it} = X_{it}\beta_{it} + \varepsilon_{it} \quad (2)$$

error term:

$$\varepsilon_{it} = \alpha_{it} + \eta_{it} \quad (3)$$

The first term in the error term, α_{it} , is called the individual effect, and varies across the cross-section unit, but is constant over time. This is the state variable in the data set. The second term, η_{it} , varies unsystematically across time and individuals. The random effects model has

an important assumption regarding the individual effect, namely that α_{it} is uncorrelated with X_{it} . It assumes that the error term is not correlated with any of the regressors and that the error variance estimates are specific to the individual variable. One can use generalized least squares (GLS) or OLS estimator to estimate the random effects model. Models with random effects help to control for the unobserved heterogeneity, when the heterogeneity is constant over time and not correlated with independent variables (Johnstone & DiNardo, 1997).

The Fixed Effects Model

The fixed effects model (FE) is different from the random effects model since the fixed effects model assumes that the individual effect, α_{it} , is correlated with X_{it} . This estimation method controls for any potential correlation between the independent variables, in addition to controlling for omitted variables by “throwing away” some of the variances that contaminates either OLS or the random effects estimator (Johnstone & DiNardo, 1997).

$$Y_{it} = X_{it}\beta + Z_i\delta + \varepsilon_{it} \quad (4)$$

error term:

$$\varepsilon_{it} = \alpha_{it} + \eta_{it} \quad (5)$$

The fixed effects model is estimated by using least squares dummy variables (LSDV) estimation.

5.2 Estimation Method

Before we choose an estimation method, we must analyze and examine the data set. In addition, we must test the variables for important factors like heteroscedasticity, multicollinearity, autocorrelation, unit root, cointegration and skewness. We can choose between pooled OLS, fixed effects model or random effects model. In general, pooled OLS may seem like the least feasible approach since this estimation method ignores the panel data structure (Johnstone & DiNardo, 1997). Also, the assumptions of the pooled OLS is equal to the regular OLS. Therefore, pooled OLS may result in heterogeneity bias (Bauer College of Business, 2015). However, pooled OLS can mitigate some of this heterogeneity bias by using time dummy variables, in addition to utilize different econometric approaches.

According to Johnstone and DiNardo, the fixed effects model is preferred to the random effects model. The reason is that when the random effects model is appropriate, fixed effect estimators will still produce consistent estimates. Nevertheless, the random effects model can sometimes

be used if we are certain that we can measure all of the time-invariant factors possibly correlated with the other regressors (Johnstone & DiNardo, 1997). For practical reasons it is very complicated to identify all of them, especially as some might not even be estimated as they might be hard to measure. Johnstone and DiNardo also state that researchers find a precisely estimated fixed effects result as more persuasive and significant than a random effects estimation model. To test whether to use fixed effects model or random effects model, one can perform the Hausman test (Johnstone & DiNardo, 1997).

A major concern for the analysis is regarding the time-constant variable *Final grade*. This variable will be omitted if we use the fixed effects estimation model. In addition to the time-constant variable, we are not certain that we can measure all of the time-invariant factors correlated with the other regressors. Consequently, we choose to use the pooled OLS estimation method with clustered individual variables (states), even though the panel data is not strongly balanced. We choose to perform a truncated regression. Truncated regressions are applied in regressions in cases where observations with values in the dependent variable is below or above a certain threshold or limit (Amemiya, 1973). This estimation method is relevant for us in the sense that mortgage complaints cannot be below zero. Hence, we execute truncated regressions with a lower limit of zero. In addition, we include time dummy variables in the regressions for each year, to control for heterogeneity in complaints across time.

5.2.1 Mortgage Complaints Analysis

As mentioned previously, we have divided the data set into five groups: main variables, education variables, state program variables, demographic variables, and supplemental variables. The mortgage complaints analysis contains five regressions. The first regression involves the main variables. Further, we control for the other groups with four extending regressions. The dependent variable, *Mortgage complaints*, is mortgage complaints per state divided by the number of mortgages per state.

Year Dummies

Year dummies capture the influence of aggregated trends (Dartmouth College, 2018). It is important to control for year effects, because panel regressions which fail to control for this, picks up the influence of exogenous trends which could influence the relationship between the dependent and independent variables. In the data for instance, we can argue that income and population are asset data, which increases with time and thereby possess an aggregated upward

trend. Such trends and other types of heterogeneity in complaints across time can cause the coefficient on the included variable to be biased, and thus lead to spurious regressions. Hence, we include year dummies in all the regression.

Main Variables

Mortgage complaints

$$\begin{aligned}
 &= \beta_0 + \beta_1 \text{final grade}_{it} + \beta_2 \text{unemployment rate}_{it} \\
 &+ \beta_3 \text{families below poverty level}_{it} + \beta_4 \text{debt-to-income}_{it} \\
 &+ \gamma_1 \delta_{2013t} + \dots + \gamma_5 \delta_{2017t} + \varepsilon_{it}
 \end{aligned} \tag{6}$$

Final grade is a variable that describe the adult financial literacy in a state. The final grades comprise the weighted average of five financial literacy categories: financial knowledge, credit, saving and spending, retirement readiness, protect and insure (Pelletier, 2016). This is one of the most important variables, given that we are trying to analyze what impact financial literacy has on mortgage complaints. Further, we include the variable *Unemployment rate* to see if states with lower employment rates complain more than others. For a socio-economic perspective, we include the variables *Families below poverty level* and *Debt-to-income*.

Education Variables

Mortgage complaints

$$\begin{aligned}
 &= \beta_0 + \beta_1 \text{final grade}_{it} + \beta_2 \text{unemployment rate}_{it} \\
 &+ \beta_3 \text{families below poverty level}_{it} + \beta_4 \text{debt-to-income}_{it} \\
 &+ \beta_5 \text{business school}_{it} + \beta_6 \text{bachelor's degree}_{it} \\
 &+ \beta_7 \text{legal occupation}_{it} + \gamma_1 \delta_{2t} + \dots + \gamma_5 \delta_{6t} + \varepsilon_{it}
 \end{aligned} \tag{7}$$

In this regression, we include education variables in addition to an occupation variable. The *Business school* variable represents the absolute number in each state that has a bachelor's degree in business, while the *Bachelor's degree* is the percentage of the population that has a bachelor's degree. *Legal occupation* represents the total number of people working within legal, law and order. The argument for including this variable is that we believe the knowledge and experience these workers possess, make them more aware of what they can and cannot complain about.

State Program Variables

Mortgage complaints

$$\begin{aligned}
&= \beta_0 + \beta_1 \text{final grade}_{it} + \beta_2 \text{unemployment rate}_{it} \\
&+ \beta_3 \text{families below poverty level}_{it} + \beta_4 \text{debt-to-income}_{it} \\
&+ \beta_5 \text{business school}_{it} + \beta_6 \text{bachelor's degree}_{it} \\
&+ \beta_7 \text{legal occupation}_{it} \\
&+ \beta_8 \text{10-year lagged economic education}_{it} \\
&+ \beta_9 \text{10-year lagged personal finance}_{it} + \gamma_1 \delta_{2t} + \dots + \gamma_5 \delta_{6t} \\
&+ \varepsilon_{it}
\end{aligned} \tag{8}$$

State programs are similar to the education variables, since state programs are education programs initiated by the states. The state program variables are scaled from one to five, based on how many of these standards that are included in the given state. The majority of millennial home buyers are around the age of 25 to 33, with a median age of 28 (Siniavskaia, 2013). Hence, we use 10-year lagged variables for the state programs. As this median age is approximately ten years after high school, we capture the effect of the state programs.

Demographic Variables

Mortgage complaints

$$\begin{aligned}
&= \beta_0 + \beta_1 \text{final grade}_{it} + \beta_2 \text{unemployment rate}_{it} \\
&+ \beta_3 \text{families below poverty level}_{it} + \beta_4 \text{debt-to-income}_{it} \\
&+ \beta_5 \text{business school}_{it} + \beta_6 \text{bachelor's degree}_{it} \\
&+ \beta_7 \text{legal occupation}_{it} \\
&+ \beta_8 \text{10-year lagged economic education}_{it} \\
&+ \beta_9 \text{10-year lagged personal finance}_{it} \\
&+ \beta_{10} \text{african american}_{it} + \beta_{11} \text{indian or alaskan}_{it} \\
&+ \beta_{12} \text{asian}_{it} + \beta_{13} \text{native hawaiian}_{it} + \beta_{14} \text{white}_{it} + \gamma_1 \delta_{2t} + \dots \\
&+ \gamma_5 \delta_{6t} + \varepsilon_{it}
\end{aligned} \tag{9}$$

As mentioned earlier, these demographic variables represent the composition of several ethnicities in each state. These are in percentages of the total population, where a high percentage represents a large proportion of that ethnicity in that given state. We include these variables in the regression to control for the different ethnicities, and to see if there are any relationships between ethnicities and number of complaints.

Supplemental Variables

Mortgage complaints

$$\begin{aligned}
&= \beta_0 + \beta_1 \text{final grade}_{it} + \beta_2 \text{unemployment rate}_{it} \\
&+ \beta_3 \text{families below poverty level}_{it} + \beta_4 \text{debt-to-income}_{it} \\
&+ \beta_5 \text{business school}_{it} + \beta_6 \text{bachelor's degree}_{it} \\
&+ \beta_7 \text{legal occupation}_{it} \\
&+ \beta_8 \text{10-year lagged economic education}_{it} \\
&+ \beta_9 \text{10-year lagged personal finance}_{it} \\
&+ \beta_{10} \text{african american}_{it} + \beta_{11} \text{indian or alaskan}_{it} \\
&+ \beta_{12} \text{asian}_{it} + \beta_{13} \text{native hawaiian}_{it} + \beta_{14} \text{white}_{it} \\
&+ \beta_{15} \text{internet users above 15 years}_{it} + \beta_{16} \text{median income}_{it} \\
&+ \gamma_1 \delta_{2t} + \dots + \gamma_5 \delta_{6t} + \varepsilon_{it}
\end{aligned} \tag{10}$$

For the last regression, we include the variables *Median income* and *Internet users above 15 years old*. The reason for including *Internet users above the age of 15* is to see if states with more internet accessibility complain more. We assume that if more people have access to the internet, there is a lower barrier to complain on your mortgage. The reason why the lower limit is 15 is simply because it was the only option including a lower limit, which obviously is more fitting for our research question. The inclusion of *Median income* is discussed in depth later.

5.2.2 Student Loan Complaints Analysis

There are many similarities between the two models. Nevertheless, there are a few alterations concerning our estimation method compared to the first analysis regarding mortgage complaints. To control for population differences across states, we divide the total number of student loan complaints by our proxy for the number of student loans. Our estimation model is a truncated pooled OLS, *without* cluster on the individual variable (states). We run four regressions, presented by each group in the following section.

Main Variables

Student loan complaints

$$\begin{aligned}
&= \beta_0 + \beta_1 \text{final grade}_{it} + \beta_2 \text{unemployment rate}_{it} \\
&+ \beta_3 \text{families below poverty level}_{it} + \beta_4 \text{debt-to-income}_{it} \\
&+ \beta_5 \text{legal occupation}_{it} + \gamma_1 \delta_{2t} + \dots + \gamma_5 \delta_{6t} + \varepsilon_{it}
\end{aligned} \tag{11}$$

As mentioned, we have only one modification in the main variables group, which is to include the variable *Legal occupation*. The intuition of including this variable is to capture the effect of the students that has parents working within legal occupations. The argument for including legal occupations in the mortgage complaints analysis, is due to the fact that individuals with such occupations might have a better grasp on laws and regulations, and might better know how to maneuver their way through a complaint process. Consequently, students with parents working within law, might to a higher degree receive help concerning complaints towards student loans.

State Program Variables

Student loan complaints

$$\begin{aligned}
 &= \beta_0 + \beta_1 \text{final grade}_{it} + \beta_2 \text{unemployment rate}_{it} \\
 &+ \beta_3 \text{families below poverty level}_{it} + \beta_4 \text{debt-to-income}_{it} \\
 &+ \beta_5 \text{legal occupation}_{it} \\
 &+ \beta_6 \text{1-year lagged economic education}_{it} \\
 &+ \beta_7 \text{1-year lagged personal finance}_{it} + \gamma_1 \delta_{2t} + \dots + \gamma_5 \delta_{6t} \\
 &+ \varepsilon_{it}
 \end{aligned} \tag{12}$$

In this regression, we choose to include the state programs concerning education within economics and personal finance in high schools. For mortgage complaints, we lagged these variables with ten years, whereas in this analysis we lag the variables with one year. It takes at least one year before high school students become college students. Hence, it makes more sense to use standards from one year ago.

Demographic Variables

Student loan complaints

$$\begin{aligned}
 &= \beta_0 + \beta_1 \text{final grade}_{it} + \beta_2 \text{unemployment rate}_{it} \\
 &+ \beta_3 \text{families below poverty level}_{it} + \beta_4 \text{debt-to-income}_{it} \\
 &+ \beta_5 \text{legal occupation}_{it} \\
 &+ \beta_6 \text{1-year lagged economic education}_{it} \\
 &+ \beta_7 \text{1-year lagged personal finance}_{it} \\
 &+ \beta_8 \text{african american}_{it} + \beta_9 \text{indian or alaskan}_{it} \\
 &+ \beta_{10} \text{asian}_{it} + \beta_{11} \text{white} + \beta_{12} \text{native hawaiian}_{it} + \gamma_1 \delta_{2t} + \dots \\
 &+ \gamma_5 \delta_{6t} + \varepsilon_{it}
 \end{aligned} \tag{13}$$

In this regression, we implement demographic variables to the regression. As might be expected, there are no changes in these variables compared to the demographic variables used in the mortgage complaints analysis.

Supplemental Variables

Student loan complaints

$$\begin{aligned}
&= \beta_0 + \beta_1 \text{final grade}_{it} + \beta_2 \text{unemployment rate}_{it} \\
&+ \beta_3 \text{families below poverty level}_{it} + \beta_4 \text{debt-to-income}_{it} \\
&+ \beta_5 \text{legal occupation}_{it} \\
&+ \beta_6 \text{1-year lagged economic education}_{it} \\
&+ \beta_7 \text{1-year lagged personal finance}_{it} \\
&+ \beta_8 \text{african american}_{it} + \beta_9 \text{indian or alaskan}_{it} \\
&+ \beta_{10} \text{asian}_{it} + \beta_{11} \text{white} + \beta_{12} \text{native hawaiian}_{it} \\
&+ \beta_{13} \text{internet users above 15 years} + \beta_{14} \text{median income} \\
&+ \beta_{15} \text{private school} + \beta_{16} \text{public school} + \gamma_1 \delta_{2t} + \dots + \gamma_5 \delta_{6t} \\
&+ \varepsilon_{it}
\end{aligned} \tag{14}$$

Lastly, we include the supplemental variables consisting of *Internet users above 15 years*, *Median income*, *Private school* and *Public school*. According to Murnane & Reardon, “private schools have increasingly served students from high-income families” (2018). Hence, the reason for including public and private school is to control for any differences towards complaining behavior concerning students from the two different school types.

6. Analysis

In this section, we will go through a correlation analysis of the data, an evaluation of our estimation methods and present and discuss our results concerning mortgage- and student loan complaints.

Correlation Analysis

As we can see from table D.1 in Appendix D, there are many strong correlation coefficients in the data set. We start by looking at the dependent variable, *Mortgage complaints*. The complaints are positively correlated with *Unemployment rate*, with a correlation coefficient of 40,19%. Also, *Debt-to-income* has a strong relationship to complaints with a correlation coefficient of 36,67%. Those two are the variables with the highest degree of correlation with complaints within this group.

Further, we look into the group of education variables. All three variables in this group has a strongly positive relationship with complaints. *Business school*, *Bachelor's degree*, and *Legal occupation* respectively has 41,86%, 31,83% and 49,72% as correlation coefficient. This implies that there should be some relationship between complaints and these three variables.

In the three next groups, state program variables, demographic variables and supplemental variables, there exist a smaller amount of correlation with *Mortgage complaints*. The strongest correlation coefficient is *White*.

If we look at the correlation between the independent variables, we will also find some strong and interesting relationships. *Final grade* has strong relationships with several independent variables. *Final grade* has a negative correlation coefficient of 83,80% with *Families below poverty rate*. Remember, *Final grade* is a variable describing financial literacy in each state. Hence, the negative correlation coefficient indicates that when the number of families below poverty level increases, the financial literacy decreases. Further, *Final grade* has a strong positive relation with *Median income* of 64,58%. Hence, when the median income increases, the financial literacy tend to increase as well. There are also some strong relationships in the group of demographic variables, for instance *African American* and *Final grade*. The correlation coefficient is -64,71% and indicates that if the number of African American increases, the financial literacy decreases.

6.1 Robustness of our Estimation Models

In this subsection, we will describe the tests and the applications added to our estimation models, the suitability of our models and finally the results from the regressions. To get to our estimation model, we test for multicollinearity, autocorrelation, unit root, cointegration, heteroscedasticity and skewness.

6.1.1 Mortgage Complaints

Multicollinearity

First, we test the data set for multicollinearity. Collinearity tends to inflate the variance of the betas. Hence, this can cause some coefficients to have the wrong sign (Johnstone & DiNardo, 1997). To test whether the variables are correlated or not, we perform the Variance Inflation Factor (VIF) test. A variable is said to be highly collinear if the variable has a higher VIF score than ten (Hair Jr., Anderson, Tatham, & Black, 1995). Table B.1 in Appendix B illustrates that there are four variables that has a VIF score above ten. Three of these four variables are demographic variables. When independent variables are collinear, they tend to share the explanatory force in the regression. Hence, the effect of some of the independent variables is already explained by some other variables in the model (Arora, 2016). Therefore, we remove the variables *White* and *Native Hawaiian* from the regression and test the model for multicollinearity again. As table B.2 (in Appendix B) illustrates, none of the variables is highly collinear post-removal. Hence, we decide to exclude the variables *White* and *Native Hawaiian* in the analysis.

Autocorrelation/Serial Correlation

Further, we test the data for autocorrelation/serial correlation. Autocorrelation is the relationship between a variable and the lagged version of itself over a time period. Autocorrelation is often found in repeating patterns, where previous levels have an impact of future levels. We use the Wooldridge test for autocorrelation in panel data to test the variables. The H_0 is that it does not exist autocorrelation in the data, whereas H_A is that there is autocorrelation. Table B.3 (in Appendix B) shows that the test is significant; hence we reject the H_0 and conclude the existence of autocorrelation. The cluster command will effectively make standard errors robust to any kind of autocorrelation.

Unit Root and Cointegration

Next, we test if the data contains unit roots, and further if the variables with unit root are cointegrating. A process contains unit root if the $\beta = 1$ (Johnstone & DiNardo, 1997). The equation is:

$$Y_{it} = \alpha + Y_{t-1} + \varepsilon_t, \quad (15)$$

and is called a *random walk with drift*. If this is the case, the Y-series are nonstationary, and shows an explosive behavior. When regressing two nonstationary variables it usually results in spurious regression (Barunik, 2010). However, if X_t and Y_t are cointegrated, there would not be any spurious regression. Therefore, if we find any unit roots in the data set, we need to test if they are cointegrating. Cointegration is the existence of a long-run relationship between two or more variables.

We use the Im-Pesaran-Shin unit-root test and the Kao test for cointegration. Table B.4 (in Appendix B) shows the existence of unit root in five variables: *Mortgage complaints*, *Median income*, *Business school*, *Bachelor's degree*, and *Internet users above 15 years*. The Kao test for cointegration consists of five tests:

1. Modified Dickey-Fuller test
2. Dickey-Fuller test
3. Augmented Dickey-Fuller test
4. Unadjusted modified Dickey-Fuller test
5. Unadjusted Dickey-Fuller test.

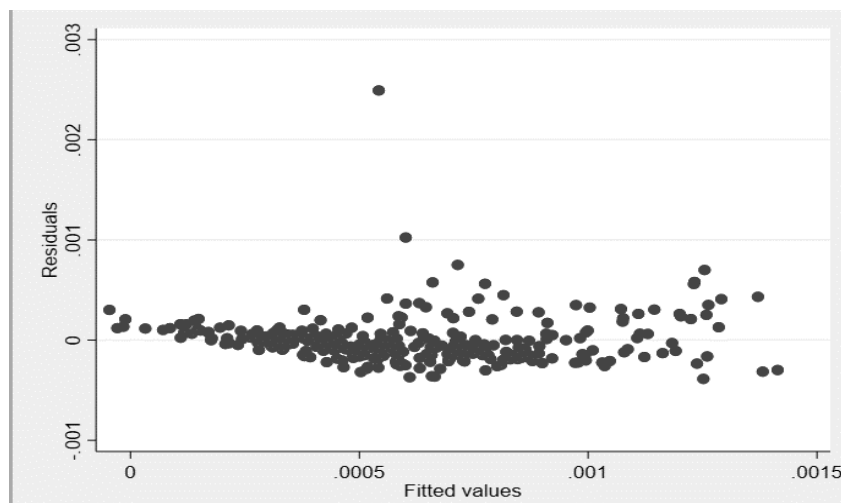
These tests differ in how the framework of the tests consider serial correlation. Since there is serial correlation in the data set, we use the modified Dickey-Fuller test option for cointegration. This option is derived by specifying a data-generating process for the dependent variable and the regressors. This specification allows the regressors to be serially correlated (Persyn & Westerlund, 2008). Table B.5 (in Appendix B) shows a p-value from the modified Dickey-Fuller test of 0.005; hence we reject the H_0 that there is no cointegration. Hence, we will not get spurious regressions, thus we do not have to control the variables for unit root.

Heteroscedasticity

Further, we test the data for heteroscedasticity. Heteroscedasticity occurs if the variance of the error terms differs across observations (University of Canterbury, 2018). Heteroscedasticity

can be a concern in the application of regression analysis, since it can invalidate statistical tests of significance that assume the existence of homoscedasticity. We use the Breusch-Pagan test for heteroscedasticity, where H_0 is that the variances are constant, while H_A is that the variances are not. Table B.3 (in Appendix B) shows that the Breusch-Pagan test is significant; hence we reject the null hypothesis that there is no heteroscedasticity. Figure 6.1 illustrates the residuals in the regression.

Figure 6.1: Residual Plot



To control for the heteroscedasticity, we use the cluster command on the individual variable, i.e. the states. By clustering the individual variable, it makes it possible to arrange the variables into homogenous clusters and therefore obtain meaningful structures (Chavent, Liquet, Kuentz, & Saracco, 2012).

Skewness

Skewness is a measure of symmetry. If there is perfect symmetry, the data are normally distributed. In this case, the left side and the right side of the center look the same. Negative skewness indicates that the data is skewed to the left, whereas positive skewness represents skewness to the right. We use the Stata normality test that tests the data set for kurtosis and skewness. The test in table B.6 (in Appendix B) indicate a skewed data set. However, as mentioned before, we utilize a truncated regression with a lower limit at zero. By doing this, we are controlling for the skewness. Figures B.1 and B.2 illustrates the differences towards the normality of using truncated regression and regular regression.

Test Summary

In this subsection, we have tested for multicollinearity, autocorrelation, unit root, cointegration, heteroscedasticity and skewness. The Variance Inflation Factor test initially detected a high degree of multicollinearity between some of the variables. Hence, we removed the variables *White* and *Native Hawaiian*. After running another VIF test the presence of multicollinearity diminished. As the Breusch-Pagan test for heteroscedasticity and Wooldridge test for autocorrelation is significant, we used the cluster command on the individual variable to control for the potential biases by making the standard errors more robust. Further, the normality test for kurtosis and skewness showed the existence of skewness in the data, and we control for this by using a truncated regression. Moreover, we discovered that some variables contained unit root, but detected that they are cointegrating.

6.1.2 Student Loan Complaints

Multicollinearity

First, we test the data set for multicollinearity. A variable is highly collinear if the variable has a VIF score higher than ten. As we can see from table C.1 in Appendix C, there are five variables that are highly collinear with the other independent variables. The two strongest ones are *Public school* and *Private school*. As illustrated in table D.2 (in Appendix D), these two variables are negatively correlated to the extreme.

Since the variables *Public school* and *White* have the highest score from the VIF test in their respective variable group, we exclude them from the analysis. We also exclude *Native Hawaiian*. By doing so, we obtain a data set with variables that score less than ten on the VIF test. This is illustrated in table C.2 (in Appendix C).

Autocorrelation/Serial Correlation

Further, we use the Wooldridge test for autocorrelation. As illustrated in table C.3 (in Appendix C) the test is insignificant, in contrast with the mortgage complaint analysis. Hence, we keep the H_0 that there is no autocorrelation in the data.

Unit Root and Cointegration

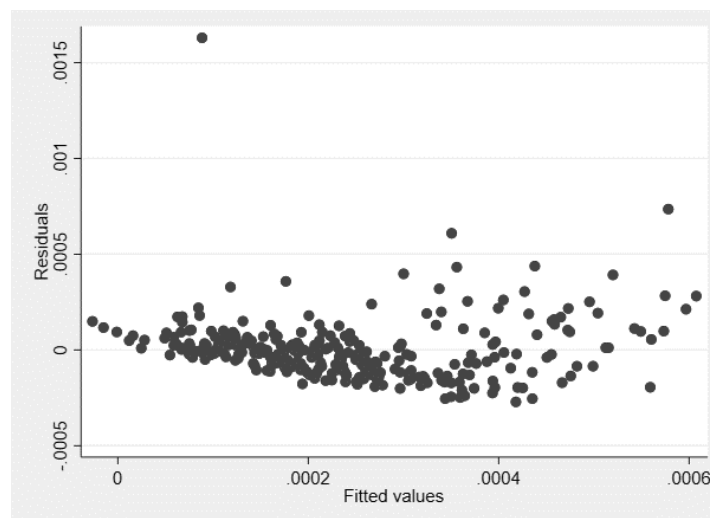
Further, we test five new variables for unit root, and to see whether any are cointegrating with each other. The Im-Pesaran-Shin unit-root test, table C.4 (in Appendix C) shows three variables with unit root; *Student loan complaints*, *Median income*, and *Internet users above*

15 years. As we did in the previous analysis, we perform the Kao test for cointegration. However, since we do not have autocorrelation in this data set, we choose the unadjusted Dickey-Fuller test, as it assumes the absence of serial correlation. As illustrated in table C.5, the p-value from the unadjusted Dickey-Fuller test is 0.000, hence we reject the H_0 that there is no cointegration.

Heteroscedasticity

In contrast to the data set for mortgage complaints, we do not discover the presence of heteroscedasticity. One might argue a visual detection of heteroscedasticity in figure 6.2, as the residuals of the variables are expanding as the fitted values increase. However, the Breusch-Pagan test for heteroscedasticity is clearly insignificant (table C.3 in Appendix C) and we cannot reject the H_0 that there is heteroscedasticity. As we can reject both the presence of heteroscedasticity and autocorrelation, we conclude that there is no usefulness for the cluster command in this model.

Figure 6.2: Residual Plot



Skewness

We test the data for skewness using the normality test for kurtosis and skewness. As shown in table C.6 (in Appendix C), the results are significant; hence we do have skewness in the data. However, this is controlled for by using the truncated regression with a lower limit of zero. We illustrate the differences towards the normality of using truncated regression and regular regression in figures C.1 and C.2. By using the truncated regression, we obtain more normally distributed residuals.

Test Summary

In this subsection, we have tested the data set for multicollinearity, autocorrelation, unit root, cointegration, heteroscedasticity and skewness. The VIF test detected a high degree of multicollinearity between some of the variables. Hence, we removed the variables *White*, *Native Hawaiian*, and *Private school*. After running another VIF test, the presence of multicollinearity diminished. In contrast to the previous analysis, the Breusch-Pagan test for heteroscedasticity and the Wooldridge test for autocorrelation is insignificant. Hence, we did not use the cluster command on the individual variable to control for the potential biases. Further, the normality test for kurtosis and skewness showed the existence of skewness in the data, and we control for this by using a truncated regression. Moreover, we discovered that some variables contain unit root, but detected that they are cointegrating.

6.2 Hypotheses

Based on the theories and findings presented in the literature review, we have formulated three hypotheses regarding our research question. The first one is based on the studies by Warland, Herrmann, & Willits (1975) and Gronhaug (1977), concerning more complaints from upscale socio-economic people, which primarily included education level, type of education and income level. The second is based both on the foregoing, in addition to the study by Rutledge (2010), which argues that financial literacy gives consumers the knowledge, skills and confidence to understand the information they receive and evaluate the risks and rewards of financial services and product. In light of this we assume that literate consumers would to a higher degree recognize dishonorable practices and potentially regulatory violations. The third hypothesis is based on the study by Stokes (1974), which suggests that people with more time on their hand complain more.

The hypothesis testing will be used to examine whether the data are in agreement with existing literature. The null hypothesis will be rejected on the 5% significance level.

1. H_0 : The effect of *Families below poverty level, Legal occupation, Bachelor's degree, Private school and Median income* = 0.
 H_A : The effect of *Families below poverty level, Legal occupation, Bachelor's degree, Private school and Median income* $\neq 0$
2. H_0 : The effect of *Final grade, Lagged state programs and Business school* = 0
 H_A : The effect of *Final grade, Lagged state programs and Business school* $\neq 0$
3. H_0 : The effect of *Unemployment rate* = 0
 H_A : The effect of *Unemployment rate* $\neq 0$

6.3 Empirical Results - Mortgage complaints

The aim of the analysis is to answer the following research question:

How do financial literacy and other socio-economic characteristics relate to mortgage and student loan complaints?

We will begin with the findings concerning mortgage complaints. The analysis is based on the previously discussed specifications. In this subsection, we present the results from the regressions, then we discuss and compare the findings with our proposed hypotheses which are based on the established theories and literature presented in section 2.

Table 6.1: Truncated Regressions of Mortgage Complaints on Financial Literacy and Other Socio-economic Factors

<i>Mortgage complaints</i>	(1)	(2)	(3)	(4)	(5)
Final grade	-.0000192** (.0000115)	-9.55e-06* (6.82e-06)	-7.88e-06* (6.25e-06)	-7.50e-06* (5.41e-06)	-8.80e-06* (5.54e-06)
Unemployment rate	.0113969*** (.0034855)	.0060413*** (.0023649)	.0074303*** (.0022309)	.0054249*** (.0019096)	.0052265*** (.0017977)
Families below poverty level	-.008375*** (.0029883)	-.0041439** (.0022841)	-.0058809*** (.0022725)	-.0048072*** (.0021539)	-.0043078** (.0023672)
Debt-to-income	.0003391** (.0001826)	.0003066*** (.0000978)	.0002934*** (.0000896)	.000336*** (.0000846)	.0003159*** (.0000911)
Business school		3.40e-10*** (7.76e-11)	2.58e-10*** (7.65e-11)	1.74e-10 (7.75e-11)	1.67e-10*** (8.39e-11)
Bachelor's degree		.0007387 (.0005325)	.0006529* (.0004924)	.0003768 (.0003687)	.0003456 (.0003251)
Legal occupation		6.58e-08*** (2.22e-08)	6.98e-08*** (1.78e-08)	6.68e-08*** (1.73e-08)	6.88e-08*** (1.61e-08)
10-year lagged economic education			.0000612*** (.0000263)	.0000516*** (.0000265)	.0000528*** (.0000271)
10-year lagged personal finance			-1.94e-08 (.0000171)	-6.80e-06 (.0000151)	-9.83e-06 (.0000154)
African American				.0004029 (.0003915)	.0004079 (.0003928)
Indian or Alaskan				-.0014962* (.0010294)	-.0014586* (.0010611)
Asian				.0016603*** (.0001977)	.0018139*** (.000301)
Internet users above 15 years					.0006802 (.0006827)
Median income					-1.09e-10 (4.57e-09)
Year dummies	Y	Y	Y	Y	Y
<i>Sigma/Z</i>	<i>8.41</i>	<i>8.93</i>	<i>10.12</i>	<i>10.33</i>	<i>10.26</i>

Dependent variable is the number of mortgage complaints over number of households with a mortgage (Consumer Financial Protection Bureau, 2018). Year dummies are included in all regressions to control for heterogeneity in complaints across time. Standard errors are clustered by state, and the regression is truncated. * denotes significance at the 20% level, ** denotes significance at the 10% level and *** denotes significance at the 5% level.

Due to the presence of heteroscedasticity and autocorrelation detected in the data set, we conduct a cluster robust regression. Because this data set contains relatively few time variables compared to individual variables, clustering will effectively make standard errors robust to

any kind of autocorrelation and heteroscedasticity. Furthermore, since mortgage complaints cannot be below zero, we perform a truncated regression with a lower limit.

Table 6.1 illustrates that there are differences in the magnitude of the coefficients between the variables and the different regressions. When interpreting the coefficients' effect on the dependent variable, all other variables shall be held constant (*ceteris paribus*). Furthermore, as many of the coefficients appear to be very small, it is important to recall that the mean value of the dependent variable is 0.0006, as illustrated in table 4.1.

Main Variables

Final Grade

When we execute a regression with the main variables, there is a weak negative relationship between *Final grade* and *Mortgage complaints*. When *Final grade* increases by one unit, *Mortgage complaints* decreases by 8.80e-06. In the fifth regression, the variable is significant at the 20% level.

As mentioned in the literature review, the study by Stokes (1974) indicated that educated and articulate people complain more. Hence, the result is inconsistent with that theory and our second hypothesis. We assumed that better financial knowledge would lead to a higher rate of complaints, as more knowledgeable consumers would be more aware of dishonorable practices and regulatory violations. Furthermore, according to Rutledge (2010), educated people, and people that possess more knowledge about financial topics, are more familiar with the different products and services. Hence, it is reasonable to presume that informed individuals have a better grasp on what they can and should complain about. However, in the studies by Moore D. L. (2003) and Lusardi & Mitchell (2014) it was proposed that the substantial variety of modern financial products and services are difficult to master for the financially unsophisticated, and that the least financially literate were likely to have costlier mortgages and other more consumer-unfriendly financial products. Thus, one might argue that products and services of low quality are more likely to lead to dissatisfaction, as well as a greater number of complaints. This might explain the fact that even though individuals of higher socio-economic status tend to complain more, they might in recent years, with a growing degree of autonomy and financial complexity, be better at distinguishing the good financial products and services from the bad, which may serve as an explanation for a tendency to complain less. This reasoning might give grounds for the insignificant variable.

Unemployment Rate

As illustrated in table 6.1, there is a strong positive relationship between *Unemployment rate* and *Mortgage complaints*. When *Unemployment rate* increases by one unit, *Mortgage complaints* increases by 0.005 in the fifth regression. The variable is significant at the 5% level across all five regressions and stands out among the variables. We can also see that once we control for other variables, the significance of the independent variable remains unaffected while the size of the coefficient decreases.

According to our results, a higher unemployment rate leads to more complaints. As complaining is a process of time-demanding nature, we proposed in our third hypothesis that people with a higher degree of spare time complains to a greater extent. This is also in line with the study of Stokes (1974), which indicated that people with more time on their hands complain more. Additionally, one could argue that unemployed people are more vulnerable to shocks to their economy, and by that more inclined to take action or complain in such events.

Families Below Poverty Level

Table 6.1 provides a strong negative relationship between *Families below poverty level* and *Mortgage complaints*. When *Families below poverty level* increases by one unit, *Mortgage complaints* decreases by 0.004 in the fifth regression. The variable is significant at the 5% level across all five regressions, except for the second and fifth regression, where it is just outside the significance cutoff point.

According to the regressions, a higher percentage of families below poverty level yields less complaints. One could argue that families below poverty level may work hard and several hours to earn enough to cover expenses for a decent living. Hence, they have less spare time, and the results are in line with the study by Stokes (1974). In addition, it is consistent with our first hypothesis and the theory proposed by Warland, Herrmann, & Willits (1975) regarding the fact that individuals of higher socio-economic status complain more frequently than others.

Debt-to-income

Table 6.1 shows a strong positive relationship between *Debt-to-income* and the *Mortgage complaints*. When *Debt-to-income* increases by one unit, *Mortgage complaints* increases by 0.0003 in the fifth regression. The variable is significant at the 5% level across regression 2-5 and stands out as one of the most significant variables. We can also see that once we control

for other variables in the extended models, the significance of *Debt-to-income* and the size of its coefficient remain unaltered.

According to these results, a higher debt-to-income level leads to more complaints. A high ratio of debt-to-income often concur with unhealthy financial conditions. Complaints about financial hardship and affordability are often linked (Financial Ombudsman Service, 2018). On many occasions, complaints originate from lending that was unaffordable from the outset and should not have been allowed by the lender in the first place. Furthermore, as in the section regarding unemployment rate one could argue that individuals with higher interest cost, and thereby lower disposable income, are more vulnerable to shocks to their economy and financial distress. Hence, they are more inclined to take action or complain if instances such as that occur. Which is in line with the findings of Hirschman (1970), as the value of the potential upside gained from a successful complaint is relatively higher.

Education Variables

Business School and Bachelor's Degree

When we execute the regression including the education variables and the main variables, there is a strong positive relationship between *Business school* and *Mortgage complaints*. When *Business school* increases by one unit, *Mortgage complaints* increases by $1.67e-10$, in the fifth regression. The variable is significant at the 5% level in all regressions, except the fourth. We also include *Bachelor's degree*. However, it has no effect, which might owe to the fact that *Business school* grasps the majority of the explanation power.

Since the results in some of the regressions show that a higher number of individuals with a business school background leads to more complaints, these results are in agreement with our first hypothesis, and the study by Warland, Herrmann, & Willits (1975). However, as the significance vanishes in the fourth regression, its effect is uncertain.

Legal Occupation

Table 6.1 illustrates a strong positive relationship between *Legal occupation* and the *Mortgage complaints*. When *Legal occupation* increases by one unit, *Mortgage complaints* increases by $6.88e-08$ in the fifth regression. The variable is significant at the 5% level across regression 2-5 and stands out among the variables as one of the most significant. We can also see that once we control for other variables, the significance of the independent variable persists, and the size of the coefficient remain rather unaltered.

According to the regression output, a higher number of legal occupations in a state leads to more complaints. As mentioned, complaining processes can be time-demanding. Additionally, such a process might be complex, especially with regards to that the justice system in the United States is complicated (Berg, 2017). In fact, the differences and obscurity in legislation and regulations have given rise to establish uniform acts to bring clarity and stability of state statutory laws (Uniform Law Commission, 2018). Therefore, legal expertise might be beneficial for complainers, and might be an argument for a higher number of complaints. Moreover, it is in line with the socio-economic perspectives as suggested.

State Program Variables

10-year Lagged Economic Education

The third regression, which introduces the state program variables, reveal that there is a strong positive relationship between *10-year lagged economic education* and *Mortgage complaints*. When *10-year lagged economic education* increases by one unit, *Mortgage complaints* increases by 0.00005 in the fifth regression. The variable is significant at the 5% level across the rest of the regressions. We can also see that once we control for other variables, the significance of the independent variable remains unaffected, and the size of the coefficient barely decreases. In the regression, we also control for the variable *10-year lagged personal finance education*. However, the variable is statistically insignificant.

The results illustrate that a higher level of high school standards and requirements, with respect to economic education, yields higher number of complaints ten years later. As discussed concerning *Final grade*, one could argue the same towards this variable. If we once again refer to the study by Stokes (1974), which indicate that educated and articulate people tend to complain more, one could argue that this is in agreement with his findings. According to these results, one might argue that states which provide standardized requirements and testing at high school over time, inhabits people which become more aware and knowledgeable of their personal economy and complains to a higher degree than in states which do not have provided the same specifications. Additionally, one can argue that the implementation of higher standards on education programs could be because of a record with inferior results. In other words, a state initiated these projects because of a history of unsatisfactory results. That being the case, one can argue that it removes the effect of already possessing the financial knowledge (which we find in the variable *Final grade*), which might affect the ability to carefully sort out the good financial products, of which the tendency to complain towards is lower. That means

that one might better capture the effect of people getting improved financial knowledge and its effect on consumer complaints. However, this is conjectures, and might be of benefit to elaborate further as a topic for future research.

Demographic Variables

As table 6.1 illustrates, the fourth regression introduces the different demographic variables. After implementing these variables, *Asian* is the only significant variable. The positive relationship tells us that if *Asian* increases by one unit, *Mortgage complaints* increases by 0.002 in the fifth regression. The variable is significant at the 5%. These results are in contrast with the findings introduced by Liu & McClure (2001), concerning cross-cultural differences in consumer complaint behavior. *African American* is insignificant, which contradicts the study by Ayres, Lingwall, & Steinway (2013).

Supplemental Variables

The fifth regression introduces the last two variables; *Median income* and *Internet users above 15 years*. The variables are insignificant, which was quite surprising due to *Median income's* inclusion in the first hypothesis. However, as illustrated in table D.1 in Appendix D, there is evidence of a strong correlation between *Median income* and *Families below poverty level*, as well as with *Bachelor's degree*. The correlation coefficients are -0.7869 and 0.6465, respectively. That being the case, it might explain why the significance of *Median income* diminishes. Furthermore, as a controlling variable it affects the significance of the variable *Families below poverty level*.

We wanted to analyze the effect of internet users, due to the findings presented by Andreassen & Streukens (2013) that the ability to complain online increases the likelihood that consumers will take action towards a provider. In addition, one might expect that it is easier to complain over the internet rather than through old methods. Hence, we anticipated that individuals with a higher grade of internet accessibility would have an easier ability to file complaints. However, this variable is insignificant as well.

6.4 Empirical Results - Student Loan Complaints

The second part of our research question is to analyze how financial literacy and socio-economic characteristics relate to student loan complaints. In this subsection, we will present and discuss the findings. The structure of the data set is similar to the data set in the analysis of mortgage complaints and is based on the aforementioned specifications.

As table 6.2 illustrates, there are many significant variables in line with the results regarding mortgage complaints. We anticipated differences between complaints regarding mortgages and student loans as there are a couple of differences between the analyses. To be specific; in the CFPB database, there are over 218,000 mortgage complaints compared to 34,000 student loan complaints. In addition, we use a proxy for the total number of student loans.

Table 6.2: Truncated Regression of Student Loan Complaints on Financial Literacy and Socio-economic Factors

<i>Student loan complaints</i>	(1)	(2)	(3)	(4)
Final grade	-7.69e-06*** (3.23e-06)	-6.79e-06*** (3.29e-06)	-4.06e-06 (3.54e-06)	-3.42e-06 (3.66e-06)
Unemployment rate	.0020147** (.0011735)	.0025209*** (.0012294)	.0022704** (.0011987)	.0024818*** (.0012022)
Families below poverty level	-.0038874*** (.0009582)	-.0039545*** (.0010458)	-.0032386*** (.0010226)	-.0033771*** (.001279)
Debt-to-income	.0000549* (.0000428)	.0000438 (.0000434)	.0000676* (.0000418)	.0001024*** (.0000463)
Legal occupation	1.23e-08* (8.52e-09)	1.25e-08* (8.66e-09)	1.33e-08* (8.25e-09)	1.33e-08* (8.16e-09)
1-year lagged economic education		-8.92e-07* (.0000126)	4.26e-06 (.0000102)	6.65e-06 (.0000102)
1-year lagged personal finance		.0000136 (.0000105)	-2.96e-06 (.0000125)	3.57e-07 (.0000125)
African American			-2.76e-06 (.0001899)	-.0001788 (.0002192)
Indian or Alaskan			-.0020926*** (.0005587)	-.0017979*** (.0005822)
Asian			-.0005134** (.0002781)	-.000678*** (.0003187)
Internet users above 15yrs				-.0000451 (.0004038)
Median income				-2.34e-09 (2.90e-09)
Private school				-.0012732*** (.0006171)
Year dummies	Y	Y	Y	Y
<i>Sigma/Z</i>	<i>14.96</i>	<i>15.01</i>	<i>15.52</i>	<i>15.66</i>

Dependent variable is the number of student loan complaints over total number of student loans (Consumer Financial Protection Bureau, 2018). Year dummies are included in all regressions to control for heterogeneity in complaints across time. The regression is truncated, and standard errors are *not* clustered by state. * denotes significance at the 20% level, ** denotes significance at the 10% level and *** denotes significance at the 5% level.

Main Variables

In the fourth regression, there are three significant variables; *Unemployment rate*, *Families below poverty level* and *Debt-to-income*. When *Unemployment rate* and *Debt-to-income* increases by one unit, *Student loan complaints* increases (ceteris paribus) by 0.0025, and 0.0001, respectively. The variables are significant at the 5% level. A unit increase in *Families below poverty level*, *Student loan complaints* decreases by 0.0034.

Concerning *Families below poverty level*, the results are in line with our reasoning concerning that families below income level works several hours to cover living expenses. Hence, they have less spare time, and could be in line with the findings presented by Stokes (1974). Furthermore, this is consistent with the theory proposed by Warland, Herrmann, & Willits, (1975) regarding the fact that individuals of higher socio-economic status complain more frequently than others.

As for *Debt-to-income* and *Unemployment rate*, one can argue that a lower disposable income leads to higher vulnerability towards financial distress and other shocks to one's economy. Thus, a reasoning of being more inclined to take action or complain in such occurrences. Which is in line with the findings of Hirschman (1970), as the value of the potential upside gained from a successful complaint is relatively higher. Furthermore, the positive significance of *Unemployment rate* is in line with the study by Stokes (1974) regarding that people with more time on their hands complain more.

Private School

There is a strong negative relationship between *Private school* and *Student loan complaints* in the fourth regression. When *Private school* increases by one unit, *Student loan complaints* decreases by 0.0013. The variable is significant at the 5% level. According to the results, a higher number of kids at private schools leads to fewer complaints in that state. As mentioned, the study by Murnane & Reardon found that "private schools have increasingly served students from high-income families" (2018). This contradicts our first hypothesis and the theory proposed by Warland, Herrmann, & Willits (1975) regarding the fact that individuals of higher socio-economic status complain more frequently than others. However, as an opposing argument one might claim that students from private schools come from backgrounds of wealthier families and does not need student loans. This reasoning might give grounds for the negatively significant variable.

Asian and Indian or Alaskan

In the fourth regression *Asian* and *Indian or Alaskan* are significant at the 5% level. When *Indian or Alaskan* increases by one unit in the fourth regression, *Student loan complaints* decreases by 0.0018, and 0.0007 when *Asian* increases by one unit. This means that the higher the proportion of *Indian or Alaskan* in a state, the lower the degree of student loan complaints. The results are opposite for the variable *Asian* and increases the amount of complaints. As opposed to the mortgage analysis this is in line with the study by Liu & McClure (2001), which found that people from collectivistic cultures tended to complain less.

Except for Alaska, the differences between states regarding the amount of people within “Indian or Alaskan” are very small. In addition, there are a lot less student complaints per student loan in Alaska compared to the other states. These two factors in conjunction might explain some of the reason behind the relationship between *Indian or Alaska* and the dependent variable.

As for *African American*, the variable is insignificant, which contrasts with the findings in the study by Ayres, Lingwall, & Steinway (2013). They found that complaint rates were significantly higher in areas with higher concentrations of African Americans.

6.5 Other Results

In addition to the main analyses, we execute supplementary regressions to obtain a broader picture of the research question and perhaps draw a better conclusion. In this subsection, we will compare the different regressions with our main estimation model. The regression outputs are illustrated in table 6.3.

Table 6.3: Different Regressions of Mortgage Complaints on Financial Literacy and Socio-economic Factors

<i>Mortgage complaints</i>	<i>Main estimation model</i>	<i>Without cluster</i>	<i>With the outlier</i>	<i>One-year regression</i>
Final grade	-8.80e-06* (5.54e-06)	-8.80e-06*** (3.64e-06)	-6.52e-06 (6.45e-06)	-7.16e-06 (5.89e-06)
Unemployment rate	.0052265*** (.0017977)	.0052265*** (.0011882)	.0069701*** (.0025391)	.0065937*** (.0017227)
Families below poverty level	-.0043078** (.0023672)	-.0043078*** (.0013145)	-.0060122** (.0031504)	-.0053516*** (.0020694)
Debt-to-income	.0003159*** (.0000911)	.0003159*** (.0000447)	.0003626*** (.0001061)	.0001665*** (.0000739)
Business school	1.67e-10*** (8.39e-11)	1.67e-10*** (5.23e-11)	1.68e-10*** (8.38e-11)	1.81e-10*** (6.80e-11)
Bachelor's degree	.0003456 (.0003251)	.0003456 (.0003284)	.0004839* (.0003775)	.00138** (.0007923)
Legal occupation	6.88e-08*** (1.61e-08)	6.88e-08*** (9.25e-09)	7.21e-08*** (1.73e-08)	5.79e-08*** (1.40e-08)
10-year lagged economic education	.0000528*** (.0000271)	.0000528*** (.0000116)	.0000591** (.0000308)	.0000107 (.0000191)
10-year lagged personal finance	-9.83e-06 (.0000154)	-9.83e-06 (9.78e-06)	-.0000104 (.0000165)	2.32e-06 (.0000169)
African American	.0004079 (.0003928)	.0004079*** (.0001865)	.000367 (.0004437)	.0006854*** (.0002746)
Indian or Alaskan	-.0014586* (.0010611)	-.0014586*** (.0005265)	-.0017982** (.0011238)	.0010715 (.0010939)
Asian	.0018139*** (.000301)	.0018139*** (.0002606)	.0018763*** (.0003165)	.0017688*** (.0003986)
Internet users above 15 years	.0006802 (.0006827)	.0006802** (.0004101)	.0005871 (.0007495)	.0006407 (.0006992)
Median income	-1.09e-10 (4.57e-09)	-1.09e-10 (3.17e-09)	-6.12e-09 (8.15e-09)	-1.10e-08** (6.38e-09)
Year dummies	Y	Y	Y	Y
<i>Sigma/Z</i>	<i>10.26</i>	<i>22.12</i>	<i>4.54</i>	<i>11.40</i>

Dependent variable is the number of mortgage complaints over number of households with a mortgage (Consumer Financial Protection Bureau, 2018). Year dummies are included in all regressions to control for heterogeneity in complaints across time. * denotes significance at the 20% level, ** denotes significance at the 10% level and *** denotes significance at the 5% level.

6.5.1 Regression Without Cluster Command

We cluster the individual variable (state) in our main estimation model. We do another regression without cluster to detect any differences. In table 6.3 we include both of the regressions to compare the two. The significance level of the coefficients from the two regressions are fairly different, as the robust standard errors in the clustered regression are

somewhat higher. This is because the cluster command takes the heteroscedasticity into account. As a result, the significance of the variables decreases.

The main difference in the two outputs, is that *Final grade* and all of the demographic variables are significant at the 5% level in the regression model without the cluster command. However, as discussed earlier, the model with the cluster command is the most appropriate model. If we do not control for heteroscedasticity and autocorrelation, we risk the variance of the coefficients to be biased.

6.5.2 Regression with Outlier

As mentioned, the data set contains one outlier in number of complaints (Maine 2016). Outliers can be legitimate observations and can sometimes be the most interesting observations. However, one should be careful when dealing with an outlier. We run a regression with the outlier to see if we get different results. As we can see from table 6.3, there are minor differences between doing the regression with and without the outlier. The biggest difference is the change in *10-year lagged economic education*. The coefficient's significance decreases from a 5% level to a 10% level.

6.5.3 One-year Regression

Since two of the variables are time-constant, *Final grade* and *Legal occupation*, we perform a one-year regression to see if we get a significantly different result regarding the variables. The *Final grade* variable is from 2016, hence the regression is done with data from the corresponding year. The data collected regarding *Legal occupation* originate from 2017. However, as we do not expect there to be any considerable year-to-year changes, we include the variable. We control for the same variables as in the previous analysis and the results are illustrated in table 6.3. As we can see from the table, there are few differences. The two noteworthy changes are the presence of the positively significant *African American* at the 5% level and the disappearance of the significance of *10-year lagged economic education*.

6.5.4 First-difference Regression

Finally, we perform a regression that analyzes the change in the variables. The dependent variable is the *Growth rate in mortgage complaints*, while the independent variables are first-difference variables. The reason for running this regression is to see if the changes in the

independent variables have a relation with the change in the mortgage complaints. We used the following formula to generate the dependent variable:

$$\text{Growth in Mortgage complaints} = \frac{\text{Mortgage complaints}_t}{\text{Mortgage complaints}_{t-1}} - 1 \quad (15)$$

All of the usual variables are included, except for *Legal occupation* and *Final grade*. Since these two variables are time-constant, they are omitted when creating the first-difference variables.

Table 6.4: First-difference Regression of Growth in Mortgage Complaints on Financial Literacy and Socio-economic Factors

<i>Growth in mortgage complaints</i>	(1)	(2)	(3)	(4)	(5)
Diff unemployment rate	.8829503 (1.408662)	.8118199 (1.423537)	.7914737 (1.424863)	.1714013 (1.442797)	.1311049 (1.451614)
Diff families below poverty level	1.911573 (7.33629)	1.890667 (7.376575)	1.291246 (7.406509)	.8094759 (7.392423)	.7484171 (7.476824)
Diff debt-to-income	.0449787 (.0916091)	.0424364 (.0928397)	.0423952 (.0929174)	.0329176 (.0935719)	.0286117 (.0950026)
Diff business school		-4.43e-07 (7.82e-07)	-4.58e-07 (7.83e-07)	-4.48e-07 (7.84e-07)	-4.69e-07 (7.88e-07)
Diff bachelor's degree		.1235737 (.2361678)	.1680867 (.2392117)	.2203377 (.2405022)	.2190128 (.2413638)
Diff 10-year lagged economic education			.0054921 (.0132628)	.0299816 (.0239004)	.0301229 (.0242226)
Diff 10-year lagged personal finance			.0284452 (.0239527)	.0046638 (.0131984)	.0045891 (.0132967)
Diff African American				2.116936 (7.499309)	2.451078 (7.560889)
Diff Indian or Alaskan				27.36792*** (11.88302)	27.85837*** (11.95302)
Diff Asian				-5.986749 (10.17299)	-5.721844 (10.24105)
Diff internet users above 15 years					.2596307 (.435965)
Diff median income					-2.03e-07 (3.34e-06)
<i>Sigma/Z</i>					

As illustrated in formula (7), the dependent variable is the *Growth in number of Mortgage complaints* (Consumer Financial Protection Bureau, 2018). Year dummies are included in all regressions to control for heterogeneity in complaints across time. * denotes significance at the 20% level, ** denotes significance at the 10% level and *** denotes significance at the 5% level.

As we can see from table 6.4, *Diff Indian or Alaskan* is the only significant variable at a 5% level in the fifth regression.

7. Conclusion

In this section, we provide a summary of the thesis and its most important findings together with some concluding remarks regarding the results of the research question. We discuss the limitations and presents direction for future research.

7.1 Summary

Ever since the 70's, where pioneers within the field of consumer complaint behavior such as Hirschman (1970) and Day & Landon Jr. (1976) started laying the foundations, there has been a vast amount of succeeding literature. The literature has predominantly focused on customer satisfaction and consumer complaint behavior. Especially, studies regarding determinants of customer satisfaction accompanied by the types of responses and actions that are carried out by dissatisfied consumers, are dominant in the literature. However, the amount of contemporary studies on how financial literacy and other socio-economic characteristics relate to complaint behavior towards financial-services providers is very limited. Consequently, with more available data and a new approach, the main purpose was to provide some information which can enlarge the existing literature. Based on the overall aim of this thesis, the following research question was formulated:

How do financial literacy and other socio-economic characteristics relate to mortgage and student loan complaints?

Based on existing literature, mostly from the mid 70's, there are several consumer characteristics which can be related to complaints. As referred to in the literature review and in our hypotheses, some of them are: upscale socio-economic and highly educated people, people with more time on their hands and people with higher income. In addition to finding similarities with existing studies, the study by Ayres, Lingwall, & Steinway (2013) found that areas with a higher proportion of African American complained more frequently.

The data consists of over 218,000 mortgage complaints and 34,000 student loan complaints originated from the Consumer Financial Protection Bureau over a time period of six years from 2012-2017. For collecting relevant and suitable data we have made use of several sources. As a result, the panel data set included a total of 22 variables corresponding to the different states across the specified years. Before the analysis was conducted, we presented

some descriptive statistics of the data. As the data set is characterized as panel data, we tested the OLS assumptions and panel data effects. After the tests, we discovered that the data, regarding the mortgage analysis, were both slightly skewed and included a presence of heteroscedasticity and autocorrelation. Consequently, we decided that it was most appropriate to perform a truncated regression with a lower limit at zero, with clustered standard errors, robust to the exposed heteroscedasticity and autocorrelation. We did not include clustered standard errors regarding the student loan analysis, as the corresponding data had no presence of heteroscedasticity and autocorrelation. After conducting and analyzing the regressions, we attempted to constitute regression models to answer our research question.

7.2 Main Findings and Concluding Remarks

The significant variables for the dependent variable, *Mortgage complaints*, were *Unemployment rate*, *Debt-to-income*, *Business school*, *Legal occupation*, *10-year lagged economic education* and *Asian*, in addition *Families below poverty level* also revealed to be significant in the majority of the regressions. As discussed to a greater extent in the analysis, it turned out that *Final grade* was insignificant, which contradicted our second hypothesis.

As for *Unemployment rate*, we found a positive relationship indicating that the higher the number of unemployed people in a state, the higher the number of complaints. The positive relationship is in line with the study by Stokes (1974), mainly because he suggested that people with more time on their hands tend to complain more.

Regarding *Legal occupation*, *Business school* and *10-year lagged economic education*, the regression outcome indicates that they had a positive relationship with the dependent variable. This implies that states with a higher degree of inhabitants with a legal occupation, more inhabitants with business school degrees and that had better state economic education programs, complain more. These results are in agreement with our first and second hypothesis, and the studies by Warland, Herrmann, & Willits (1975) and Stokes (1974), which suggest that individuals of higher socio-economic status and higher education levels complain more frequently than others.

Families below poverty level had a negative relationship with *Mortgage complaints*. This indicates that states with a higher degree of families of this status yield fewer complaints. *Families below poverty level* were very close to the 5% significance cutoff point, and *Debt-*

to-income had a positive significant relationship. Presumably, both these results are because consumers with a lower disposable income, are more vulnerable to financial distress and shocks to their economy. Hence, they could be more inclined to complain or take action if such unfortunate instances should occur.

The results we obtained from the analysis are generally consistent with the existing studies, especially the variables *Unemployment rate*, *Families below poverty level*, *Business school*, *10-year lagged economic education* and *Legal occupation*. This implies that areas with (1) more upscale socio-economic consumers, (2) more highly educated consumers and (3) consumers with more time on their hands tend to complain more. Financial knowledge on the other hand, ended up insignificant. The results contradicted our second hypothesis, as we assumed that better financial knowledge would lead to a higher rate of complaints, as better-informed consumers would to a higher degree recognize dishonorable practices and potentially regulatory violations. However, the insignificant result may owe to the fact that they are better at distinguishing good financial products and services from bad, which may neutralize the tendency to complain.

Despite the difference in sample size and the usage of a proxy in the dependent variable, we found several comparable results between the analyses. Especially, regarding the variables *Unemployment rate*, *Debt-to-income* and *Families below poverty level*, which were in agreement with our hypotheses and theories presented in the literature review. The result regarding the significance of the variable *Private school* was unexpected. Which as discussed, were in contrast with the theory proposed by Warland, Herrmann, & Willits (1975) regarding the fact that individuals of higher socio-economic status complain more frequently than others, in addition to our first hypothesis. However, as an opposing argument one might claim that students from private schools come from backgrounds of wealthier families and does not need student loans. This reasoning might give grounds for the negatively significant variable.

7.3 Limitations

First of all, our results are not at any individual level. Hence, our results may not be able to reveal any consumers' reasons to complain, or which factors that may influence the consumers' choice of making a complaint or not. Furthermore, it is also debatable whether we can find a causality regarding whether the reason for more complaints is caused by the combination of different consumer groups or that the consumer groups are being mistreated

by lenders in different states. However, on the state level, the regressions have certainly unveiled interesting findings.

The CFPB is relatively young and has not collected consumer complaints for many years. Hence, the time period covered in this study and the size of our model could be a limitation, especially considering that some of the estimators may suffer from small finite sample biases (Héroux, 2018). However, we have performed different econometric approaches that mitigate different issues.

The data consists of 22 different variables corresponding to the different states across the specified years. As our research question was regarding financial literacy, we collected different types of suitable data. Since these types of data exclusively existed at the state level, and the fact that there could be considerable differences regarding financial literacy within each state, one might argue that it might have exerted influence towards our results, as opposed to if we had data available at a deeper geographical level. In other words, it would have been valuable to the study if there were financial literacy data available on a more detailed level.

As mentioned in section 4, the variable *Final grade* is time-constant as it has only been conducted on one occasion. Hence, the variable does not change across the years, and could contain lower degree of explanation power. Furthermore, the time-constant variable had severe implications for the thesis. For instance, we could not run the fixed effect estimation model because the time-constant variables would have been omitted. However, as the Director of Center for Financial Literacy at Champlain College, John Pelletier, wrote to us in an email “We have only done the adult report card once in 2016, and hope to do it again in 2019 or 2020... A yearly report does not make much sense - a lot would not change.” (Pelletier, Personal communication, 2018) The same remain valid for the time-constant variable *Legal occupation*, as well as for the variables concerning the state programs, which only change every second year. In these variables one would not expect considerable year-to-year changes. Nevertheless, these variables cause the panel data to be less balanced.

Finally, we expect that there is a large set of other variables that may influence the amount of consumer complaints. For practical reasons it is very complicated to identify all of them, especially as some might not even be estimated as they might be hard to measure.

7.4 Recommendation for Future Research

This thesis has focused on financial literacy and other determinants of consumer complaint behavior. For the future, it can be interesting to investigate how to capture the impact of people improving their financial knowledge and its impact on consumer complaints. In other words, as improved financial knowledge might lead individuals to better distinguish good financial products and services from the bad, which consumers has the tendency to complain less towards, it can be interesting to study how to capture the effect of improved financial knowledge.

As the CFPB is relatively young and an initiative among other regulations and actions taken subsequent to the Great Recession, as well as conceptions from the media, it has probably influenced consumer behavior. Hence, it would be interesting to investigate complaint behavior in the financial markets in the past and after the Great Recession, as this changed opinions and attitudes towards the industry (Kurt, 2018). Furthermore, an interesting study would be to study the effect of the CFPB itself. How has this agency reshaped consumer complaint handling, behavior and the financial system itself?

The variable *Debt-to-income* proved to be significant in our results. According to the Financial Ombudsman Service (2018) there are a lot of complaints regarding affordability which often is linked with financial hardship. On many occasions, complaints originate from lending that was unaffordable from the outset and should not have been allowed by the lender in the first place. Thus, an interesting study would be to investigate whether differences between state regulations towards lending practices, or differences between financial institution lending practices would affect consumer complaint behavior.

Since Andreassen & Streukens (2013) suggest that internet accessibility increases the likelihood that consumers will take action towards a supplier, a study regarding the effect of internet availability and the evolution of social media platforms with respect to consumer complaint behavior, would be an interesting topic.

8. References

- Amadeo, K. (2018, November 14). *Dodd-Frank Wall Street Reform Act*. Retrieved from thebalance.com: <https://www.thebalance.com/dodd-frank-wall-street-reform-act-3305688>
- Amemiya, T. (1973). Regression Analysis when the Dependent Variable is Truncated Normal. *Econometrica*, vol 41, issue 6, Pages: 997-1016.
- Anderson, E. W., & Sullivan, M. W. (1993). The Antecedents and Consequences of Customer Satisfaction for Firms. *Marketing Science*, Vol. 12, No. 2, Pages: 125-143.
- Andreassen, T. W., & Streukens, S. (2013). On-line Complaining: Understanding the Adoption Process and the Role of Individual and Situational Characteristics. *Journal of Service Theory and Practice*.
- Andrews, E. L. (2009, June 30). *Banks Balk at Agency Meant to Aid Consumers*. Retrieved from nytimes.com: <https://www.nytimes.com/2009/07/01/business/economy/01regulate.html>
- Arora, N. (2016). *In statistics, what is the difference between collinearity and correlation?* Retrieved from quora.com: <https://www.quora.com/In-statistics-what-is-the-difference-between-collinearity-and-correlation>
- Ayres, I., Lingwall, J., & Steinway, S. (2013). Skeletons in the Database: An Early Analysis of the CFPB's Consumer Complaints. *John M. Olin Center for Studies in Law, Economics, and Public Policy Research Paper No. 475*.
- Barunik, J. (2010). Introduction to Cointegration. *Applied Econometrics*. Charles University.
- Bauer College of Business. (2015). Panel Data Methods. Houston, Texas.
- Bearden, W. O., & Teel, J. E. (1983). Selected Determinants of Consumer Satisfaction and Complaint Reports. *Journal of Marketing Research*, Vol. 20, No. 1, Pages: 21-28.
- Berg, O. T. (2017, January 29). *Rettsvesen i USA*. Retrieved from snl.no: https://snl.no/Rettsvesen_i_USA

-
- Bureau of Labor Statistics. (2018, December 1). *Occupational Employment Statistics Query System*. Retrieved from data.bls.gov: <https://data.bls.gov/oes/#/home>
- Cappeli, P., Barankay, I., & Lewin, D. (2018, September 10). *How the Great Recession Changed American Workers*. Retrieved from knowledge.wharton.upenn.edu: <http://knowledge.wharton.upenn.edu/article/great-recession-american-dream/>
- Chavent, M., Liquet, B., Kuentz, V., & Saracco, J. (2012). ClustOfVar: An R Package for the Clustering of. *Journal of Statistical Software*.
- Churchill Jr. , G., & Surprenant, C. (1982). An Investigation into the Determinants of Customer Satisfaction. *Journal of Marketing Research (JMR)*, 19, Pages: 491-504.
- Consumer Financial Protection Bureau. (2018, 30 November). *consumerfinance.gov*. Retrieved from Consumer Complaint Database: <https://www.consumerfinance.gov/data-research/consumer-complaints/search/>
- Council for Economic Education . (2016). *The 2016 National State of Financial & Economic Education*. Retrieved from surveyofthestates.com: <http://www.surveyofthestates.com/#debt>
- Dartmouth College. (2018). Always Control for Year Effects in Panel Regressions. Hanover, New Hampshire.
- Day, R. L., & Bodur, M. (1978). "Consumer Response to Dissatisfaction With Services and Intangibles". *NA - Advances in Consumer Research Volume 05*, eds. Kent Hunt, Ann Abor, MI : Association for Consumer Research, Pages: 263-272.
- Day, R. L., & Landon Jr., E. (1976). Collecting Comprehensive Consumer Compliant Data By Survey Research. *NA - Advances in Consumer Research Volume 03*, eds. Beverlee B. Anderson, Cincinnati, OH : Association for Consumer Research, Pages: 263-268.
- Emmons, W. R. (2005, January). Consumer-Finance Myths and Other Obstacles to Financial Literacy. *Saint Louis University public law review*, Pages: 1-33.
- Financial Ombudsman Service. (2018, December 5). *financial-ombudsman.rg.uk*. Retrieved from Financial hardship and unaffordable lending: <https://www.financial->

ombudsman.org.uk/publications/technical_notes/financial-hardship-unaffordable-lending.htm

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2005). *Applied Longitudinal Analysis*. Hoboken: John Wiley & Sons.

Grace-Martin, K. (2018, November 27). *Outliers: To Drop or Not to Drop*. Retrieved from theanalysisfactor.com: <https://www.theanalysisfactor.com/outliers-to-drop-or-not-to-drop/>

Gronhaug, K. (1977). Exploring Consumer Complaining Behavior: a Model and Some Empirical Results. in *NA - Advances in Consumer Research Volume 04*, eds. William D. Perreault, Jr., Atlanta, GA : Association for Consumer Research, Pages: 159-165.

Grubbs, F. E. (1969). Procedures for Detecting Outlying Observations in Samples. *Technometrics*, Vol. 11, No. 1. Retrieved from *Technometrics*, Vol. 11, No. 1, Pages: 1-21.

Hair Jr., J., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate Data Analysis (3rd ed)*. New York: Macmillan.

Hart, C. W., Heskett, J. L., & Sasser Jr., E. W. (1990). The Profitable Art of Service Recovery. *Harvard Business Review*, Vol. 68 Issue 4, Pages: 148-156.

Héroux, M. (2018, April 19). *SMALL SAMPLE SIZES AND THE BIAS OF SMALL NUMBERS*. Retrieved from scientificallysound.org: <https://scientificallysound.org/2018/04/19/small-sample-sizes-and-the-bias-of-small-numbers/>

Hess, A. (2017, July 3). *This is the age most Americans pay off their student loans*. Retrieved from cnbc.com: <https://www.cnbc.com/2017/07/03/this-is-the-age-most-americans-pay-off-their-student-loans.html>

Hirschman, A. O. (1970). Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States. *Cambridge: Harvard University Press*.

-
- Investopedia. (2018, April 11). *Consumer Financial Protection Bureau - CFPB*. Retrieved from investopedia.com: <https://www.investopedia.com/terms/c/consumer-financial-protection-bureau-cfpb.asp>
- Johnston, J., & DiNardo, J. (1997). *Econometric Methods - 4th Edition*. The McGraw-Hill Companies, Inc.
- Kasman, M., Heuberger, B., & Hammond, R. A. (2018). *A Review of Large-Scale Youth Financial Literacy Education Policies and Programs*. The Brookings Institution.
- Kurt, D. (2018, September 13). *How the Financial Crisis Affected Millennials*. Retrieved from investopedia.com: <https://www.investopedia.com/insights/how-financial-crisis-affected-millennials/>
- Landon Jr., E. (1980). The Direction of Consumer Complaint Research. *NA - Advances in Consumer Research Volume 07*, eds. Jerry C. Olson, Ann Abor, MI : Association for Consumer Research, Pages: 335-338.
- Levesque, T., & McDougall, G. H. (1996). Determinants of customer satisfaction in retail banking. *International Journal of Bank Marketing*, Vol. 14 Issue: 7, Pages: 12-20.
- Liu, R. R., & McClure, P. (2001). Recognizing cross-cultural differences in consumer complaint behavior and intentions: an empirical examination. *Journal of Consumer Marketing*, Vol. 18, Issue: 1, Pages: 54-75.
- Lusardi, A., & Mitchell, O. S. (2014). The Economic Importance of Financial Literacy: Theory and Evidence. *J Econ Lit*, 52(1), Pages: 5-44.
- Maute, M. F., & Forrester Jr., W. R. (1993, June). The structure and determinants of consumer complaint intentions and behavior. *Journal of Economic Psychology*.
- Mayer, M. (2010, June). Unbalanced Panel Data Models. Vienna, Austria: Department of Economics - University of Vienna.
- Meier, S., & Sprenger, C. (2010). Present-Biased Preferences and Credit Card Borrowing. *American Economic Journal: Applied Economics*, Vol. 2, No. 1, Pages: 193-210.

- Moore, D. L. (2003). Survey of Financial Literacy in Washington State: Knowledge, behavior, Attitudes, and Experiences. *Knowledge, Behavior, Attitudes and Experiences. Washington State University Social and Economic Sciences Research Center Technical Report*, Pages: 3-39.
- Moore, E. M., & Shuptrine, F. (1984). Disconfirmation Effects on Consumer Satisfaction and Decision Making Processes. *NA - Advances in Consumer Research Volume 11*, eds. Thomas C. Kinnear, Provo, UT : Association for Consumer Research, Pages: 299-304.
- Murnane, R. J., & Reardon, S. F. (2018, January). Long-Term Trends in Private School Enrollments by Family Income.
- National Telecommunications and Information Administration. (2018, June 6). *Digital Nation Data Explorer*. Retrieved from [ntia.doc.gov: https://www.ntia.doc.gov/data/digital-nation-data-explorer#sel=adultInternetUser&disp=map](https://www.ntia.doc.gov/data/digital-nation-data-explorer#sel=adultInternetUser&disp=map)
- O'Donoghue, T., & Rabin, M. (1999). Doing It Now or Later. *The American Economic Review*, Vol. 89, No. 1, Pages: 103-124.
- Oliver, R. L. (1977). Effect of expectation and disconfirmation on postexposure product evaluations: An alternative interpretation. *Journal of Applied Psychology*, Vol 62(4), Pages: 480-486.
- Pelletier, J. (2016). *2016 National Report Card on Adult Financial Literacy*. Center for Financial Literacy Champlain College.
- Pelletier, J. (2018, November 9). Personal communication.
- Persyn, D., & Westerlund, J. (2008). Error-correction–based cointegration tests for panel data. *The Stata Journal* 8, Number 2, Pages: 232-341.
- Pesaran, H. (1987). Econometrics. In *The New Palgrave Dictionary of Economics*, Pages: 1-25).
- Rutledge, S. L. (2010). Consumer Protection and Financial Literacy - Lessons from Nine Country Studies. *The World Bank Policy Research Working Paper 5326*.

-
- Siniavskaia, N. S. (2013). Homes Millennials Buy - 2013 American Housing Survey.
- Stokes, R. C. (1974). *Consumer Complaints and Consumer Dissatisfaction*. Phoenix: The Food and Drug Law Institute.
- Thaler, R. H., & Shefrin, H. M. (1981). An Economic Theory of Self-Control. *Journal of Political Economy*, vol. 89, no. 2, Pages: 392-406.
- The Consumer Financial Protection Bureau. (2017). *The CFPB strategic plan, budget, and performance plan and report*. Washington, D.C.
- The Consumer Financial Protection Bureau. (2018, November 15). *About Us*. Retrieved from consumerfinance.gov: <https://www.consumerfinance.gov/about-us/>
- The Consumer Financial Protection Bureau. (2018). *Bureau of Consumer Financial Protection - Annual performance plan and report*. Washington D.C.
- The Consumer Financial Protection Bureau. (2018, November 10). *Complaints*. Retrieved from consumerfinance.gov: <https://www.consumerfinance.gov/complaint/>
- The Consumer Financial Protection Bureau. (2018, November 13). *Creating the Consumer Bureau*. Retrieved from consumerfinance.gov: <https://www.consumerfinance.gov/about-us/the-bureau/creatingthebureau/>
- The Consumer Financial Protection Bureau. (2018). *Semi-Annual Report of the Bureau of Consumer Financial Protection*. Washington D.C.
- The Consumer Financial Protection Bureau. (2018, November 10). *The Bureau*. Retrieved from consumerfinance.gov: <https://www.consumerfinance.gov/about-us/the-bureau/>
- The Federal Reserve. (2018, October 4). *EFA: Household Debt*. Retrieved from federalreserve.gov: https://www.federalreserve.gov/releases/z1/dataviz/household_debt/state/table/
- Trochim, W. M. (2018, November 9). *Reliability*. Retrieved from socialresearchmethods.net: <http://www.socialresearchmethods.net/kb/reliable.php>

- Uniform Law Commission. (2018, November 28). *About the ULC*. Retrieved from uniformlaws.org:
<http://www.uniformlaws.org/Narrative.aspx?title=About%20the%20ULC>
- United States Census Bureau. (2018, 8 December). *American FactFinder*. Retrieved from factfinder.census.gov:
<https://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t>
- United States Census Bureau. (2018, September 6). *Poverty Thresholds*. Retrieved from census.gov: <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html>
- University of Canterbury. (2018). Heteroscedasticity. Retrieved from <http://www.math.canterbury.ac.nz/~m.reale/econ324/Topic2.pdf>
- Warland, R. H., Herrmann, R. O., & Willits, J. (1975). Dissatisfied consumers: Who gets upset and who takes action. *The Journal of Consumer Affairs*, 9(2), Pages: 148-163.
- World Bank. (2005). *Financial Liberalization: What Went Right, What Went Wrong?* Washington D.C: World Bank.

Appendix

Appendix A

Table A.1: Snapshot of the Collected Data

Year	State	Mortgage complaints	Final grade	Unemployment rate	Families below poverty level	Debt-to-income
2012	Alabama	0,00036084	64,5	0,08	0,139	1,46
2013	Alabama	0,000230907	64,5	0,072	0,143	1,46
2014	Alabama	0,001001332	64,5	0,068	0,146	1,46
2015	Alabama	0,000251819	64,5	0,072	0,145	1,37
2016	Alabama	0,001511092	64,5	0,08	0,14	1,37
2017	Alabama	0,000638069	64,5	0,044	0,14	1,37
2012	Alaska	0,000731401	84,5	0,071	0,066	1,565
2013	Alaska	0,000975819	84,5	0,07	0,068	1,565
2014	Alaska	0,001440432	84,5	0,069	0,07	1,565
2015	Alaska	0,001382819	84,5	0,07	0,07	1,685
2016	Alaska	0,000775467	84,5	0,071	0,07	1,82
2017	Alaska	0,000454919	84,5	0,072	0,07	1,685
2012	Arizona	0,000541442	74,5	0,083	0,124	1,82
2013	Arizona	0,00026406	74,5	0,077	0,13	1,82
2014	Arizona	0,000175894	74,5	0,068	0,133	1,82
2015	Arizona	0,000307398	74,5	0,077	0,133	1,685
2016	Arizona	0,000295885	74,5	0,083	0,129	1,82
2017	Arizona	0,000389266	74,5	0,049	0,129	1,82
2012	Arkansas	0,00049038	64,5	0,076	0,141	1,285
2013	Arkansas	0,001372109	64,5	0,072	0,144	1,285
2014	Arkansas	0,000742477	64,5	0,06	0,143	1,285
2015	Arkansas	0,000824662	64,5	0,072	0,142	1,165
2016	Arkansas	0,000461207	64,5	0,076	0,138	1,285
2017	Arkansas	0,000274392	64,5	0,037	0,138	1,285
2012	California	0,000491654	78	0,104	0,115	2,01
2013	California	0,000365106	78	0,089	0,12	2,01
2014	California	0,000335395	78	0,075	0,123	2,01
2015	California	0,000972586	78	0,089	0,122	1,685
2016	California	0,001625932	78	0,104	0,118	1,82
2017	California	0,00099602	78	0,048	0,118	1,685

The table illustrates the structure of the data set. The data set contains of 300 observations from 50 states from the period 2012-2017, in addition to 22 different variables.

Appendix B

Tables B.1 & B.2: Variance Inflation Factor Tests

VIF-test	VIF score	VIF-test	VIF score
White	53.11	Families below poverty level	7.67
Asian	50.08	Final grade	5.72
African American	23.32	Median income	4.42
Native Hawaiian	15.72	Internet users above 15	2.27
Families below poverty level	9.00	Bachelor's degree	2.19
Indian or Alaskan	6.14	African American	2.17
Final grade	6.01	10-year lagged economic education	1.91
Median income	5.14	Business school	1.83
Business school	2.84	Legal occupation	1.69
Internet users above 15 years	2.37	Unemployment rate	1.54
Bachelor's degree	2.21	Asian	1.46
10-year lagged economic education	1.92	Indian or Alaskan	1.40
Legal occupation	1.83	10-year lagged personal finance	1.31
Unemployment rate	1.68	Debt-to-income	1.30
10-year lagged personal finance	1.34		
Debt-to-income	1.31		
Mean VIF score	11.50	Mean VIF score	2.63

These tables show the VIF score from the multicollinearity test on the mortgage complaints variables. A VIF score above ten implies that the variable is highly collinear with other independent variables.

Table B.3: Breusch-Pagan & Wooldridge Test Results

Breusch-Pagan test for heteroscedasticity	Value
Chi2(1)	43.63
Prob > Chi(2)	0.00***
Wooldridge test for serial correlation	Value
F(1, 49)	20.49
Prob > F	0.00***

This table shows the results from the Breusch-Pagan test for heteroscedasticity and the Wooldridge test for serial correlation.

* denotes significance at the 20% level, ** denotes significance at the 10% level and *** denotes significance at the 5% level.

Table B.4: Results from Im-Pesaran-Shin Unit-root Tests

Mortgage complaints	Final grade	Unemployment rate	Median income	Business school	Internet users above 15 years	10-year lagged economic education
0.9974	*	0.0005	0.9998	1	0.9958	*
Bachelor's degree	Legal occupation	Debt-to-income	African American	Indian or Alaskan	Families below poverty level	10-year lagged personal finance
0.7485	*	*	0.0571	0.0031	0.1086	*
Asian						
0.4159						

This table shows the p-values from the Im-Pesaran-Shin unit-root test for each variable. A p-value below 0.1 denotes a significance at the 10% level, a p-value below 0.05 denotes a significance at the 5% level, and a p-value below 0.01 denotes a significance at the 1% level. * denotes either that the variable is time-constant, or does not always include year-to-year changes.

Table B.5: Cointegration Test Results

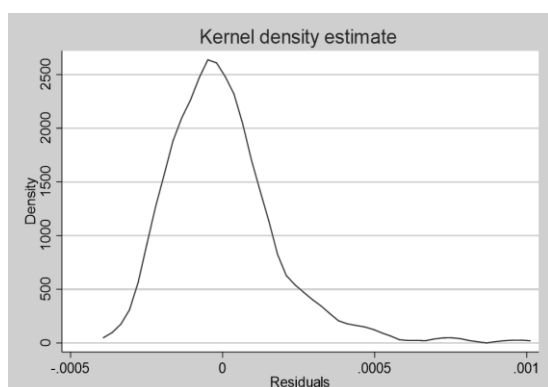
Test	Statistic	P-value
Modified Dickey-Fuller test	2.5709	0.005***
Dickey-Fuller test	-1.0564	0.145
Augmented Dickey-Fuller test	0.6616	0.254
Unadjusted modified Dickey-Fuller test	0.1180	0.453
Unadjusted Dickey-Fuller test	-3.2971	0.001***

This table shows the Kao test for cointegration. * denotes significance at the 20% level, ** denotes significance at the 10% level and *** denotes significance at the 5% level.

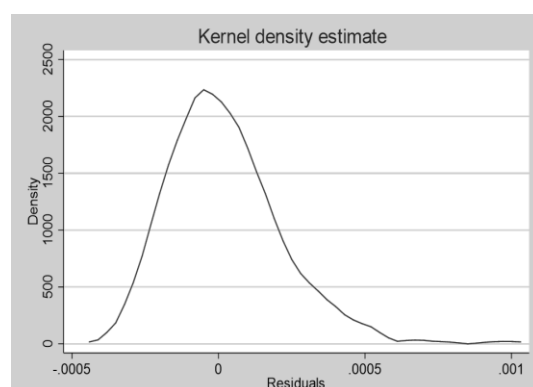
Table B.6: Skewness Table

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	Adj Chi2(2)	Prob > Chi2
Mortgage complaints	300	0.0000	0.0007	49.71	0.0000***
Final grade	300	0.6942	0.0000	60.06	0.0000***
Unemployment rate	300	0.9150	0.0016	9.10	0.0106***
Families below poverty level	300	0.0025	0.0059	14.37	0.0008***
Debt-to-income	300	0.3257	0.1685	2.88	0.2371
Business school	300	0.0000	0.0000	-	0.0000***
Bachelor's degree	300	0.0000	0.0006	25.29	0.0000***
Legal occupation	300	0.0000	0.0146	33.08	0.0000***
10-year lagged economic education	300	0.3087	0.0000	51.89	0.0000***
10-year lagged personal finance	300	0.0000	0.9819	19.72	0.0001***
African American	300	0.0000	0.1189	37.26	0.0000***
Indian or Alaskan	300	0.0000	0.0000	-	0.0000***
Asian	300	0.0000	0.0000	-	0.0000***
Internet users above 15 years	300	0.1280	0.0005	12.72	0.0017***
Median Income	300	0.1464	0.0088	8.34	0.0155***

This table illustrates the skewness of the variables in the data set. * denotes significance at the 20% level, ** denotes significance at the 10% level and *** denotes significance at the 5% level.

Figures B.1 & B.2: Kernel Density Estimates

This figure illustrates the Kernel density estimate with a regular regression



This figure illustrates the Kernel density estimate with a truncated regression

Appendix C

Tables C.1 & C.2: Variance Inflation Factor Test

VIF-test	VIF score	VIF-test	VIF score
Private school	222.01	Families below poverty level	7.56
Public school	221.94	Final grade	5.85
White	56.37	Median income	3.76
Asian	48.51	African American	3.01
African American	27.66	Private school	2.28
Native Hawaiian	10.33	Internet users above 15	2.20
Families below poverty level	8.72	1-year lagged economic education	2.04
Final grade	6.13	Asian	1.84
Indian or Alaskan	6.03	Unemployment rate	1.55
Median income	4.87	Legal occupation	1.45
Internet users above 15 years	2.30	Indian or Alaskan	1.35
1-year lagged economic education	2.20	Debt-to-income	1.34
Legal occupation	1.75	1-year lagged personal finance	1.32
Unemployment rate	1.72		
1-year lagged personal finance	1.49		
Debt-to-income	1.33		
Mean VIF score	38.96	Mean VIF score	2.73

These tables show the VIF score from the multicollinearity tests on the student loan complaints variables. A VIF score above ten implies that the variable is highly collinear with other independent variables.

Table C.3: Breusch-Pagan & Wooldridge Test Results

Breusch-Pagan test for heteroscedasticity	Value
Chi2(1)	1.22
Prob > Chi(2)	0.27
Wooldridge test for serial correlation	Value
F(1, 49)	0.003
Prob > F	0.9578

This table shows the results from the Breusch-Pagan test for heteroscedasticity and the Wooldridge test for serial correlation.

* denotes significance at the 20% level, ** denotes significance at the 10% level and *** denotes significance at the 5% level.

Table C.4: Results from Im-Pesaran-Shin Unit-root Tests

Student loan complaints	Final grade	Debt-to-income	African American	Legal occupation	Families below poverty level	1-year lagged economic education
1	*	*	0.0571	*	0.1086	*
Indian or Alaskan	Asian	Private school	Unemployment rate	Median income	Internet users above 15 years	1-year lagged personal finance
0.0031	0.4159	*	0.0005	0.9998	0.9958	*

This table shows the p-values from the Im-Pesaran-Shin unit-root test for each variable. A p-value below 0.1 denotes a significance at the 10% level, a p-value below 0.05 denotes a significance at the 5% level, and a p-value below 0.01 denotes a significance at the 1% level. * denotes either that the variable is time-constant, or does not always include year-to-year changes.

Table C.5: Cointegration Test Results

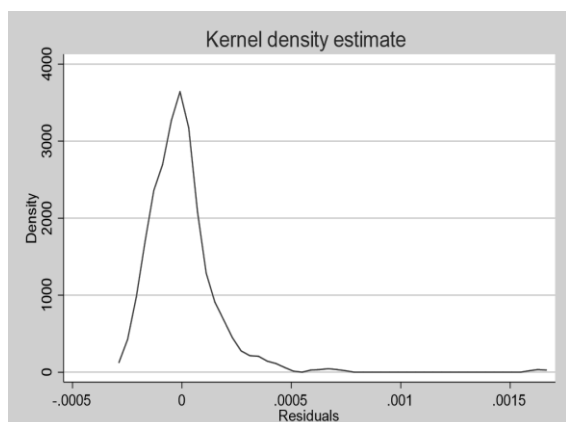
Tests	Statistic	P-value
Modified Dickey-Fuller test	3.0727	0.0011***
Dickey-Fuller test	0.7478	0.2273
Augmented Dickey-Fuller test	-0.2190	0.4133
Unadjusted modified Dickey-Fuller test	-2.7202	0.0033***
Unadjusted Dickey-Fuller test	-4.3927	0.0000***

This table shows the Kao test for cointegration.. * denotes significance at the 20% level, ** denotes significance at the 10% level and *** denotes significance at the 5% level.

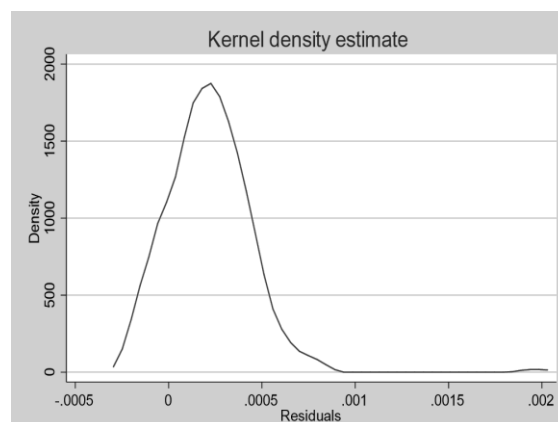
Table C.6: Skewness Table

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	Adj Chi2(2)	Prob > Chi2
Student loan complaints	300	0.0000	0.0000	-	0.0000***
Final grade	300	0.6942	0.0000	60.06	0.0000***
Unemployment rate	300	0.9150	0.0016	9.10	0.0106***
Families below poverty level	300	0.0025	0.0059	14.37	0.0008***
Debt-to-income	300	0.3257	0.1685	2.88	0.2371
Legal occupation	300	0.0000	0.0146	33.08	0.0000***
1-year lagged economic education	300	0.0691	0.0000	23.91	0.0000***
1-year lagged personal finance	300	0.9894	0.0000	-	0.0000***
African American	300	0.1280	0.0005	12.72	0.0017***
Indian or Alaskan	300	0.1464	0.0088	8.34	0.0155***
Asian	300	0.0000	0.1189	37.26	0.0000***
Internet users above 15 years	300	0.0000	0.0000	-	0.0000***
Median income	300	0.0000	0.0000	-	0.0000***
Private school	300	0.0001	0.1071	15.24	0.0000***

This table illustrates the results from the normality test for skewness and kurtosis. * denotes significance the 20% level, ** denotes significance at the 10% level and *** denotes significance at the 5% level.

Figures C.1 & C.2: Kernel Density Estimates

This figure illustrates the Kernel density estimate with a regular regression



This figure illustrates the Kernel density estimate with a truncated regression

Appendix D

Table D.1: Correlation matrix – Mortgage complaints

	Mortgage complaints	Final grade	Unemployment rate	Families below poverty level	Debt-to-income	Business school	Bachelor's degree	Legal occupation	Economic education	Personal finance	White	African American	Indian or Alaskan	Asian	Internet users above 15 years	Median Income
Mortgage complaints	1.0000															
Final grade	-0.1218	1.0000														
Unemployment rate	0.4168	-0.4521	1.0000													
Families below poverty level	-0.0554	-0.8380	0.3719	1.0000												
Debt-to-income	0.3838	0.0121	0.2364	-0.0433	1.0000											
Business school	0.4665	-0.2227	0.2594	0.1919	0.0402	1.0000										
Bachelor's degree	0.3402	0.3930	-0.1129	-0.5222	0.2256	0.2208	1.0000									
Legal occupation	0.5458	-0.1266	0.1868	-0.1077	0.0871	0.3851	0.4052	1.0000								
10-year lagged economic education	0.1456	-0.4986	0.1674	0.5809	-0.0057	0.3284	-0.2159	-0.0316	1.0000							
10-year lagged personal finance	-0.0697	-0.1112	0.0503	0.2521	0.0026	0.0078	-0.0759	-0.0892	0.3028	1.0000						
White	-0.4386	0.2874	-0.2474	-0.2187	-0.0651	-0.2921	-0.0674	-0.2659	-0.2265	-0.0094	1.0000					
African American	0.2764	-0.6467	0.3210	0.4928	-0.0336	0.2118	-0.1159	0.1830	0.4570	0.1253	-0.5376	1.0000				
Indian or Alaskan	-0.2707	0.1322	-0.1521	-0.0418	0.0672	-0.2678	-0.1462	-0.1064	-0.1205	-0.1992	0.0194	-0.3348	1.0000			
Asian	0.3570	0.2505	0.0229	-0.2340	0.0397	0.2075	0.2306	0.1556	-0.1218	-0.0502	-0.7132	-0.0991	-0.1047	1.0000		
Internet users above 15 years	-0.0729	0.5938	-0.2791	-0.6235	0.2216	-0.1287	0.4013	-0.0486	-0.3642	0.0165	0.3148	-0.3800	0.0033	-0.0375	1.0000	
Median Income	0.1838	0.6458	-0.3003	-0.7869	0.1763	0.0865	0.6465	0.2080	-0.3970	-0.1596	-0.0696	-0.3128	0.0114	0.3682	0.5843	1.0000

This table shows the correlation coefficients between the variables in the mortgage complaints data set.

Table D.2: Correlation matrix - Student loan complaints

	Student loan complaints	Final grade	Unemployment rate	Families below poverty level	Debt-to-income	Legal occupation	Personal finance	Economic education	White	African American	Indian or Alaskan	Asian	Internet users above 15 years	Median income	Private school	Public school
Student loan complaints	1.0000															
Final grade	0.0166	1.0000														
Unemployment rate	-0.3190	-0.4521	1.0000													
Families below poverty level	-0.1352	-0.8380	0.3719	1.0000												
Debt-to-income	0.0464	0.0121	0.2364	-0.0433	1.0000											
Legal occupation	0.1308	-0.1266	0.1868	-0.1077	0.0871	1.0000										
1-year lagged personal finance	0.0323	-0.2294	-0.0434	0.2512	0.0525	-0.0330	1.0000									
1-year lagged economic education	-0.0163	-0.4436	0.1790	0.5701	0.0690	0.0346	0.4359	1.0000								
White	0.0079	0.2874	-0.2474	-0.2187	-0.0651	-0.2659	0.0204	-0.2680	1.0000							
African American	0.0398	-0.6467	0.3210	0.4928	-0.0336	0.1830	0.3060	0.4829	-0.5376	1.0000						
Indian or Alaskan	-0.1471	0.1322	-0.1521	-0.0418	0.0672	-0.1064	-0.1654	-0.0961	0.0194	-0.3348	1.0000					
Asian	-0.0005	0.2505	0.0229	-0.2340	0.0397	0.1556	-0.2321	-0.0908	-0.7132	-0.0991	-0.1047	1.0000				
Internet users above 15 years	0.2737	0.5938	-0.2791	-0.6235	0.2216	-0.0486	-0.0838	-0.3564	0.3148	-0.3800	0.0033	-0.0375	1.0000			
Median income	0.3197	0.6458	-0.3003	-0.7869	0.1763	0.2080	-0.2489	-0.3911	-0.0696	-0.3128	0.0114	0.3682	0.5843	1.0000		
Private school	-0.1137	-0.0851	-0.0303	0.2390	0.1649	-0.2230	0.1172	0.1079	0.4141	-0.3140	0.3907	-0.4511	0.0315	-0.2089	1.0000	
Public school	0.1049	0.0870	0.0325	-0.2375	-0.1635	0.2288	-0.1194	-0.1113	-0.4186	0.3069	-0.3823	0.4597	-0.0336	0.2039	-0.9975	1.0000

This table shows the correlation coefficients between the variables in the student loan complaints data set.