ESSAYS ON THE SOCIETAL IMPLICATIONS OF ONLINE LENDING PLATFORMS

A Dissertation
Presented to
The Academic Faculty

by

Hongchang Wang

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
Scheller College of Business

Georgia Institute of Technology
August 2019

ESSAYS ON THE SOCIETAL IMPLICATIONS OF ONLINE LENDING PLATFORMS

Approved by:

Dr. Eric Overby, Advisor
Scheller College of Business
Georgia Institute of Technology

Dr. Marius Florin Niculescu
Scheller College of Business
Georgia Institute of Technology

Dr. Sabyasachi Mitra
Scheller College of Business
Georgia Institute of Technology

Dr. Mingfeng Lin
Scheller College of Business
Georgia Institute of Technology

Dr. Chris Forman
Dyson School of Applied Economics and Management
Cornell University

Date Approved: July 08, 2019

To my wife Jie and sons Allen and Evan.

ACKNOWLEDGEMENTS

Pursuing a second Ph.D. was a hard decision for me back to 5 years ago, however, thanks to the ITM group at Georgia Tech and especially the members of my dissertation committee, I went through this journey and am moving to the next stage of my career.

I am foremost grateful to my advisor Dr. Eric Overby, who is more than knowledgeable, insightful, supportive, and patient. He can not only help me shape and sharpen research questions at the high level but also works me through detailed technical challenges. He is so unique that he combines the abilities of thinking the big picture and working the tedious details. Under his supervision, I eventually grow to an independent researcher. He is also nice and encouraging, which is his mentoring principle as a Ph.D. advisor, and now it also becomes my belief.

I am extremely grateful to Dr. Saby Mitra, who can always provide constructive insights and feasible suggestions. His passion in research and carefulness in logical thinking pushes me to think harder and deeper. I am forever grateful to Dr. Marius Florin Niculescu, who is superb acknowledgeable in platform business models and smart in thinking through all types of research setting. He also provides many insights from the practical side of online lending markets, which deepen my understanding of this dissertation's research context. I also want to express my gratitude to Dr. Mingfeng Lin and Dr. Chris Forman. Dr. Lin's expertise in online lending and empirical research helps me shape my research questions and dip into underlying mechanisms. Dr. Forman's lectures and wisdoms in econometrics and identification are always excellent and help my build empirical skills. Special thanks go to Dr. DJ Wu, who always inspires me to explore for innovative but fundamental research topics and creates many research opportunities for me. Working with Dr. Wu brings me a higher level of overview of research and a deeper understanding of innovation. Dr Michael Smith serves as the mentor for me on teaching and shares with me many tips in managing an effective classroom.

Many thanks to all the other ITM faculty members, i.e. Dr. Sri Narasimhan, Dr. Yu (Jeffery) Hu, Dr Han Zhang, and Dr. Lizhen Xu. I am loving so much the diversity of the Georgia Tech ITM group and the seminars taught by these excellent researchers. I am forever grateful to the marketing professor Dr. Ajay Kohli and the strategy professor Dr. Alex Oettl and would like recommend their seminars to future ITM students. The Scheller staff Ursula Reynolds and Shannon Smith provide excellent support to Ph.D. students and endeavor to make my life easy. The fellow Ph.D. students provide tremendous supports in my journey and make me not to feel lonely.

Final thanks and the biggest credits go to my family, especially my wife Dr. Jie Liu. Without her love and support I have no way to complete this Ph.D. Absolutely no way. It is her taking the responsibility for the family so I can spend more time in my research. My in-laws and my parents not only provide monetary support for my study but also help take care of my two kids. The degree is under my name, but it should be indebted to my whole family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Information systems have always had great potential to disrupt industries and affect social welfare. This dissertation studies the societal implications of online lending platforms, which are enabled and supported by recent developments in two-sided online markets and automated underwriting technologies. These platforms underwrite borrowers automatically and match borrowers directly with investors willing to lend their capital by making borrowers' credit information transparent to investors. Online lending has the potential to not only increase lending market efficiency due to its nature of disintermediation and automation but also provide access to capital to traditionally underserved individuals and small businesses. However, various concerns still exist about online lending on several aspects, e.g. the risk of involving individual investors, the risk of automated underwriting and investment decision making, and the ultimate outcomes of online lending. This dissertation includes three interrelated essays which investigate the outcomes of online lending platforms on both borrowers and investors. More specifically, the first one looks at the borrowers' side and investigates how online lending influences bankruptcies of borrowers, the second one looks at the investors' side and investigates how the use of algorithmic trading in online lending platforms influences investing opportunities of individual investors, and the third one looks at the whole online lending market and investigates how political ideology and political distance influence investors' behaviors and market efficiency. This dissertation contributes to multiple IS research areas, including economics of online platforms, online behavioral bias, automated data-driven technologies, etc. This dissertation also provides insights to regulators on online lending

regulation and to practitioners on platform design. The specific research questions and main findings of each essay/chapter are summarized below.

Essay 1: **How Does Online Lending Influence Bankruptcy Filings?**

By providing quick and easy access to credit, online lending platforms may help borrowers overcome financial setbacks and/or refinance high-interest debt, thereby decreasing bankruptcy filings. On the other hand, these platforms may cause borrowers to overextend themselves financially, leading to a "debt trap" and increasing bankruptcy filings. To investigate the impact of online lending on bankruptcy filings, we leverage variation in when state regulators granted approval for a major online lending platform – Lending Club – to issue peer-to-peer loans. Using a difference-in-differences approach, we find that state approval of Lending Club leads to an increase in bankruptcy filings. A complementary instrumental variable analysis using loan-level data yields similar results. We find suggestive evidence that the ease of receiving a Lending Club loan causes some borrowers to overextend themselves financially, leading to bankruptcy. We also find that "strategic" borrowing – in which borrowers who are considering bankruptcy use a Lending Club loan to restructure their debt or to engage in last-minute consumption before they file – may play a role. Our results suggest that recent initiatives from online lending platforms to control how borrowers use loans, such as Lending Club's Direct Pay program that sends loan funds directly to creditors, can help these platforms provide safe and affordable credit. Our study adds to the literature that examines how online platforms influence society and the economy; it contributes to the literature that examines how financial products, services, and regulations influence bankruptcy filings; and it has policy implications for online lending design and regulation.

Essay 2: **How Does Algorithmic Trading Influence Investor Participation in Peer-to-Peer Online Lending Markets?**

Algorithmic trading has reshaped equity markets and had significant effects on market performance. In this paper, we examine the effect of algorithmic trading in online peer-to-peer lending markets. As the “peer-to-peer” label suggests, these markets were originally designed to be accessible to individual investors. However, because algorithmic trading is typically used by institutional investors with substantial resources, advances in algorithmic trading threaten to shut individual investors out of the market. Ironically, this could exacerbate inequalities in the financial system that peer-to-peer lending markets were designed to help eliminate. To study the effects of algorithmic trading, we examine the effect of an API upgrade on Prosper.com that facilitated algorithmic trading. Using a difference-in-differences strategy, we find that individual “manual” investors were crowded out of the most quickly-funded and typically best-performing loans after the API upgrade. However, the API upgrade may have increased the size of the market, thereby allowing individual investors to continue investing in the market, albeit for somewhat lower quality loans. Our study contributes to several emerging research areas, including online lending, algorithmic trading, data-driven decision making, and the effect of technology on social and financial inequality.

Essay 3: **Do Political Differences Decrease Market Efficiency? An Investigation in the Context of Online Lending**

We study whether political differences – which are becoming increasingly acute among Americans – inhibit market efficiency by examining whether investors in online lending markets are less likely to lend to borrowers whose political ideology (i.e., liberal or conservative) is likely to be different from their own. We leverage state-level legalization of

same-sex marriage as a natural experiment to investigate how investors in online lending markets respond to this signal of a state's "liberalness". Results of a difference-in-differences analysis show that: (1) investors make more bids (loan offers) to borrowers in states that legalize same-sex marriage in the days immediately after passage of the law; and (2) investors from politically similar states contribute more to this increase than do investors from politically dissimilar states. This suggests that political differences influence lending decisions in online lending markets, potentially preventing beneficial investor/borrower matches from being formed. To test the generalizability of these findings, we use all U.S. states and measure the number of bids from investors in each state to borrowers in each state. We use a gravity model to examine how political differences across states influence bidding behaviors. Results are consistent with the difference-in-differences analysis. Investors have a general preference for borrowers from liberal states, but this dissipates (and sometimes disappears) as the political distance between the investor and borrower states grows, particularly when the investor state is more conservative than the borrower state. We also investigate the mechanism driving the effects. We find evidence that investors' preference for borrowers from liberal states is because investors view "liberalness" as a sign of low credit risk. But we also find evidence that the negative effect of political distance on investor / borrower matching is purely preference-based, perhaps reflecting an in-group bias. Given the fast growth of online lending as well as the rapid increase in political polarization, understanding the impact of political differences on market outcomes yields important theoretical and practical implications.

CHAPTER 1. INTRODUCTION

Information systems have always had great potential to disrupt industries and affect social welfare. This dissertation studies the societal implications of online lending platforms, which are enabled and supported by recent developments in two-sided online markets and automated underwriting technologies. Online lending is an emerging business model that directly matches borrowers and investors. Online lending platforms post the information of borrowers who are seeking for money online for investors to make lending decisions. Online lending platforms have been established as a mainstay source of alternative funding (the total dollar amount of loans issued grew from $1.99 billion in 2010 to $15.91 billion in 2014.) and are expected to provide 8% of total unsecured consumer lending by 2020 (Demyanyk et al. 2017). As is often the case with new IT-enabled business models, online lending has potentially massive societal implications. Online lending might directly improve social welfare by expanding access to capital to previously underserved individuals and small businesses. It might also indirectly improve social welfare by reshaping the lending/credit industry and increasing lending market efficiency. However, empirical evidence about the outcomes of online lending is still understudied by researchers and unclear to regulators, which is why the U.S. Department of Treasury requested for information on online lending on 2015 and characterized the industry as "untested" in a follow-up white paper on 2016. Motivated by the great potential of online lending, this dissertation investigates the societal implications of online lending and explores for the mechanisms.

Different from traditional lenders, online lending platforms have combined two innovative features of modern developments in digital business models and information technologies. First, online lending utilizes the power of two-sided or peer-to-peer platforms and reduces transaction cost (service cost, searching cost, contract cost, etc.) by reducing information asymmetry and improving matching efficiency (See Figure 1-1).

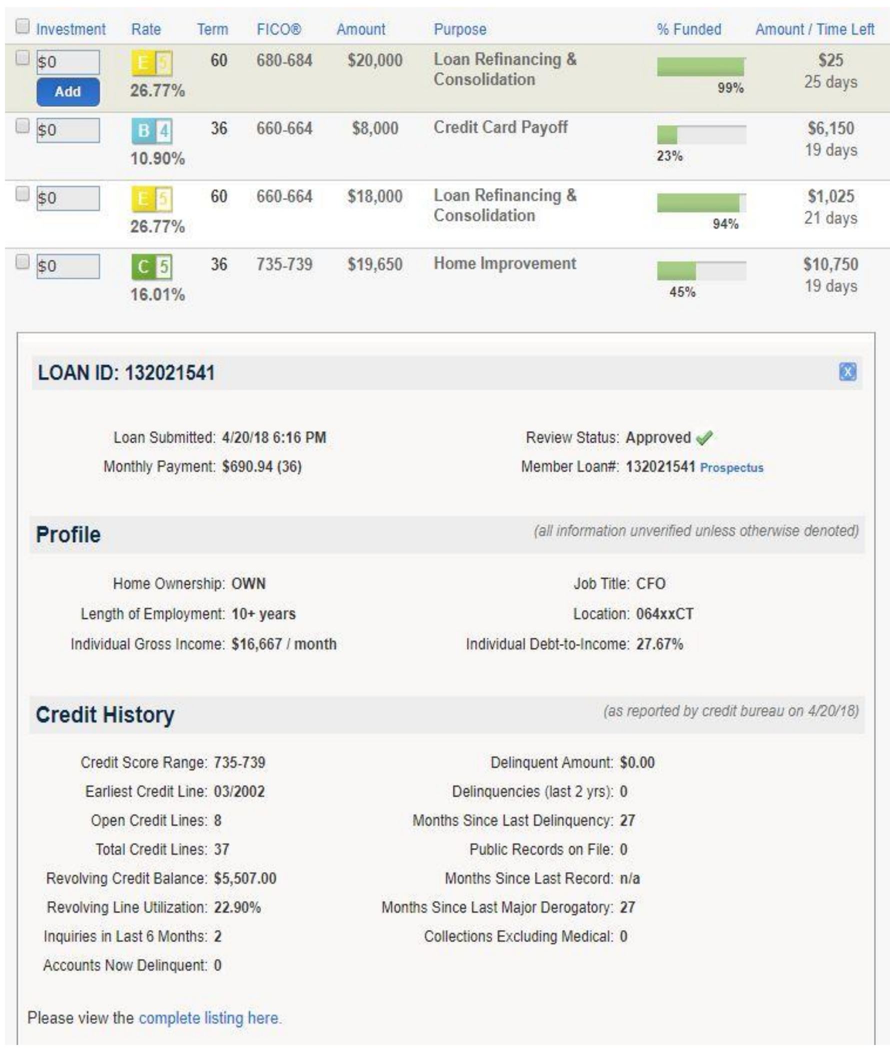


Figure 1-1. Loan listing from Lending Club

Second, online lending platforms use information technologies and big-data enabled algorithms to automate lending procedures, including borrower information collection, loan underwriting, loan issuance, and payment collection. These platforms also provide automated lending tools to their investors (See Figure 1-2). Automated lending not only reduces service cost but also provides capital to underserved populations who are previously restricted by thin credit profiles or inconvenient locations.

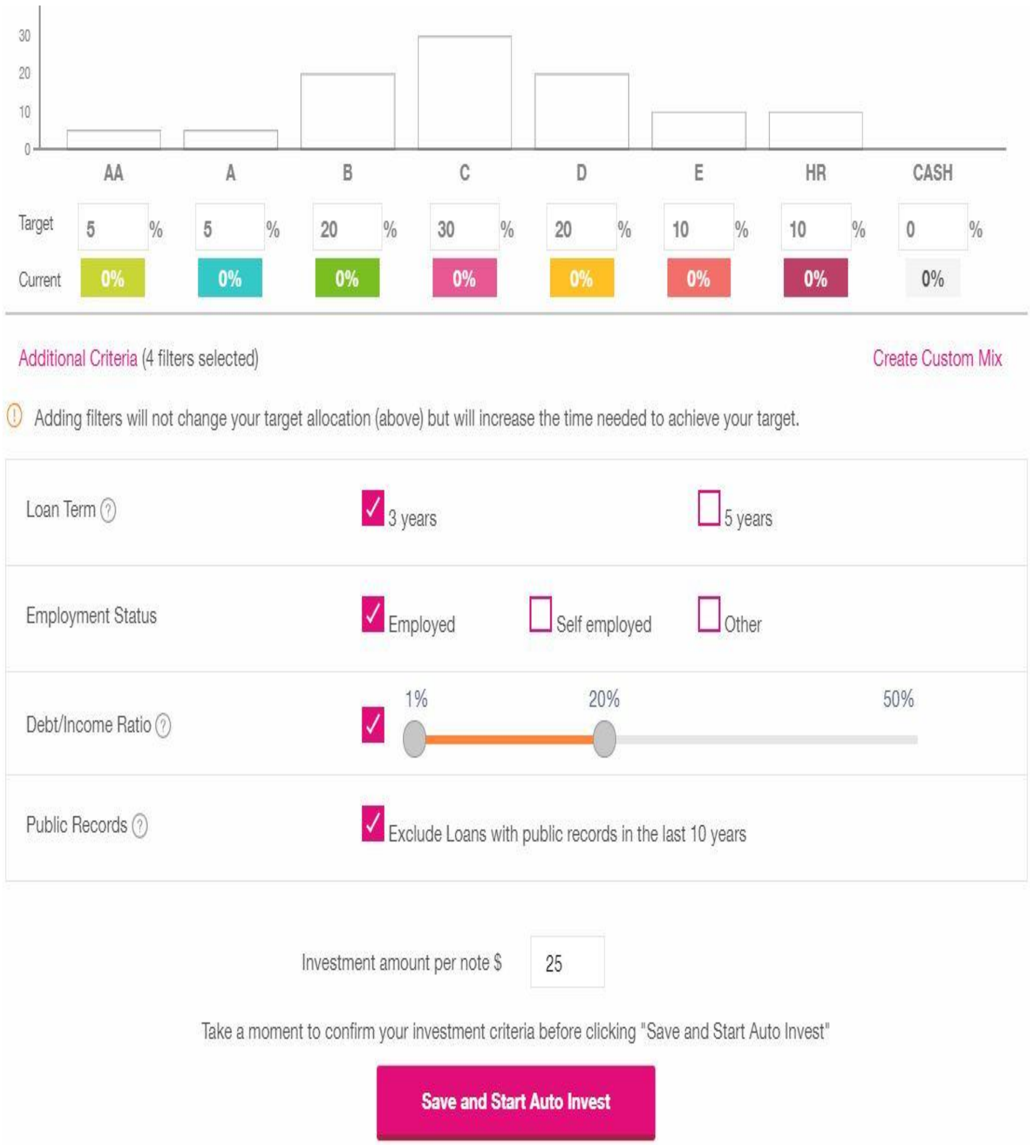


Figure 1-2 Auto invest setting from Prosper.com

These two features make online lending a significant player in the finance industry, but the societal implications are far beyond the success of online lending platforms themselves. The most important implication is the large social welfare improvement. Since online lending provides digital platforms for borrowers and removes geographical restrictions, it can expand access to capital to previously underserved populations who are restricted geographically. What’s more, online lending platforms are using automated algorithms to underwrite/price borrowers, which is likely to lead to low operational cost and high-quality lending decisions. This feature enables online lending to serve individuals and small businesses who are previously restricted by thin credit profiles or bad credit indicators. Due to these two channels, it is widely believed that online lending has the potential to directly improve social welfare. Another potential implication is industry disruption as well as the indirect social welfare improvement brought in by incumbent lenders. Incumbent lenders, either big or small, are learning from or responding to the online lending threats. Big lenders have started to build in-house automated technologies to support their own lending business, including but not limited to consumer lending, small business lending, and mortgage. Small lenders have started to cooperate with online lending platforms by directing their consumers to these platforms or underwriting their consumers with the automated lending technologies provided by these platforms. These disruptions and improvements, which are enlightened and triggered by online lending business, are reshaping the whole lending industry by transforming how the lending market works. It is reasonable to expect that this transformation can not only increase lending market efficiency but also further improve social welfare.

Despite the aforementioned potentials of online lending, there is no guarantee that online lending works in the expected way or can fulfill these potentials. Concerns and critics coexist with these potentials from the first day of the introduction of online lending. These concerns cover a broad range of topics, including the risk of matching naïve investors with uncreditworthy borrowers, the risk of automated underwriting technologies and lending decision tools, borrower monitoring and default risks, etc. Some anecdotal evidence also keeps reminding researchers and regulators that online lending may go wrong for different reasons. The exciting potentials and the reasonable concerns create an important and urgent research topic for researchers, i.e. how online lending really works and what its societal implications would be. This dissertation focuses on empirical approaches to investigate the societal implications of online lending. Essentially, this is to examine how the two features of online lending, i.e. two-sided matching marketplace and automated lending technology, work in the online lending context and what the outcomes they might bring in.

This dissertation is made up of three interrelated essays. Essay I focuses on an important indicator of financial health, i.e. bankruptcy filings, and investigates how online lending influences bankruptcy filings. This essay directly answers what economic/social outcomes online lending might bring in. Because the outcome of online lending is partially determined by whether investors can screen good borrowers from bad borrowers and whether investors can allocate their capital wisely, Essay II and III investigate how investors select borrowers. Due to the two features of online lending, two different approaches for lending have evolved across time. The peer-to-peer feature enables investors to manually evaluate borrowers, i.e. investors can make a decision based on all the information of a borrower/listing, while the automated lending feature enables investors to rely on lending tools or algorithms to select borrowers, i.e. investors are supported or even totally replaced by lending tools in selecting investors. The former one is the dominating approach in the early time of online lending. As online lending markets have attracted more borrowers and investors, the investing side changes from manual lending to automated lending. Regarding to the automated lending approach, Essay II focuses on automated lending tools and studies how these tools influence lending performance and market efficiency. Regarding to the manual lending approach, Essay III focuses on one important but understudied decision factor--political ideology and political distance, and studies how this information influences investors lending behavior and market efficiency.

This dissertation is motivated by the potential of emerging online lending platforms and investigates how these platforms work and what outcomes they might bring in. This dissertation provides not only managerial insights for practitioners and regulators but also theoretical implications for IS research. This dissertation adds to the growing IS literature of online lending and crowdfunding. The three essays contribute to multiple IS research areas, including economics of online platforms, dark side of IS, online behavioral bias, automated data-driven technologies, etc.

CHAPTER 2. HOW DOES ONLINE LENDING INFLUENCE BANKRUPTCY FILINGS?

How do you begin a new chapter using this template? What must you do to get the page numbers to act correctly? Below are the steps for making a new chapter.

2.1 Introduction

Online lending platforms match borrowers with investors willing to lend their capital. These platforms are sometimes referred to as peer-to-peer lending platforms because investors and borrowers are often individuals. Online lending platforms have the potential to expand access to credit to borrowers who are underserved by traditional credit sources such as banks as well as to provide better terms to all borrowers. Online lending is growing rapidly and represents as much as one-third of the U.S. market for personal loans.[1] Despite this, regulators are unsure about the impact of online lending and whether it should be encouraged. For example, the U.S. Department of Treasury issued a white paper in May 2016 that characterized the industry as "untested" and called for greater regulation.[2]

To provide some insight into the impact of online lending, we study the effect of online lending on bankruptcy filings. Online lending could plausibly decrease or increase bankruptcy filings. The optimistic view (i.e., that online lending decreases bankruptcy) is that the increased access to credit that online lending provides will help borrowers handle unanticipated financial setbacks and stave off bankruptcy. For some borrowers, including those who are traditionally underserved, an online loan may help them remain financially solvent during times of financial need. The relatively low interest rate of online loans

[1] See https://www.ft.com/content/4cf113a4-bf39-11e7-b8a3-38a6e068f464.
[2] See https://www.treasury.gov/connect/blog/Pages/Opportunities-and-Challenges-in-Online-Marketplace-Lending.aspx.

(compared to credit cards) may also help borrowers refinance existing debt, thereby reducing their debt burden and helping them avoid bankruptcy. The pessimistic view (i.e., that online lending increases bankruptcy) is that online lending will cause borrowers to take on more debt than they can service, driving them into bankruptcy. This could occur if online loans are issued to unqualified borrowers who cannot repay the loans. It could also occur if the quickness and convenience of obtaining an online loan causes otherwise qualified borrowers to overextend themselves financially, leading them into a "debt trap" and subsequent bankruptcy. Overall, the effect of online lending platforms on bankruptcy filings is theoretically ambiguous and warrants empirical examination.

We focus our analysis on Lending Club, which is the largest online lending platform and is representative of other online lending platforms such as Prosper and Funding Circle. We use two complementary identification strategies to examine Lending Club's impact on bankruptcy filings. First, we conduct a difference-in-differences analysis in which we exploit variation in when state regulators granted approval for Lending Club to issue peer-to-peer loans. We find that Lending Club approval leads to an increase in bankruptcy filings. Second, we use micro-level loan data published by Lending Club to examine the relationship between lending activity and bankruptcy filings. We use an instrumental variables approach to improve the causal interpretation of our results. We find that a one standard deviation increase in Lending Club loans is associated with an approximately 3% increase in bankruptcy filings. We find suggestive evidence that the increase in bankruptcy filings is because some borrowers become overextended financially after receiving a Lending Club loan (as opposed to their being inherently uncreditworthy). Some of the increase in bankruptcy may also be due to strategic borrowing in which borrowers who are considering bankruptcy use a Lending Club loan to restructure their debt or to engage in last-minute consumption (e.g., taking a vacation) before they file.

Our study contributes to the online lending literature as well as to the bankruptcy and household finance literature. First, as an increasing number of online lending platforms/technologies are created, it is important to study their effects on access to capital, funding allocation efficiency, and household financial stability (Burtch and Chan 2018; Butler et al. 2016; Kim and Hann 2018; Mollick and Robb 2016; Wei and Lin 2017). Our findings suggest that online lending platforms may have harmful effects and that design and/or regulatory changes may help these platforms provide safe and affordable credit. Indeed, near the end of our study period, Lending Club launched its Direct Pay program, which sends loan funds directly to the borrowers' creditor(s). This type of program could address (at least partially) the effect that we document. Second, our study contributes to the bankruptcy and household finance literature by investigating how online lending influences bankruptcy. This adds to existing studies that have examined how financial products/services such as payday loans and credit cards as well as regulatory changes such as bankruptcy reform and interstate banking deregulation influence bankruptcy filings (Dick and Lehnert 2010; Hynes 2012).

2.2 Background, Literature Review, and Motivation

We first discuss the literatures on bankruptcy and online lending. We then discuss how online lending platforms might affect bankruptcy.

2.2.1 Bankruptcy

Bankruptcy is a legal process used by individuals and businesses to resolve unpaid debts. In the United States, debtors can file under different chapters. Chapter 7 and Chapter 13 filings are the most common (Dobbie and Song 2015). Under Chapter 7, debtors liquidate nonexempt assets (e.g., their house may be foreclosed upon) in exchange for the discharge of most debts and protection from collection actions such as lawsuits, wage garnishment, and telephone calls. Under Chapter 13, debtors can avoid liquidating assets but must enter into a repayment plan for all or part of their debt.

Reasons for and implications of bankruptcy. Different theories have been proposed to explain why debtors file for bankruptcy, including strategic motive theory and adverse events theory. Strategic motive theory argues that debtors are motivated by financial benefits to file bankruptcy. These benefits include discharging (some) debt and stopping collection activities of creditors, including collection letters/phone calls/visits, wage garnishment, and other court orders (Dawsey et al. 2013; Lefgren and Mclntyre 2009). The surge of bankruptcy filings prior to the Bankruptcy Abuse Prevention and Consumer Protection Act (BAPCPA) in 2005 (which reduced the benefits of filing for bankruptcy for many) provides support for the strategic motive perspective (White 2009). Adverse events theory considers bankruptcy to be a consequence of growing financial distress, potentially driven by unemployment, increasing housing and medical costs, divorce, credit card debt, and/or unfair and abusive practices by lenders (Dick and Lehnert 2010). Several studies support this theory by showing that unemployment is a major contributing factor to bankruptcy (Himmelstein et al. 2005; Zhu 2011). Whatever the reasons for bankruptcy filings, the downsides of bankruptcy are substantial and discourage filing, including harm to the debtor's credit score (that may take years to repair) and the cost of filing (both financially and in terms of social stigma).

Programs, policies, etc. that affect bankruptcy. Given the negatives of filing bankruptcy, researchers have studied financial products, regulations, and market activities that might affect bankruptcy. One stream examines the effect of payday loans, credit cards, and online platforms. The effect of payday loans is inconclusive. There is evidence that payday loans increase consumer bankruptcy filings and that Chapter 13 bankruptcies decrease after payday loan bans (Morgan et al. 2012). However, the legalization of payday loans has been shown to reduce bankruptcy filings in counties with large military populations (Hynes 2012). Research on the effect of credit cards suggests that expanded credit card debt contributed to an increase in bankruptcy from 1980 to 2004 (White 2007). This may be because the pool of consumers who were issued credit cards became riskier over time (Livshits et al. 2016). Research on the effect of online platforms is nascent (with our study being one of the first), although a recent study indicates that a medical crowdfunding platform (GiveForward) reduces bankruptcy by helping individuals cover unexpected medical costs (Burtch and Chan 2018). Another stream examines the effect of banking deregulation and bank mergers on bankruptcy. Research has shown that competition brought about by deregulation prompted banks to adopt sophisticated credit rating technology, which they used to expand lending to previously excluded (and typically riskier) households. This explained at least 10% of the rise in bankruptcy rates between 1980 and 1994 (Dick and Lehnert 2010). Bank mergers have been shown to increase consumer bankruptcies because they destroy interpersonal, relational knowledge that lenders use to identify creditworthy borrowers and to shepherd them through financial difficulties (Allen et al. 2016).

2.2.2 Online Lending Platforms

Online lending platforms match borrowers with investors for personal or small business loans. Online lending is also referred to as peer-to-peer lending, loan-based crowdfunding, and marketplace lending (Morse 2015). To illustrate how online lending works, we describe the typical model pioneered by Lending Club. First, a borrower requests a loan by providing personal information and the desired loan amount. Second, the online lending platform analyzes the borrower's information to assess risk and to assign an interest rate. Third, lenders/investors choose which loans to fund, using their own capital. They decide the amount they want to lend and can invest as little as $25 in each loan. If enough investors want to lend money to the borrower, then s/he can get the loan. There is a growing body of research about online lending, including how it affects access to capital, how the design and operation of online lending platforms affect lending outcomes, and how investors behave.

Access to capital. Because online lending is an alternative funding source compared to banks, it has the potential to democratize access to capital (Mollick and Robb 2016). Indeed, research has concluded that online lending has penetrated areas likely to benefit from increased access to capital, including those with highly concentrated (i.e. weakly competitive) banking markets, those that are losing bank branches, and those of low socioeconomic status (Jagtiani and Lemieux 2017; Kim and Hann 2017). Online lending may also help borrowers with good access to traditional capital secure loans with attractive terms. For example, online lending borrowers from areas with good access to bank finance seek loans with low interest rates and (perhaps as a result) are more likely to prepay (Alyakoob et al. 2017; Butler et al. 2016).

Platform design, operation, and outcomes. Research has studied how the design and operation of online lending platforms affects lending outcomes. For example, underwriting for online loans is based on more information and is faster than traditional underwriting (Buchak et al. 2018; Jagtiani and Lemieux 2017). This allows more borrowers to receive credit at favorable terms; Jagtiani and Lemieux (2017) show that online lending technology is more likely to classify a subprime borrower into a better loan grade compared to traditional lenders. It also allows loans to be issued more quickly. Other research has studied whether online lending platforms should assign interest rates to borrowers (i.e., the posted price regime) or allow investors to propose their own interest rates to borrowers (i.e., the auction regime). Wei and Lin (2017) found that the posted price regime yields more matches between borrowers and investors but also yields higher default rates.

Investor behavior. Online lending investors rely on both traditional financial information and "soft" information to make lending decisions (Iyer et al. 2016). In addition to traditional factors such as credit scores, the decision process is influenced (and potentially biased) by several factors, including other investors' decisions, loan descriptions, borrowers' friendship networks, borrowers' demographics (including gender, race, and overall "appearance"), and the distance (geographical and cultural) between investor and borrower (Burtch et al. 2014; Duarte et al. 2012; Galak et al. 2011; Greenberg and Molick 2017; Harkness 2016; Hildebrand et al. 2017; Lin et al. 2013; Lin and Viswanathan 2016; Pope and Sydnor 2011; Younkin and Kuppuswamy 2017; Zhang and Liu 2012). Sophisticated and less sophisticated investors often rely on different information when determining whom to fund, but their investment returns are often similar (Lin et al. 2018; Mollick and Nanda 2016).

2.2.3 How Online Lending Might Affect Bankruptcy

There are multiple mechanisms through which online lending could affect bankruptcy filings (most of which we examine empirically in our analysis). These mechanisms relate to the characteristics of online lending borrowers and how they use online loans.

Characteristics of online lending borrowers. It is possible that many online loans are issued to high-risk borrowers who lack access to traditional capital. If these borrowers are inherently uncreditworthy and unable to repay the loan, then online lending should increase bankruptcy filings. This might occur if investors' biases (such as those noted above) cause them to fund high-risk borrowers. On the other hand, online lending may provide the capital necessary for these borrowers to handle unanticipated financial setbacks and to remain solvent during times of financial need, which could decrease bankruptcy filings.

Use of online loans. Regardless of whether a borrower is high-risk or not, how borrowers use online loans should influence bankruptcy filings. If borrowers use the loans to refinance high interest debt, then online lending would reduce their debt burden, leading to fewer bankruptcies. On the other hand, if borrowers use the loans to add to existing debt, then online lending would increase their debt burden, leading to more bankruptcies. To illustrate, consider the following example. Assume that person Z has an average credit risk profile and has $13,000 in credit card debt at a 20% interest rate. (According to the 2007 Consumer Bankruptcy Project, median credit card debt was $13,279 for bankruptcy filers.) Assume that Z gets a $13,000 online loan with a 13% interest rate. If Z pays off his credit card debt with the online loan, then he will have $13,000 in debt at a 13% rate instead of at a 20% rate. This improves his financial situation, helping him avoid a potential bankruptcy. However, if Z does not pay off his credit card debt and instead uses the online loan for purchases/vacation/etc., then he will have $13,000 in credit card debt at a 20% rate plus $13,000 in online loan debt at a 13% rate. This worsens his financial situation, potentially driving him into bankruptcy.

A recent working paper (Chava and Paradkar 2018) based on proprietary credit bureau data suggests a related possibility. It shows that many borrowers use online loans to pay off credit card debt, as intended. Because this increases borrowers' credit ratings, they receive – and often accept – additional offers of credit. Ironically, this often leads to greater aggregate credit card debt and subsequent default – and potentially to bankruptcy.

Another possibility is that online lending attracts borrowers who are considering bankruptcy. The relative ease and speed of receiving an online loan might tempt these borrowers to take a loan and then declare bankruptcy shortly thereafter. For example, a borrower might use an online loan to pay off his/her car loan – thereby swapping secured debt for unsecured debt – and then file for bankruptcy. This type of debt restructuring may be tempting, because it may protect a borrower's property (e.g., car) from immediate repossession after declaring bankruptcy. Or, a borrower considering bankruptcy might use an online loan to engage in last-minute consumption, such as taking a vacation, before declaring bankruptcy. If this "strategic" borrowing occurs, then online lending would increase bankruptcy filings.

Overall, it is unclear a priori whether online lending has a positive or negative effect on bankruptcy filings, or which mechanisms are responsible for any effect. This motivates our empirical examination.

2.3 Empirical Setting and Overview of Empirical Strategy

The online lending platform we investigate is Lending Club. We chose this platform for three reasons: 1) it is the largest online lending platform, 2) it was approved to issue peer-to-peer loans in different states at different times, which provides a natural experiment that we leverage to examine its effect on bankruptcy filings, and 3) it publishes micro-level loan data. We conduct two complementary analyses. First, we leverage the staggered approval of Lending Club across states in a difference-in-differences ("DD") analysis to examine its impact on bankruptcy. This strategy of exploiting staggered approval/entry has been implemented in several studies that investigate the impact of regulatory change and platform implementation (Bertrand et al. 2004; Burtch et al. 2016; Chan and Ghose 2014; Greenwood and Agarwal 2016). Second, we use Lending Club's loan data to examine the relationship between the level of Lending Club loans and bankruptcy filings, using instrumental variables to improve the causal interpretation of our results.

2.3.1 Difference-in-differences Analysis: Analyzing the Effect of Lending Club Regulatory Approval

Overview of Approach

Lending Club launched its platform in 2007. In April 2008, Lending Club entered a "quiet" period in which it suspended peer-to-peer lending until it registered with federal and state regulators as a licensed lender (or loan broker). During the quiet period, Lending Club funded some loans with its own money (instead of with investors' money), but these loans were few (see Appendix A-1 for an illustration). Lending Club pursued regulatory approval to resume peer-to-peer lending in all 50 states. By October 2008, it had received approval in 40 states plus the District of Columbia (DC). For 9 states, it received approval at different times between 2010 and 2016. For 1 state (Iowa), it had not received approval as of November 2018. Table 1 shows the quarter in which Lending Club received regulatory approval in each state. We gathered this information from Lending Club's blog, from news about Lending Club, and by using Lending Club's loan data to examine lending activity in each state over time. The variation in when states allowed Lending Club to resume peer-to-peer lending provides a natural experiment that we exploit to examine the impact of Lending Club on bankruptcy filings.

Table 2-1 Lending Club Approval by State

State	Approval Quarter	Approval Year (as coded for county-year analysis)	Approval Quarter (as coded for state-quarter analysis)
All states, except those listed below	2008-Q4	Not included	2009-Q1
Kansas	2010-Q4	2011	2011-Q1
North Carolina	2010-Q4	2011	2011-Q1
Indiana	2012-Q4	2013	2013-Q1
Tennessee	2013-Q1	2013	2013-Q2
Mississippi	2014-Q2	n/a; see text	2014-Q3
Nebraska	2015-Q2	2015	2015-Q3
North Dakota	2015-Q2	2015	2015-Q3
Maine	2015-Q3	2016	2015-Q4
Idaho	2016-Q1	2016	2016-Q2
Iowa	Not approved as of 2018-Q4	Not approved as of 2018-Q4	Not approved as of 2018-Q4

We constructed two panel data sets with different units of analysis: (1) a county-year panel, and (2) a state-quarter panel. We chose 2006 as the initial year for our analysis because the Bankruptcy Abuse Prevention and Consumer Protection Act (BAPCPA) took effect in October 2005. Many debtors rushed to file bankruptcy before this act took effect because it introduced an income test that limited which borrowers could file Chapter 7 bankruptcy.

For the county-year analysis, we limit our analysis to counties in the states that approved Lending Club no earlier than 2010. We study the period from 2006 to 2014. During this period, counties within 4 states (Kansas, North Carolina, Indiana, Tennessee) were "treated" with Lending Club approval while counties within 5 "control" states (Nebraska, North Dakota, Maine, Idaho, Iowa) were not.[3] Because bankruptcy filings and many of our control variables are only available at the county level on a yearly basis, we use a yearly panel. We estimate the effect of Lending Club' approval using a difference-in-differences approach, with the counties in the states in which Lending Club hadn't yet been approved or was not approved serving as the counterfactuals for the counties in the states in which Lending Club was approved. We use this sample for the following reasons. First, we observe each county for at least 4 years before Lending Club approval. This helps us assess whether pre-existing trends in bankruptcy filings might confound the effect of Lending Club approval. (Our pre-treatment observation window is shorter for the other 40 states.) Second, Lending Club was relatively well-established and likely to be known by prospective borrowers by 2010, when the first states in this analysis were treated. This increases the likelihood that Lending Club approval will have a detectable effect. Third, analysis at the county level allows us to control for county-level demographic and economic variables, thereby improving the precision of our estimate of Lending Club's

[3] We exclude Mississippi because its bankruptcy trends differ from the other states in this analysis. In the 9 focal states, bankruptcy filings per capita declined year-over-year from 2010 to 2014. This is not true for Mississippi, which experienced a pronounced increase in bankruptcy filings in 2013. This suggests a possible policy change or economic shock – specific to Mississippi – that could confound our estimation of the effect of Lending Club.

impact. Fourth, analysis at the county level improves our identification because Lending Club was approved at the state (not county) level. Thus, even if unobserved state-level factors influence both Lending Club approval and bankruptcy filings (thereby creating endogeneity concerns), these factors may not apply at the county level.

For the state-quarter analysis, we use data from all 50 states plus DC. This analysis complements the county-year analysis and addresses some of its shortcomings. First, it allows us to assess whether our results are idiosyncratic to the 9 states in the county-year panel. Second, because bankruptcy filings and control variables are available for each state (but not each county) on a quarterly basis, we can conduct this analysis by quarter rather than by year. This permits a more precise measure of when Lending Club was approved in each state. Third, we are able to extend the study period to 2015; we use 2014 as the final year in the county-year analysis in order to preserve a clear distinction between treated and control counties (see Table 2-1).

In both the county-year and state-quarter analyses, all of the control states (except for Iowa) approved Lending Club by at least 2016 (see Table 2-1). This suggests that there may be no dramatic difference between the control and treated states in terms of their overall attitude to online lending, only differences in how long it took Lending Club to receive the necessary regulatory approvals. This increases the likelihood that the control states are valid counterfactuals for the treated states.

Data and Variables

Bankruptcy filings. The key dependent variable is *bankruptcy filings per capita*, which is the number of bankruptcy filings per 1,000 people in county *i* in year *t* (or in state *i* in quarter *t*). We obtained bankruptcy filing data from the Administrative Office of the U.S. Courts website.[4] The U.S. Courts data distinguishes between non-business and business bankruptcy. We include both in our study because individuals might use Lending Club loans to fund their businesses. The data also distinguishes between chapters of bankruptcy (e.g., Chapter 7, 13, etc.) In addition to *bankruptcy filings per capita*, we also use the raw number of bankruptcy filings and the natural log of the raw number (plus 1 to account for values of 0). These measures are widely used in bankruptcy studies (Burtch and Chan 2018; Dick and Lehnert 2010).

Lending Club approval. The key independent variable is the *Lending Club available* dummy. For the county-year panel, this variable is 1 if Lending Club is available to borrowers in county *i* in year *t* and 0 otherwise. If Lending Club was approved during the first half of the year, then we consider it to be available that year and all subsequent years. If Lending Club was approved during the *second* half of the year, then we consider it to be *un*available that year but available all subsequent years. For robustness, we used an alternative coding rule (see below), which does not affect our results. For the state-quarter panel, we defined Lending Club as available in the first full quarter after Lending Club approval.

Demographic and socioeconomic information. We include several demographic and socioeconomic control variables gathered from the U.S. Census and the U.S. Bureau of Labor Statistics. These include population, demographic mix, unemployment rate, and

[4] Because federal courts have jurisdiction over personal and business bankruptcy cases, the U.S. Courts data fully represent the bankruptcy activity of individuals and businesses in the U.S. Bankruptcy filing data are available from http://www.uscourts.gov/statistics-reports/caseload-statistics-data-tables.

median household income. This allows us to control for alternative explanations and improves the precision of our estimate of the effect of Lending Club.

Table 2-2 Variables and Descriptive Statistics for the State-Quarter Sample

Variable	Source	Min	Max	Mean	Median	St. Dev
Bankruptcy variables						
Bankruptcy filings per capita	U.S. Bankruptcy Courts	0.11	3.10	0.83	0.75	0.43
Bankruptcy filings - raw	U.S. Bankruptcy Courts	80	69359	5406	3403	6962
Bankruptcy filings - natural log	U.S. Bankruptcy Courts	4.39	11.15	7.90	8.13	1.32
Lending Club variable						
Lending Club available (binary variable)	Lending Club data, news reports	0	1	0.60	1	0.49
Demographic control variables (from U.S. Census)						
Population (in millions)	Population Estimates Program	0.52	39.14	6.08	4.33	6.82
% age 60 & above	American Community Survey	10.7	24.9	18.6	18.7	2.3
% white	American Community Survey	24.9	96.3	77.4	79.3	13.7
Socioeconomic control variables (from U.S. Census; *Unemployment rate* from Bureau of Labor Statistics)						
Number of employed individuals (in millions)	County Business Patterns	0.21	14.36	2.24	1.50	2.46
Average monthly earnings (in thousands)	Quarterly Workforce Indicators	2.54	7.50	3.80	3.65	0.71
Median household income (in thousands)	Small Area Income and Poverty Estimates	34.47	74.55	52.06	50.26	8.67
Unemployment rate (percent)	Local Area Unemployment Statistics	2	15.4	6.5	6.2	2.3
% below high school attainment	American Community Survey	7.2	22.1	13.2	12.5	3.5
% housing units rented	American Community Survey	23.7	58.8	32.7	31.5	5.5
% housing units with mortgage	American Community Survey	31.8	54.8	44.5	44.8	4.7

Each variable is available at both the county and state levels. Most are available quarterly. For those available only yearly, we use the yearly value to proxy for quarterly values in the state-quarter analysis. These variables, their sources, and descriptive statistics are listed in Table 2-2 (state-quarter) and Table 2-3 (county-year). Because the last three control variables in Table 2-2 are not always reported for small counties, we include them only in the state-quarter analysis. The results of the state-quarter analysis are similar if we drop these variables.

In the interest of transparency and so that others can replicate our results, we provide the data and regression commands for most of our analyses. See Appendix A-2 for details.

Empirical Strategy

Our baseline specification is (1). We describe the specification for the county-year analysis; the description is analogous for the state-quarter analysis.

$$Y_{it} = \alpha + \beta LC_{it} + T_t + S_i + \gamma X_{it} + \varepsilon_{it} \tag{1}$$

In (1), Y_{it} is the number of bankruptcy filings per capita in county i in year t. LC_{it} is a dummy variable equal to 1 if Lending Club is available to borrowers in county i during year t and 0 otherwise. α is a constant term, T_t are year fixed effects, S_i are county fixed effects, X_{it} are control variables, γ are associated coefficients, and ε_{it} is the error term, which is clustered at the county level (and alternatively at the state level, which does not affect our results). The year fixed effects account for general changes over time that affected bankruptcy filings, which were substantial during the study time period because of the Great Recession. The county fixed effects account for unobserved time-invariant characteristics of each county. The control variables help us better estimate the effect of Lending Club. The parameter of interest is β, which represents the average treatment effect of Lending Club approval on bankruptcy filings.

Ideally, our sample would include treated and control counties that had parallel bankruptcy trends prior to Lending Club's approval in the treated counties. This would increase the likelihood that a change in bankruptcy filings in the treated counties following Lending Club approval was caused by Lending Club approval. Pre-treatment bankruptcy trends are parallel in the state-quarter analysis (as we will show below) but not in the county-year analysis. The left panel of Figure 2-1 shows that for the county-year analysis, the bankruptcy trends before 2011 (when the first treated state was treated) are not parallel. To account for this, we used coarsened exact matching (CEM) to build a matched sample for the county-year analysis in which the pre-treatment trends for treated and control counties are parallel. We matched treated and control counties based on their annual values of *bankruptcy filings per capita* and *unemployment rate* from 2006 to 2010, i.e., prior to Lending Club approval. We coarsened these variables into equally spaced bins and only matched treated and control counties within the same bins. We also matched on *population* (in 2006). The matching yielded 42 matched strata that each contained at least one treated and one control county. These strata contained 259 counties in total, which comprise the matched sample: 97 treated counties matched to 162 control counties (see Appendix A-3 for the list). Because we observe each county for 9 years, this yielded a panel of 2,331 observations. A characteristic of matching procedures, reflected in our study, is that a treated observation is sometimes matched to more than one control observation and vice versa. To accommodate this, the CEM algorithm generates weights for each county, which we use in our analysis (Iacus et al. 2012). (We also matched each treated county to a single control county, which affects our sample size but not our results.) We checked the balance between treated and control counties in the matched sample by running several regressions of the form $Y_i = \alpha + \beta Treated_i + \varepsilon_i$. Y_i is one of the matching variables (e.g., *bankruptcy filings per capita* in 2006, *bankruptcy filings per capita* in 2007, etc.) and $Treated_i = 1$ for treated counties and 0 for control counties. We included the 259 counties in the matched sample and used the weights generated by the CEM procedure. Appendix A-4 shows that we achieved good balance on not only the matching variables but also on variables not included in the matching procedure. Table 2-3 shows the descriptive statistics of the county-year matched sample.

Table 2-3 Variables and Descriptive Statistics for the County-Year Matched Sample

Variable	Min	Max	Mean	Median	St. Dev
Bankruptcy variables					
Bankruptcy filings per capita	0	8.11	2.23	2.09	1.14
Bankruptcy filings - raw	0	2413	98.32	28	220.68
Bankruptcy filings - natural log	0	7.79	3.48	3.37	1.46
Lending Club variable					
Lending Club available (binary variable)	0	1	0.15	0	0.36
Demographic control variables (from U.S. Census)					
Population (in thousands)	1.28	543.99	35.49	13.90	63.94
% age 60 & above	9.81	36.95	23.92	23.87	5.19
% white[a]	42.62	99.11	94.35	96.26	6.15
Socioeconomic control variables (from U.S. Census; *Unemployment rate* from Bureau of Labor Statistics)					
Number of employed individuals (in thousands)	0	318.69	13.62	3.85	31.42
Average monthly earnings (in thousands)	1.61	6.30	2.79	2.72	0.54
Median household income (in thousands)	25.25	97.94	46.00	45.13	7.71
Unemployment rate (percent)	1.30	14.80	4.97	4.50	1.76

Notes: Means and standard deviations are calculated using the CEM weights.
[a] The mean and median are higher than might be expected because the unit of analysis is county-year. This means that a small county (which is likely to have high percentage of white people) contributes similarly (depending on the CEM weights) to the statistics as does a large county (which is less likely to have high percentage of white people).

The right panel of Figure 2-1 shows that the pre-treatment bankruptcy trends for treated and control counties in the county-year matched sample are parallel; we provide more formal evidence of parallel pre-treatment trends below. Figure 2-1 also foreshadows our conclusion that Lending Club approval leads to more bankruptcies than would otherwise occur.

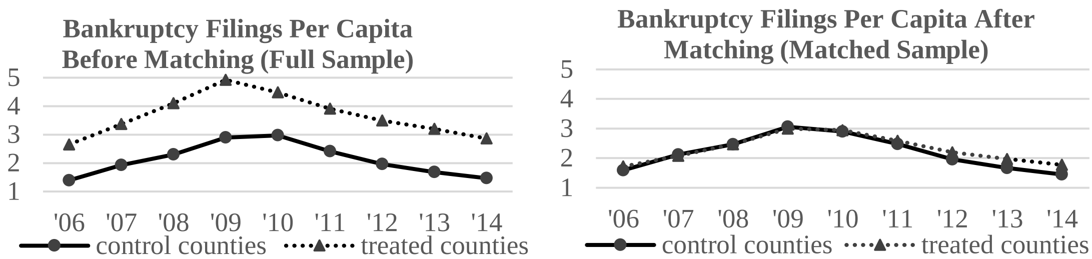


Figure 2-1 Trends in Bankruptcy Filing per Capita for the County-Year Analysis, Before and After Matching

Leads/lags model. To examine more formally whether the treated and control counties (or states) follow parallel pre-treatment trends, we implemented a leads/lags model, shown in specification (2) (Autor 2003). This approach is widely used in studies that use a DD strategy (Chan and Ghose 2014; Greenwood and Agarwal 2016). As above, we describe the specification for the county-year analysis. The description is analogous for the state-quarter analysis, with one change noted below.

$$Y_{it} = \alpha + \sum_{\tau=-5}^{-2} \rho_\tau \, LC_{it+\tau} + \sum_{\tau=0}^{3} \rho_\tau \, LC_{it+\tau} + T_t + S_i + \gamma X_{it} + \varepsilon_{it} \qquad (2)$$

Specification (2) mirrors (1) except that we replace βLC_{it} with $\sum_{\tau=-5}^{-2} \rho_\tau \, LC_{it+\tau} + \sum_{\tau=0}^{3} \rho_\tau \, LC_{it+\tau}$. $LC_{it+\tau}$ is a dummy variable equal to 1 for observations in year t if year t is τ years after Lending Club approval (or for $\tau < 0$, $-\tau$ years before Lending Club approval). We withhold LC_{it-1} to avoid the "dummy variable trap". For example, we coded Lending Club as being approved in North Carolina in 2011. Thus, for any county i in North Carolina, $LC_{it-4} = 1$ for the year 2007 observations and 0 otherwise, $LC_{it+0} = 1$ for the year 2011 observations and 0 otherwise, $LC_{it+1} = 1$ for the year 2012 observations and 0 otherwise, etc. Because $LC_{it+\tau}$ always equals 0 for the control counties, the coefficient ($\rho_{-\tau}$) represents the average difference in *bankruptcy filings per capita* between treated and control counties

τ years prior to Lending Club receiving approval in the treated counties. Because LC_{it-7} and LC_{it-6} are rarely defined during our sample period, we collapse them into LC_{it-5} (results are robust if we estimate them separately). For the state-quarter analysis, we consider 5 pre-treatment and 5 post-treatment quarters. For states in which we observe more than 5 pre-treatment and/or post-treatment periods, we collapse the preceding and/or following quarters into the -5 and/or +5 time periods.

Results. Table 2-4 shows the results of our baseline model (specification 1) for the county-year matched sample and the state-quarter sample with both bankruptcy filings per capita and bankruptcy filings - natural log as the dependent variables. Lending Club approval has a positive and significant impact on bankruptcy filings. The per capita model indicates that Lending Club approval increases bankruptcy filings by 0.179 per thousand people in the county-year matched sample and 0.057 per thousand people in the state-quarter sample. This represents increases of 8.0% and 6.8%. The smaller effect size in the state-quarter analysis may be because this analysis contains treated observations from years in which Lending Club was very new (e.g., 2009 to 2010). Appendix A-5 reports the results of a Poisson model and a negative binominal model using bankruptcy filings – raw as the dependent variable. Results are similar. Across all models, the coefficient for unemployment rate is positive and significant, which is consistent with the bankruptcy literature that unemployment increases bankruptcy. The unreported time fixed effects also follow intuition: bankruptcy filings increased until 2009 and then decreased, likely due to the Great Recession.

Table 2-5 shows the results from the leads/lags model (specification 2) for *bankruptcy filings per capita*. We find evidence of parallel pre-treatment trends for both the county-year and state-quarter samples. In both samples, the coefficients for *Lending Club(−4)* (i.e., LC_{it-4}), *Lending Club(−3)*, and *Lending Club(−2)* are insignificant (recall that the omitted "baseline" dummy variable is *Lending Club(−1)*). This indicates that there is little to no meaningful difference in bankruptcy trends between the treated and control counties in the 4 time periods before Lending Club approval. Although the coefficient for *Lending Club(−5)* is significant in the county-year analysis, this difference existed long before treatment and should therefore not confound our results.

Table 2-4 Regressions Results for Bankruptcy Filings (Per Capita and Natural Log) for the County-Year Matched Sample and the State-Quarter Sample

Sample	**County-year matched sample**		**State-quarter sample**	
Dependent variable	**Bankruptcy filings per capita**	**Bankruptcy filings - natural log**	**Bankruptcy filings per capita**	**Bankruptcy filings - natural log**
Lending Club available	0.179 (0.059) ***	0.116 (0.028) ***	0.057 (0.031) *	0.108 (0.041) **
Population	-0.000 (0.004)	0.006 (0.002) ***	0.036 (0.042)	0.027 (0.061)
Number of employed individuals	-0.004 (0.007)	0.007 (0.003) **	-0.184 (0.086) **	-0.152 (0.112)
Average monthly earnings	0.143 (0.157)	0.001 (0.131)	0.085 (0.044) *	0.047 (0.060)
Unemployment rate	0.139 (0.019) ***	0.044 (0.011) ***	0.080 (0.016) ***	0.045 (0.011) ***
Median household income	-0.015 (0.008) *	-0.010 (0.004) **	0.001 (0.009)	-0.015 (0.014)
% age 60 & above	0.031 (0.024)	0.027 (0.014) *	0.001 (0.034)	-0.049 (0.048)
% white	0.034 (0.041)	0.040 (0.027)	0.025 (0.021)	0.029 (0.018)
% below high school attainment	n/a	n/a	0.061 (0.032) *	0.084 (0.038) **
% housing units rented	n/a	n/a	0.010 (0.023)	0.058 (0.033) *
% housing units with mortgage	n/a	n/a	0.031 (0.020)	0.030 (0.030)
County (or state) fixed effects	✓	✓	✓	✓
Year (or quarter) fixed effects	✓	✓	✓	✓
n (counties/states)	259	259	51	51
n (observations)	2,331	2,331	2,039	2,039
R^2, including fixed effects	0.799	0.967	0.912	0.989

Notes: Regressions weighted using the CEM weights in the county-year matched sample. Standard errors (in parentheses) are clustered by county for the county-year matched sample and by state for the state-quarter sample. *** p<0.01, **p<0.05, *p<0.1.

The county-year results show that the effect of Lending Club on bankruptcy filings only becomes significant in years 1, 2, and 3 after approval (i.e., only the coefficients for *Lending Club(+1)*, *Lending Club(+2)*, and *Lending Club(+3)* are significant). Furthermore, the effect grows in magnitude over time. The *Lending Club(+3)* coefficient is significantly larger than the *Lending Club(+1)* coefficient (p=0.033). This may be because usage of Lending Club is small in the approval year and doesn't become large enough to have a measurable effect on bankruptcy until a year or two after approval. We see a similar pattern in the state-quarter analysis, where the effect is not significant until the third quarter after approval. (Note that *LendingClub(+3)* refers to the third quarter – not year – after approval in the state-quarter panel.) In unreported analysis, we included dummy variables for *LendingClub(+6)* through *LendingClub(+27)* in the state-quarter panel. These coefficients are similar to that for *LendingClub(+5)* reported in Table 2-5, which suggests that the effect stabilizes. Appendix A-6 plots the lead and lag coefficients with their 95% confidence intervals.

Table 2-5 Regressions Results for Bankruptcy Filings Per Capita for the County-Year Matched Sample and the State-Quarter Sample: Leads/Lags Model

Sample	**County-year matched sample**	**State-quarter sample**
Dependent variable	**Bankruptcy filings per capita**	
Lending Club(−5)	0.211 (0.104) **	0.024 (0.027)
Lending Club(−4)	0.156 (0.101)	0.009 (0.022)
Lending Club(−3)	0.125 (0.089)	0.023 (0.024)
Lending Club(−2)	-0.090 (0.099)	0.002 (0.019)
Lending Club(−1)	omitted baseline period	
Lending Club(0)	0.148 (0.098)	0.017 (0.020)
Lending Club(+1)	0.224 (0.087) **	0.036 (0.025)
Lending Club(+2)	0.279 (0.114) **	0.037 (0.034)
Lending Club(+3)	0.395 (0.113) ***	0.060 (0.028) **
Lending Club(+4)	n/a	0.101 (0.035) ***
Lending Club(+5)	n/a	0.095 (0.033) ***
Population	0.002 (0.004)	0.035 (0.042)
Number of employed individuals	-0.003 (0.007)	-0.190 (0.086) **
Average monthly earnings	0.159 (0.157)	0.088 (0.044) *
Unemployment rate	0.149 (0.019) ***	0.079 (0.016) ***
Median household income	-0.017 (0.008) **	0.001 (0.009)
% age 60 & above	0.036 (0.042)	-0.002 (0.034)
% white	0.026 (0.025)	0.026 (0.021)
% below high school attainment	n/a	0.061 (0.032) *
% housing units rented	n/a	0.006 (0.023)
% housing units with mortgage	n/a	0.030 (0.021)
County (or state) fixed effects	✓	✓
Year (or quarter) fixed effects	✓	✓
n (counties/states)	259	51
n (observations)	2,331	2,039
R^2, including fixed effects	0.801	0.912

Notes: Regressions weighted using the CEM weights in the county-year matched sample. Standard errors (in parentheses) are clustered by county for the county-year matched sample and by state for the state-quarter sample. *** p<0.01, **p<0.05, *p<0.1.

We conducted several robustness checks, including: 1) using an alternative coding rule in which we considered Lending Club to be available throughout its approval year, 2) clustering the standard errors by state, 3) matching on raw and logged values of pre-treatment bankruptcy filings (instead of on bankruptcy filings per capita), and 4) including additional control variables. Results remain robust. We also ran a placebo test in which we randomly assigned *Lending Club availability* within the county-year panel. This placement assignment yielded no significant effect.

Falsification Test: Non-Business vs. Business Bankruptcy Filings

To enhance the causal interpretation of our findings, we conducted a falsification test based on non-business vs. business bankruptcy filings. Given that the maximum amount of Lending Club loans during our study period was relatively small ($35,000), we hypothesize that Lending Club has a larger impact on the financial health – and bankruptcy prospects – of non-businesses than of businesses (for which larger amounts are likely necessary to prevent – or cause – bankruptcy). We reran specifications (1) and (2) with *non-business bankruptcy filings per capita* and *business bankruptcy filings per capita* as the dependent variables. In both the county-year matched sample and the state-quarter sample, we find parallel pre-treatment trends for both non-business and business bankruptcies (results available from authors). Table 2-6 shows that Lending Club approval has a significant effect on non-business bankruptcy filings but not on business bankruptcy filings, which is consistent with our hypothesis and supports our causal interpretation. We also applied seemingly unrelated regression to estimate the models for both dependent variables simultaneously, which allowed us to verify that the Lending Club coefficient for non-business bankruptcies is significantly larger than the Lending Club coefficient for business bankruptcies.

Table 2-6 Regressions Results for Bankruptcy Filings Per Capita for the County-Year Matched Sample and the State-Quarter Sample: Non-Business vs. Business Bankruptcy

Sample	County-year matched sample		State-quarter sample	
Dependent variable	Non-business Bankruptcy	Business Bankruptcy	Non-business Bankruptcy	Business Bankruptcy
Lending Club available	0.182 (0.057) ***	-0.003 (0.011)	0.053 (0.031) *	0.004 (0.003)
Population	-0.001 (0.004)	0.000 (0.001)	0.039 (0.040)	-0.003 (0.002)
Number of employed individuals	-0.004 (0.006)	-0.000 (0.001)	-0.190 (0.084) **	0.006 (0.007)
Average monthly earnings	0.130 (0.148)	0.013 (0.021)	0.088 (0.043) **	-0.003 (0.004)
Unemployment rate	0.129 (0.019) ***	0.010 (0.004) **	0.076 (0.015) ***	0.004 (0.001) ***
Median household income	-0.016 (0.007) **	0.000 (0.003)	0.001 (0.009)	0.001 (0.001)
% age 60 & above	0.025 (0.025)	0.007 (0.008)	-0.001 (0.033)	0.003 (0.004)
% white	0.033 (0.040)	0.001 (0.007)	0.025 (0.020)	0.001 (0.001)
% below high school attainment	n/a	n/a	0.062 (0.031) *	-0.001 (0.002)
% housing units rented	n/a	n/a	0.006 (0.021)	0.004 (0.004)
% housing units with mortgage	n/a	n/a	0.027 (0.020)	0.004 (0.003)
County (or state) fixed effects	✓	✓	✓	✓
Year (or quarter) fixed effects	✓	✓	✓	✓
n (counties/states)	259	259	51	51
n (observations)	2,331	2,331	2,039	2,039
R^2, including fixed effects	0.806	0.113	0.913	0.679

Notes: Regressions weighted using the CEM weights in the county-year matched sample. Standard errors (in parentheses) are clustered by county for the county-year matched sample and by state for the state-quarter sample. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Potential Concurrent (and Confounding) Events

Our analysis thus far indicates that bankruptcy filings increase after Lending Club approval. This could be due to the causal effect of Lending Club, but it could also be due to any other event or policy change that occurred in the treated states at the same time as Lending Club approval. We investigate this possibility both theoretically and empirically. Theoretically, we looked for state-level policy changes (particularly those related to bankruptcy exemptions and payday lending) that might have influenced bankruptcy filings during our time period. We could not find any major changes that coincided with Lending

Club approval in the treated states but not in the controls.[5] Furthermore, we examined whether states granted Lending Club's license as part of a broader set of regulations/policies that might explain the rise in bankruptcy filings. We found no evidence for this. One indication that Lending Club approval was distinct from other policy changes is that its competitor Prosper.com (which has a similar business model) received regulatory approval in all states except for Iowa, Maine, and North Dakota in 2009.[6]

One feature of our setting that supports our causal interpretation is that Lending Club received approval at different times in the county-year analysis: 2011 for counties in Kansas and North Carolina and 2013 for counties in Indiana and Tennessee (see Table 2-1 and Appendix A-2). Thus, if an unobserved change is responsible for the effect that we attribute to Lending Club, then the change would have to have occurred at (or around) these specific times in the treated states while not occurring in the control states. Although we believe this to be unlikely, we implemented two additional analyses to improve the evidence that the change in bankruptcy was caused by Lending Club approval rather than an unobserved event: 1) a difference-in-difference-in-differences analysis, and 2) subsample analyses.

Difference-in-difference-in-differences. We exploited variation in the level of internet access across counties (and states) to conduct the difference-in-difference-in-differences ("DDD") analysis. As above, we describe this strategy using the county-year analysis; it applies to the state-quarter analysis in a similar way. Because Lending Club is

[5] For example, we reviewed payday loan state statutes from 2011 to 2015 at http://www.ncsl.org/research/financial-services-and-commerce/payday-lending-state-statutes.aspx.
[6] See Page 1 of https://www.prosper.com/Downloads/Legal/prosper10k123109%20(3.31.2010%20final).pdf.

an online platform, its approval should have a greater effect on bankruptcy filings in treated counties in which internet access is widespread than in those in which internet access is limited. We examine this via the DDD model, the intuition for which is as follows. Consider two separate difference-in-differences (DD) analyses. The first, denoted DD_I, calculates the difference-in-differences for treated and control counties i with widespread internet access (i.e., $I_i = 1$). The second, denoted $DD_{\sim I}$, calculates the difference-in-differences for treated and control counties i with limited internet access ($I_i = 0$). If Lending Club approval (or some other concurrent internet related event) increases bankruptcy, then we should see it primarily in DD_I. If there is an unobserved, non-internet related factor that increases bankruptcy in treated states after treatment, then we should see it in both DD_I and $DD_{\sim I}$. Calculating $DD_I - DD_{\sim I}$, which is the difference-in-difference-in-differences (DDD), helps us separate the Lending Club factor from the unobserved factor. The DDD regression model appears as specification (3).

$$Y_{it} = \alpha + \beta LC_{it} + \delta(LC_{it}*I_i) + T_t + \lambda(T_t*I_i) + S_i + \gamma X_{it} + \rho(X_{it}*I_i) + \varepsilon_{it} \tag{3}$$

Specification (3) mirrors (1) except for the inclusion of I_i, I_i's interactions with other variables, and associated coefficients (δ, λ, and ρ). The parameter of interest is δ, which represents the DDD estimate of the effect of Lending Club. We measured I_i for each county via its annual scores from the FCC's Form 477 County Data on Internet Access Services. We set $I_i = 1$ for counties whose average score from 2011 to 2014 was above the median (which is 3.75) and set $I_i = 0$ otherwise.[7] Because we modeled I_i as a fixed characteristic of

[7] We used this time period because 2011 was when Lending Club first became available to counties in our sample. We checked whether each county's level of internet access was consistent before and after 2011 by creating another I_i dummy variable for the 2008-2010 period. For 205 of the 259 counties in our matched sample, the 2008-2010 and 2011-2014 dummy variables are the same. The correlation between these two dummy variables is 0.73.

county *i*, its main effect is absorbed by the county fixed effects. Results for specification (3) appear in Table 2-7. The *Lending Club available* coefficient represents the average effect of Lending Club in counties in which internet access is not widespread (i.e., $I_i = 0$). Adding this coefficient to the *Lending Club available * Widespread internet access* coefficient yields the average effect for counties with widespread internet access. Results indicate that Lending Club has an effect in counties with widespread internet access but not otherwise. This supports the casual interpretation of our results.

Table 2-7 Regressions Results for Bankruptcy Filings Per Capita for the County-Year Matched Sample and State-Quarter Sample: Difference-in-Difference-in-Differences Model

Sample	**County-year matched sample**	**State-quarter sample**
Dependent variable	**Bankruptcy filings per capita**	
Lending Club available	0.002 (0.092)	-0.022 (0.034)
Lending Club available * widespread internet access	0.343 (0.115) ***	0.171 (0.068) **
Population	0.004 (0.014)	0.163 (0.060) ***
Population * widespread internet access	-0.002 (0.003)	0.083 (0.074)
Number of employed individuals	0.058 (0.053)	-0.521 (0.124) ***
Number of employed individuals * widespread internet access	-0.004 (0.006)	-0.091 (0.057)
Average monthly earnings	-0.298 (0.249)	-0.096 (0.112)
Average monthly earnings * widespread internet access	0.336 (0.107) ***	0.033 (0.029)
Unemployment rate	0.134 (0.027) ***	0.060 (0.016) ***
Unemployment rate * widespread internet access	0.157 (0.030) ***	0.074 (0.029) **
Median household income	-0.012 (0.010)	0.014 (0.009)
Median household income * widespread internet access	-0.019 (0.011)	0.006 (0.016)
% age 60 & above	0.033 (0.035)	-0.016 (0.026)
% age 60 & above * widespread internet access	0.030 (0.025)	0.035 (0.058)
% white	-0.002 (0.057)	0.014 (0.016)
% white * widespread internet access	0.042 (0.062)	0.070 (0.035) *
Additional state controls	n/a	✓
Additional state controls * widespread internet access	n/a	✓
County (or state) fixed effects	✓	✓
Year (or quarter) fixed effects	✓	✓
Year (or quarter) fixed effects * widespread internet access	✓	✓
n (counties/states)	259	51
n (observations)	2,331	2,039
R^2, including fixed effects	0.801	0.922

Notes: Regressions weighted using the CEM weights in the county-year matched sample. Standard errors (in parentheses) are clustered by county for the county-year matched sample and by state for the state-quarter sample. *** p<0.01, **p<0.05, *p<0.1.

We applied the DDD strategy to the state-quarter sample by measuring state-level internet access as the proportion of households with a minimum upload speed of 376kbps and a minimum download speed of 3Mbps, the data for which are available from the FCC's Form 477 State Data on Internet Access Services. We classified the 50 states plus DC as having either widespread internet access or not (via a median split) and ran specification (3) using the state-quarter sample. The results (see Table 2-7) indicate that Lending Club has an effect in states with widespread internet access but not otherwise.

Subsample Analyses. For the county-year analysis, we reran our matching and focal DD models twice to correspond to the two periods in which Lending Club was approved in the treated states: 2011 for Kansas and North Carolina and 2013 for Indiana and Tennessee. The first analysis (referred to as the 2011 treatment analysis) spans 2006 to 2014 and tests the effect of Lending Club approval in Kansas and North Carolina counties in 2011, using Idaho, Iowa, Maine, Nebraska, and North Dakota counties as controls. The second analysis (referred to as the 2013 treatment analysis) spans 2006 to 2015 and tests the effect of Lending Club approval in Indiana and Tennessee counties in 2013, using Idaho, Iowa, and Maine counties as controls (we excluded Nebraska and North Dakota because they were treated in the first half of 2015). We used 2015 as the ending year for the 2013 treatment analysis so as to include three years of post-treatment observations, which is important because the effect of Lending Club may evolve over time. (We used 2014 as the ending year in the 2011 treatment analysis so that we could keep Nebraska and North Dakota as controls.) We created separate matched samples for both analyses by matching treated and control counties on the values of *bankruptcy filings per capita* in the five years preceding treatment and on *population* (in 2006). This yielded 316 matched counties for the 2011 analysis and 83 matched counties for the 2013 analysis. (We could not conduct a similar subsample analysis for the state-quarter sample because there are not enough states.) Figure 2-2 shows the trends of bankruptcy filings and provides visual evidence of parallel pre-treatment trends for both analyses. After treatment, the rate of bankruptcy filings decreases at a slower rate in treated counties than in control counties, suggesting that Lending Club approval increases bankruptcy filings.

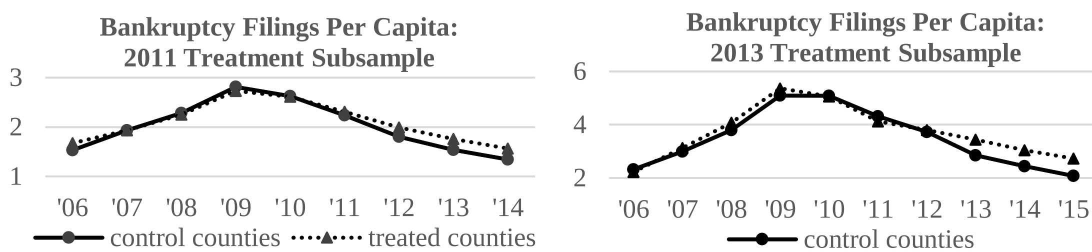


Figure 2-2 Bankruptcy Filing Per Capita Trends: Subsample Analysis

Results of the leads/lags models for both analyses are shown in Table 2-8. The pre-treatment trends are parallel, and Lending Club approval is associated with an increase in bankruptcy filings in both analyses. Thus, if an unobserved change – concurrent with Lending Club approval – is responsible for the increase in bankruptcy filings, then this change would have to have occurred in 2011 in Kansas and North Carolina (and not the control states) and in 2013 in Indiana and Tennessee (and not the control states). We believe this would be an unlikely coincidence, lending further support to our causal interpretation. Notice that the *Lending Club(0)* coefficient is positive and significant in the 2013 treatment analysis but not the 2011 treatment analysis. This may reflect increased awareness and use of Lending Club in 2013 vis-à-vis 2011. Indeed, Appendix A-1 shows a faster growth rate for loans in Indiana and Tennessee in 2013 than for loans in Kansas and North Carolina in 2011.

Table 2-8 Regressions Results for Bankruptcy Filings Per Capita for the County-Year Sample: 2011 Treatment for Kansas and North Carolina and 2013 Treatment for Indiana and Tennessee

	2011 Treatment (Kansas and North Carolina: 2006-2014)	2013 Treatment (Indiana and Tennessee: 2006-2015)
Dependent variable	**Bankruptcy filings per capita**	
Lending Club(−5)	0.295 (0.097) ***	-0.030 (0.167)
Lending Club(−4)	0.157 (0.101)	0.011 (0.203)
Lending Club(−3)	0.103 (0.091)	-0.214 (0.258)
Lending Club(−2)	-0.028 (0.085)	-0.276 (0.194)
Lending Club(−1)	omitted baseline period	
Lending Club(0)	0.077 (0.084)	0.326 (0.154) **
Lending Club(+1)	0.209 (0.080) ***	0.399 (0.153) **
Lending Club(+2)	0.264 (0.088) ***	0.499 (0.194) **
Lending Club(+3)	0.307 (0.095) ***	n/a
Additional county controls	✓	✓
County fixed effects	✓	✓
Year fixed effects	✓	✓
n (counties)	316	83
n (observations)	2,844	830
R^2, including fixed effects	0.751	0.882

Notes: Regressions weighted using the CEM weights. Standard errors (in parentheses) are clustered by county. *** p<0.01, **p<0.05, *p<0.1.

Effect of Similar Online Lending Platforms (e.g., Prosper.com)

To extend our results, we examined the effect on bankruptcy filings of Prosper.com, which is an online lending platform similar to Lending Club. Like Lending Club, Prosper.com had to cease originating peer-to-peer loans temporarily to seek regulatory approval, after which they resumed origination. Unlike Lending Club, Prosper.com received regulatory approval from all states except Iowa, Maine and North Dakota in 2009. Thus, there is less variation in Prosper.com's availability across counties and states, making it harder to identify its effect on bankruptcy. Nevertheless, we ran two analyses to analyze the effect of Prosper.com. First, we reran specification (1) on both samples (county-year and state-quarter) after adding a *Prosper.com available* dummy variable as a control. As shown in columns 1 and 2 of Table 2-9, the *Lending Club available* coefficient remains positive and significant. Because Prosper.com received approval before Lending Club in

Kansas, North Carolina, Indiana, and Tennessee (the treated states in the county-year analysis), the *Lending Club available* coefficient in column 1 can be interpreted as the additional impact of Lending Club approval. Second, we restricted our analysis to counties in the five states in which Lending Club had not been approved by the end of 2014, thereby allowing us to study the effect of Prosper.com approval without potential contamination by Lending Club approval. Prosper.com was approved in two of these states in 2009 (Idaho and Nebraska), but not in the other three (Iowa, Maine, and North Dakota) by the end of 2014. Using data from 2006 to 2014, we reran specification (1) with *Prosper.com availability* in place of *Lending Club availability*. We used all counties from these states (the full sample) as well as a matched sample, using the matching process from above. Results are shown in columns 3 and 4 of Table 2-9 and show that Prosper.com approval had a positive effect on bankruptcy filings.

Table 2-9 Results of Regressions Examining the Effect of Prosper.com

	Same as Focal Analyses		**Restricted to 5 states (ID, NE, IA, ME, ND)**	
Sample	**County-year matched sample**	**State-quarter sample**	**County-year full sample**	**County-year matched sample**
Dependent variable		**Bankruptcy filings per capita**		
Lending Club available	0.124 (0.074) *	0.057 (0.030) *	n/a	n/a
Prosper.com available	0.113 (0.069)	0.021 (0.034)	0.309 (0.064) ***	0.262 (0.117) **
Population	-0.001 (0.004)	0.037 (0.042)	0.001 (0.005)	-0.044 (0.042)
Number of employed individuals	-0.005 (0.007)	-0.189 (0.086) **	-0.031 (0.014) **	0.034 (0.019) *
Average monthly earnings	0.162 (0.153)	0.088 (0.044) *	0.131 (0.103)	0.089 (0.144)
Unemployment rate	0.133 (0.019) ***	0.079 (0.016) ***	0.165 (0.030) ***	0.210 (0.047) ***
Median household income	-0.015 (0.008) *	0.002 (0.009)	-0.004 (0.008)	0.004 (0.013)
% age 60 & above	0.030 (0.024)	0.001 (0.034)	0.001 (0.020)	0.013 (0.043)
% white	0.025 (0.041)	0.025 (0.021)	0.001 (0.037)	-0.029 (0.073)
Additional state controls		✓		
County (or state) fixed effects	✓	✓	✓	✓
Year (or quarter) fixed effects	✓	✓	✓	✓
n (counties/states)	259	51	305	75
n (observations)	2,331	2,039	2,745	675
R^2, including fixed effects	0.800	0.912	0.685	0.665

Notes: Regressions weighted using the CEM weights in the county-year matched sample. Standard errors (in parentheses) are clustered by county for the county-year sample and by state for the state-quarter sample. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

2.3.2 Instrumental Variables Analysis: Analyzing the Effect of Lending Club Loan Activity

We extend our analysis by using micro-level loan data from Lending Club to examine the relationship between the number of Lending Club loans and the number of bankruptcies in a state.

Data and Variables

Lending Club publishes loan data via its web site (https://www.lendingclub.com/info/download-data.action). We downloaded the data from 2007 to 2015 (n=877,440 loans) to mirror the time period of the difference-in-differences analysis (Lending Club was not operating in 2006). The data describe each borrower (e.g., state of residence, self-reported income and debt-to-income ratio, and FICO credit score) and loan (e.g., grade assigned by Lending Club, origination date, size (i.e., amount), principal amount paid, term (36- or 60-month), purpose, and last payment date). We coded each loan as "outstanding" or "default" in each quarter as follows. We coded mature loans (i.e., those whose terms had expired) that Lending Club marked as paid as "outstanding" (i.e., current) in each quarter from loan origination to payoff. We coded immature loans listed as current or late as "outstanding" in each quarter from loan origination to the 4th quarter of 2015 (when our data collection stopped). We coded loans that Lending Club marked as in default or charged off as "outstanding" in each quarter from loan origination to the last payment quarter and then "default" in the subsequent quarter. We counted the number of loans outstanding in state i in quarter t and merged them with the state-quarter panel. We conducted this analysis at the state-quarter level because Lending Club does not publish borrowers' counties.

Empirical Strategy

Our baseline specification is shown below.

$$Y_{it} = \alpha + \beta LoansOutstandingPerCapita_{it} + T_t + S_i + \gamma X_{it} + \varepsilon_{it} \quad (4)$$

As in the difference-in-differences analysis, Y_{it} is bankruptcy filings per capita. The key independent variable is Lending Club loans outstanding per capita in state i in quarter t. T_t are quarter fixed effects, S_i are state fixed effects, X_{it} are control variables, and ε_{it} is the error term, which is clustered at the state level. β is the coefficient of interest. OLS estimation will yield a biased estimate of β due to endogeneity issues. Potential reasons for endogeneity include omitted variables that affect both the number of loans and bankruptcy filings and simultaneity (or reverse causality) by which bankruptcy filings influence the number of loans. To address this, we used instrumental variables and two-stage least squares (2SLS).

Selection of instrumental variables. We identified two instrumental variables. Each instrument relies on different sources of (plausibly) exogenous variation, and using both allows us to check the consistency of our results across instruments (Murray 2006). The first instrument is LC availability / maturity. This measures whether Lending Club was available to borrowers in state i in quarter t and, if so, the maturity of Lending Club in quarter t. LC availability / maturity is the product of Lending Club available and Time since 2008-Q4 (squared). Lending Club available reflects whether Lending Club was available to borrowers in state i in quarter t; it is the key dummy variable used in the DD analysis. Time since 2008-Q4 (squared) is the square of the number of quarters between quarter t and the 4th quarter of 2008. This captures how Lending Club matured over time after it received regulatory approval from most states in 2008-Q4 (see Table 2-1). We used a squared term (rather than a linear term) to mirror the non-linear increase over time in Lending Club loans. LC availability / maturity should be correlated with loans outstanding per capita because states in which Lending Club is available to borrowers will have a larger number of loans than states in which Lending Club is not available, particularly as Lending Club becomes more mature. LC availability / maturity also captures within-state variation because loans outstanding per capita increases non-linearly for each treated state. The second instrument, debt-to-income policy change, is based on a policy change made by Lending Club in the 2nd quarter of 2012. Lending Club does not issue loans if an applicant's debt-to-income ("DTI") ratio exceeds a threshold. In the 2nd quarter of 2012, Lending Club raised this threshold from 30 to 35 (technically from 0.3 to 0.35). This change should be associated with an increase in loans outstanding per capita because it expands the pool of qualified borrowers. Debt-to-income policy change is the product of Lending Club available, Post DTI Policy Change (a dummy variable equal to 1 for all quarters from the 2nd quarter of 2012 forward), and Time since 2012-Q1 (squared). The first two variables ensure that debt-to-income policy change is 0 before the policy change and in states in which Lending Club was not available. The third variable causes the instrument to increase in magnitude each quarter the policy has been in place. This reflects the quarter-by-quarter accumulated growth in the pool of qualified borrowers, which should correlate with growth in loans in states in which Lending Club was available.

Each instrument will be exogenous to bankruptcy filings if the factors that comprise it are exogenous. Because Time since 2008-Q4 (squared) and Time since 2012-Q1 (squared) are deterministic, they are exogenous to bankruptcy filings, particularly because we include quarter fixed effects (which capture any time trend influencing bankruptcy filings). Post DTI Policy Change should be exogenous to bankruptcy filings in a state because it reflects a blanket policy change that Lending Club applied across all states, i.e., it does not explain which states had higher (or lower) bankruptcy levels. Of greater concern is whether Lending Club available is exogenous. The state-quarter DD analysis (which uses the same panel as the instrumental variables analysis) provides evidence that Lending Club available is exogenous to bankruptcy filings at the state level. This is because there is no significant difference in the pre-treatment trends in bankruptcy filings between treated and control states, as shown in Table 2-5. To further assess the exogeneity of Lending Club available, we used a Cox proportional hazards model to examine whether a state's bankruptcy filings affect when Lending Club became available to borrowers. Table 2-10 shows that neither the level nor the change rate of bankruptcy filings per capita have a significant effect.

Table 2-10 Proportional Hazards Model of When Lending Club Became Available to Borrowers

	Lending Club approval			
Bankruptcy per capita (previous 1 quarter)	1.000 (0.232)			
Bankruptcy per capita (previous 1 year)		1.006 (0.064)		
Bankruptcy per capita change rate (1 quarter change rate)			1.580 (0.912)	
Bankruptcy per capita change rate (1 year change rate)				0.512 (0.320)
Population	1.179 (0.105) *	1.181 (0.103) *	1.165 (0.105) *	1.226 (0.138) *
Number of employed individuals	0.643 (0.157) *	0.641 (0.154) *	0.664 (0.164) *	0.577 (0.178) *
Average monthly earnings	1.262 (0.156) *	1.264 (0.153) *	1.233 (0.157) *	1.327 (0.192) *
Unemployment rate	1.083 (0.077)	1.082 (0.076)	1.092 (0.073)	1.060 (0.083)
Median household income	1.020 (0.016)	1.020 (0.016)	1.021 (0.015)	1.016 (0.016)
% age 60 & above	0.978 (0.027)	0.978 (0.026)	0.974 (0.025)	0.973 (0.028)
% white	0.997 (0.004)	0.997 (0.004)	0.996 (0.004)	0.997 (0.004)
% below high school attainment	1.013 (0.026)	1.011 (0.025)	1.013 (0.027)	1.013 (0.027)
% housing units rented	0.970 (0.025)	0.970 (0.025)	0.967 (0.024)	0.966 (0.023)
% housing units with mortgage	0.972 (0.036)	0.971 (0.035)	0.972 (0.032)	0.968 (0.030)
n (states)	51	51	51	51
n (observations)	767	614	716	563
Wald chi^2 P value	0.423	0.427	0.234	0.454
Log Likelihood	-174.33	-174.32	-174.29	-174.22

Notes: Standard errors (in parentheses) are clustered by state. *** p<0.01, **p<0.05, *p<0.1.

Results

Appendix A-7 reports the first-stage results of the 2SLS regressions. The instruments are relevant and not weak. Table 2-11 lists statistics showing the strength of the instruments.

Table 2-11 Instrument Relevance/Strength Statistics

Instrumental	First-stage Adjusted R^2	Partial R^2 (attributable to instrument)	F (significance of the instrument), with p-value
LC availability / maturity	0.93	0.37	70.59 (p<0.01)
Debt-to-income policy change	0.94	0.49	449.25 (p<0.01)

Table 2-12 shows the second-stage results, along with the ordinary least squares results. The OLS coefficient for loans outstanding per capita, shown in the first column, is negative and insignificant. The negative coefficient is likely because of a simultaneity / reverse causality issue: namely, that a high level of bankruptcy filings is likely to reduce the number of Lending Club loans because bankrupt individuals (or highly risky borrowers) will not qualify for loans. This highlights the need to use instrumental variables. The other columns show the 2SLS results when using each of the instruments individually (Columns 2 and 3) and jointly (Column 4). The 2SLS coefficients for loans outstanding per capita coefficients are positive, significant, and similar in magnitude across models. When using both instruments, we conduct a test of overidentifying restrictions (see Murray (2006) for details). We are unable to reject the null hypothesis that the instruments are exogenous (Hansen's J statistic = 2.08, p=0.15), which supports the instruments' validity. The 2SLS coefficient when using both instruments is 0.039 (p=0.059), which is the most conservative estimate. Using this coefficient, a one standard deviation increase in loans outstanding per capita (δ=0.67) is associated with a 0.026 increase in bankruptcy filings per capita. Because the mean of bankruptcy filings per capita is 0.833, this represents a 3.1% increase. The coefficient also implies that an increase of 100 loans per capita in a quarter is associated with an increase of 3.9 bankruptcies per capita. This is a fairly large effect size, which we discuss below.

Table 2-12 2SLS Regressions for Bankruptcy Filings Per Capita: Second-Stage Results

Dependent Variable	Bankruptcy filings per capita			
Model	**OLS**	**IV: LC availability / maturity**	**IV: Debt-to-income policy change**	**IV: Both instruments**
Loans outstanding per capita	-0.029 (0.040)	0.058 (0.029) **	0.041 (0.021) *	0.039 (0.021) *
Population	0.041 (0.041)	0.035 (0.041)	0.036 (0.040)	0.036 (0.040)
Number of employed individuals	-0.178 (0.090) *	-0.198 (0.086) **	-0.194 (0.086) **	-0.193 (0.086) **
Average monthly earnings	0.032 (0.046) *	0.092 (0.044) **	0.091 (0.044) **	0.090 (0.044) **
Unemployment rate	0.081 (0.017) ***	0.080 (0.016) ***	0.080 (0.016) ***	0.080 (0.016) ***
Median household income	0.003 (0.010)	-0.000 (0.009)	0.000 (0.009)	0.000 (0.009)
% age 60 & above	0.004 (0.034)	-0.001 (0.035)	0.000 (0.034)	0.000 (0.034)
% white	0.024 (0.020)	0.027 (0.021)	0.026 (0.021)	0.026 (0.021)
% below high school attainment	0.061 (0.033) *	0.056 (0.031) *	0.057 (0.032) *	0.057 (0.032) *
% housing units rented	0.018 (0.022)	0.007 (0.023)	0.009 (0.024)	0.009 (0.024)
% housing units with mortgage	0.031 (0.020)	0.032 (0.021)	0.031 (0.021)	0.031 (0.021)
State fixed effects	✓	✓	✓	✓
Quarter fixed effects	✓	✓	✓	✓
n (states)	51	51	51	51
n (observations)	2,039	2,039	2,039	2,039
R^2 (centered)	0.78	0.77	0.77	0.77
Hansen's J statistic	n/a	n/a	n/a	2.08 (p=0.15)

Notes: Standard errors (in parentheses) are clustered by state. *** p<0.01, **p<0.05, *p<0.1.

We conducted several additional analyses and robustness checks. First, we reran the regressions using lagged values of *loans outstanding per capita* and the instruments. (For lags, we used the previous quarter, the previous 2 quarters, the previous 3 quarters, and the previous 4 quarters.) This reflects the possibility that the effect of Lending Club loans on bankruptcy filings does not show up immediately. These results are consistent, with the magnitude of the coefficients increasing slightly with lag length. Second, we confirmed that the results are robust after we removed potentially fraudulent loans that Lending Club allegedly loaned to itself to boost its loan volume statistics.[8] Third, we confirmed that the results are robust when using alternative independent variables, including the natural log of loans per capita and the value of loans (instead of the number of loans).

[8] We define fraudulent loans as those that are fully paid within two billing cycles. The ratio of fraudulent loans to total loans is 246 to 887,440.

2.3.3 Mechanisms for the Lending Club Effect

As described in the Background, Literature Review, and Motivation section, there are several mechanisms that could explain the link between Lending Club approval and increased bankruptcy filings. One potential mechanism, referred to as the "credit risk" mechanism, is that Lending Club borrowers are high credit risks who do not have access to traditional capital. If so, then Lending Club loans could be going to people who are inherently uncreditworthy, thereby leading them into bankruptcy. A second potential mechanism is that Lending Club borrowers are normal credit risks who have access to traditional capital. However, the ease of getting a Lending Club loan may cause these borrowers to become over-extended financially, leading to bankruptcy. We refer to this as the "debt trap" mechanism. A third potential mechanism is that "strategic" borrowers who are considering bankruptcy use a Lending Club to restructure their debt (e.g., swapping secured debt for unsecured debt) or to engage in last-minute consumption (e.g., taking a vacation) before they file. We examine these three mechanisms empirically.

We examined the credit risk mechanism by reviewing the FICO scores of Lending Club borrowers from the micro-level loan data. The mean FICO score range was 695 to 699 (the median range was 690 to 694), which is similar to the mean FICO scores reported in credit bureau data (Jagtiani and Lemieux 2017). Jagtiani and Lemieux (2017) also find that based on observable credit features, borrowers get lower interest rates from Lending Club than from traditional lenders. Based on this, it does not appear that Lending Club attracts and issues loans to borrowers who are systematic credit risky. However, because FICO score is only one measure of creditworthiness, we examined the credit risk mechanism, along with the debt trap mechanism, further by exploiting differences between

Chapter 7 and Chapter 13 bankruptcy. Chapter 7 requires filers to have little to no disposable income, whereas Chapter 13 is available to filers who have disposable income left over after paying household expenses. If Lending Club approval primarily increases Chapter 7 filings, then its effect is likely to be concentrated among low-income people who might not be creditworthy, which would support the credit risk mechanism. On the other hand, if Lending Club approval primarily increases Chapter 13 filings, then its effect is likely concentrated among people who may otherwise have access to credit but who become over-extended after using Lending Club, which would support the debt trap mechanism.

To explore this, we reran specification (1) using the natural log of Chapter 7 bankruptcy filings per capita and Chapter 13 bankruptcy filings per capita as the dependent variables. Taking the natural log allows us to compare the percentage increases in Chapter 7 and 13 bankruptcies attributable to Lending Club, which is important given that the base rates of these bankruptcies differ significantly. Table 2-13 shows that Lending Club has a larger effect on Chapter 13 bankruptcies. This suggests that it is more likely that Lending Club borrowers become overextended and fall into a debt trap as opposed to being inherent credit risks. The self-reported income of Lending Club borrowers (from the Lending Club loan data) provides further support for the debt trap mechanism. As shown in Appendix A-8, Lending Club borrowers, including those who default on their loans, report substantially higher-than-average incomes.

We further examined the debt trap mechanism by using the micro-level loan data to analyze what factors predict whether a Lending Club borrower will default on his/her loan (coded as Loan default = 1). In our first analysis, we included only mature loans. This ensured that if a borrower defaulted during the term of the loan that we would observe it. We observe individual loan defaults, but we cannot observe individual bankruptcies (the bankruptcy data is only available in aggregate form). However, the two should be related, because borrowers who do not repay their loans are more likely to file for bankruptcy than borrowers who repay. Thus, we assume that factors that lead to loan default also lead to bankruptcy. The key independent variable in this analysis is debt expansion, which is a dummy variable that indicates whether the purpose of the loan was to expand the borrower's overall debt or to consolidate it (e.g., by using a Lending Club loan with a 10% interest rate to pay off credit card debt with a 22% interest rate). For each loan application, the borrower selects the purpose of the loan from a pre-defined list. Two choices indicate debt consolidation: "credit card" and "debt consolidation". We coded debt expansion = 1 for loans with a purpose other than these two; examples include home improvement, major purchase, and vacation. We included several other variables published by Lending Club as controls. We used a linear probability model to predict loan default, although results from a logistic regression are similar. The results are shown in Column 1 of Table 2-14. Control variables have the expected signs: debt-to-income ratio and loan amount are positively correlated with default, while FICO score is negatively correlated. The coefficient for debt expansion (β=0.014) is positive and significant. Because 13.2% of loans in the sample ultimately default (the quarterly default rate is approximately 1.1%), the debt expansion coefficient represents a 10.6% increase in the probability of default. In the second approach, we used a Cox proportional hazards model to assess what factors, including loan purpose, affect the "hazard" of a loan defaulting. Because hazard models accommodate right censoring, we used immature loans in this analysis. Column 2 of Table 2-14 shows that debt expansion loans are more likely to default. Overall, these results indicate that Lending Club loans that increase a borrower's overall debt are positively associated with default (and potentially bankruptcy). This provides additional support for the debt trap mechanism.

Table 2-13 Regressions Results for Bankruptcy Filings: Chapter 7 vs Chapter 13 Bankruptcies

Sample	**County-year matched sample**		**State-quarter sample**	
	Bankruptcy Filings per Capita – Natural Log			
Dependent Variable	**Chapter 7**	**Chapter 13**	**Chapter 7**	**Chapter 13**
Lending Club available	0.032 (0.030)	0.215 (0.041) ***	0.088 (0.042) **	0.135 (0.045) ***
Population	0.006 (0.003) **	0.000 (0.003)	-0.003 (0.063)	0.163 (0.084) *
Number of employed individuals	0.009 (0.004) **	0.001 (0.005)	-0.155 (0.098)	-0.203 (0.158)
Average monthly earnings	0.071 (0.136)	-0.078 (0.071)	0.046 (0.053)	0.085 (0.086)
Unemployment rate	0.049 (0.012) ***	0.024 (0.017)	0.056 (0.010) ***	0.022 (0.014)
Median household income	-0.015 (0.006) **	0.005 (0.006)	0.001 (0.017)	-0.046 (0.021) **
% age 60 & above	0.022 (0.017)	-0.005 (0.013)	-0.033 (0.054)	-0.045 (0.062)
% white	0.027 (0.030)	0.027 (0.028)	0.025 (0.020)	0.015 (0.032)
Additional state controls			✓	✓
County (or state) fixed effects	✓	✓	✓	✓
Year (or quarter) fixed effects	✓	✓	✓	✓
n (counties/states)	259	259	51	51
n (observations)	2,331	2,331	2,039	2,039
R^2, including fixed effects	0.959	0.926	0.988	0.985

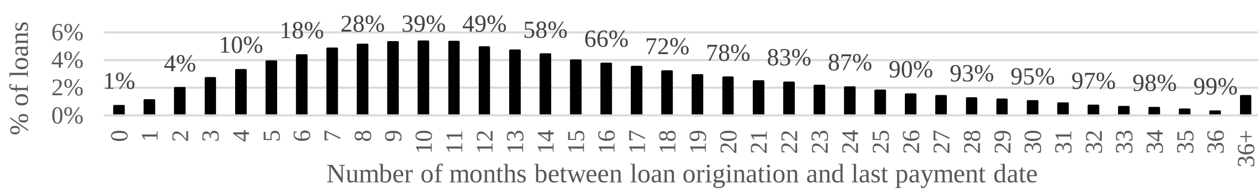
Notes: Regressions weighted using the CEM weights in the county-year matched sample. Standard errors (in parentheses) are clustered by county for the county-year matched sample and by state for the state-quarter sample. *** p<0.01, **p<0.05, *p<0.1.

Table 2-14 Regressions Results for Loan Default: Loan-Level Analysis

Dependent Variable	**Loan default**	
Model	**Linear probability**	**Hazard**
Debt expansion	0.014 (0.002) ***	1.169 (0.012) ***
Debt to income ratio (pre-loan)	0.002 (0.000) ***	1.001 (0.000) ***
Loan amount	0.0003 (0.0001) ***	1.012 (0.001) ***
Annual borrower income	-0.0002 (0.0000) ***	0.996 (0.001) ***
36 month loan term (=1 if 36-month term and =0 if 60-month term)	-0.064 (0.004) ***	1.364 (0.013) ***
FICO score (low end of range)	-0.0002 (0.0000) ***	0.988 (0.000) ***
Loan grade dummy variables (A1-G5, assigned by Lending Club)	✓	
Borrower's state of residence dummy variables	✓	
Month-year of loan origination dummy variables	✓	
n (observations)	209,882	677,554
R^2, including fixed effects	0.039	n/a

Notes: Standard errors (in parentheses) are clustered by month of loan origination. *** p<0.01, **p<0.05, *p<0.1.

We examined the "strategic" borrowing mechanism by analyzing when borrowers who default stopped repaying their loans. If borrowers use a Lending Club loan to pay off another (likely secured) loan or to engage in last-minute consumption with no intention of repaying the (unsecured) Lending Club loan, then they should stop paying the Lending Club loan relatively quickly. For each defaulted loan, we computed the number of months between loan origination and the last payment date. Figure 2-3 shows that for approximately 18% (6.5%) of loans that default, the borrowers ceased repayment after 6 (3) months. This suggests that strategic borrowing may explain some of the effect of Lending Club on bankruptcy filings. The finding that Lending Club has a larger effect on Chapter 13 vs. Chapter 7 bankruptcies also supports the strategic borrowing mechanism. This is because Chapter 13 allows debtors to avoid liquidating assets, such as those a strategic borrower would theoretically pay off via a Lending Club loan.



Note: The numbers listed at the top of the data columns are cumulative percentages.

Figure 2-3 Histogram of how many months borrowers repay loans that end in default

Other, indirect mechanisms may also contribute to the effect. One possibility is that traditional lenders respond to competition from Lending Club loans by issuing riskier loans (in order to maintain consistent loan volumes) that result in borrowers' bankruptcy (Wolfe and Yoo 2018). Similar risky behavior – and increased bankruptcies – occurred when competition among banks increased after deregulation (Dick and Lehnert 2010; see the Literature Review section). We explored this by considering how much of the effect size implied by the instrumental variables analysis – that an increase of 100 loans per capita in a quarter is associated with an increase of 3.9 bankruptcies per capita – can be explained by the direct mechanisms discussed above. If every loan that defaulted because of the debt trap or strategic borrower mechanisms resulted in a bankruptcy, then the default rate would need to be 3.9% per quarter to account for the effect size. The default rate of loans published in the micro-data is only 1.1% per quarter. One potential explanation for this gap is that the published Lending Club loan data systematically undercounts the number of Lending Club loans. This occurs because Lending Club also operates loan programs outside of its standard marketplace (e.g., see Lending Club's 2015 10-K). We compared the total amount loaned from 2007 to 2015 as reported in Lending Club's 2015 10-K (https://ir.lendingclub.com/Cache/33047201.pdf, p.3) to the total amount of loans published in the micro-data. The total amount in the micro-data accounts for approximately 81% of the total amount in the 10-K; thus, our measure of *loans outstanding per capita* is undercounted. If the non-published loans yield the same outcomes (on average) as the published loans, then our coefficient (and effect size) will be inflated. To examine this, we reran specification (4) after multiplying *loans outstanding per capita* by 1.23 (i.e. 100/81). This narrows the gap by reducing the *loans outstanding per capita* coefficient from 0.039 to 0.031 ($p<0.10$).[9] Another potential explanation for the gap is that some bankruptcies may occur without a corresponding loan default. For example, borrowers who file Chapter 13 bankruptcy (and most of the effect we document is specific to Chapter 13; see above) may continue paying off their Lending Club loan as part of their repayment plan. 10% of mature loans not listed as charged off have a principal amount paid that is less than the loan amount. This suggests that these borrowers had a revised payment plan, potentially as part of Chapter 13 bankruptcy. Despite these explanations, the direct mechanisms are unlikely to explain

[9] A more formal way to see this relationship is via the simple regression coefficient formula: $\beta = \sum_1^n (x_i - \bar{x})\,(y_i - \bar{y}) / \sum_1^n (x_i - \bar{x})^2$. If *x* is really 2*x*, such that $(x_i - \bar{x})$ is really $(2x_i - 2\bar{x})$, then β will be half as large.

all of the effect. This suggests that part of the Lending Club effect is due to indirect mechanisms such as change in traditional lenders' practices in response to competition from Lending Club loans.

2.4 Discussion and Implications

Using different identification strategies, data samples, and levels of analysis, we consistently find a positive relationship between online lending and bankruptcy filings. Although we cannot fully characterize the reason for this effect, our analysis suggests the following. The difference-in-differences analysis, particularly the leads/lags analysis reported in Table 2-5, indicates that Lending Club approval has both short-term and long-term effects. The short-term effect may be due (in part) to strategic borrowing in which borrowers who were considering bankruptcy before receiving a Lending Club loan use the loan to restructure their debt or to engage in last-minute consumption before they file. The long-term effect is likely due (in part) to both the strategic borrowing and debt trap mechanisms, given that it takes time for the debt trap mechanism to operate: borrowers must receive a Lending Club loan, have the additional debt burden drive them to bankruptcy, and then file bankruptcy. Indirect mechanisms, such as competitive responses to Lending Club loans, also appear to play a role.

The strategic borrowing mechanism represents adverse selection. This is because there is information asymmetry between strategic borrowers and investors before the loan is issued: the borrowers know that they have no intention to repay the Lending Club loan (making them "adverse") but withhold this information from investors. There is also a possibility of moral hazard, in the sense that borrowers may decide after they receive the loan that there are minimal consequences if they choose to declare bankruptcy rather than to repay the loan. However, we find this unlikely, because there are consequences to declaring bankruptcy. Furthermore, we believe that most borrowers (other than strategic borrowers) are sincere in their intention to repay, but that some are unable to, perhaps due to unforeseen circumstances or because they misjudged their financial status.

The effect of online lending on bankruptcy filings can be mitigated by policy changes by online lending platforms; indeed, Lending Club has recently made several such changes. For example, in 2017, Lending Club removed the riskiest loans (grades F and G) from its platform, citing their high delinquency rates, and it launched a hardship program to help struggling borrowers. Also, Lending Club introduced its Direct Pay program shortly after our study period. This program sends a portion of a loan directly to the borrower's creditor(s), which essentially requires that the borrower use the loan (or at least a portion thereof) for debt consolidation instead of debt expansion. Programs like this, particularly if implemented for borrowers most likely to default, could diminish or potentially eliminate the effect that we document. Studying the effect of these programs is a promising avenue for future research. Increased education about financial management, which could be provided by online lending platforms to borrowers during the loan application process, may also help.

2.5 Conclusion

Online lending platforms have great potential to improve individuals' financial health and security by providing easy access to affordable credit. However, they could also lure borrowers into a debt trap that leads to bankruptcy. We exploit variation in when states granted approval for Lending Club to issue peer-to-peer loans to examine Lending Club's effect on bankruptcy filings. We conduct a difference-in-differences analysis at both the state-quarter and county-year levels. We also conduct an instrumental variables analysis to examine the relationship between the number of Lending Club loans and bankruptcy filings. We consistently find that Lending Club increases bankruptcy filings. We identify possible mechanisms for the effect, including: 1) the ease of receiving a Lending Club loan lures borrowers into a debt trap that results in bankruptcy, and 2) "strategic" borrowers who are considering bankruptcy use Lending Club loans to restructure their debt or to engage in last-minute consumption before filing.

Our study contributes to both the online platforms literature and the bankruptcy literature. Online platforms are growing quickly, and regulators and researchers are unsure of their impacts on society and the economy. Our findings point to a potential dark side of online lending platforms: increased bankruptcy. To be clear, we are not arguing that online lending platforms are "bad" for the economy or for society. They likely have positive effects that we do not explore. Instead, we identify a specific issue that could potentially be addressed via platform design or regulation. Indeed, recent initiatives by Lending Club may address the issue, at least in part.

Our study has limitations. First, we cannot be sure that our results apply to all online lending platforms, although Lending Club is the largest platform and our results also appear to hold for Prosper.com. Second, our results may be specific to the time period of our sample, which is a common limitation of empirical studies. It is possible that recent online loans are less (or more) likely to contribute to bankruptcy. Third, we are unable to empirically examine indirect mechanisms by which Lending Club may affect bankruptcy, such as changes in traditional banks' lending practices caused by increased competition.

This represents an opportunity for future research. Fourth, because Lending Club does not identify individual borrowers, we cannot connect individual Lending Club borrowers to bankruptcy records. Instead, we rely on aggregate data (at the county and state levels) to identify the impact of Lending Club. This lack of individual-level data is common in studies such as ours that investigate the societal impacts of online platforms (Burtch et al. 2016; Chan and Ghose 2014; Greenwood and Agarwal 2015; Greenwood and Wattal 2017; Seamans and Zhu 2014). Future research can leverage individual-level data (if available) to provide additional insight into the mechanisms by which online lending platforms affect financial well-being.

2.6 References

Allen J, Damar HE, Martinez-Miera D (2016) Consumer bankruptcy, bank mergers, and information. *Rev. Finance* 20(4):1289-1320.

Alyakoob M, Rahman MS, Wei ZY (2018) Where you live matters: The impact of local financial market competition in managing online peer-to-peer loans. Working paper available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2985099.

Autor DH (2003) Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *J. Labor Econom.* 21(1):1-42.

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? *Quart. J. Econom.* 119(1):249-275.

Buchak G, Matvos G, Piskorski T, Seru A (2018) Fintech, regulatory arbitrage, and the rise of shadow banks. *J. Financial Econom.* 130(3):453-483.

Burtch G, Carnahan S, Greenwood B (2018) Can you gig it? An empirical examination of the gig-economy and entrepreneurial activity. Forthcoming in *Management Sci.*

Burtch G, Ghose A, Wattal S (2014) Cultural differences and geography as determinants of online prosocial lending. MIS Quart. 38(3):773-794.

Burtch G, Chan J (2018) Investigating the relationship between medical crowdfunding and personal bankruptcy in the United States: Evidence of a digital divide. Forthcoming in *MIS*

Quart.

Butler AW, Cornaggia J, Gurun UG (2017) Do local capital market conditions affect consumers' borrowing decisions? *Management Sci*. 63(12):4175-4187.

Chan J, Ghose A (2014) Internet's dirty secret: Assessing the impact of online intermediaries on HIV transmission. *MIS Quart.* 38(4):955-975.

Chava, S, Paradkar, N (2018) Winners and losers of marketplace lending: Evidence from borrower credit dynamics. *Georgia Tech Scheller College of Business Research Paper No. 18-16*, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3178322.

Dawsey AE, Hynes RM, Ausubel LM (2013) Non-judicial debt collection and the consumer's choice among repayment, bankruptcy and informal bankruptcy. *Amer. Bankruptcy Law J.* 87(1):1-26.

Dick AA, Lehnert A (2010) Personal bankruptcy and credit market competition. *J. Finance* 65(2):655-686.

Dobbie W, Song J (2015) Debt relief and debtor outcomes: Measuring the effects of consumer bankruptcy protection. *Amer. Econom. Rev.* 105(3):1272-1311.

Duarte J, Siegel S, Young L (2012) Trust and credit: The role of appearance in peer-to-peer lending. *Rev. Financial Stud.* 25(8):2455-2484.

Galak J, Small D, Stephen AT (2011) Microfinance decision making a field study of prosocial lending. *J. Marketing Res.* 48:S130-S137.

Greenberg J, Mollick E (2017) Activist choice homophily and the crowdfunding of female founders. *Administrative Sci. Quart.* 62(2):341-374.

Greenwood BN, Agarwal R (2016). Matching platforms and HIV incidence: An empirical investigation of race, gender, and socioeconomic status. *Management Sci.* 62(8):2281-2303.

Greenwood BN, Wattal S (2017) Show me the way to go home: An empirical investigation of ride sharing and alcohol related motor vehicle homicide. *MIS Quart.* 41(1):163-187.

Harkness SK (2016) Discrimination in lending markets: Status and the intersections of gender and race. *Social Psychology Quart.* 79(1):81-93.

Hildebrand T, Puri M, Rocholl J (2017) Adverse incentives in crowdfunding. *Management Sci.* 63(3):587-608.

Himmelstein DU, Warren E, Thorne D, Woolhandler S (2005) Illness and injury as contributions to bankruptcy. *Health Affairs* 24: 63-73.

Hynes R (2012) Payday lending, bankruptcy, and insolvency. *Washington & Lee Law Rev.* 69(2):607-648.

Iacus SM, King G, Porro G (2012) Causal inference without balance checking: Coarsened exact matching. *Political Anal.* 20(1):1-24.

Iyer R, Khwaja AI, Luttmer EFP, Shue K (2016) Screening peers softly: Inferring the quality of small borrowers. *Management Sci.* 62(6):1554-1577.

Jagtiani, J, Lemieux, C (2017) Fintech lending: Financial inclusion, risk pricing, and alternative information. Federal Reserve Bank of Philadelphia Working Paper No. 17-17. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3005260 .

Kim K, Hann I (2018) Disillusion of the democratization of finance: Housing prices and crowdfunding. Forthcoming in *Inform. Systems Res.*

Lefgren L, Mclntyre F (2009) Explaining the puzzle of cross-state differences in bankruptcy rates. *J. Law and Econom.* 52(2):367-393.

Lin MF, Prahala NR, Viswanathan S (2013) Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. *Management Sci.* 59(1):17-35.

Lin MF, Sias RW, Wei ZY (2018) The survival of noise traders: Evidence from peer-to-peer lending. Working paper available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3185608.

Lin MF, Viswanathan S (2016) Home bias in online investments: An empirical study of an online crowdfunding market. *Management Sci.* 62(5):1393-1414.

Livshits I, Macgee JC, Tertilt M (2016) The democratization of credit and the rise in consumer bankruptcies. *Rev. Econom. Stud.* 83(4):1673-1710.

Mollick E, Nanda R (2016) Wisdom or madness? Comparing crowds with expert evaluation in funding the arts. *Management Sci.* 62(6):1533-1553.

Mollick E, Robb A (2016) Democratizing innovation and capital access: The role of crowdfunding. *California Management Rev.* 58(2):72-87.

Morgan DP, Strain MR, Seblani I (2012) How payday credit access affects overdrafts and other outcomes. *J. Money, Credit and Banking* 44(2-3):519-531.

Morse A (2015) Peer-to-peer crowdfunding: Information and the potential for disruption in consumer lending. *Annual Rev. Financial Econom.* 7:463-482.

Murray MP (2006) Avoiding invalid instruments and coping with weak instruments. *J. Econom. Perspectives* 20(4):111-132.

Pope DG, Sydnor JR (2011) What's in a picture? Evidence of discrimination from Prosper.com. *J. Human Resources* 46(1):53-92.

Seamans R, Zhu F (2014) Responses to entry in multi-sided markets: The impact of Craigslist on local newspapers. *Management Sci.* 60(2):476-493.

Wei ZY, Lin MF (2017) Market mechanisms in online peer-to-peer lending. *Management Sci.* 63(12):4236-4257.

White MJ (2007) Bankruptcy reform and credit cards. *J. Econom. Perspectives* 21(4):175-199.

White MJ (2009) Bankruptcy: Past puzzles, recent reforms, and the mortgage crisis. *Amer. Law and Econom. Rev.* 11(1):1-23.

Wolfe, B, Yoo, W (2018) Crowding out banks: Credit substitution by peer-to-peer lending. Working paper, SUNY at Buffalo, Buffalo. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3000593.

Younkin P, Kuppuswamy V (2018) The colorblind crowd? Founder race and performance in crowdfunding. *Management Sci.* 64(7):3269-3287.

Zhang JJ, Liu P (2012) Rational herding in microloan markets. *Management Sci.* 58(5):892-912.

Zhu N (2011) Household consumption and personal bankruptcy. *J. Legal Stud.* 40(1): 1-37.

CHAPTER 3. HOW DOES ALGORITHMIC TRADING INFLUENCE INVESTOR PARTICIPATION IN PEER-TO-PEER ONLINE LENDING MARKETS?

3.1 Introduction

Decisions that were previously made by humans are increasingly being made by information systems. One example is algorithmic trading, which we examine in this study. Algorithmic trading, which can be loosely defined as "the use of computer systems to execute trading strategies" (Weller 2018), has reshaped equity markets and had significant implications for market performance (Menkveld 2016). Algorithmic trading also raises questions about fairness, given that some market participants do not have the expertise or sophistication to use the technology. This may create an uneven playing field in which sophisticated investors who engage in algorithmic trading crowd out unsophisticated investors who do not.

We investigate the implications of algorithmic trading in the context of peer-to-peer lending. In peer-to-peer lending, borrowers seeking loans create listings on web sites such as Prosper.com and LendingClub. Investors choose which of these borrowers to fund based on these listings. If a borrower's listing attracts enough investors, then s/he can receive the loan. Peer-to-peer lending is an interesting context for our analysis for two reasons. First, as reflected by the "peer-to-peer" label, online lending markets were originally designed to connect individual (and presumably non-sophisticated) investors with borrowers. The original model was that these individual investors would access the web site manually to review listings and identify which of their "peers" they wanted to lend to. This model is changing as institutional investors fund an increasing percentage of online loans, largely via algorithms that select the loans automatically (and very quickly). This threatens to upend the traditional model upon which online lending originally flourished. Second, this context allows us to extend prior research on the effects of algorithmic trading. In most contexts in which algorithmic trading has been studied (e.g., the stock market), the same assets (e.g., stocks) may be bought and sold multiple times. This means that there are always opportunities for non-sophisticated investors to purchase or sell assets. This is not true of online loans. Once these loans are funded by an investor(s), they are no longer available to other investors. Thus, it is possible that algorithmic trading could allow sophisticated investors to capture the entire online lending market. Further, the ultimate impact of algorithmic trading in these markets is unclear. "Manual" investors argue that they are crowded out of the market because they cannot match the speed advantage of the algorithmic investors. Algorithmic investors argue that the algorithms help satisfy borrowers' needs more quickly and efficiently, which leads to market growth and more opportunity for all investors, including manual investors.

We use data from Prosper.com to study the effects of algorithmic trading. Because we cannot directly observe which investors use algorithmic trading technologies, we study the effect of a policy change that facilitated algorithmic trading. On March 11, 2013, Prosper.com released a major upgrade to its API (Application Programming Interface). The new API made it easier to algorithmically select loans to fund by providing more data fields and improving response time. If the API helped institutional investors "crowd out" manual investors via algorithmic trading strategies, then we should see the followings after the release of the new API: high-quality loans being funded very quickly (too quickly to be funded manually) and by a relatively small number of investors who loan large amounts. Using a difference-in-differences strategy, we find precisely that. This suggests that manual investors are being "crowded" out of the market by algorithmic trading. However, there is some evidence that algorithmic trading has led to market growth, such that opportunities remain for manual investors, although these are typically for lower quality loans as measured by default risk and yield.

3.2 Literature Review and Theoretical Foundation

3.2.1 Investor Decision Making in Online Peer-to-Peer Lending Markets

Manual investors tend to rely on both traditional financial information and "soft" information to make investment decisions (Iyer et al. 2016). In addition to traditional financial information, the decision making process is influenced (and biased sometimes) by several factors, including peer decisions, borrowers' friendship networks, loan descriptions, geographical distance, cultural distance, political distance, and borrowers' appearance, gender, and race (Burtch et al. 2014; Galak et al. 2011; Greenberg and Molick 2017; Harkness 2016; Hildebrand et al. 2017; Lin et al. 2013; Lin and Viswanathan 2016; Liu et al. 2015; Pope and Sydnor 2011; Wang and Overby 2018; Younkin and Kuppuswamy 2017; Zhang and Liu 2012). Some recent studies distinguish sophisticated investors from manual investors and find that they rely on different information to screen loans or projects, but their investing performance is not necessarily different (Lin et al. 2018; Mollick and Nanda 2016; Vallee and Zeng 2018). Although there is some evidence that sophisticated investors might evaluate borrowers differently, it is unclear how this difference affects market opportunities for different types of investors.

3.2.2 Algorithmic Trading in Equity Markets

Algorithmic trading refers to "the use of computer systems to execute trading strategies" (Weller 2018) or "any form of trading using sophisticated algorithms (programmed systems) to automate all or some part of the trade cycle" (Treleaven et al. 2013). Algorithmic traders may be both faster-acting and better-informed than manual traders (Menkveld 2016). Algorithmic traders can act faster because their trades are executed automatically based on decision rules, and they may be better informed because trades are based on statistical models fed by rich market data. Algorithmic trading improves price efficiency, reduces price discovery and information acquisition, and increases market liquidity (Hendershott et al. 2011; Weller 2018). These findings stem largely from stock markets where algorithmic investors can provide liquidity by buying and selling stocks without necessarily eliminating opportunities for manual investors to also buy and sell. However, it is unclear what the impact of algorithmic trading would be in a market where algorithmic investors compete with manual investors for a fixed set of assets, such as for loans in the peer-to-peer lending context. Accordingly, we investigate the implications of algorithmic trading on an understudied outcome, i.e. investors' participation.

3.2.3 How Algorithmic Lending Might Influence Investor Participation

The basic components of algorithmic trading are automation and information (Menkveld 2016). The automation component means that most or all of the trade is executed by automated systems and technologies. In the context of online lending, automation enables investors to automate their investment decisions rather than manually logging into the platform, picking loans, and placing orders. Therefore, automation should speed up funding time, which becomes our first hypothesis: algorithmic trading can decrease loan funding time. The information component means that investors use models/algorithms to analyze information to evaluate and select borrowers. Because it is widely expected that data-driven statistical models can improve decision efficiency and accuracy, we hypothesize that algorithmic trading can improve investment performance, i.e. the performance of loans selected by algorithms should be higher. Combining H1 and H2, it is reasonable to expect that manual investors would be crowded out of the best, most quickly-funded loans. We thus propose the third hypothesis: algorithmic trading decreases the number of investors of "flash" loans, which we use to denote the top 20% of loans in terms of funding time, from fastest to slowest.

Besides a direct impact, algorithmic trading might also indirectly influence the whole market. This might occur if algorithmic trading decreases funding time and increases decision efficiency. Decreased funding time might retain/attract more borrowers to online lending platforms. Increased decision efficiency (either real or perceived) might increase investors' confidence and motivate them to fund more borrowers, especially those risky borrowers in whom they otherwise won't invest. These two effects could lead to a larger market size, which is our fourth hypothesis. Because H3 suggests less investor participation while H4 suggests more, there is no prior theoretical expectation for the overall impact of algorithmic trading on investor participation. Therefore, we propose our final hypothesis in a non-directional way, saying algorithmic trading might either increase or decrease investor participation. Table 3-1 summarizes these five hypotheses.

Table 3-1 Research Hypotheses

H1	Algorithmic trading reduces loan funding time.
H2	Algorithmic trading increases lending performance.
H3	Algorithmic trading reduces number of investors.
H4	Algorithmic trading increases market size.
H5	Algorithmic trading increases or decreases the overall investor participation.

3.3 Empirical Setting and Data

3.3.1 Empirical Setting

The focal online lending platform that we study is Prosper.com. Prosper.com operates in the following way (since 2011): (1) borrowers submit their loan requests and personal financial information to Prosper.com; (2) Prosper.com underwrites the requests (to set the interest rate) and posts the loan requests (i.e., “listings”); (3) investors choose which borrowers to fund and how much to fund; and (4) borrowers receive their loans if they attract enough investment (either 70% or 100% of the requested amount, depending on borrower’s choice). Initially, step 3 was conducted manually by investors. Currently, much of step 3 is conducted via algorithmic trading in which sophisticated investors use data provided by Prosper.com (perhaps combined with other data) to automatically select and fund loans.

3.3.2 Data and Variables

We gathered a dataset of 63,706 loans funded through Prosper.com from 2011 to 2013. We created several variables to describe each loan. Funding time is the difference between when the loan was first posted and when it was funded. We created two measures of investor concentration per loan: Number of investors is the number of investors per loan, and average funding amount is the average amount that each investor invested in the loan.

We measured each loan's performance via compound annual growth rate (CAGR) and internal rate of return (IRR) as well as whether the borrower defaulted on the loan (default status). We use the following as control variables: loan interest rate, amount borrowed, monthly payment, and several variables about the borrower, including monthly income, debt-to-income ratio (including the loan), months employed, number of credit inquires in the last 6 months, and open credit lines. Table 3-2 provides descriptive statistics of this dataset.

Table 3-2 Descriptive Statistics of Loan Dataset

Variable	**Units**	**Min**	**Max**	**Mean**	**Median**	**Std. Dev**
Loan Funding Variables						
Funding time (in hours)	Continuous	0.00	355.26	48.48	2.18	80.92
Funding time (in logged seconds)	Continuous	1.39	14.06	8.55	8.97	4.06
Number of investors	Count	1	779	66.66	34	90.06
Average funding amount per investor	Continuous	32.26	35,000	3,227.27	153.85	5,902.76
Amount borrowed	Continuous	2,000	35,000	9,045.99	7,500	6,147.36
Interest rate	Continuous	0.050	0.330	0.202	0.198	0.072
Loan Performance Variables						
Compound annual growth rate (CAGR)	Continuous	-1	0.660	0.023	0.057	0.141
Internal rate of return (IRR)	Continuous	-1	0.702	0.050	0.091	0.185
Default status	1 if defaulted	0	1	0.204	0	0.403
Loan Credit Variables						
FICO score	Continuous	600	835	697	690	38.751
Monthly payment (in $1,000)	Continuous	0.041	2.252	0.294	0.252	0.183
Borrower stated monthly income (in $1,000)	Continuous	0	1,750	5.921	5	9.161
Debt-to-income ratio	Continuous	0	2	0.401	0.240	0.512
Months employed	Count	0	755	104.433	75	97.415
Inquires in last 6 months	Count	0	37	0.981	1	1.427
Open credit lines	Count	0	49	9.426	9	4.912

Notes: This dataset contains 63,706 loans that are (1) eventually issued and (2) listed on Prosper platform from 2011 to 2013. Debt-to-income ratio is capped at 2.

We also gathered a complementary dataset of all listings that appeared on Prosper.com during the same period. Not all listings become funded loans. Listings may expire (if a listing fails to pass the funding threshold), be cancelled by Prosper.com (if a listing is incomplete or contains incorrect information), or be withdrawn (if the borrower withdraws the loan application due to personal reasons). This dataset allows us examine changes in the size of the market. Table 3-3 provides descriptive statistics of this dataset.

Table 3-3 Descriptive Statistics of Listing Dataset

Variable	Units	Min	Max	Mean	Median	Std. Dev
Loan Funding Variables						
Funding time (in hours)	Continuous	0	358.91	66.51	4.20	106.00
Funding time (in logged seconds)	Continuous	0	14.07	9.62	8.91	4.08
Listing amount	Continuous	2,000	35,000	9,488.54	8,000	6,345.67
Interest rate	Continuous	0.050	0.330	0.204	0.201	0.073
Percent funded	Between [0, 1]	0	1	0.849	1	0.331
Passing funding threshold	1 if yes	0	1	0.824	1	0.381
Loan Credit Variables						
FICO score	Continuous	600	835	700	690	40.901
Monthly payment (in $1,000)	Continuous	0.041	2.260	0.309	0.269	0.190
Borrower stated monthly income (in $1,000)	Continuous	0	5,416.66	6.515	5	28.796
Debt-to-income ratio	Continuous	0	2	0.416	0.23	0.546
Months employed	Count	0	1,271	102.577	73	99.822
Inquires in last 6 months	Count	0	29	1.057	1	1.512
Open credit lines	Count	0	49	9.108	8	4.954

Notes: This dataset contains 113,500 listings that have appeared on Prosper platform between 2011 and 2013. 11,607 listings are withdrawn by borrowers, 5,793 listings expire, 30,772 listings are canceled by Prosper, and the rest 65,328 listings are funded and issued. Due to the availability of loan performance data, 63,706 out of 65,328 loans are recorded in the loan dataset. Debt-to-income ratio is capped at 2.

3.4 Empirical Strategy, Analysis, and Results

Because we cannot directly observe investors' use of algorithmic trading technologies, we study the effect of a major upgrade to Prosper's API on March 11, 2013 that facilitated algorithmic trading.[10] This upgrade yielded three key improvements. First, the data structure exposed by the API became more user-friendly, such that investors could use the API more easily. In addition, the API became faster and more responsive. Second,

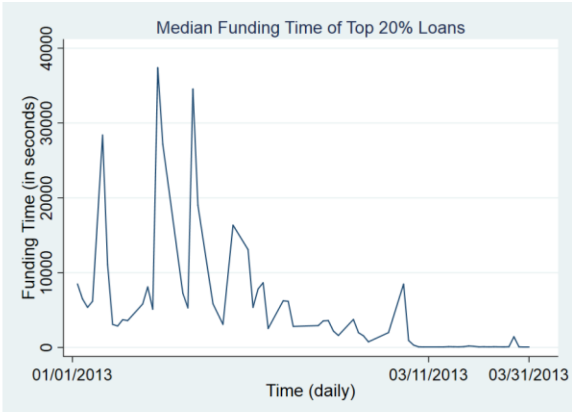
[10] See http://web.archive.org/web/20130315105454/http:/blog.prosper.com:80/2013/03/11/prosper-announces-new-api-for-lenders/ and https://www.nsrinvest.com/prospers-new-restful-api/ for more details.

approximately 460 new data elements were made available through the API, thereby permitting more sophisticated loan-selection models. Third, the API allowed investors to use third-party tools, thereby providing more options for investors to use API.

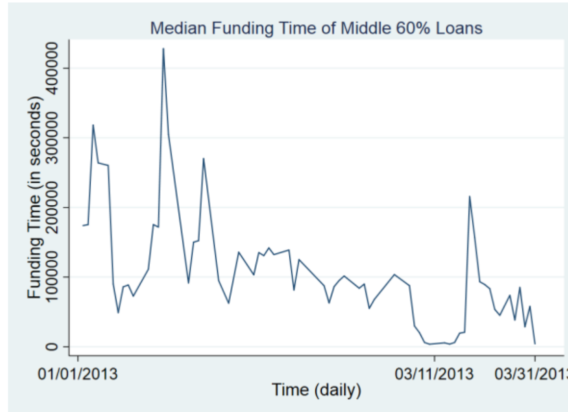
3.4.1 Model Free Evidence

We first investigated the funding time of loans across time. We classified funded loans into three categories based on how fast they got funded: top 20% ("flash" loans), the middle 60%, and the bottom 20% ("leftover" loans). Panels A, B, and C of Figure 3-1 show the median funding time of each group day by day from January 1, 2013 to March 31, 2013. A notable pattern is that funding time of "flash" loans drops significantly after the new API was released on March 11, but the funding time of "leftover" loans is almost unchanged. Panel D of Figure 3-1 shows the percentage of loans that get funded within 60 seconds, 10 minutes, and 1 hour across time. By the end of March 2013, about 10% of loans are funded in 60 seconds and 30% of loans are funded in 10 minutes. This suggests that these loans are being funded algorithmically, given that manual investors are unlikely to be able to fund loans this quickly.

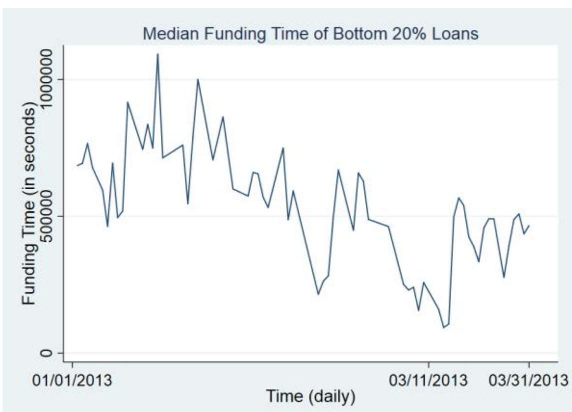
We next investigated investor concentration in loans over time. Panel A of Figure 3-2 shows the median number of investors per loan of each group from January 1, 2013 to March 31, 2013 while Panels B, C, and D of Figure 3-2 show the median average funding amount per investor per loan. After the new API was released on March 11, fewer investors share in "flash" loans or in the middle 60% of loans. This suggests that these loans are being funded by institutional investors who provide a larger portion of the loan fund.



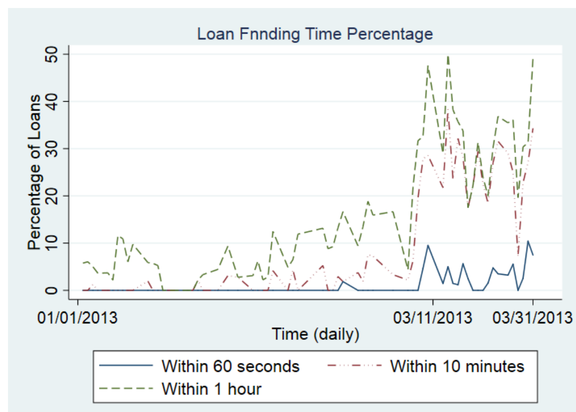
Panel A: Median Funding Time of Flash Loans



Panel B: Median Funding Time of Middle 60% Loans

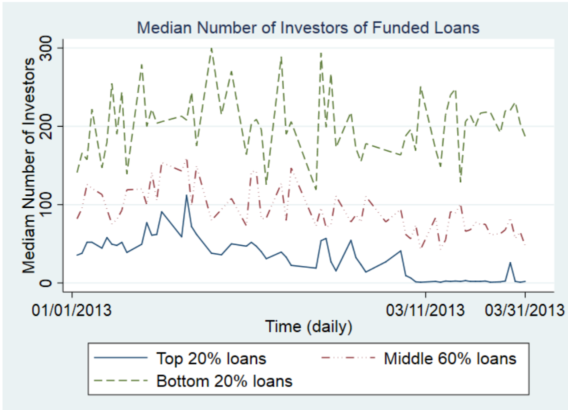


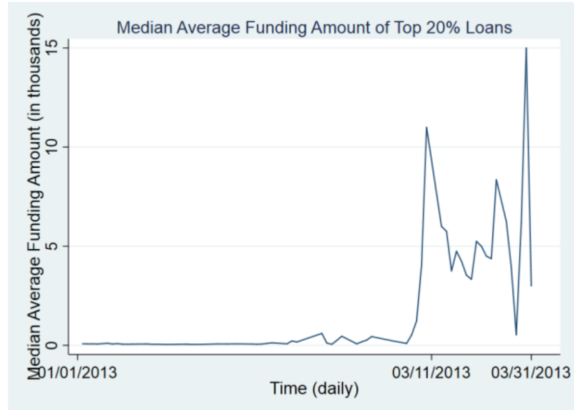
Panel C: Median Funding Time of Leftover Loans



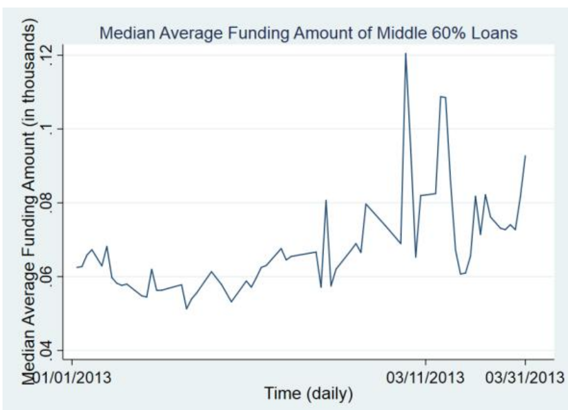
Panel D: Loan Funding Time Percentage

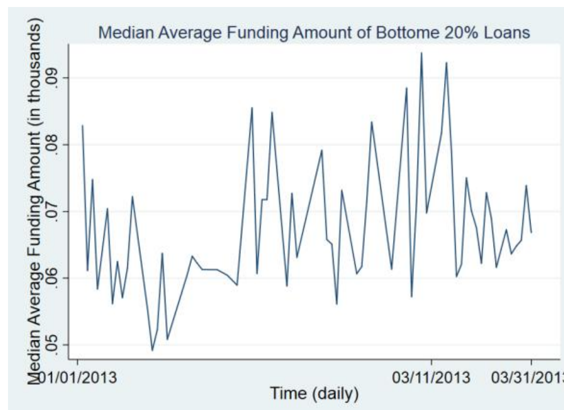
Figure 3-1 Funding Time by Group across Time





Panel A: Number of Investors of Funded Loans   Panel B: Average Funding Amount of Flash Loans





Panel C: Average Funding Amount of Middle 60%   Panel D: Average Funding Amount of Leftover Loans

Figure 3-2 Funding Time and Investor Concentration

3.4.2 Empirical Strategy

To better identify the effect of the API upgrade, we use a difference-in-differences strategy, which is widely used in studies about platform implementation, policy change or technology change (Bertrand et al. 2004; Greenwood and Agarwal 2016; Wang and Overby

2017). As shown in the model-free analysis, the “leftover” loans (i.e., those funded last) were largely unaffected by the API upgrade. We believe that this is because these loans are typically funded by manual investors who are not influenced by the API upgrade (because they don’t use the API). Thus, we use “leftover” loans as the control group. Conversely, the “flash” loans appear to be affected by the API upgrade, and we consider them to be the treated group. The difference-in-differences approach allows us to separate the effect of any general supply or demand shock or macro-economic trend (which should affect both “flash” and “leftover” loans) from the effect of the API upgrade (which should not affect “leftover” loans). We use the middle 60% of loans as a secondary treatment group. Including them allows us to check the proposed mechanism, because we expect them to be affected in the same direction as the “flash” loans but with a smaller magnitude.

To avoid conflating our analysis with other policy changes made by Prosper.com (such as its introduction of the “whole loan” program in April 2013), we restricted our analysis to the days close to the March 11, 2013 treatment date. At the month level, we define February 2013 as the pre-treatment period while March 2013 as the post-treatment period. At the day level, we define 1st-24th February 2013 as the pre-treatment period and 11th-31st March 2013 as the post-treatment period; we left a 14-day gap to avoid overlapping listings. Flash (i.e. top 20%), middle 60%, and leftover (i.e. bottom 20%) are defined based on funding time ranking within each month when analysis is at monthly level and within each day when analysis is at daily level. The basic DID model is shown in specification (1).

$$Y_{it} = \alpha + \beta_1 LoanGroup_i + \beta_2 PostTreatment_t + \beta_3 DID_{it} + \gamma_1 X_i + \gamma_2 Time_t + \varepsilon_{it} \tag{1}$$

Y_{it} indicates the outcome variables including funding time, number of investors, average funding amount per investor and loan performance measures such as CAGR, IRR, and default. Due to the skewed distribution and the between-group variation of funding time, number of investors, and average funding amount per investor, we use the natural log of each as the dependent variable. $LoanGroup_i$ indicates whether loan i belongs to the top 20% in terms of funding time ("flash" loans), the middle 60%, or the bottom 20% ("leftover" loans). When estimating the model, we use the "leftover" loans as the baseline group. $PostTreatment_t$ is defined as 1 if a loan was listed after the treatment and 0 otherwise. DID_{it} is the interaction of $LoanGroup_i$ and $PostTreatment_t$. X_i controls for other factors that might influence the dependent variables (see Tables 3-2 and 3-3). $Time_t$ are time fixed effects; they are used only in the daily analysis (note that they are collinear with $PostTreatment_t$ in the monthly analysis).

Specification (1) is designed to test hypotheses 1 to 3, which are about the direct impacts of algorithmic trading on loan outcomes. We test hypothesis 4 via conditional correlations rather than a DID model, given that we lack a control group.

$$Y_{it} = \alpha + \beta_1 PostTreatment_t + \gamma_1 X_i + \gamma_2 Time_t + \varepsilon_{it} \tag{2}$$

Y_{it} is an indicator of listing success, measured as whether the listing passes the funding threshold (passing funding threshold) and the actual funding percentage (funding percentage). This model simply tests whether listings are more likely to become loans after the API upgrade. This allows us to test whether the API upgrade increases the number of available loans, thereby increasing the market size. We consider alternative explanations for market growth (e.g., increased number of investors) via robustness checks. We find that the API upgrade also changes the "taste" of investors, which is not likely to be a consequence of increased investor money.

3.4.3 Main Results

The results of our tests of H1 to H3 are shown in Tables 4 and 5. They are estimated based on specification (1). H1 is supported at both the monthly and daily levels. Columns 1 and 4 of Table 3-4 show that after the API upgrade, "flash" loans get funded much faster. The coefficient for the "flash" loan group after the API treatment is -3.179 and significant. This reduces the funding time of "flash" loans by 95.8% ($e^{-3.179}$-1); the corresponding reduction for middle 60% loans is 62.7% ($e^{-0.985}$-1). Considering the mean funding time for "flash" loans is 4,032 seconds pre-treatment, the effect amounts to eliminating 3,862 seconds. H3 is also supported at both the monthly and daily levels. Columns 2, 3, 5 and 6 in Table 3-4 show that after the API upgrade, both "flash" loans and middle 60% loans have fewer investors and a larger average funding amount per investor. For "flash" loans, the number of investors decreases by 93.7% ($e^{-2.760}$-1) and the average funding amount per investor increases by 1603% ($e^{2.835}$-1). Considering the mean number of investors is 40 pre-treatment, the decrease implies that approximately 37 investors are crowded out of "flash" loans after the API upgrade. Middle 60% loans also experience a 34.2% decrease in number of investors and a 48.3% increase in average funding amount per investor.

Table 3-4 DID Analysis on Funding Time and Investor Concentration

Model Sample	Monthly Analysis			Daily Analysis		
Outcome Variable	Log (Funding Time in Seconds)	Log (Number of Investors)	Log (Average Funding Amount)	Log (Funding Time in Seconds)	Log (Number of Investors)	Log (Average Funding Amount)
Post treatment	-0.303*** (0.037)	0.012 (0.031)	0.000 (0.036)	-1.371*** (0.161)	-0.551*** (0.059)	0.603*** (0.060)
Loan group: Flash loans	-5.016*** (0.110)	-1.501*** (0.075)	1.399*** (0.077)	-4.837*** (0.238)	-1.655*** (0.141)	1.557*** (0.146)
Loan group: Middle 60%	-1.557*** (0.056)	-0.348*** (0.039)	0.350*** (0.041)	-1.665*** (0.102)	-0.521*** (0.074)	0.527*** (0.075)
Loan group: Flash loans * post treatment	-3.179*** (0.114)	-2.760*** (0.089)	2.835*** (0.093)	-3.066*** (0.300)	-2.519*** (0.149)	2.588*** (0.162)
Loan group: Middle 60% * post treatment	-0.985*** (0.077)	-0.418*** (0.045)	0.394*** (0.047)	-1.210*** (0.186)	-0.622*** (0.082)	0.600*** (0.079)
Interest rate	3.603* (2.110)	1.493 (1.201)	-2.735** (1.178)	0.083 (2.727)	0.518 (1.702)	-1.618 (1.766)
FICO score	0.000 (0.001)	0.000 (0.000)	-0.000 (0.000)	-0.001 (0.001)	-0.000 (0.001)	0.000 (0.001)
Amount borrowed	0.043*** (0.012)	0.045*** (0.009)	0.022*** (0.009)	0.022 (0.022)	0.033* (0.016)	0.032* (0.017)
Monthly payment	0.207 (0.360)	-0.008 (0.268)	0.898*** (0.261)	0.002 (0.634)	-0.193 (0.416)	1.165*** (0.424)
Borrower stated monthly income	-0.004 (0.003)	-0.002 (0.002)	0.001 (0.002)	-0.006 (0.005)	-0.003 (0.003)	0.002 (0.004)
Debt-to-income ratio	0.239*** (0.054)	0.143*** (0.023)	-0.133*** (0.021)	0.430*** (0.064)	0.225*** (0.048)	-0.215*** (0.047)
Months employed	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)
Inquires in last 6 months	-0.006 (0.020)	0.026** (0.011)	-0.028** (0.011)	0.036 (0.035)	0.048** (0.018)	-0.050*** (0.018)
Open credit lines	-0.008 (0.005)	0.001 (0.003)	-0.000 (0.003)	0.003 (0.007)	0.006 (0.004)	-0.005 (0.005)
Loan Term Fixed Effects	√	√	√	√	√	√
Loan Grade Fixed Effects	√	√	√	√	√	√
Time fixed effects				√	√	√
# of Observations	2,581	2,581	2,581	1,979	1,979	1,979
R^2	0.8109	0.8049	0.8045	0.7426	0.7027	0.7040
F	3150.46	538.83	388.44	NA	NA	NA

Note: *** p<0.01; ** p<0.05; * p<0.1. For columns (1)-(3), robust standard errors are reported in parentheses. For columns (4)-(6), standard errors in parentheses are clustered at daily level. Monthly analysis groups loans based on funding time ranking within each month while daily analysis groups loans based on funding time ranking within each day. All results are consistent without control variables. Results on *number of investors* and *average funding amount* are consistent when raw values are used.

Table 3-5 shows the results of our tests of H2. Because algorithmic investors might have different loan preferences than manual investors (e.g. algorithmic investors might invest in only loans of grade C, D, E, and HR while manual investors might invest in only loans of grade AA, A, B, C, and D), we test H2 by examining both absolute performance (columns 1-3) and within loan grade performance (columns 4-6). We find weak to no evidence for H2. The API upgrade doesn't significantly affect the performance difference between "flash" loans and "leftover" loans. As shown below, this is likely because "flash" loans have always outperformed "leftover" loans: the API upgrade did not increase this performance gap.

Table 3-6 shows the results of our tests of H4, which were estimated from specification (2). This analysis is based on the listing dataset rather than the loan dataset used for testing H1 to H3. The coefficients for Post treatment are always positive and significant, indicating that after the API upgrade more loans are funded. Considering the pre-treatment means of funding percentage and passing funding threshold are 0.871 and 0.847, the 0.029 coefficient in column 1 represents a 3.3% increase in funding percentage while the 0.034 coefficient in column 2 represents a 4.0% increase in funding likelihood. This suggests that the API upgrade corresponds to an increase in market size (holding the number of loan supply fixed). It is possible that other events (besides the API upgrade) that occurred at a similar time as the API upgrade might explain this finding. However, the relatively narrow time window that we use helps to make this less likely.

Table 3-5 DID Analysis on Lending Performance (at Monthly Level)

Model Specification	Absolute Performance			Within Loan Grade Performance		
Outcome Variable	CAGR	IRR	Default	CAGR	IRR	Default
Post treatment	-0.006 (0.010)	-0.011 (0.013)	0.022 (0.036)	-0.002 (0.010)	-0.004 (0.014)	0.016 (0.037)
Loan group: Flash loans	0.001 (0.012)	-0.001 (0.015)	-0.021 (0.038)	0.001 (0.012)	-0.000 (0.016)	0.000 (0.042)
Loan group: Middle 60%	-0.000 (0.009)	0.000 (0.013)	-0.028 (0.032)	-0.000 (0.001)	-0.001 (0.013)	-0.020 (0.035)
Loan group: Flash loans * post treatment	0.005 (0.015)	0.015 (0.020)	-0.065 (0.049)	0.001 (0.015)	0.006 (0.020)	-0.070 (0.051)
Loan group: Middle 60% * post treatment	0.018[a] (0.012)	0.026* (0.016)	-0.028 (0.041)	0.012 (0.012)	0.018 (0.016)	-0.023 (0.042)
Interest rate				0.106 (0.198)	0.253 (0.263)	1.759*** (0.639)
FICO score				-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)
Amount borrowed				-0.001 (0.001)	-0.002 (0.001)	0.006** (0.003)
Monthly payment				0.025 (0.027)	0.035 (0.030)	-0.073 (0.056)
Borrower stated monthly income				0.000 (0.000)	0.000 (0.000)	-0.001 (0.001)
Debt-to-income ratio				-0.006 (0.005)	-0.009 (0.007)	0.038** (0.017)
Months employed				0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)
Inquires in last 6 months				-0.004** (0.002)	-0.007*** (0.003)	0.004 (0.006)
Open credit lines				0.001 (0.000)	0.001 (0.001)	0.001 (0.002)
Loan Term Fixed Effects			√	√	√	√
Loan Grade Fixed Effects			√	√	√	√
# of Observations	2,626	2,626	2,626	2,581	2,581	2,581
R^2	0.0029	0.0032	0.0504	0.0173	0.0187	0.0580
F	1.74	1.84	16.53	3.17	2.86	11.85

Note: *** p<0.01; ** p<0.05; * p<0.1. Robust standard errors are reported in parentheses.
a=0.119. The results are still mixed at daily level.

A naïve check of the market size corroborates this conclusion. The number of funded loans increases from 1,034 in February 2013 to 1,592 in March 2013 (a 54.0% increase) when the number of listings increases from 1,789 to 2,564 (a 43.3% increase).

Table 3-6 Before/After Treatment Analysis on Funding Percentage

Model Sample	**Monthly Analysis**		**Daily Analysis**	
Outcome Variable	**Funding Percentage**	**Passing Funding Threshold**	**Funding Percentage**	**Passing Funding Threshold**
Post treatment	0.029*** (0.009)	0.034*** (0.011)	0.036*** (0.011)	0.039*** (0.013)
Interest rate	0.250 (0.380)	0.078 (0.442)	0.348 (0.440)	0.110 (0.520)
FICO score	-0.0002* (0.0001)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)
Listing amount	-0.007*** (0.002)	-0.010*** (0.002)	-0.009*** (0.002)	-0.013*** (0.003)
Monthly payment	0.034 (0.052)	0.023 (0.060)	0.088 (0.060)	0.075 (0.071)
Borrower stated monthly income	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Debt-to-income ratio	0.022** (0.010)	0.024** (0.011)	0.020* (0.012)	0.022 (0.014)
Months employed	0.000* (0.000)	0.000 (0.000)	0.000** (0.000)	0.000* (0.000)
Inquires in last 6 months	0.003 (0.004)	0.003 (0.004)	0.004 (0.004)	0.004 (0.005)
Open credit lines	0.002 (0.001)	0.002* (0.001)	0.002** (0.001)	0.003** (0.001)
Loan Term Fixed Effects	√	√	√	√
Loan Grade Fixed Effects	√	√	√	√
# of Observations	4,248	4,248	3,254	3,254
Adjusted R^2	0.0206	0.0311	0.0237	0.0355
F	5.97	8.57	5.38	7.66

Note: *** p<0.01; ** p<0.05; * p<0.1. For columns (1)-(2), robust standard errors are reported in parentheses. For columns (3)-(4), standard errors in parentheses are clustered at daily level.

3.4.4 Robustness Checks and Additional Analysis

The previous results suggest that algorithmic trading reduces loan funding time, increases investor concentration, and increases the market size. Although the DID design can rule out the influence of trends that affect all loan groups equally, it is still possible that there are some confounding events that influence these groups differently. One major concern is the inflow of large amounts of investor money, which might yield similar results as those that we find. To rule out this alternative explanation, we analyzed two events in which large amount of investor money became available in the market. First, in May 2011, a large institutional investor made a $150 million investment commitment to Prosper. Second, in February 2013, investors in Michigan were allowed to invest in Prosper loans. We viewed each of these events as treatments that affected the Prosper investment pool and replicated our DID analysis using the months around each treatment. The results are shown in Table 3-7. Columns 1-3 in Table 7 show that the treatment effects are similar to those that reported in Table 4 (except for the funding time of "flash" loans). However, the magnitude is far smaller: for "flash" loans, there is no significant change in funding time, a 15.0% decrease in number of investors, and a 17.6% increase in average funding amount per investor. The corresponding magnitudes from Table 4 are 95.8%, 93.7%, and 1603%. Therefore, even if the API upgrade happened to coincide with a huge inflow of investor money, our effect is unlikely to be explained solely by the inflow of investor money. Columns 4-6 in Table 3-7 show a similar impact, with the impact limited to only the "flash" loans.

Table 3-7 Impacts of the Inflow of Large Amounts of Investor Money

Model Sample	May 2011: Institutional Investor Makes $150 Million Commitment			February 2013: Investors From Michigan Are Permitted		
Outcome Variable	**Log (Funding Time in Seconds)**	**Log (Number of Investors)**	**Log (Average Funding Amount)**	**Log (Funding Time in Seconds)**	**Log (Number of Investors)**	**Log (Average Funding Amount)**
Post treatment	-0.012 (0.023)	-0.127*** (0.028)	0.125*** (0.028)	-0.185* (0.111)	-0.075 (0.072)	0.114 (0.085)
Loan group: Flash loans	-3.415*** (0.108)	-1.008*** (0.056)	0.884*** (0.056)	-4.337*** (0.193)	-0.794*** (0.105)	0.649*** (0.107)
Loan group: Middle 60%	-0.589*** (0.027)	-0.358*** (0.025)	0.345*** (0.024)	-1.414*** (0.151)	-0.186** (0.086)	0.246** (0.097)
Loan group: Flash loans * post treatment	0.021 (0.141)	-0.163** (0.072)	0.162** (0.072)	-0.720*** (0.201)	-0.590** (0.110)	0.604*** (0.113)
Loan group: Middle 60% * post treatment	-0.092*** (0.035)	-0.119*** (0.035)	0.099*** (0.035)	-0.139 (0.148)	-0.060 (0.082)	-0.013 (0.094)
Interest rate	5.938*** (1.367)	-0.164 (0.748)	-1.218 (0.755)	-0.469 (2.167)	1.198 (1.292)	-2.949** (0.1.294)
FICO score	NA	NA	NA	0.000 (0.000)	-0.001 (0.000)	0.000 (0.000)
Amount borrowed	0.006 (0.018)	0.051*** (0.010)	-0.017* (0.010)	0.050*** (0.011)	0.060*** (0.007)	0.008 (0.008)
Monthly payment	1.482*** (0.514)	1.065*** (0.272)	2.007*** (0.270)	-0.145 (0.270)	-0.019 (0.144)	0.809*** (0.174)
Borrower stated monthly income	0.000 (0.001)	-0.001 (0.001)	0.000 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.000 (0.001)
Debt-to-income ratio	0.031 (0.028)	-0.019 (0.019)	0.021 (0.019)	0.073 (0.051)	0.043* (0.026)	-0.043* (0.026)
Months employed	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000* (0.000)	-0.000** (0.000)
Inquires in last 6 months	0.011 (0.009)	-0.002 (0.007)	0.002 (0.007)	-0.023 (0.021)	0.014 (0.015)	-0.004 (0.015)
Open credit lines	-0.008** (0.004)	0.004* (0.002)	-0.002 (0.002)	-0.005 (0.005)	-0.001 (0.003)	0.002 (0.003)
Loan Term Fixed Effects	√	√	√	√	√	√
Loan Grade Fixed Effects	√	√	√	√	√	√
# of Observations	2,818	2,818	2,818	1,102	1,102	1,102
R^2	0.7034	0.6546	0.3485	0.8130	0.7217	0.4612
F	265.01	326.83	77.59	269.98	135.56	20.04

Note: *** p<0.01; ** p<0.05; * p<0.1. Robust standard errors are reported in parentheses. This analysis is conducted at monthly level.

Regarding investment performance, we find no support that algorithmic trading improves investment performance. To explore this, we conducted an additional analysis. It is possible that "flash" loans have always outperformed "leftover" loans, and this performance difference is not enlarged by the API upgrade. In Columns 1-3 in Table 3-8 we simply compare the performance of "flash" loans, middle 60% loans, and "leftover" loans in the year 2013. "Flash" loans (as well as middle 60% loans) outperform "leftover" loans: they have a 26.8% (0.009/0.038) advantage in CAGR, a 23.1% (0.015/0.065) advantage in IRR, and an 11.1% (0.028/0.252) advantage in default rate (i.e. lower default rate). When funding time is used as the key independent variable in columns 4-6, the results are the same: slower funding time predicts poorer performance. Given these findings, we conclude that the API upgrade doesn't bring in an additional performance advantage for "flash" loans compared to "leftover" loans (at least not in the short term around the API upgrade).

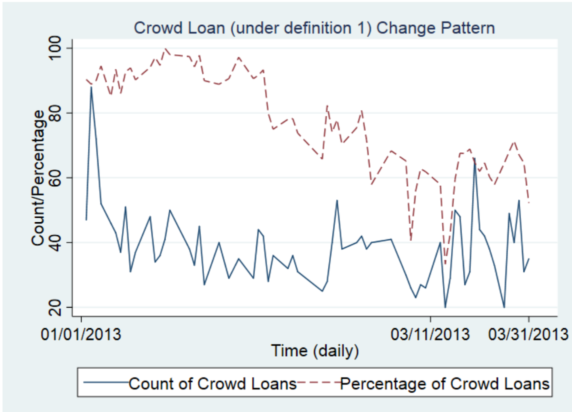
Table 3-8 Funding Time, Algorithmic Trading, and Lending Performance

Model Sample	**Year 2013 Fully Funded Loans with Loan Term of 36-Months**					
Outcome Variable	**CAGR**	**IRR**	**Default**	**CAGR**	**IRR**	**Default**
Loan group: Flash loans	0.009^{**} (0.003)	0.015^{***} (0.004)	-0.028^{**} (0.010)			
Loan group: Middle 60%	0.007^{***} (0.001)	0.015^{***} (0.001)	-0.019^{***} (0.005)			
Log (funding time in seconds)				-0.001^{***} (0.000)	-0.002^{***} (0.000)	0.003^{**} (0.001)
Constant	0.038^{***} (0.001)	0.065^{***} (0.001)	0.252^{***} (0.018)	0.055^{***} (0.003)	0.098^{***} (0.004)	0.201^{***} (0.017)
Loan Grade Fixed Effects			√			√
Time Fixed Effects	√	√	√	√	√	√
# of Observations	20,183	20,183	20,183	20,182	20,183	20,183
R^2	0.0034	0.0051	0.0355	0.0035	0.0050	0.0354

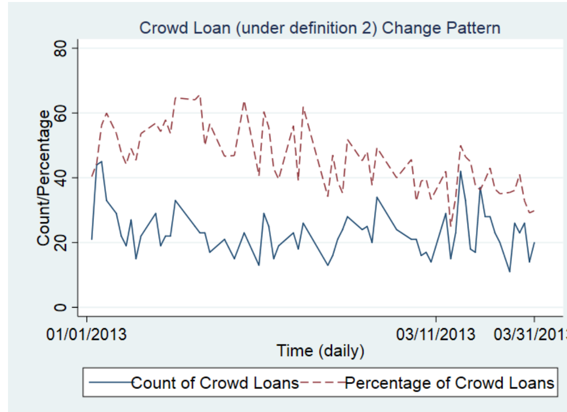
Note: *** p<0.01; ** p<0.05; * p<0.1. Standard errors in parentheses are clustered at month level. All models don't include credit information (to check absolute performance). The results are similar when 60-month loans are included.

3.4.5 Implications for Investor Participation

Our results indicate that algorithmic trading crowds out manual investors from "flash" loans and middle 60% loans but also increases the market size. Thus, it is not clear whether the absolute number of loans available to manual investors increases or decreases. Unfortunately, we cannot observe the total investment amount from manual investors. Instead, we create two proxy indicators for "crowd" loans, i.e., loans funded by manual investors. In the first approach, we define a loan as a "crowd" loan when the average funding amount per investor is less than $100. We then count both the number of "crowd" loans and the percentage of "crowd" loans per day before and after the API upgrade. Panel A of Figure 3-3 shows the result. The second approach (the results of which are shown in Panel B of Figure 3-3) is similar, except that we define a loan as a "crowd" loan if the number of investors exceeds 100. With both measures, the number of "crowd" loans increases after the API upgrade, but the percentage of "crowd" loans decreases. This suggests that even though manual investors are being crowded out of some loans, market growth allows them to continue to invest in the market.



Panel A: “Crowd” Loans under Definition 1



Panel B: “Crowd” Loans under Definition 2

Figure 3-3 Total Number and Share of “Crowd” Loans

Although the *quantity* of loans available to manual investors may not suffer from algorithmic trading, the *quality* of loans might suffer. Because manual investors are crowded out of “flash” loans and middle 60% loans, and because “flash” loans and middle 60% loans always outperform “leftover” loans, it is reasonable to expect that loans available to manual investors are inferior. We test this directly in Table 3-9. In columns 1-3, the *number of investors* is the key independent variable while in columns 4-6 the *crowd loan* dummy variable (using definition 1 from above) is the key independent variable. Under both approaches the “crowd” loans available to manual investors almost always perform worse than loans funded by fewer investors. Based on the results in column 2, compared with a loan funded by 1 investor, a loan funded by 101 investors has a lower IRR with a 0.008 (0.00008*100) rate difference, indicating a 14.5% (0.008/0.055) performance decline. Column 4 shows a similar result, implying a 19.2% (0.01/0.052) performance decline. In addition, “crowd” loans experience an additional performance decline after the API upgrade, although the decline is not always significant.

Table 3-9 Performance of "Crowd" Loans

Model Specification	Number of Investor as Key Independent Variable			"Crowd" Loan Dummy as Key Independent Variable		
Outcome Variable	**CAGR**	**IRR**	**Default**	**CAGR**	**IRR**	**Default**
Post treatment	0.004*** (0.001)	-0.003** (0.001)	-0.041*** (0.005)	0.005*** (0.001)	-0.001 (0.001)	-0.046*** (0.003)
Number of investor	-0.00002* (0.00001)	-0.00008*** (0.00002)	0.0002*** (0.0000)			
Number of investor * post treatment	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)			
Crowd loan				-0.001 (0.003)	-0.010*** (0.003)	0.022*** (0.007)
Crowd loan * post treatment				-0.009** (0.003)	-0.012** (0.004)	0.009 (0.008)
Constant	0.021*** (0.001)	0.055*** (0.001)	0.185*** (0.008)	0.020*** (0.001)	0.052*** (0.001)	0.194*** (0.008)
Loan Term Fixed Effects			√			√
Loan Grade Fixed Effects			√			√
Time Fixed Effects	√	√	√	√	√	√
# of Observations	29,657	29,657	29,657	29,657	29,657	29,657
R^2	0.0035	0.0043	0.0540	0.0035	0.0040	0.0539

Note: *** p<0.01; ** p<0.05; * p<0.1. Standard errors in parentheses are clustered at month level. Credit controls are excluded to check overall performance.

To summarize the overall implications of algorithmic lending on investor participation, we compare four basic statistics of "flash" loans, middle 60% loans, and "leftover" loans before and after the API upgrade. In Figure 3-4, the x-axis represents the natural log of funding time in seconds and the y-axis represents the number of investors. Circle size is proportional to the number of funded loans and a darker color indicates better loan performance. Arrows show the changes triggered by algorithmic trading. Loosely speaking, deep-pocket investors are more likely to be algorithmic investors, so they can funds loans in a faster and smarter way. As a result, they fund the majority of "flash" loans and middle 60% loans, which typically perform better than "leftover" loans. Individual investors are more likely to be manual investors who are crowded out of "flash" and middle

60% loans by algorithmic investors. Although manual investors may still be able to fully allocate their money, they can only select from “leftover” loans, which harms their investment performance. In a nutshell, algorithmic trading creates a new lending environment: algorithmic investors are likely to achieve high returns at the expense of manual investors, who are likely to receive lower (although perhaps still acceptable) returns.

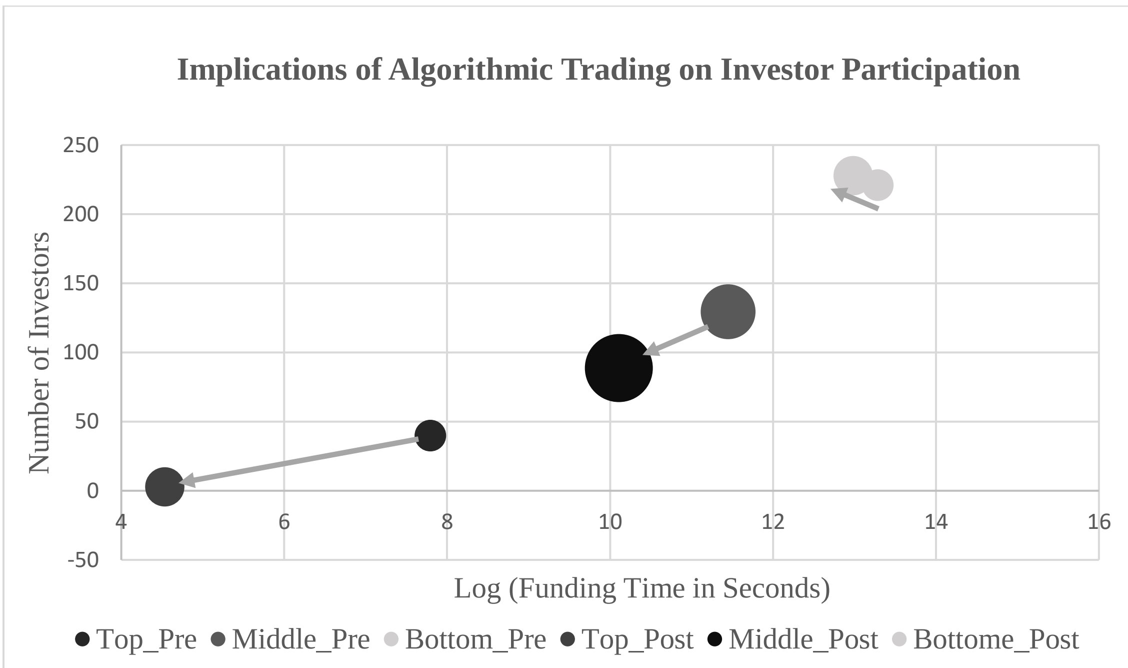


Figure 3-4 Pattern Change of Investor Participation (Pre and Post API Upgrade)

3.5 Conclusion

Algorithmic trading has great potential to affect investor participation in online lending markets, although the effect is unclear a priori. Leveraging a policy change likely to facilitate algorithmic trading (viz., an API upgrade), we identify the impacts of algorithmic trading on funding time, lending performance, investor concentration, and market size. We find that algorithmic trading significantly reduces the funding time of loans. As a result, manual investors are crowded out of the most quickly-funded and best performing loans. However, manual investors are able to continue investing in the market, given that algorithmic trading appears to increase the market size. However, manual investors are restricted to "leftover" loans, which normally perform worse than other loans. Our findings reveal several promises of algorithmic trading, but they also suggest that algorithmic trading may exacerbate inequality among market participants. Indeed, Prosper.com has launched initiatives to make automated trading tools more accessible to all investors as well as limiting API functionality to reduce the advantages enjoyed by institutional investors. We explored the impact of these initiatives and found that both appear to "level the playing field". This study contributes to several emerging research areas, including online lending, algorithmic trading, data-driven decision making, and more broadly the economics of artificial intelligence. This study also explores means to alleviate the "disparate impact" of algorithmic trading on less sophisticated investors.

3.6 References

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

Burtch, G., Ghose, A., and Wattal, S. 2014. "Cultural Differences and Geography as Determinants of Online Prosocial Lending," *MIS Quarterly* 38(3), pp. 773-794.

Galak, J., Small, D., and Stephen, A.T. 2011. "Microfinance Decision Making: A Field Study of Prosocial Lending," *Journal of Marketing Research* 48, pp. S130-S137.

Greenberg, J., and Mollick, E. 2017. "Activist Choice Homophily and the Crowdfunding of Female Founders," *Administrative Science Quarterly* 62(2), pp. 341-374.

Greenwood, B.N., and Agarwal, R. 2016. "Matching Platforms and HIV Incidence: An Empirical Investigation of Race, Gender, and Socioeconomic Status," *Management Science* 62(8), pp. 2281-2303.

Harkness, S.K. 2016. "Discrimination in Lending Markets: Status and the Intersections of Gender and Race," *Social Psychology Quarterly* 79(1), pp. 81-93.

Hendershott, T., Jones, C.M., and Menkveld, A.J. 2011. "Does Algorithmic Trading Improve Liquidity?" *The Journal of Finance* 66(1), pp. 1-33.

Hildebrand, T., Puri, M., and Rocholl, J. 2017. "Adverse Incentives in Crowdfunding," *Management Science* 63(3), pp. 587-608.

Iyer, R., Khwaja, A.I., Luttmer, E., and Shue, K. 2016. "Screening Peers Softly: Inferring the Quality of Small Borrowers," *Management Science* 62(6), pp. 1554-1577.

Lin, M.F., Prabhala, N.R., and Viswanathan, S. 2013. "Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending," *Management Science* 59(1), pp. 17-35.

Lin, M.F., and Viswanathan, S. 2016. "Home Bias in Online Investments: An Empirical Study of an Online Crowdfunding Market," *Management Science* 62(5), pp. 1393-1414.

Lin, M.F., Sias, R.W., and Wei, Z.Y. 2018. "The Survival of Noise Traders: Evidence from Peer-to-peer Lending," *SSRN working paper* available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3185608 .

Liu, D., Brass, D.J., Lu, Y., and Chen, D.Y. 2015. "Friendship in Online Peer-to-Peer Lending: Pipes, Prisms, and Relational Herding," *MIS Quarterly* 39(3), pp. 729-742.

Mollick, E., and Nanda, R. 2016. "Wisdom or Madness? Comparing Crowds with Expert Evaluation in Funding the Arts," *Management Science* 62(6), pp. 1533-1553.

Menkveld, A.J. 2016. "The Economics of High-frequency Trading: Taking Stock," *Annual Review of Financial Economics* 8, pp. 1-24.

Pope, D.G., and Sydnor, J.R. 2011. "What's in a Picture? Evidence of Discrimination from Prosper.com," *The Journal of Human Resources* 46(1), pp. 53-92.

Treleaven, P., Galas, M., and Lalchand, V. 2013. "Algorithmic Trading Review," *Communications of the ACM* 56(11), pp. 76-85.

Valle, B., and Zeng, Y. 2018. "Marketplace Lending: a New Banking Paradigm?" *SSRN working paper* available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3102984 .

Wang, H.C., and Overby, E. 2017. “How Does Online Lending Influence Bankruptcy Filings? Evidence from a Natural Experiment,” *SSRN working paper* available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2958916 .

Wang, H.C., and Overby, E. 2018. “Does Political Polarization Decrease Market Efficiency? An Investigation in the Context of Online Lending,” *working paper*.

Weller, B.M. 2018. “Does Algorithmic Trading Reduce Information Acquisition?” *The Review of Financial Studies* 31(6), pp. 2184-2226.

Younkin, P., and Kuppuswamy, V. 2017. “The Colorblind Crowd? Founder Race and Performance in Crowdfunding,” *Management Science* Articles in Advance.

Zhang, J.J., and Liu, P. 2012. “Rational Herding in Microloan Markets,” *Management Science* 58(5), pp. 892-912.

CHAPTER 4. DO POLITICAL DIFFERENCES DECREASE MARKET EFFICIENCY? AN INVESTIGATION IN THE CONTEXT OF ONLINE LENDING

4.1 Introduction

The United States is becoming increasingly polarized politically. In many cases, those with opposing political ideologies can't even agree on basic facts. This has several negative effects, including a downturn in civil discourse and an increase in political conflict. It may also have negative implications for markets. We pose the following research question: do political differences harm market efficiency by preventing transactions that might otherwise occur? We study this in the context of online lending. We investigate whether political differences inhibit market efficiency by examining whether investors in online lending markets are less likely to lend to borrowers whose political ideology is likely to be different from their own. Although prior research has examined how political differences across countries influence international trade, there is little research on how political differences between individuals may influence their economic transactions.

We use data from the first peer-to-peer online lending market in the United States: Prosper.com. This market matches borrowers seeking loans to investors willing to fund them. We approximate the political ideology of investors and borrowers by using state-level measures of political ideology drawn from the political science literature (Berry et al. 1998; 2010). We use two complementary empirical approaches to examine our research question. First, we use a difference-in-differences (DID) approach to study how investors respond to legalization of same-sex marriage in California, Connecticut, and New York. We examine how investors react to this signal of a state's (liberal) political ideology, including how this varies based on the investors' likely political ideology. Second, we estimate a gravity model – using data from all 50 U.S. states – to examine which factors, including likely differences in political ideology, influence whether investors in state i fund borrowers in state j. Both sets of analysis yield two findings. First, investors are more likely to fund borrowers in liberal states than in conservative states. This appears to be because investors view a state's liberalness as a proxy for the creditworthiness of borrowers in that state. Second, investors prefer to fund borrowers in states whose political ideology is likely to match their own. Lending activity between states with similar political ideologies is as much as 6.4% higher than that between states on polar ends of the ideological spectrum. This appears to reflect investor preferences for borrowers of their own "type", i.e., a type of in-group bias. Of these two effects, the "political distance" effect often overwhelms the general preference for borrowers in liberal states. This suggests that differences in political ideology decrease market efficiency by preventing matches between investors and borrowers that might otherwise be beneficial.

A key challenge for our analysis is measuring the political ideology of borrowers and investors using Prosper.com. As noted above, we use state-level measures as a proxy in some of our analyses. A potential concern is that the state-level measure may not match the ideology of an individual in that state. We address this in multiple ways, which we discuss fully in the "Potential Measurement Error" section below. These include: 1) estimating models that do not rely on a state-level measure of political ideology (e.g., the main DID model); 2) showing that our use of a state-level measure of borrower ideology is likely to mirror what investors do, given that investors can observe the ideology of the borrower's state but not that of the borrower himself; 3) implementing robustness checks in which the potential measurement error is minimized, and 4) showing that using group measures to proxy for individual measures is common in research in this stream (e.g., Blum and Goldfarb 2006, Burtch et al. 2014, Dajud 2013, Decker and Lim 2009, Hortacsu et al. 2009, Morrow et al. 1998, Siegel et al. 2013).

Our study contributes to research on factors that create frictions and that inhibit efficient transactions in markets. The potential of online markets to eliminate traditional market frictions is an important and enduring research stream in the Information Systems field. Because online markets eliminate transportation costs for digital goods and reduce information search costs for all goods, these markets should experience fewer frictions and be highly efficient (Hortacsu et al. 2009). This is potentially true for online lending markets, because loan information (including a borrower's credit rating and other financial information) is provided online and the loan transaction can be conducted (almost) fully digitally. However, studies show that several frictions persist in online markets, including those stemming from geographic and cultural distance (Burtch et al. 2014; Agrawal et al. 2015; Lin and Viswanathan 2016; Senney 2016). We contribute to this research stream by showing that political distance represents another friction, along with insights into why. Our findings also have practical implications. Given that many online markets are essentially two-sided markets that rely on matching to facilitate transactions/activities, understanding factors that inhibit matching is critical for the design and operation of these markets (Einav et al. 2016; Wei and Lin 2017). Because we show that political differences impact matching, designers of online lending (and other online) markets might experiment with different information revelation policies. One potential policy is to reduce information, in order to make it difficult for investors to condition their decisions on signals of borrowers' political ideology. Another policy is to increase information, in order to "drown out" potential signals of political ideology with information more relevant to online lending transactions.

4.2 Background, Literature Review, and Motivation

4.2.1 Political Ideology and Political Distance

Political ideology is "an interrelated set of moral and political attitudes that possess cognitive, affective, and motivational components" (Jost 2006, p. 653). It is usually characterized along a continuum between liberalism and conservatism. At the individual level, liberals and conservatives embrace different core beliefs and central values that manifest not only in political events but also in everyday behaviours and underlying desires (Jost 2006; Feldman and Johnson 2014). For example, conservatives are more rigid, close-minded, organized, and uncertainty averse than are liberals (Jost 2006). Conservatives generally have a stronger need than do liberals to reduce uncertainty, ambiguity, and threat (Jost and Amodio 2012). At the state level, liberal states have policies that involve greater government regulation and welfare provision than do conservative states (Caughey and Warshaw 2016). Liberal states tend to have minimal restrictions on abortion, ban the death penalty, regulate guns more tightly, offer generous welfare benefits, and have progressive tax systems (Caughey and Warshaw 2016). Jost (2006) summarized the most meaningful and enduring differences between liberal and conservative ideologies as: (1) attitudes toward inequality, and (2) attitudes toward social change versus tradition. Political distance represents the difference in political ideology between two individuals, groups, states, countries, etc.

Research has shown that political ideology and political distance influence interactions among individuals, firms, countries, etc. in often profound ways. For example, Twitter users are more likely to connect and communicate (e.g. retweet and comment) with others who have similar political ideologies (Barbera et al. 2015; Boutyline and Willer 2017). Mutual fund managers are more likely to allocate assets to firms managed by people who share their political affiliations, which is mainly due to in-group favoritism rather than possible offline connections or familiarity (Wintoki and Xi 2017). Similar political ideology between top management and independent directors is negatively associated with performance, likely because this alignment creates high empathy and leads to less monitoring (Lee et al. 2014). Political distance also creates frictions in international trade and foreign direct investment (Morrow et al. 1998; Siegel et al. 2013). Countries with dissimilar political systems trade less than countries with similar systems (Decker and Lim 2009; Dajud 2013). Possible explanations are that political distance increases the cost of negotiating trade agreements and/or that consumers prefer products from politically similar countries (Dajud 2013).

These studies provide valuable insights about the effect of political distance as well as the underlying mechanisms. However, many of these mechanisms (reduced monitoring, low negotiating costs, offline connections, etc.) are unlikely to operate in online markets. Therefore, it is important to study whether – and in what ways – political distance affects online transactions, with online lending serving as our empirical context.

4.2.2 Behavioral Biases and Market Efficiency

If political distance influences online lending behaviors, it might represent a type of behavioral bias. Online market participants have displayed several types of behavioral bias. For example, African Americans are consistently discriminated against in online e-commerce markets, online accommodation markets, and online lending markets (Pope and Sydnor 2011; Doleac and Stein 2013; Edelman et al. 2017). Furthermore, males are less preferred than females in crowdfunding markets (Greenberg and Mollick 2017). These biases can harm market efficiency by preventing the formation of matches that would benefit both parties.

Behavioral bias can operate unconditionally or conditionally. Unconditional bias occurs when members of a given race or gender are universally discriminated against, even by those of the same race or gender. Conditional bias occurs when the nature of bias is determined/moderated by the relationship between two parties. For example, conditional bias might occur if members of a given race or gender are discriminated against, but only by members of a different race or gender. One type of conditional bias is in-group bias, which occurs when people are biased against others outside of their group, which may be defined by geography, culture, political beliefs, etc. One well-established type of in-group bias is home bias, which occurs when traders in geographically distributed markets trade with those who are geographically nearby. Research on home bias has shown that institutional investors prefer same-state private equity, employers prefer same-country workers, and individual investors prefer same-state borrowers (Hochberg and Rauh 2013; Lin and Viswanathan 2016; Galperin and Greppi 2017). Another form of in-group bias relates to culture. Research has shown that lenders tend not to lend money to borrowers in countries with different cultural values (Burtch et al. 2014). In this study, we examine whether political ideology represents another form of in-group bias, whereby online lending investors prefer borrowers with ideologies likely to be similar to theirs.

4.2.3 Investors' Decision Making in Online Lending

Online lending investors rely on both traditional credit information and "soft" information to make lending decisions (Iyer et al. 2016). In addition to traditional factors, such as credit scores, income, debt-to-income ratio, the decision making process is influenced (and potentially biased) by several factors, including other investors' decisions, loan descriptions, borrowers' friendship networks, and borrowers' demographics (including gender, race, and overall "appearance") (Duarte et al. 2012; Galak et al. 2011; Greenberg and Molick 2017; Harkness 2016; Hildebrand et al. 2017; Lin et al. 2013; Pope and Sydnor 2011; Younkin and Kuppuswamy 2017; Zhang and Liu 2012). Some of these factors may not be correlated with loan default rate or expected yield, but investors still consider them. In addition to characteristics of borrowers, relationships between investors and borrowers also influence investors' decisions. Lin and Viswanathan (2016) find that investors prefer borrowers from their state, and Burtch et al. (2014) reach a similar finding with respect to cultural values, as noted above.

4.2.4 Why Political Ideology and Political Distance Might Influence Online Lending

In online lending markets, investors choose which borrowers to fund. This decision is likely to be based on multiple pieces of information provided by the online lending platform, including a borrower's credit profile and the borrower's reason for seeking the loan. Investors are also likely to infer information about borrowers that is not provided by the platform. One way that investors are likely to do this is by using a borrower's state of residence, which is the only location information provided by Prosper.com during the time frame of our analysis. For example, investors may infer that a borrower is likely to be conservative if s/he lives in Alabama (which is consistently regarded as a conservative state) and likely to be liberal if s/he lives in Massachusetts (which is consistently regarded as a liberal state). We believe that these inferences are likely because a state's political ideology is often one of its most visible characteristics to outsiders due to media coverage of state and national elections. For example, average Americans are more likely to know that Vermont is a fairly liberal state than to know that it has below average GDP per capita. We posit that investors consider a borrower's (inferred) political ideology when making lending decisions, which is plausible because most Americans "think, feel, and behave in ideologically meaningful and interpretable terms" (Jost 2006). A key goal of our paper is to test whether investors behave in a way that is consistent with this hypothesis.

There are two theoretical mechanism by which political ideology might influence investors' decisions. The first, which we label the rationality-based mechanism, is that political ideology proxies for other characteristics of a borrower that affect his/her ability or intention to repay the loan. The second, which we label the preference-based mechanism, is that investors simply prefer borrowers of a given political ideology for reasons unrelated to their ability or intention to repay. We discuss each in turn.

If present, the rationality-based mechanism would operate as follows. Assume that investors believe that liberal borrowers are better at managing debt (which signals their ability to repay) and/or are more trustworthy (which signals their intention to repay) compared to conservative borrowers. In this case, investors will prefer liberal borrowers because they rationally expect that they will get a positive return on their loan, not because they prefer liberal borrowers for their own sake. This would yield an overarching preference for liberal borrowers from both liberal and conservative investors, i.e., political differences would not matter in investors' decisions. (By the same logic, investors could believe that conservatives are better at managing debt and/or more trustworthy compared to liberal borrowers, leading to an overarching preference for conservative borrowers.) It is also possible that investors do not view a borrower's likely political ideology as an unconditional signal of his/her creditworthiness, but rather view it conditionally based on the investor's own political ideology. For example, it may be that liberal investors view liberal borrowers as being highly creditworthy, while conservative investors view conservative borrowers the same way. If investors have these beliefs, then they would tend to fund borrowers whose political ideology was likely to be similar to their own. In this case, political differences would matter in investors' decisions. We posit that one way that investors will infer a borrower's political ideology is by his/her state of residence. This is consistent with statistical discrimination theory (Phelps 1972; Fang and Moro 2011), which states that when a decision-maker lacks information about an individual (in this case, an individual borrower's political ideology), s/he will rationally substitute group averages (in this case, the political ideology of the borrower's state).

The preference-based mechanism would operate differently. In this case, liberal investors would still prefer liberal borrowers, and conservative investors would still prefer conservative borrowers. However, these preferences would not be based on the belief that political ideology signals a borrower's creditworthiness. Instead, liberal investors might simply prefer to support liberal borrowers because they are similar to them, because they are likely to share their worldview, because they like them or wish to support them, etc. (Hirshleifer 2015). The same logic may be true for conservative investors and conservative borrowers. This type of in-group political preference has been documented in several contexts. The magnitude of this in-group preference may be different for liberal and conservative investors. This is because liberals and conservatives exhibit different levels of tolerance for those that are different, with liberals typically being viewed as more tolerant (Boutyline and Willer 2017). Thus, it is important to examine whether liberal and conservative investors respond to political difference in the same way.

Both the preference-based and the rationality-based mechanisms could generate in-group favoritism, but for different reasons. In our empirical analysis below, we examine not only whether investors prefer borrowers who are likely to share their political ideology (i.e., who are within their "group"), but also which of the mechanisms appears to drive the effect.

4.3 Setting, Data, and Empirical Strategy

4.3.1 Empirical Context and Data

The online lending market that we analyze is Prosper.com, which is the first peer-to-peer online lending platform in the United States. We use data from 2008 to 2011. During this time period, Prosper.com was essentially a peer-to-peer market in which investors lent funds to borrowers. (More recently, Prosper.com has attracted many institutional investors whose decision-making processes may differ from those of individual investors.) Borrowers seeking a loan use Prosper.com's online platform to create a listing, which shows the requested loan amount along with the borrower's credit information (including credit grade, debt-to-income ratio, etc.) and state of residence. Investors search for and select borrowers to whom they want to lend money. If an investor is interested in funding a borrower, s/he places a bid, which can be as low as $25 and as high as the entire loan amount. If a borrower's listing attracts enough bids, then the loan will be issued. Prosper.com used an auction system until the end of 2010 and a posted price system since then. Under the auction system, only investors with accepted bids own a portion of the loan. Under the posted price system, each investor who places a bid owns a portion of the loan.

During the study period, Prosper.com published investor and borrower data (including their states of residence), listing data, and bid data, including which investor bids on which listings. We select years 2008, 2010, and 2011 for our analysis; each represents a distinct period for Prosper.com. We omit 2009 because Prosper.com was not operational for part of the year after the U.S. Securities and Exchange Commission (SEC) intervened to require Prosper.com to register with the SEC and with each state as a lender or loan broker. 2008 is the last full year of Prosper.com's operation before the SEC intervened. 2010 is the first full year after registration and the last year it used the auction system. 2011 is the first full year after switching from the auction system to the posted price system.

We supplement this Prosper.com dataset with data on state-level political ideology from Berry et al. (1998; 2010), who calculate a political ideology score for each state annually from 1960 to 2017 (see https://rcfording.wordpress.com/state-ideology-data/). Berry et al.'s measure is "the mean position on a liberal-conservative continuum of the active electorate in a state" (Berry et al. 1998, p. 327). Political ideology ranges from 0 to

100, with higher values indicating a more liberal leaning. We measure political distance as the absolute difference between the political ideology of the borrower's state and the political ideology of the investor's state. For example, if the political ideology of the investor's state is 60 and the political ideology of the borrower's state is 50, then political distance is 10. We use state-level political ideology as a proxy for the political ideology of individual borrowers and investors. This introduces potential measurement error into our models, although perhaps not as much as it may first appear. We discuss this in-depth in the "Potential Measurement Error" section below.

We also collect demographic and economic data from the U.S. Census and other public data sources. We use this to construct variables such as the geographic and economic distance between investor and borrower states. We measure geographic distance as the great circle distance between the investor's and borrower's state capitals. We measure economic distance as the absolute value of the difference between the real GDPs per capita of the investor's and borrower's states.

4.3.2 Empirical Strategy, Models and Results

We use three complementary approaches to study how political ideology and political distance influence online lending: 1) model free analysis, 2) a difference-in-differences approach that leverages quasi-natural experiments, and 3) a state-dyad gravity model. Using multiple approaches increases our confidence in the findings, as each approach addresses weaknesses of the other approaches. The first approach provides an overall picture of the correlation between online lending activities and political ideology, the second approach focuses on causal inference and identification, and the third approach addresses the generalizability of the findings.

Model Free Analysis. We begin with model-free analysis (Figure 1) to investigate whether investors prefer borrowers who are likely to share their political ideology. For this analysis, we use the data from 2008, although analysis using the 2010 and 2011 data yields similar results. Each bubble in Figure 4-1 depicts a state j, with the size of the bubble reflecting the number of bids placed in 2008 by investors from state j. State j's position on the x-axis reflects its political ideology in 2008. (States' political ideology measures shift over time; see https://rcfording.wordpress.com/state-ideology-data/ for measures through 2017.) State j's position on the y-axis reflects the estimated political ideology of the borrowers funded by state j's investors (estimated borrower ideology). We measure this as the weighted average of the borrower states' political ideology, with the weights determined by the number of bids from investors in state j to borrowers in state i. For example, assume that investors in state j issued 300 bids to borrowers in state A (with political ideology = 70) and 100 bids to borrowers in state B (with political ideology = 30). This would yield an estimated borrower ideology of 60 for investor state j ((300*70+100*30)/400=60). The upward-sloping dotted line shows the trend line between a state's political ideology and its estimated borrower ideology, which is weighted by the number of bids placed by investors from state j. (The unweighted trend line is similar.) The dashed horizontal line in Figure 4-1 shows estimated borrower ideology if we assume that investors from a state j fund borrowers at random. We calculated this by computing the weighted average of all states' political ideology, with the weights determined by the number of borrower listings per state.

The model-free analysis suggests two findings that we explore more formally below. First, investors prefer to fund borrowers from (relatively) liberal states: all states (except for Kansas) appear above the dashed horizontal line, which can be thought of as the baseline if investors funded borrowers randomly. We explore below whether this is due to a general preference for "liberalness" or whether "liberalness" is a proxy for creditworthiness. Second, investors in liberal states tend to prefer borrowers in liberal states (including their own state), with the analogous holding for investors in conservative states, as can be seen by the upward-sloping trend line. Because all investors have access to the same pool of borrowers, this suggests that political distance affects investors' decision-making. However, it is possible that political distance is merely a proxy for some other factor that affects investors' decision-making, which we explore below.

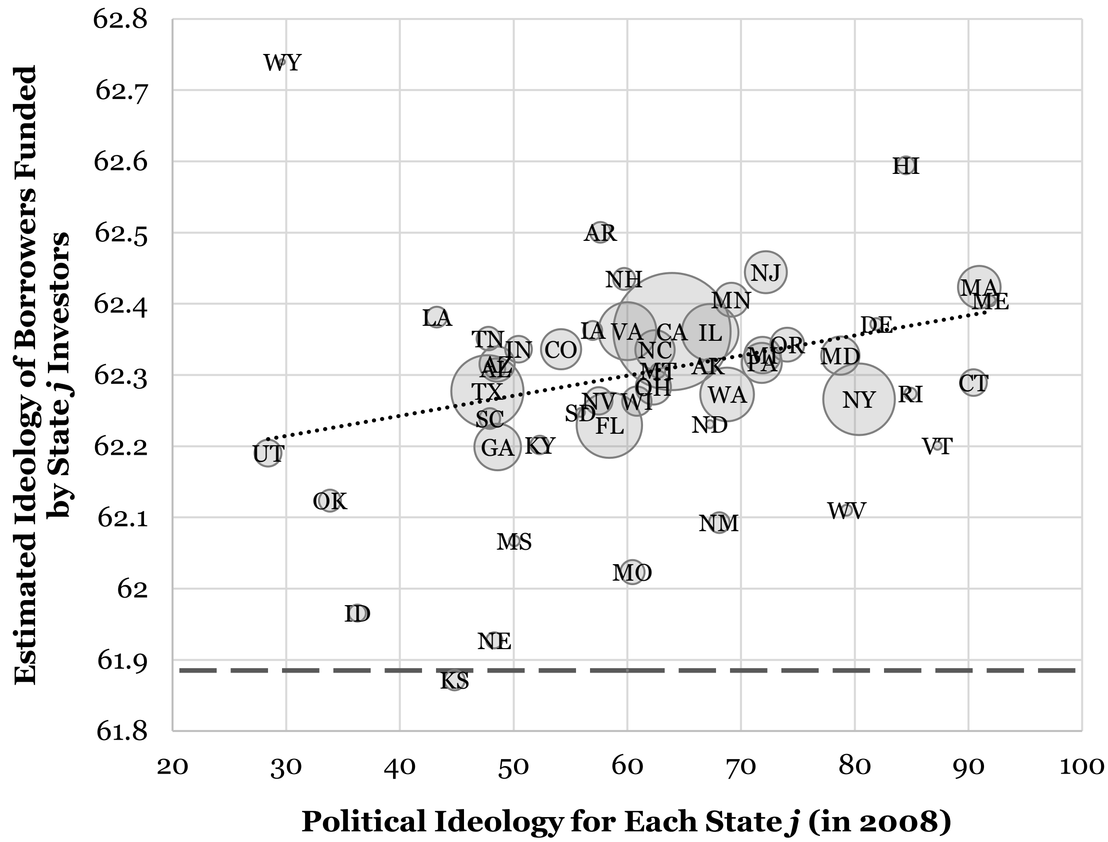


Figure Notes: Each bubble depicts a state *j*, with the bubble's size reflecting the number of bids issued by investors from state *j* in 2008. The x-axis depicts state *j*'s political ideology in 2008. The y-axis depicts estimated borrower ideology, which is the weighted average of the political ideology of the states of the borrowers that investors from state *j* invested in. The trend line shows that state *j*'s political ideology and its estimated borrower ideology are positively correlated. The dashed horizontal line depicts a state's estimated borrower ideology if the investors from that state funded borrowers randomly.

Figure 4-1 Investor Bidding Behavior by State in 2008

Difference-in-differences Model. To explore the causal effect of political ideology on investor behavior, we would ideally assign borrowers' political ideology randomly and measure how this affects investor behavior. Unfortunately, such an experiment is not feasible. Another approach is to see how investors react to changes in political ideology over time. This is also difficult because political ideology is relatively constant and slow to change (Brace et al. 2004). We overcome these issues by recognizing that the relevant consideration for investors is not a borrower's political ideology per se, but rather the investors' *perception* of a borrower's political ideology. Thus, if we could identify an event(s) that shifted investors' perceptions of borrowers' ideology in an exogenous way, then we could identify the effect of borrowers' political ideology on investor behavior. We identified state-level legalization of same-sex marriage as a suitable event for the following reasons. First, it provides a clear signal of a state's relatively liberal political ideology, given that support for lesbian, gay, bisexual, and transgender rights is typically a liberal cause (Lewis and Gossett 2008) and because a state's policy typically reflects its ideology (Brady and Sniderman 1985). Second, a state's legalization of same-sex marriage was (and remains) controversial and newsworthy, such that people across the United States are likely to notice (and therefore react to) the legalization event. There were three same-sex marriage legalization events during our study period: in California on May 15th 2008, in Connecticut on October 10th 2008, and in New York on June 24th 2011.[11]

Our basic strategy is to test whether borrowers in states that were "treated" by legalization of same-sex marriage receive more bids than do borrowers in "control" states that were not treated. If so, then this would suggest that investors prefer borrowers they perceive as being more liberal. To see why this is the case, recall that investors do not have information about individual borrowers' political ideology. Instead, they likely infer this based on the borrowers' state of residence, as suggested by statistical discrimination theory (as noted above). Investors are likely to view borrowers in "treated" states as liberal, given

[11] The state Supreme Courts for California and Connecticut ruled that same-sex couples had a right to marry on
May 15, 2008 and October 10, 2008 (Liptak 2008; McFadden 2008). The New York state legislature voted to legalize same-sex marriage on June 24, 2011 (Confessore and Barbaro 2011).

the liberal signal that same-sex marriage legalization represents. We also examine how the treatment effect varies based on investors' likely political ideology. For example, if investors from liberal states respond positively to legalization while investors from conservative states respond negatively, then that would suggest that investors prefer borrowers who are likely to share their political ideology.

We conducted a separate DID analysis for each of the 3 legalization events. Due to data availability (discussed below), we use the California event as the primary analysis, the Connecticut event as a secondary analysis, and the New York event as a tertiary analysis. We constructed our analysis sample as follows. First, because each borrower listing was available on Prosper.com for 7 days during our study period, we collected all borrower listings that were posted exactly 3 days before the event. For example, there were 484 borrowers listings posted on May 12, 2008, which is 3 days before the California event. Of these 484 listings, 56 were for borrowers from California (the "treated" listings) and 428 were for borrowers from other states (the "control" states). This allowed us to examine bids placed both 3 days before and 3 days after the legalization event for both treated and control listings. If the legalization event caused listings to receive more (or fewer) bids, then we should see an increase (or decrease) in bids for the treated listings relative to the control listings after the legalization event. Second, active investors on Prosper.com consider multiple listings when deciding which to bid on. To examine how the legalization event affected their behavior, we limited our analysis to bids placed by investors who were active on Prosper.com during the 7-day window. We defined an investor to be "active" if s/he placed at least one bid (on any listing) both before and after the legalization event. We counted the number of bids for each listing on each day from active investors in each state.

This yields a panel with investor state-listing-day as the unit of analysis. Specification (1) is our basic DID model.

$$Bid_{ijt} = \alpha + \beta Treated_{jt} + Dyad_{ij} + Time_t + \varepsilon_{ijt} \tag{1}$$

Bid_{ijt} is the number of bids from investors in state *i* to listing *j* on day *t*. This means that for each listing j on day t we have 50 observations, one for each investor state. The key explanatory variable is $Treated_{jt}$, which is set to 1 for listing-days in the state that legalized same-sex marriage after the law was passed and 0 otherwise. $Dyad_{ij}$ are investor state-listing dyad fixed effects; all time-invariant factors (i.e., features of investor states, features of listings such as the borrower's profile and loan request, and features of state-listing dyads) are captured by these fixed effects. $Time_t$ are fixed effects for each day in the 7-day window; these control for unobserved daily shocks common to all listings that influence the number of bids. We ran the analysis on the full sample and also on a matched sample. Using Coarsened Exact Matching (Iacus et al. 2011), we matched on the bids received by control and treated listings on each day before the treatment as well as on listing features, including *loan amount requested*, *the borrower's debt-to-income ratio*, *interest rate*, *monthly payment*, and whether the listing had an *image*. Descriptive statistics for each legalization event are reported in Table 4-1 for both the full and matched samples. As shown in Table 1, there are 56 treated listings in California (i.e., 56 listings were posted 3 days before the California legalization event), 4 treated listings in Connecticut, and 1 treated listing in New York. The low number of treated listings is why we consider Connecticut and New York analyses to be secondary and tertiary, as noted above.

Table 4-1 Descriptive Statistics of the Samples Used in the DID Analysis

Legalization Event	California Event		Connecticut Event		New York Event	
Event Date	May 15, 2008		Oct. 10, 2008		June 24, 2011	
Full or Matched Sample	Full	Matched	Full	Matched	Full	Matched
Number of Listings	484	299	314	113	19	4
Number of Treated Listings	56	41	4	4	1	1
Number of State-Listing Dyads	24,200	14,950	15,700	5,650	950	200
Average Daily Bids	0.040	0.020	0.040	0.005	0.065	0.066
Range of Investor State Political Ideology	[28.40, 91.90]	[28.40, 91.90]	[28.40, 91.90]	[28.40, 91.90]	[21.73, 84.50]	[21.73, 84.50]
Average Political Distance	15.95	15.65	16.01	15.77	13.64	14.93

DID Model Results. We first show the basic DID model results from specification (1) in Table 4-2. For all 3 legalization events, the treatment effect is positive and significant for both the full and matched samples. We focus on the California analysis, given that the number of treated listings is small for the Connecticut (n=4) and New York analyses (n=1). Using the matched sample for California, the estimated treatment effect of the same-sex marriage legalization event is 0.019. To gauge the magnitude of this effect, we calculated the mean number of bids for the control listings after the legalization event. This value is 0.021, such that the treatment represents a 90% increase. Another way to think about this is that treated listings received 0.019 more bids from investors in each state *i* in each day *t* on or after treatment than do control listings. When averaged across all 50 states and all 4 days on or after treatment, this suggests that treated listings received approximately 4 more bids (i.e., 0.019 * 50 states * 4 days) than did control listings (specifically, approximately 8 bids vs. approximately 4 bids). Also, because several listings receive 0 bids, we restricted the matched sample to only those listings that received at least 1 bid before the legalization event (n=14 listings, yielding n=4,900 state-listing-day observations). For this subset, the mean number of bids for the controls after legalization is 0.143 and the treatment effect is 0.129 (p=0.05), which is a 90% increase. When averaged across all 50 states and all 4 days on or after treatment, this suggests that treated listings in the sub-sample received approximately 26 more bids than did control listings (or approximately 54 bids vs. 28 bids).

Table 4-2 Results of the DID Analysis

Event/Treatment	California Event		Connecticut Event		New York Event	
Full or Matched Sample	Full	Matched	Full	Matched	Full	Matched
Treated	0.023^{***}	0.019^{*}	0.045^{***}	0.021^{**}	0.041^{*}	0.098^{***}
	(0.008)	(0.011)	(0.010)	(0.009)	(0.024)	(0.033)
Time Fixed Effects	√	√	√	√	√	√
State-Listing Dyad Fixed Effects	√	√	√	√	√	√
# of Observations (State-Listing-Days)	169,400	104,650	109,900	39,550	6,650	1,400
# of Groups (State-Listings)	24,200	14,950	15,700	5,650	950	200
Mean # of Bids for Control Listings After Legalization Event	0.040	0.020	0.040	0.005	0.065	0.066
Adjusted R^2	0.4298	0.4792	0.3331	0.0189	0.2329	0.1427

Notes: *** p<0.01; ** p<0.05; * p<0.1. Standard errors in parentheses are clustered at the state-listing level.

For DID analysis to be valid, the control and treated listings must follow parallel trends in terms of the number of bids they received before the treatment. We use the leads/lags model shown in specification (2) to check this.

$$Bid_{ijt} = \alpha + \sum_{\tau=-3}^{-2} \beta_\tau Treated_{jt+\tau} + \sum_{\tau=0}^{3} \beta_\tau Treated_{jt+\tau} + Dyad_{ij} + Time_t + \varepsilon_{ijt} \quad (2)$$

Specification (2) mirrors (1) except that we replace $\beta Treated_{jt}$ with $\sum_{\tau=-3}^{-2} \beta_\tau Treated_{jt+\tau} + \sum_{\tau=0}^{3} \beta_\tau Treated_{jt+\tau}$. $Treated_{jt+\tau}$ is a dummy variable equal to 1 for treated observations if day t is τ days after the legalization event (or for $\tau < 0$, $-\tau$ days before the event). The β_τ coefficients measure whether there is a difference in the number of bids on treated and control listings on the days before legalization, the legalization date, and the days after legalization; we use the −1 period as the baseline to avoid the "dummy variable trap”. If treated and control listings follow pre-treatment parallel trends, then β_{-3} and β_{-2} will be insignificant.

Table 4-3 shows that pre-treatment trends are often not parallel when using the full samples. This illustrates the importance of the matched samples, in which pre-treatment trends appear to be parallel: β_{-3} and β_{-2} are insignificant. Accordingly, we focus our analysis on these results. For the California analysis (which we consider as focal as noted above), the effect appears to grow larger each day, as evidenced by the increasing magnitude of the coefficients for *Treated (0)* through *Treated (3)*. This may be because it takes a few days for investors to assimilate and act on the news.

Table 4-3 Results of the DID Analysis, Including Lead and Lag Terms

Event/Treatment	California Event		Connecticut Event		New York State Event	
Full or Matched Sample	Full	Matched	Full	Matched	Full	Matched
Treated (t−3)	-0.019^{***} (0.006)	0.002 (0.006)	-0.029^{**} (0.013)	-0.000 (0.015)	-0.006 (0.059)	-0.120 (0.093)
Treated (t−2)	-0.010^{**} (0.006)	0.000 (0.005)	0.023^{**} (0.011)	0.010 (0.013)	0.014 (0.049)	0.020 (0.051)
Treated (t−1)	Baseline					
Treated (t0)	0.013 (0.008)	0.011 (0.010)	0.020^{*} (0.012)	0.005 (0.014)	-0.014 (0.054)	-0.027 (0.057)
Treated (t+1)	0.008 (0.009)	0.019^{*} (0.011)	0.032^{**} (0.012)	0.000 (0.015)	-0.060 (0.051)	-0.047 (0.053)
Treated (t+2)	0.022^{**} (0.010)	0.022^{*} (0.013)	0.047^{***} (0.016)	0.020 (0.017)	-0.076^{*} (0.045)	-0.033 (0.045)
Treated (t+3)	0.009 (0.012)	0.031^{*} (0.016)	0.072^{**} (0.032)	0.071^{**} (0.033)	0.326^{*} (0.173)	0.367^{**} (0.175)
Time Fixed Effects	√	√	√	√	√	√
State-Listing Dyad Fixed Effects	√	√	√	√	√	√
# of Observations (State-Listing-Days)	169,400	104,650	109,900	39,550	6,650	1,400
# of Groups (State-Listings)	24,200	14,950	15,700	5,650	950	200
Adjusted R^2	0.4298	0.4792	0.3331	0.0219	0.2377	0.1612

Notes: *** p<0.01; ** p<0.05; * p<0.1. Standard errors in parentheses are clustered at the state-listing level.

We next explore treatment effect heterogeneity. If investors prefer borrowers whose ideology is likely to mirror theirs, then we should see a stronger treatment effect for investors from states whose political ideology is similar to that of the treated state (i.e., California, Connecticut, or New York). To test this, we classify investors into five groups (much more liberal, somewhat more liberal, similar, more conservative, much more conservative) based on the difference between their state's political ideology and the treated state's political ideology. For the California analysis, we defined these groups as follows, where *CA* = California's political ideology and σ = the standard deviation of state political ideology: 1) much more liberal (investor state's ideology is within [CA+1.5σ, 100]), 2) more liberal [CA+0.5σ, CA+1.5σ), 3) similar [CA−0.5σ, CA+0.5σ), 4) more conservative [CA−1.5σ, CA−0.5σ), and 5) much more conservative [0, CA−1.5σ).[12] Using the matched samples, we reran specification (1) after interacting the $Treated_{jt}$ dummy variable with dummy variables for four of the five groups; we withheld the interaction term for the "similar" group to use it as a baseline and to avoid the dummy variable trap. Results are reported in Columns 1, 2, and 3 of Table 4-4. The coefficient for $Treated_{jt}$ is positive and significant, indicating that investors from states with a similar political ideology as the treated state respond positively to the liberal signal. The coefficients for the interaction terms measure the differential effect based on whether the investors' state is more liberal or more conservative. All of these coefficients are negative, and those for the "much more conservative" group are significant for the California and Connecticut analysis. This indicates that investors from states with different political ideologies react less positively (or not at all) to the legalization event. In other words, political distance appears to matter:

[12] We adjusted the group definitions for Connecticut and New York, given that there are few states with more a liberal political ideology (and many with a more conservative ideology). Let *CT* = Connecticut's political ideology and σ = the standard deviation of state political ideology. The group definitions are: 1) much more liberal (investor state's ideology is within [CT+1.5σ, 100]), 2) more liberal [CT+0.5σ, CT+1.5σ), 3) similar [CT−1σ, CT+0.5σ), 4) more conservative [CT−3σ, CT−1σ), and 5) much more conservative [0, CT−3σ). We used the same group definitions for Connecticut.

it often cancels out the main effect.[13] This suggests that the positive average treatment effect reported in Table 4-2 comes mainly from investors from politically similar states. For robustness, we also used a linear interaction term (instead of the five dummy variables) and found similar results (see Appendix B-2). This limits the risk that our results are an artifact of how we defined the categories.

Table 4-4 Results of the DID Analysis, Including Treatment Effect Heterogeneity Based on Political Difference

Measure of Political Ideology	Focal Measure			Alternative Measure: Obama Advantage		
Event/Treatment	California	Connecticut	New York	California	Connecticut	New York
Treated * Investor State Much More Liberal	-0.016 (0.040)	n/a	n/a	n/a	n/a	n/a
Treated * Investor State More Liberal	-0.025 (0.028)	n/a	-0.056 (0.044)	-0.045** (0.020)	-0.037** (0.016)	-0.071* (0.040)
Treated	0.039* (0.023)	0.030* (0.015)	0.114** (0.049)	0.035* (0.020)	0.042** (0.016)	0.130*** (0.046)
Treated * Investor State More Conservative	-0.030 (0.026)	-0.008 (0.018)	-0.020 (0.048)	-0.025 (0.022)	-0.039** (0.017)	-0.051 (0.044)
Treated * Investor State Much More Conservative	-0.053** (0.024)	-0.025* (0.015)	-0.056 (0.044)	-0.041** (0.020)	-0.037** (0.016)	-0.071* (0.040)
Time Fixed Effects	√	√	√	√	√	√
State-Listing Dyad Fixed Effects	√	√	√	√	√	√
# of Observations (State-Listing-Days)	104,650	39,550	1,400	104,650	39,550	1,400
# of Groups (State-Listings)	14,950	5,650	200	14,950	5,650	200
Adjusted R^2	0.4793	0.0190	0.1407	0.4793	0.0195	0.1409

Notes: *** p<0.01; ** p<0.05; * p<0.1. Standard errors (in parentheses) are clustered at the state-listing level. There are few states that are more liberal than Connecticut and New York, leading to "n/a" for some results.

We use *Obama Advantage* as an alternative measurement of political difference and report the results in Columns 4, 5, and 6 of Table 4-4. *Obama Advantage* is the percentage of voters in each state who voted for Barack Obama (Democratic candidate) minus the percentage who voted for John McCain (Republican candidate) in the 2008 presidential election. Generally speaking, a high value of *Obama Advantage* indicates a

[13] Notice that the effect for each of the investor state groups is reflected by the sum of the $Treated_{jt}$ coefficient and the relevant interaction term.

liberal leaning. We created investor groups (much more liberal, more liberal, etc.) using the same approach as for our focal measure. These results are similar to those obtained with our focal measure, although the coefficients for the interaction terms are more often significant.

It is possible that another event occurred at the same time as the treated states legalized same-sex marriage and that this (confounding) event could generate the effect that we see. This is unlikely for several reasons. First, such an event (or set of events) would need to occur in three different states at three different times – namely 5/15/08 in California, 10/10/08 in Connecticut, and 6/24/11 in New York – but nowhere else and not at any other time. This would be an improbable coincidence. Second, any confounding event would need to explain not only the main treatment effect but also why it varies based on likely political differences between investors and borrowers. This would also be improbable. Third, we searched for other key events that occurred on the legalization event date in each state that might confound our results. In each case, the legalization event was prominently displayed while no other (potentially confounding) event was.

Gravity Model. The DID model helps us identify the causal impacts of political ideology and political distance on investor behavior. However, generalizability is a concern because all three events in the DID analysis cover short time periods and include a small group of investors and listings. To explore the more general and longer-term impacts, we use a gravity model to assess the factors that influence the number of bids placed by investors in state *i* to borrowers in state *j* across multiple years. Gravity models are widely used in international trade, foreign direct investment, and migration studies to investigate the impact of geographic distance and other types of distance on trade or migration patterns

(McCallum 1995; Wolf 2000; Anderson and Wincoop 2003; Santos Silva and Tenreyro 2006). Gravity models have also been used to study transaction patterns in online markets (Hortacsu et al. 2009; Burtch et al. 2014). For our setting, we use the following model:

$$\ln Bids_{ijt} = \ln Investors_{it} + \ln Listings_{jt} + \ln GeographicDistance_{ij} + \ln EconomicDistance_{ijt} + ListingFeatures_{jt} + InvestorStateFeatures_{it} + Political\ Ideology\ and\ Distance\ Measures_{ijt} + \varepsilon_{ijt} \quad (3)$$

The dependent variable in gravity models is typically a measure of transaction volume between two locations. Our key dependent variable is $\ln Bids_{ijt}$, which is the natural log of the number of bids from investors in state *i* to borrowers in state *j* in year *t*. For each year, this yields 2,500 observations (i.e., 50 states crossed with 50 states). The main independent variables in gravity models are measures of the mass/size of the two locations and measures of distance between them. As our "mass" variables, we use $lnInvestors_{it}$, which is the natural log of the number of investors in state *i* in year *t* and $lnListings_{jt}$, which is the natural log of the number of listings from borrowers in state *j* in year *t*. We include well-established "distance" variables, including $lnGeographicDistance_{ij}$ and $lnEconomicDistance_{ijt}$. We control for several other factors likely to influence the number of bids for each state dyad. We control for the quality of listings in the borrowers' state by including $ListingFeatures_{jt}$, which are state-level averages for borrowers' *Credit Score*, *Debt-to-Income (DTI) Ratio,* and *Estimated Monthly Payment*. We control for the potential lending power of investors by including $InvestorStateFeatures_{it}$, which includes *Median Household Income*. The independent variables of primary interest are included in $Political\ Ideology\ and\ Distance\ Measures_{ijt}$. These include *Political Difference* (which is the investor state's political ideology minus the borrower state's ideology), *Political Distance* (the absolute value of *Political Difference*), and *Borrower State Political Ideology*. As is typical with gravity models, we do not include state fixed effects because the distance measures (geographic, economic, and political) are sometimes perfectly determined when state fixed effects are included. Including state fixed effects would also "absorb" the impact of state political ideology. Table 4-5 provides summary statistics for the 2008 data used in the gravity model. Statistics for the 2010 and 2011 data are available upon request.

Table 4-5 Descriptive Statistics for 2008 Data Used in the Gravity Model

Variables	Obs.	Mean	Median	S.D.	Min	Max
Dependent Variable						
Number of Bids from Investor State to Borrower State	2,500	1,043	265	3,129	0	90,401
Key Independent Variables						
Political Distance	2,500	17.84	15.05	13.28	0	63.51
Borrower State Political Ideology	2,500	61.74	60.64	15.73	28.40	91.90
Political Difference	2,500					
Other Independent Variables						
Number of Investors	2,500	780	408	1,063	59	6,604
Number of Listings	2,500	2,365	1,310	2,782	0	16,278
Geographic Distance (Miles)	2,500	1,187	985	895	0	5,110
Economic Distance ($)	2,500	9,458	7,554	7,660	0	36,054
Average Credit Score	2,450	607.0	607.0	7.216	591.2	621.3
Average DTI Ratio	2,450	0.403	0.400	0.054	0.289	0.561
Average Monthly Payment	2,450	269.4	269.2	23.70	225.3	325.1
Investor State Median Household Income ($)	2,500	51,966	50,170	8,425	37,528	70,482

Gravity Model Estimation and Main Results. We conduct cross-sectional, year-by-year analysis as our primary approach and use panel model estimation as a robustness check. We apply both PPML (Poisson pseudo-maximum likelihood) and OLS estimation methods as suggested by previous studies (Anderson and Wincoop 2003; Santos Silva and

Tenreyro 2006). Specification (3) is written in the format for OLS estimation, which is conventional for gravity models. However, when PPML is used, the raw number of bids rather than the natural log of bids serves as the dependent variable. In general, OLS results are consistent with PPML results; some OLS results are omitted due to space limitations.

The first set of regressions (results shown in Table 4-6) test whether political distance provides additional explanatory power beyond commonly used distance variables. We begin by running the regressions without the *Political Distance* variable. The third and fourth columns show the results with *Political Distance*. A likelihood ratio test shows that adding *Political Distance* results in a statistically significant improvement in model fit (Column 3 > Column 2 > Column 1), meaning that *Political Distance* adds explanatory power. The impact of political distance is negative and significant.

Table 4-6 Results of the Gravity Model: Influence of Political Distance

Sample Year	2008	2008	2008	2008
Estimation Method	PPML	PPML	PPML	OLS
lnInvestors	1.063***	1.064***	1.061***	1.027***
	(0.009)	(0.009)	(0.009)	(0.007)
lnListings	0.986***	0.987***	0.984***	1.011***
	(0.012)	(0.012)	(0.012)	(0.007)
Average Credit Score	0.023***	0.023***	0.024***	0.021***
	(0.002)	(0.002)	(0.002)	(0.001)
Average DTI Ratio	0.662***	0.687***	0.674***	0.622***
	(0.138)	(0.142)	(0.141)	(0.135)
Average Monthly Payment	0.005***	0.005***	0.005***	0.004***
	(0.001)	(0.001)	(0.001)	(0.000)
lnMedianHouseholdIncome	0.080	0.081	0.093*	-0.092*
	(0.055)	(0.055)	(0.055)	(0.047)
lnGeographicDistance	-0.007*	-0.014*	-0.014*	-0.016**
	(0.004)	(0.007)	(0.007)	(0.007)
lnEconomicDistance		0.007	0.009	-0.014**
		(0.006)	(0.006)	(0.005)
PoliticalDistance			-0.001*	-0.001*
			(0.001)	(0.001)
# of Observations	2,450	2,450	2,450	2,450
Pseudo Log-likelihood	-76675.748	-76600.674	-76354.334	
Adjusted R^2				0.9581

Notes: *** $p<0.01$; ** $p<0.05$; * $p<0.1$. Heteroskedasticity robust standard errors in parentheses.

Because political distance can be interpreted as a percentage measure, the coefficients in Columns 3 and 4 indicate that a 1 percentage point increase in political distance leads to 0.1% decrease in total number of received bids. For states with extreme political ideologies, the impact can be as high as 6.4%.

In our next set of regressions, we add political ideology of the borrower state, i.e. *Borrower Political Ideology*, into the model. This allows us to check for any overarching preference for political ideology. We also assess whether the impact of political distance is symmetric by distinguishing whether investor states are more liberal than borrower states. We do this by including *Political Difference* in the model rather than *Political Distance*. A negative value means that the investors' state is more conservative than the borrowers' state. Then, we represent *Political Difference* via a set of dummy variables for different ranges. The ranges are: 1) much more liberal (investor state's ideology is 30 to 100 points higher than the borrower state's ideology, 2) more liberal [10, 30), 3) similar [−10, 10), 4) more conservative [−30, −10), and 5) much more conservative [−100, −30). The "similar" group serves as the omitted baseline in our regressions. One benefit of using dummy variables is that it allows us to examine a potential non-linear effect. Table 4-7 shows the results. For robustness, we also used a linear interaction term (instead of the five dummy variables) and found similar results (see Appendix B-2 for details).

The coefficient for *Borrower Political Ideology* is positive and significant (in 5 of 6 models), indicating that investors tend to prefer borrowers in liberal states. Using the result from column 1, a one percent increase in *Borrower Political Ideology* leads to a 0.2% increase in bids. The standard deviation of *Borrower Political Ideology* in 2008 is 15.73. Thus, a 0.2% effect implies that a two standard deviation increase in political ideology yields approximately 6.3% more bids. The coefficients for *Political Difference in [−100, −30)* (i.e., the investor state is much more conservative than the borrower state) are generally negative and significant: 5 of 6 coefficients are negative (with 4 significant). The coefficients for *Political Difference in [−30, −10)* are similar: 6 of 6 coefficients are negative (with 4 significant). This indicates that conservative investors prefer *not* to lend to borrowers in states that are comparatively more liberal. In this case, the "political difference effect" often overwhelms the general preference for borrowers from liberal states. The coefficients for *Political Difference in [30, 100]* (i.e., the investor state is much more liberal than the borrower state) are inconsistent: 3 coefficients are positive (with 1 significant), and 3 coefficients are negative (with 2 significant). The coefficients for *Political Difference in [10, 30)* (i.e., the investor state is more liberal than the borrower state) are generally positive but typically insignificant: 5 of 6 coefficients are positive (with 2 significant). Overall, it appears that investors from liberal states are less influenced by political distance than are investors from conservative states. Thus, the effect of political distance appears to be asymmetric.

Results Summary. The model free analysis, DID models, and gravity models consistently show that borrowers attract more bids from politically similar investors than from politically dissimilar investors. The gravity model results also suggest that comparatively liberal investors care less about political distance than comparatively conservative investors, although we do not find evidence of this in the DID analysis.

Table 4-7 Results of the Gravity Model: Influence of Political Ideology and Political Difference

Sample Year	2008	2010	2011	2008	2010	2011
Estimation Method	PPML	PPML	PPML	OLS	OLS	OLS
lnInvestors	1.064***	1.061***	1.094***	1.033***	1.047***	1.061***
	(0.009)	(0.006)	(0.006)	(0.007)	(0.009)	(0.007)
lnListings	0.987***	0.990***	1.018***	1.007***	0.915***	0.805***
	(0.011)	(0.008)	(0.009)	(0.008)	(0.016)	(0.015)
Average Credit Score	0.023***	0.004***	0.004***	0.022***	-0.001	0.006**
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Average DTI Ratio	0.826***	-0.245	0.197	0.522***	0.751**	-0.321
	(0.177)	(0.200)	(0.152)	(0.143)	(0.326)	(0.354)
Average Monthly Payment	0.005***	0.001	0.004***	0.004***	0.000	0.004***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
lnMedianHouseholdIncome	0.027	0.536***	0.642***	-0.213***	0.598***	0.761***
	(0.065)	(0.055)	(0.058)	(0.057)	(0.105)	(0.090)
lnGeographicDistance	-0.013*	-0.000	0.018***	-0.016**	-0.018	0.015
	(0.007)	(0.006)	(0.005)	(0.008)	(0.014)	(0.013)
lnEconomicDistance	0.007	0.000	-0.023***	-0.015***	0.014	-0.007
	(0.006)	(0.005)	(0.004)	(0.006)	(0.010)	(0.010)
PoliticalDifference in [30, 100]:	0.036	0.021	-0.076***	0.091***	-0.089	-0.174***
Investor State Much More Liberal	(0.039)	(0.035)	(0.026)	(0.030)	(0.065)	(0.051)
PoliticalDifference in [10, 30):	-0.003	0.048**	0.025	0.045**	0.049	0.038
Investor State More Liberal	(0.027)	(0.019)	(0.016)	(0.020)	(0.042)	(0.037)
PoliticalDifference in [-10, 10):			Omitted Baseline			
Investor State Similar						
PoliticalDifference in [-30, -10):	-0.031	-0.052**	-0.041**	-0.052**	-0.001	-0.118***
Investor State More Conservative	(0.020)	(0.021)	(0.021)	(0.021)	(0.043)	(0.041)
PoliticalDifference in [-100, -30):	-0.110***	-0.164***	0.058	-0.156***	-0.074	-0.119*
Investor State Much More Conservative	(0.034)	(0.044)	(0.047)	(0.031)	(0.073)	(0.071)
Borrower Political Ideology	0.002**	0.003***	0.001	0.002**	0.004**	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
# of Observations	2,450	2,350	2,350	2,450	2,350	2,350
Pseudo Log-likelihood	-75878.482	-29296.776	-24522.043			
Adjusted R^2				0.9590	0.9175	0.9403

Notes: *** p<0.01; ** p<0.05; * p<0.1. Heteroskedasticity robust standard errors in parentheses.

4.4 Robustness Checks and Potential Measurement Error

4.4.1 DID Model Robustness Checks

Here, we present several robustness checks to provide more evidence for the validity of our DID models. First, it is possible that listings in the treated states generally receive more bids at the end of their first listing week than do listings in the control states, perhaps because treated states have special characteristics, such as high average incomes or higher media visibility. To examine this alternative explanation, we used alternative treatment dates for each treated state. We reconstructed our “California” data set so that it was centered around New York’s treatment date (i.e., we assumed that California was treated on the day that New York was treated), created a matched sample (as above), and reran the analysis. We did the analogous using Connecticut’s treatment date for California as well as randomizing the New York and Connecticut treatment dates. If the alternative explanation is correct, then we should see a positive treatment effect, regardless of when we assume treatment to occur. Results are shown in Table 4-8. None of the coefficients of the “alternative date” treatment effects is significant, indicating that this alternative explanation is unlikely.

Table 4-8 Results of the DID Analysis, using Alternative Treatment Dates

Assumed Treatment Date	California’s Treatment Date		Connecticut’s Treatment Date		New York’s Treatment Date
State Assumed to be Treated	Connecticut	New York	California	New York	California
“Fake” Treated	-0.001 (0.003)	0.001 (0.002)	0.001 (0.002)	-0.002 (0.003)	0.085 (0.060)
Time Fixed Effects	√	√	√	√	√
State-Listing Dyad Fixed Effects	√	√	√	√	√
# of Observations (State-Listing-Days)	5,600	47,250	81,550	38,850	700
# of Groups (State-Listings)	800	6,750	11,650	5,550	100
Adjusted R^2	0.0047	0.0629	0.1011	0.1580	0.3048

Notes: *** p<0.01; ** p<0.05; * p<0.1. Standard errors in parentheses are clustered at the state-listing level. Because there were no listings from Connecticut on New York’s treatment date, there is no “New York’s Treatment Date” / “Connecticut” column. We use matched samples for each analysis.

Another concern is that legalization of same-sex marriage likely increases investors’ awareness of the treated states. In other words, it is possible that investors increase their bids in the treated states simply because the treated state is “top of mind” and not because it has issued a signal of its liberalness. The treatment heterogeneity shown in Table 4-4 suggests that this is unlikely. If the effect is purely due to awareness, then we should *not* see investor states in different groups respond differently to legalization. We also test this "awareness effect" alternative mechanism by testing the treatment effect of the occurrence of national sports events that are likely to increase awareness of a state without sending an ideological signal. We used three events: 1) the final game of the NFL playoffs between the New York Giants and the New England Patriots on Feb. 3, 2008; 2) the final game of the NHL playoffs between the Detroit Red Wings and the Pittsburgh Penguins on June 4, 2008, and 3) the final game of the NBA playoffs between the Boston Celtics and the Los Angeles Lakers on June 17, 2008. If our findings are due to a general awareness effect, then these events should generate a positive and significant treatment effect for listings in the states whose teams were participating (i.e., New York and Massachusetts for the Super Bowl, Michigan and Pennsylvania for the NHL playoffs, and Massachusetts and California for the NBA playoffs).

We use the basic DID specification and report the results in Table 4-9. We confirm that the control and treated listings follow pre-treatment parallel trends via the leads/lags specification. The treatment effects of these sport events are not significant. This suggests that our main treatment effect is unlikely to be driven by a general awareness effect.

Table 4-9 Results of the DID Analysis using National Sport Events

Event/Treatment	2008 NFL Final	2008 NHL Final	2008 NBA Final
Treated	-0.155	0.008	-0.009
	(0.140)	(0.010)	(0.006)
Time Fixed Effects	√	√	√
State-Listing Dyad Fixed Effects	√	√	√
# of Observations (State-Listing-Days)	66,150	78,050	66,150
# of Groups (State-Listings)	9,450	11,150	9,450
Adjusted R^2	0.1224	0.1205	0.1116

Notes: *** $p<0.01$; ** $p<0.05$; * $p<0.1$. Standard errors in parentheses are clustered at the state-listing level.

We conducted other robustness checks, including checking whether our results are driven by home bias. If investors prefer to fund borrowers in their same state purely due to home bias, then this could explain our finding that the effect is largest when investors and borrowers share similar political ideologies (see Table 4-4). We replicated the treatment effect heterogeneity analysis for California in 2008 (shown in Table 4-4) after excluding: a) observations involving California investors, and b) observations involving all same-state pairs. The results are shown in columns 1 and 2 of Table B-1 (see the appendix) and are similar to those shown in Table 4-4. Other robustness checks include focusing on large states only, using bid amount (instead of bid count) as the dependent variable, and using count models instead of OLS. Findings (available upon request) are similar.

4.4.2 Gravity Model Robustness Checks

As discussed above, our working theory is that investors infer borrowers' political ideology based on their state of residence and that this influences who they choose to fund. However, it is possible that political ideology and political distance are proxies for other social or economic factors. If so, then our gravity model results might be (incorrectly) attributing the effects of those factors to the political measures. We examined this by adding to the gravity model (specification 3) other social and economic factors that investors might perceive as affecting the creditworthiness of borrowers, all gathered from the U.S. Census. For each borrower state, we included 5 demographic variables (percentage of white citizens, percentage of male citizens, percentage of citizens over sixty age, percentage of divorced or separated citizens, and percentage of citizens in poverty) and 5 economic/social variables (unemployment rate, percentage of citizens with at least high school attainment, credit card delinquency rate, median household income, and monthly salary).

Table 4-10 Results of the Gravity Model with Additional Demographic and Economic Control Variables

Model Specification	Gravity Model with Additional Borrower State Variables			Gravity Model with Additional Distance Variables		
Sample Year	2008	2010	2011	2008	2010	2011
Estimation Method	PPML	PPML	PPML	PPML	PPML	PPML
PoliticalDifference in Range (30, 100)	0.104*** (0.029)	0.014 (0.029)	-0.074*** (0.021)	0.035 (0.038)	0.008 (0.035)	-0.080*** (0.027)
PoliticalDifference in Range (10, 30)	0.028 (0.020)	0.039** (0.016)	0.024** (0.012)	-0.004 (0.027)	0.040** (0.019)	0.024 (0.016)
PoliticalDifference in Range (-10, 10)			Omitted Baseline			
PoliticalDifference in Range (-30, -10)	-0.016 (0.016)	-0.065*** (0.018)	-0.040** (0.018)	-0.028 (0.020)	-0.060*** (0.022)	-0.046** (0.021)
PoliticalDifference in Range (-100, -30)	-0.044[a] (0.031)	-0.151*** (0.040)	0.048 (0.049)	-0.101*** (0.036)	-0.160*** (0.037)	0.042 (0.049)
Borrower Political Ideology	0.003*** (0.001)	0.002*** (0.001)	0.002* (0.001)	0.002** (0.001)	0.003*** (0.001)	0.001 (0.001)
Borrower State Percentage of White	-0.003*** (0.001)	-0.002 (0.001)	-0.002*** (0.001)			
Borrower State Percentage of Male	-0.076*** (0.015)	-0.121*** (0.019)	-0.068*** (0.016)			
Borrower State Percentage of People over Sixty Age	-0.006 (0.008)	-0.026*** (0.005)	-0.022*** (0.005)			
Borrower State Percentage of Divorced or Separated	-0.005 (0.006)	0.023*** (0.006)	-0.001 (0.006)			
Borrower State Percentage of Poverty	-0.056*** (0.011)	-0.027** (0.012)	-0.047*** (0.010)			
Borrower State Unemployment Rate	0.045*** (0.008)	0.024*** (0.007)	0.004 (0.005)			
Borrower State Percentage of Citizens with at least High School Attainment	-0.006 (0.005)	0.002 (0.005)	-0.010*** (0.004)			
Borrower State Credit Card Delinquency Rate	-0.070*** (0.008)	-0.004 (0.004)	0.010*** (0.003)			
Borrower State Median Household Income	-0.028*** (0.004)	-0.007** (0.004)	-0.020*** (0.003)			
Borrower State Stable Monthly salary	-0.042*** (0.006)	0.057*** (0.018)	0.027* (0.014)			
Percentage of White Distance				0.000 (0.001)	-0.001 (0.001)	-0.000 (0.001)
Percentage of Male Distance				-0.004 (0.020)	0.023 (0.015)	0.049*** (0.012)
Unemployment Rate Distance				0.015* (0.008)	-0.005 (0.007)	0.003 (0.006)
Less than High School Attainment Distance				-0.010*** (0.004)	-0.003 (0.002)	0.001 (0.002)
Median Household Income Distance				-0.000 (0.002)	0.004*** (0.001)	-0.001 (0.001)
Other Control Variables Appearing in Tables 6 and 7	√	√	√	√	√	√
# of Observations	2,450	2,300	2,350	2,450	2,350	2,350
Pseudo Log-likelihood	-56,525	-25,599	-22,764	-74,968	-28,905	-24,227

We also used these variables to compute additional distance measures between investor and borrower states, in case these might explain the impact that we attribute to political distance. These include percentage of white distance, percentage of male distance, unemployment rate distance, less than high school attainment distance, and median household income distance. Results are shown in Table 4-10 and are similar to our focal results. This suggests that investors consider the political measures above and beyond these other factors.

We also conducted other robustness checks. A common concern of gravity models is that all dyads are weighted equally, even if some dyads are much more impactful than others. Another concern is that including same-state pair leads may cause the impact of political distance to be upwardly biased. To address these concerns, we reran the gravity model: 1) after weighting each dyad by the product of the number of investors and the number of listings, 2) after retaining only dyads in which both the number of investors and borrowers exceed 100, and 3) after excluding same-state dyads. The results are reported in Appendix Table B-3 and are similar to our focal results. As other robustness checks, we used a linear interaction term for the effect of political distance and estimated a panel data model instead of the cross-sectional models by year. The results are available upon request and are largely consistent with our focal results.

4.4.3 Potential Measurement Error

One concern about our analysis is whether our measurement of political ideology is accurate enough, given that we use state-level measures. We begin this discussion by pointing out when this potential measurement error issue is not a concern. First, the main effect uncovered in the DID analysis does not rely on state-level measures of political ideology. Instead, it shows how investors respond to a liberal signal issued by a borrower's state. Second, using state-level political ideology to approximate borrowers' individual political ideology is likely to be an accurate reflection of what investors do. This is because Prosper.com did not provide information on borrowers' political ideology during the time period of our analysis. Instead, investors likely infer borrowers' ideology based on their state of residence, as suggested by statistical discrimination theory and the political science literature (Phelps 1972; Erikson et al. 1987; Fang and Moro 2011). This literature suggests that decision-makers who lack information about an individual (in this case, the borrower's political ideology) will instead use the group average (in this case, the political ideology of the borrower's state). Thus, our model likely reflects the decision-making process that investors use.

Measurement error is more likely to be a problem when we use state-level measures to approximate investors' political ideology. For example, because we rely on state-level measures of investors' political ideology in the DID model when we explore treatment effect heterogeneity based on political distance, these results could be impacted by measurement error. The gravity models also have this characteristic. We address this in several ways. First, we aggregate data to the state level for analysis, which can "wash out" measurement error across individuals (Cameron and Trivedi 2005, p. 899). In other words, although one investor might be more liberal than average for his state, another is likely to be more conservative. Thus, aggregating to the state level allows us to approximate the behavior of an average investor from each state. Second, measurement error only leads to inconsistent estimates if the error term is correlated with the (potentially) mismeasured variable (Wooldridge 2002, p. 305). It is not clear if this is an issue in our models. Third, we use alternative measures of political ideology such as Obama Advantage in our models and find similar results, which shows that our findings are not generated by measurement error specific to our focal measure. In unreported analysis, we also measured political ideology using state government ideology and state political policy ideology, gathered from and found similar results (available upon request). Fourth, it is common practice in studies investigating political distance and/or cultural distance to use aggregate-level (e.g., state, country, etc.) measures (e.g., Blum and Goldfarb 2006, Burtch et al. 2014, Dajud 2013, Decker and Lim 2009, Hortacsu et al. 2009, Morrow et al. 1998, Siegel et al. 2013). For example, Burtch et al. (2014) showed that cultural differences between countries significantly reduced lending actions between individuals in those countries. They assumed that their country-level measure of cultural values was a valid proxy for the cultural values of individuals in those countries, despite that most countries are multicultural. We make a similar assumption, although our measurement is more granular (state-level vs. country-level).

We also conduct several supplemental analyses to further explore the possibility of measurement error. Our first approach is to identify those states for which state-level political ideology was most likely to match individual-level political ideology. We do this by calculating the Obama Advantage for each county in each state, assuming that county-level ideology is a closer match for individual-level ideology than is state-level ideology. (Our focal measure of political ideology is not available at the county level.) We then compute the standard deviation of the county-level Obama Advantage variables for each state. We rerun both the DID model and the gravity model after excluding the 5 (and 10) investor states with the highest standard deviation, because these states are those for which state-level ideology is least likely to match individual-level ideology. Column 1 in Table 4-11 (4-12) reproduces the results from Column 1 in Table 4-4 (4-7) and serves as a baseline. Columns 2-3 in Tables 4-11 and 4-12 report the results after removing the investor states with the highest standard deviation of Obama Advantage. The main findings are consistent across models.

Table 4-11 DID Model Results After Excluding States with the Highest Standard Deviations of Obama Advantage

Model Specification	Matched Sample with All States	Matched Sample Excluding 5 Highest	Matched Sample Excluding 10 Highest
Legalization Event	California, 2008	California, 2008	California, 2008
Treated * Investor State Much More Liberal	-0.016 (0.040)	-0.018 (0.041)	0.001 (0.036)
Treated * Investor State More Liberal	-0.025 (0.028)	-0.027 (0.029)	-0.008 (0.021)
Treated	0.039* (0.023)	0.040* (0.024)	0.022* (0.013)
Treated * Investor State More Conservative	-0.030 (0.026)	-0.026 (0.028)	-0.017 (0.015)
Treated * Investor State Much More Conservative	-0.053** (0.024)	-0.055** (0.025)	-0.036** (0.014)
Time Fixed Effects	√	√	√
State-Listing Dyad Fixed Effects	√	√	√
# of Observations (State-Listing-Days)	104,650	94,185	83,720
# of Groups (State-Listings)	14,950	13,455	11,960
Adjusted R^2	0.4793	0.4871	0.3482

Notes: *** p<0.01; ** p<0.05; * p<0.1. Standard errors in parentheses are clustered at the state-listing level.

Table 4-12 Gravity Model Results After Excluding States with the Highest Standard Deviations of Obama Advantage

Model Specification	All States	Excluding 5 Highest	Excluding 10 Highest
Sample Year	2008	2008	2008
Estimation Method	PPML	PPML	PPML
lnInvestors	1.064***	1.066***	1.052***
	(0.009)	(0.009)	(0.011)
lnListings	0.987***	0.989***	0.990***
	(0.011)	(0.012)	(0.011)
Average Credit Score	0.023***	0.023***	0.023***
	(0.002)	(0.002)	(0.002)
Average DTI Ratio	0.826***	0.830***	0.979***
	(0.177)	(0.182)	(0.186)
Average Monthly Payment	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)
lnMedianHouseholdIncome	0.027	0.131*	0.107
	(0.065)	(0.071)	(0.074)
lnGeographicDistance	-0.013*	-0.016**	-0.028***
	(0.007)	(0.008)	(0.008)
lnEconomicDistance	0.007	0.010	0.011
	(0.006)	(0.006)	(0.007)
PoliticalDifference in [30, 100]: Investor State Much More Liberal	0.036 (0.039)	0.050 (0.040)	0.071* (0.042)
PoliticalDifference in [10, 30): Investor State More Liberal	-0.003 (0.027)	-0.000 (0.028)	0.017 (0.027)
PoliticalDifference in [-10, 10): Investor State Similar	Omitted Baseline		
PoliticalDifference in [-30, -10): Investor State More Conservative	-0.031 (0.020)	-0.025 (0.020)	-0.022 (0.023)
PoliticalDifference in [-100, -30): Investor State Much More Conservative	-0.110*** (0.034)	-0.099*** (0.036)	-0.082** (0.042)
Borrower Political Ideology	0.002**	0.002**	0.002**
	(0.001)	(0.001)	(0.001)
# of Observations	2,450	2,205	1,960
Pseudo Log-likelihood	-75,878.482	-68,127.800	-53,792.386

Notes: *** p<0.01; ** p<0.05; * p<0.1. Heteroskedasticity robust standard errors in parentheses.

Our second approach is to directly use county-level ideology instead of state-level ideology in our analysis. Nearly 10-15% of investors in our sample report their city of residence, from which we determined their county of residence. This yielded 843 investor counties. For this subset of investors, we reran the gravity model using investor-county/borrower-state dyads (instead of investor-state/borrower-state dyads). We created the political difference measures using the *Obama Advantage* measure, because our focal measure is not available at the county level. Results are shown in Table 4-13. Because the scale of the *Obama Advantage* measure differs from that of our focal measure, we redefined the groups for political difference (much more liberal, more liberal, etc.). Columns 3 and 4 show the results after excluding small counties (defined as bottom 25% counties in terms of the number of voters) and after excluding same-state dyads (technically, dyads in which the investor-county is within the borrower-state). Results are similar to our focal analysis, with two exceptions. First, not only is the "much more conservative" coefficient negative and significant (as in our focal analysis), but so also is the "much more liberal" coefficient. However, the effect size appears larger for the "much more conservative" group (although the two confidence intervals overlap slightly), suggesting that the effect is asymmetric: specifically, political distance appears to deter investors the most when they are far more conservative than borrowers. Second, the borrower political ideology coefficient remains positive (as in the focal analysis) but is insignificant.

Table 4-13 Gravity Model Analysis: Investor-County/Borrower-State Dyads

Sample Year	2008	2008	2008	2008
Model Specification	Gravity Model, with Political Distance	Gravity Model with Political Difference	Gravity Model with Political Difference	Gravity Model with Political Difference
Sample	Full Sample	Full Sample	Full Sample Excluding Small Counties	Full Sample Excluding Same-State Pairs
Political Distance	-0.002^{**} (0.001)			
PoliticalDifference in [70, 200]: Investor State Much More Liberal		-0.258^{***} (0.081)	-0.264^{***} (0.081)	-0.263^{***} (0.083)
PoliticalDifference in [20, 70): Investor State More Liberal		-0.060^{*} (0.035)	-0.059 (0.036)	-0.064^{*} (0.036)
PoliticalDifference in [-20, 20): Investor State Similar		Baseline	Baseline	Baseline
PoliticalDifference in [-70, -20): Investor State More Conservative		-0.008 (0.043)	-0.013 (0.049)	-0.002 (0.044)
PoliticalDifference in [-200, -70): Investor State Much More Conservative		-0.508^{***} (0.094)	-0.552^{***} (0.124)	-0.502^{***} (0.093)
Borrower Political Ideology		0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Other Gravity Model Controls	√	√	√	√
# of Observations	38,073	38,073	30,919	37,300
# of Counties	843	843	634	843
Pseudo Log-likelihood	-132,902	-133,035	-120,933	-128,101

Notes: *** p<0.01; ** p<0.05; * p<0.1. Heteroskedasticity robust standard errors in parentheses.

4.5 Exploration of Underlying Mechanisms

As discussed in the "Background, Literature Review and Motivation" section, our results may reflect investors' rationality or their preferences/tastes. The results would reflect rationality if investors believed political ideology to be a signal of borrowers' credit risk; otherwise, they would reflect preferences. We used a treatment effect heterogeneity analysis, based on the credit risk of the borrower, to examine this.

The intuition for this analysis is as follows. Consider two pools of borrowers, both of whom are a mix of liberals and conservatives: one with relatively low credit risk uncertainty and one with relatively high credit risk uncertainty. Consider an investor who tends to fund liberals. If the investor funds liberals because they share his political ideology (preference-based mechanism), then this would be evident in both pools. However, if the investor funds liberals not simply because he likes them but because he believes liberalness to be a signal that a borrower is an acceptable credit risk (rationality-based mechanism), then this would be evident only in the second pool. We used the DID analysis to examine which of these mechanisms was more likely. Using the matched sample for the California legalization event, we created a "low risk" and a "high risk" group of listings. The low risk group consists of listings with a requested loan amount below $5,000 and a borrower's debt-to-income (DTI) ratio below 0.26; the high risk group is analogous. Because these are the same thresholds we used when matching listings for the matched sample, each group is well-balanced and satisfies the pre-treatment parallel trends assumption.

We reran the DID analyses for each group. The results of the main DID specification are shown in columns 1 and 2 of Table 4-14. The results of the DID specification with treatment effect heterogeneity based on political differences are shown in columns 3 and 4 of Table 4-14. The overall treatment effect is only apparent for the high risk group, which supports the rationality-based mechanism. (The coefficients for the two groups are statistically different, with that for the high risk group approximately 30x larger.) Figure 4-2 plots the estimated effect sizes for each political difference group (much more liberal, more liberal, etc.) for the low and high risk listing groups (the effect size for each political difference group is the sum of the Treated coefficient and the coefficient for the relevant interaction term). The heterogeneity pattern is similar for both groups, which provides evidence that the preference-based mechanism may be at work. However, the effect sizes for the low risk group are close to 0 (note that scale of the right-hand y-axis in

Figure 4-2 is 1/10th that of the left-hand y-axis), which provides further evidence for the rationality-based mechanism.

Table 4-14 Underlying Mechanism Analysis based on DID Models (California Legalization Event)

Model Specification	Basic DID		DID with Treatment Effect Heterogeneity	
Sample	"Low Risk" Listings	"High Risk" Listings	"Low Risk" Listings	"High Risk" Listings
Treated * Investor State Much More Liberal			-0.009* (0.005)	-0.051 (0.136)
Treated * Investor State More Liberal			-0.012** (0.005)	-0.080 (0.091)
Treated	0.003 (0.002)	0.092*** (0.033)	0.009* (0.005)	0.157** (0.077)
Treated * Investor State More Conservative			-0.007 (0.005)	-0.106 (0.086)
Treated * Investor State Much More Conservative			-0.009* (0.005)	-0.186** (0.079)
Time Fixed Effects	√	√	√	√
State-Listing Dyad Fixed Effects	√	√	√	√
# of Observations (State-Listing-Days)	38,150	28,350	38,150	28,350
# of Groups (State-Listings)	5,450	4,050	5,450	4,050
Adjusted R^2	0.0318	0.5302	0.0326	0.5307

Notes: *** p<0.01; ** p<0.05; * p<0.1. Standard errors in parentheses are clustered at the state-listing level.

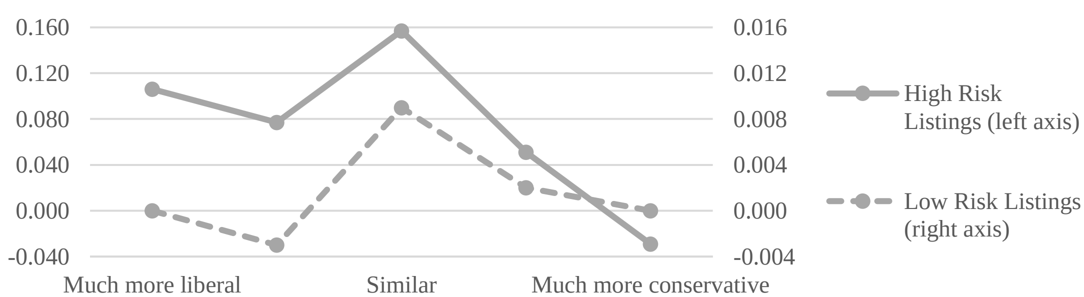


Figure 4-2 Estimated Effect Sizes By Political Difference Group for the Low and High Risk Listing Groups

We conducted a similar analysis using the gravity model. We interacted *borrower political ideology* with both *borrower credit score* and *borrower DTI ratio* (each normalized to aid interpretation) and reran the gravity model, using the 2008 data. Column 1 of Table 4-15 reproduces the results from Table 4-7, column 2 includes the *borrower credit score* interaction term, and column 3 includes the *borrower DTI ratio* interaction term. As credit score increases and/or as DTI ratio decreases, the impact of political ideology lessens. This suggests that investors pay less attention to borrowers' likely political ideology when credit risk is relatively low. This is consistent with the above result from the DID analysis and provides further evidence for the rationality-based mechanism.

We also interacted *political distance* with *borrower credit score* and *borrower DTI ratio*. Columns 4-6 of Table 4-15 show the results. Although the signs of the interaction term coefficients suggest that political distance matters less when credit scores are high and DTI ratios are low (which is consistent with the rationality-based mechanism), these coefficients are not significant (which provides evidence of the preference-based mechanism). This is consistent with the DID results reported in Table 4-14.

Overall, we conclude that both mechanisms play a role. We find a general preference for liberal investors in both the DID and gravity model analyses. This appears to be driven by a belief that liberalness signals low credit risk (i.e., the rationality-based mechanism). We also find that political distance affects investors' decisions, particularly when the investor is likely to be more conservative than the investor. This appears to be driven by a general preference for borrowers with similar political ideology (i.e., the preference-based mechanism).

Table 4-15 Underlying Mechanism Analysis based on Gravity Models

Model Specification	Gravity Model, Including Political Ideology and Difference			Gravity Model, Including Political Distance		
Borrower Political Ideology	0.0023** (0.0009)	0.0035*** (0.0010)	0.0023*** (0.0009)			
Borrower Political Ideology * Credit Score		-0.0024*** (0.0007)				
Borrower Political Ideology * DTI Ratio			0.0010*** (0.0004)			
Political Distance				-0.0014* (0.0008)	-0.0016* (0.0010)	-0.0015* (0.0008)
Political Distance * Credit Score					0.0003 (0.0009)	
Political Distance * DTI Ratio						-0.0004 (0.0005)
PoliticalDifference in [30, 100]: Investor State Much More Liberal	0.0356 (0.0392)	0.0285 (0.0395)	0.0553 (0.0392)			
PoliticalDifference in [10, 30): Investor State More Liberal	-0.0030 (0.0270)	0.0024 (0.0262)	-0.0012 (0.2679)			
PoliticalDifference in [-10, 10): Investor State Similar	baseline	baseline	baseline			
PoliticalDifference in [-30, -10): Investor State More Conservative	-0.0313a (0.0201)	-0.0331* (0.0198)	-0.0270 (0.0202)			
PoliticalDifference in [-100, -30): Investor State Much More Conservative	-0.1098*** (0.0340)	-0.1092*** (0.0336)	-0.0970*** (0.0347)			
Other Gravity Model Controls	√	√	√	√	√	√
# of Observations	2,450	2,450	2,450	2,450	2,450	2,450
Pseudo Log-likelihood	-75878.482	-75129.2	-75633.822	-76354.334	-76341.441	-76330.318

Notes: *** $p<0.01$; ** $p<0.05$; * $p<0.1$. Heteroskedasticity robust standard errors in parentheses. a: p value = 0.118.

4.6 Conclusion

Political differences are becoming increasingly stark in contemporary society, leading to a downturn in civil discourse. This paper demonstrates that political ideology and political difference also play an important role in how markets operate, with a specific focus on online lending markets. Although online lending markets eliminate geographic frictions and improve information transparency, investors still show significant behavioral biases. We collect data from Prosper.com and apply multiple models to investigate the impact of political ideology and political difference. We find that investors on average prefer liberal borrowers to conservative borrowers, which is likely due to that investors consider liberalness as an indicator of creditworthy. We also find that investors tend to prefer borrowers whose political ideology is likely to be similar to theirs. Additional analyses show that this reluctance to fund borrowers of differing political ideology is due to simple preference and/or bias. This suggests that some beneficial matches between borrowers and investors are not being made because of the negative effect of political difference, thereby reducing the market's efficiency.

Our finding that political difference serves as a barrier to online lending transactions is consistent with other types of in-group bias. In general, borrowers are more likely to attract investors from states with similar political ideology than investors from states with different political ideology. In other words, political distance deters online investment. We rule out alternative explanations related to race, gender, unemployment, education, and income. Future research can provide deeper look at the mechanisms.

Our study has implications that suggest opportunities for future research. The first implication is about platform design. For example, is it possible to remove hints about political ideology on online markets? Although we don't analyze the pros and cons of eliminating the political information, this might be interesting for platform designers. Another implication is about data analytics tools and automated investing algorithms given investors increasingly rely on these tools and algorithms to invest. Although the initial purpose of providing these tools might not be to reduce decision bias or discrimination, these tools and algorithms have the potential to get rid of many types of bias, including but not limited to political bias. Investigating the impact of automated tools on well-known behavioral biases is a promising research topic.

The findings from the DID models are largely consistent with the model-free evidence. A state's signal of liberalness attracts bids from investors in states whose political ideology is similar. Investors in states with dissimilar political ideology respond less positively (or not at all) to the signal. This indicates that political distance deters online lending.

4.7 References

Anderson, James. E. and Eric Van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." The American Economic Review 93(1): 170-192.

Agrawal, Ajay, Christian Catalini, and Avi Goldfarb. 2015. "Crowdfunding: Geography, Social Networks, and The Timing of Investment Decisions." Journal of Economics & Management Strategy 24(2): 253-274.

Bakos, J. Yannis. 1997. "Reducing Buyer Search Costs: Implications for Electronic Marketplaces." Management Science 43(12): 1676-1692.

Baltagi, Badi H., Peter Egger, and Michael Pfaffermayr. 2015. "Panel Data Gravity Model of International Trade." The Oxford Handbook of Panel Data. Oxford University Press, Oxford. Chapter 20. pp: 608-642.

Barbera, Pablo, John T. Jost, Jonathan nagler, Jpshua A. Tucker, and Richard Bonneau. 2015. "Tweeing from Left to Right: Is Online Political Communication More Than an Echo Chamber?" Psychological Science 26(10): 1531-1542.

Berry, William D, Evan J. Ringquist, Richard C. Fording, and Russell L. Hanson. 1998. "Measuring Citizen and Government Ideology in the American States, 1960-93." American Journal of Political Science 42(1): 327-348.

Berry, William D., Richard C. Fording, Evan J. Ringquist, Russell L. Hanson and Carl Klarner. 2010. "Measuring Citizen and Government Ideology in the American States: A Re-appraisal." State Politics and Policy Quarterly 10: 117-35.

Boutyline, Andrei, and Robb Willer. 2017. "The Social Structure of Political Echo Chambers: Variation in Ideological Homophily in Online Networks." Political Psychology 38(3): 551-569.

Brace, Paul, Kevin Arceneaux, Martin Johnson, and Stacy G. Ulbig. 2004. “Does State Political Ideology Change over Time?” Political Research Quarterly 57(4): 529-540.

Brandt, Mark J., Christine Reyna, John R. Chambers, Jarret T. Crawford, and Geoffrey Wetherell. 2014. “The Ideological-Conflict Hypothesis: Intolerance Among Both Liberals and Conservatives.” Psychological Science 23(1): 27-34.

Brown, Jeffery R., Joshua M. Pollet, Scott J. Weisbenner. 2014. “The In-State Equity Bias of State Pension Plans.” Available at SSRN link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2578848 .

Burtch, Gordon, Anindya Ghose, and Sunil Wattal. 2014. “Cultural Differences and Geography as Determinants of Online Social Lending.” MIS Quarterly 38(3): 773-794.

Caughey, Devin, and Christopher Warshaw. 2016. “The Dynamics of State Policy Liberalism, 1936-2014.” American Journal of Political Science 60(4): 899-913.

Chin, M.K., Donald C. Hambrick, and Linda K. Trevino. 2013. “Political Ideologies of CEOs: The Influence of Exxcutives’ Values on Corporate Social Responsibility.” Administrative Science Quarterly 58(2): 197-232.

Confessore, N., M. Barbaro 2011. New York Allows Same-Sex Marriage, Becoming Largest State to Pass Law, New York Times, June 24, 2011.

Dajud, C. Umana. 2013. “Political Proximity and International Trade.” Economics & Politics 25(3): 283-312

Decker, Jessica Henson, and Jamus Jerome Lim. 2009. “Democracy and trade: an empirical study.” Economics of Governance 10(2): 165-186.

Dixon, William J., and Bruce E. Moon. 1993. “Political Similarity and American Foreign Trade Patterns.” Political Research Quarterly 46(1): 5-25.

Doleac, Jennifer L., and Luke C.D. Stein. 2013. “The Visible Hand: Race and Online Market Outcomes.” Economic Journal 123(572): 469-492.

Edelman, Benjamin, Michael Luca, and Dan Svirsky. 2017. “Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment.” American Economic Journal: Applied Economics 9(2): 1-22.

Einav, Liran., Chiara Farronato, and Jonathan Levin. 2016. “Peer-to-Peer Markets.” Annual Review of Economics 8: 615-635.

Erikson, Robert S., Gerald C. Wright, and John P. Mclver. 1993. Statehouse Democracy: Public Opinion and Policy in the American States. Cambridge, MA. Cambridge University Press.

Erikson, Robert S., Gerald C. Wright, and John P. Mclver. 2007. "Measuring the Public's Ideological Preferences in the 50 States: Survey Responses versus Roll Call Data." State Politics & Policy Quarterly 7(2): 141-151.

Feldman, Stanley, and Christopher Johnson. 2014. "Understanding the Determinants of Political Ideology: Implications of Structural Complexity." Political Psychology 35(3): 337-358.

Galperin, Hernan, and Catrihel Greppi. 2017. "Geographical Discrimination in the Gig Economy." Available at SSRN link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2922874 .

Ghemawat, Pankaj. 2001. "Distance Still Matters: The Hard Reality of Global Expansion." Harvard Business Review 79(8): 137-147.

Greenberg, Jason, and Ethan Mollick. 2017. "Activist Choice Homophily and the Crowdfunding of Female Founders." Administrative Science Quarterly 62(2): 341-374.

Hirshleifer, David. 2015. "Behavioral Finance." Annual Review of Financial Economics 7:133-159.

Hochberg, Yael V., and Joshua D. Rauh, 2013. "Local Overweighting and Underperformance: Evidence from Limited Partner Private Equity Investments." Review of Financial Studies 26(2): 403–451.

Holbrook-Provow, Thomas M., and Steven C. Poe. 1987. "Measuring State Political Ideology." American Politics Quarterly 15(3): 399-416.

Hortacsu, Ali, F. Asis Martinez-Jerez, and Jason Douglas. 2009. "The Geography of Trade in Online Transactions: Evidence from eBay and MercadoLibre." American Economic Journal: Microeconomics 1(1): 53-74.

Hutton, Irena, Danling Jiang, and Alok Kumar. 2014. "Corporate Policies of Republican Managers." Journal of Financial and Quantitative Analysis 49(5/6): 1279-1310.

Kaustic, Markku, and Sami Torstila. 2011. "Stock Market Aversion? Political Preferences and Stock Market Participation." Journal of Financial Economics 100: 98-112.

Jost, John T. 2006. "The End of the End of Ideology." American Psychologist 61(7): 651-670.

Jost, John T., Christopher M. Federico, and Jaime L. Napier. 2009. "Political Ideology: Its Structure, Functions, and Elective Affinities." Annual Review of Psychology 60: 307-337.

Jost, John T., and David M. Amodio. 2012. “Political Ideology as Motivated Social Cognition: Behavioral and Neuroscientific Evidence.” Motivation and Emotion 36(1): 55-64.

Lax, Jeffery R., and Justin H. Phillips. 2012. “The Democratic Deficit in the States.” American Journal of Political Science 56(1): 148-166.

Lee, Jongsub, Kwang J. Lee, Nandu J. Nagarajan. 2014. “Birds of a feather: Value implications of political alignment between top management and directors.” Journal of Financial Economics 112(2): 232-250.

Lin, Mingfeng and Siva Viswanathan. 2016. “Home Bias in Online Investments: An Empirical Study of an Online Crowdfunding Market.” Management Science 62(5): 1393-1414.

Liptak, A., 2008. California Supreme Court Overturns Gay Marriage Ban, New York Times, May 16, 2008.

Marom, Dan, Alicia Robb, and Orly Sade. 2016. “Gender Dynamics in Crowdfunding (Kickstarter): Evidence on Entrepreneurs, Investors, Deals and Taste-Based Discrimination.” Available at SSRN link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2442954 .

McCallum, John. 1995. “National Borders Matter: Canada-U.S. Regional Trade Patterns.” The American Economic Review 85(3): 615-623.

McFadden, R., 2008. Gay Marriage Is Ruled Legal in Connecticut, New York Times, Oct. 10, 2008.

McPherson, M., L. Smith-Lovin, and J. M. Cook 2001 ‘‘Birds of a Feather: Homophily in Social Networks.’’ Annual Review of Sociology 27: 415–444.

Morrow, James D., Randolph M. Siverson, and Tressa E. Tabares. 1998. “The Political Determinants of International Trade: The Major Powers, 1907–1990.” American Political Science Review 92(3): 649-661.

Pool, Veronika K., Noah Stoffman, and Scott E. Yonker, 2012. “No Place like Home: Familiarity in Mutual Fund Manager Portfolio Choice.” Review of Financial Studies 25(8): 2563-2599.

Pope, Devin G., and Justin R. Sydnor. 2011. “What’s in a Picture? Evidence of Discrimination from Prosper.com.” Journal of Human Resources 46(1): 53-92.

Redding, Stephen J., and Esteban Rossi-Hansberg. 2017. “Quantitative Spatial Economics.” Annual Review of Economics 9: 21-58.

Santos Silva, J.M.C. and Silvana Tenreyro. 2006. “The Log of Gravity.” The Review of Economics and Statistics 88(4): 641-658.

Senney, Garrett T. 2016. "The Geography of Bidder behavior in Peer-to-Peer Lending Markets." Available at SSRN link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2721756 .

Sialm, Clemens, Zheng Sun, and Lu Zheng. 2014. "Home Bias and Local Contagion: Evidence from Funds of Hedge Funds." Available at SSRN link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2023851 .

Siegel, Jordan I., Amir N. Licht, and Shalom H. Schwartz, 2013 "Egalitarianism, Cultural Distance, and Foreign Direct Investment: A New Approach." Organization Science 24(4):1174-1194.

Sorens, Jason, Fait Muedini, and William P. Ruger. 2008. "U. S. State and Local Public Policies in 2006: A New Database." State Politics & Policy Quarterly 8(3): 309-326.

Sullivan, John L., and Daniel Richard Minns. 1976. "Ideological Distance between Candidates: An Empirical Examination." American Journal of Political Science 20(3): 439-468.

Tajfel, Henri, 1982. "Social Psychology of Intergroup Relations." Annual Review of Psychology 33(1): 1-39.

Wei, Zaiyan, and Mingfeng Lin. 2016. "Market Mechanisms in Online Peer-to-Peer Lending." Management Science Published online in Articles in Advance 07 Sep 2016.

Wintoki, M. Babajide, and Yaoyi Xi. 2017. "Political Partisan Bias in Mutual Fund Portfolios." Available at SSRN link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2933270 .

Wolf, Holger C. 2000. "Intranational Home Bias in Trade." The Review of Economics and Statistics 82(4): 555-563.

Wooldridge, Jeffery M. 2002. Introductory Econometrics: A Modern Approach. 2nd Edition. South-western, Mason, Ohio.

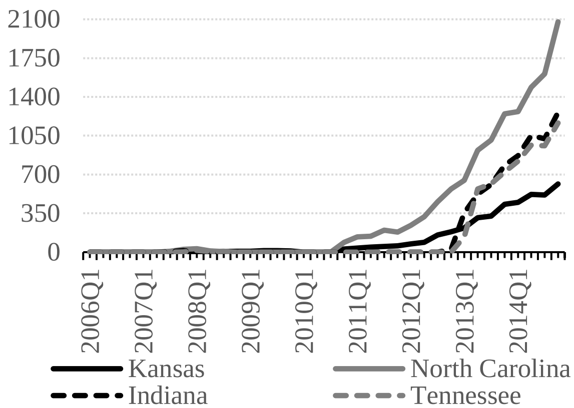
Wright, Gerald C., Robert S. Erikson, and John P. Mclver. 1985. "Measuring State Partisanship and Ideology with Survey Data." Journal of Politics 47:469-489.

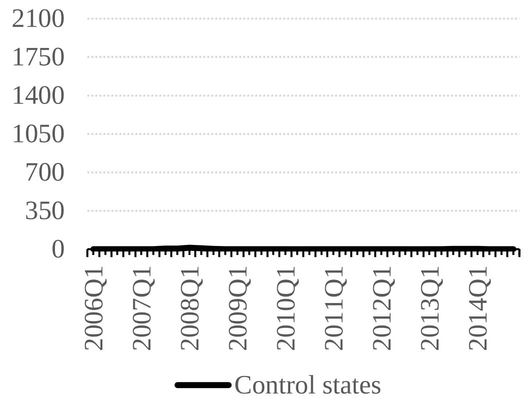
Wright, Gerald C., Robert S. Erikson, and John P. Mclver. 1987. "Political Opinion and Policy Liberalism in the American States." American Journal of Political Science 31:980-1001.

Yen, Steven T., and Ernest M. Zampelli. 2014. "What Drivers Charitable Donations of Time and Money? The Roles of Political Ideology, Religiosity, and Involvement." Journal of Behavioral and Experimental Economics 50: 58-67.

APPENDIX A. ADDITIONAL ANALYSES FOR CHAPTER 2

Appendix A-1: of Lending Club Loans Over Time in the Treated and Control States Used in the County-Year Analysis





Loan data from https://www.lendingclub.com/info/download-data.action. We determine the approval quarter for Kansas, North Carolina, Indiana, and Tennessee based on the HTML time stamp or press release date from the following pages: http://blog.lendingclub.com/welcome-kansans-lending-club-personal-loans-are-now-available-in-kansas/, http://blog.lendingclub.com/welcome-north-carolinians-lending-club-personal-loans-are-now-available-in-north-carolina/, https://www.lendingclub.com/public/lending-club-press-2012-12-06.action, and https://www.lendingclub.com/public/lending-club-press-2013-02-07.action .

Appendix A-2: List of Counties in the County-Year Matched Sample

** Please see text for description of the matching procedure, why the matches are valid for our analysis, and how observations from each county are weighted in our analysis. Although some of the matches are between counties in states that seem dissimilar (e.g., a county in North Carolina matched to one in Maine), Appendix A-4 illustrates that matched counties are quite similar in terms of pre-treatment bankruptcy and unemployment trends, income levels, and demographics.

Stra-tum	**Treated Counties**	**Matched Control Counties**
1	Crawford (Kansas); Lane (Kansas); Ness (Kansas); Riley (Kansas); Seward (Kansas); Smith (Kansas); Stanton (Kansas)	Ida (Iowa); Kossuth (Iowa); Mitchell (Iowa); Winneshiek (Iowa); Greeley (Nebraska); Barnes (North Dakota); Benson (North Dakota); Bottineau (North Dakota); Dunn (North Dakota)
2	Ellis (Kansas); Gove (Kansas)	Latah (Idaho); Chickasaw (Iowa); Floyd (Iowa); Hancock (Maine); Knox (Maine); Lincoln (Maine)
3	Alleghany (North Carolina)	Boundary (Idaho); Idaho (Idaho)
4	Rawlins (Kansas)	Dickey (North Dakota)
5	Graham (Kansas); Wichita (Kansas); Camden (North Carolina); Watauga (North Carolina)	Palo Alto (Iowa); Pierce (North Dakota)
6	Allen (Kansas); Carteret (North Carolina); Transylvania (North Carolina)	Custer (Idaho); Grundy (Iowa); Sagadahoc (Maine); Boyd (Nebraska); Dixon (Nebraska); Knox (Nebraska); Kidder (North Dakota); McHenry (North Dakota)

7	Currituck (North Carolina)	Ransom (North Dakota)
8	Logan (Kansas); Osborne (Kansas); Phillips (Kansas); Wallace (Kansas)	Sioux (Nebraska); McLean (North Dakota); Oliver (North Dakota)
9	Cherokee (Kansas); Finney (Kansas); Labette (Kansas)	Caribou (Idaho); Adams (Iowa); Bremer (Iowa); Butler (Iowa); Cherokee (Iowa); O'Brien (Iowa); Dawes (Nebraska); Eddy (North Dakota)
10	Cheyenne (Kansas); Mitchell (Kansas)	Crawford (Iowa); Delaware (Iowa); Franklin (Iowa); Osceola (Iowa); LaMoure (North Dakota); Ward (North Dakota)
11	Haskell (Kansas)	Cassia (Idaho); Audubon (Iowa); Clayton (Iowa); Dickinson (Iowa); Emmet (Iowa); Pocahontas (Iowa); Pierce (Nebraska); Sheridan (Nebraska); Ramsey (North Dakota)
12	Daviess (Indiana); Bourbon (Kansas)	Cerro Gordo (Iowa); Davis (Iowa); Greene (Iowa); Guthrie (Iowa); Iowa (Iowa); Page (Iowa); Ringgold (Iowa); Washington (Iowa); Winnebago (Iowa); Wright (Iowa); Kennebec (Maine); Penobscot (Maine); Waldo (Maine); Merrick (Nebraska); Morton (North Dakota)
13	New Hanover (North Carolina)	York (Maine)
14	Russell (Kansas); Trego (Kansas)	Adair (Iowa); Jackson (Iowa)
15	Decatur (Kansas); Ford (Kansas); Rooks (Kansas)	Tama (Iowa); Richland (North Dakota)
16	Norton (Kansas)	Humboldt (Iowa)
17	Comanche (Kansas); Nemaha (Kansas)	Jones (Iowa); Wells (North Dakota)
18	Kearny (Kansas)	Benton (Iowa); Calhoun (Iowa)
19	Barton (Kansas); Cloud (Kansas); Coffey (Kansas); Ellsworth (Kansas); Geary (Kansas); Grant (Kansas); Marshall (Kansas); Meade (Kansas); Pratt (Kansas); Chatham (North Carolina)	Elmore (Idaho); Black Hawk (Iowa); Buchanan (Iowa); Cedar (Iowa); Fayette (Iowa); Montgomery (Iowa); Plymouth (Iowa); Sac (Iowa); Shelby (Iowa); Burleigh (North Dakota); Grand Forks (North Dakota); Pembina (North Dakota); Stutsman (North Dakota); Traill (North Dakota); Walsh (North Dakota)
20	Clay (Kansas)	Emmons (North Dakota)
21	Anderson (Kansas); Elk (Kansas); Greenwood (Kansas); Harvey (Kansas); McPherson (Kansas); Woodson (Kansas); Henderson (North Carolina)	Minidoka (Idaho); Appanoose (Iowa); Cass (Iowa); Clinton (Iowa); Dallas (Iowa); Decatur (Iowa); Dubuque (Iowa); Lucas (Iowa); Mahaska (Iowa); Marshall (Iowa); Van Buren (Iowa); Woodbury (Iowa); Madison (Nebraska)
22	Buncombe (North Carolina)	Linn (Iowa)
23	Williamson (Tennessee)	Bannock (Idaho); Nez Perce (Idaho); Twin Falls (Idaho); Keokuk (Iowa); Mills (Iowa); Adams (Nebraska); Box Butte (Nebraska); Seward (Nebraska); Stanton (Nebraska)
24	Stewart (Tennessee)	Clearwater (Idaho)
25	Kingman (Kansas)	Henry (Iowa); Otoe (Nebraska)
26	Martin (Indiana)	Marion (Iowa); Dawson (Nebraska); Hall (Nebraska); Saunders (Nebraska)
27	Brown (Kansas); Ottawa (Kansas)	Nelson (North Dakota)
28	Atchison (Kansas); Cowley (Kansas); Dickinson (Kansas); Saline (Kansas); Moore (Tennessee)	Fremont (Iowa); Harrison (Iowa); Louisa (Iowa); Webster (Iowa); Clay (Nebraska)
29	Lyon (Kansas); Sumner (Kansas)	Scott (Iowa)
30	Monroe (Indiana); Tippecanoe (Indiana)	Clarke (Iowa); Muscatine (Iowa); Wayne (Iowa)

31	Dubois (Indiana); Leavenworth (Kansas); Lincoln (Tennessee); Washington (Tennessee)	Bingham (Idaho); Boone (Iowa); Madison (Iowa); Wapello (Iowa); Warren (Iowa); Scotts Bluff (Nebraska)
32	Marion (Kansas)	Fremont (Idaho); Clay (Iowa)
33	Sampson (North Carolina)	Monroe (Iowa); Poweshiek (Iowa)
34	Pottawatomie (Kansas)	Union (Iowa); Pawnee (Nebraska)
35	Posey (Indiana)	Jefferson (Iowa)
36	Wayne (North Carolina)	Monona (Iowa)
37	Bartholomew (Indiana); Butler (Kansas); Miami (Kansas); Wabaunsee (Kansas); Montgomery (Tennessee)	Des Moines (Iowa); Cass (Nebraska); Johnson (Nebraska); Sarpy (Nebraska); Washington (Nebraska)
38	Johnston (North Carolina)	Payette (Idaho)
39	Sedgwick (Kansas)	Polk (Iowa); Douglas (Nebraska)
40	Harper (Kansas); Jefferson (Kansas); Reno (Kansas)	Pottawattamie (Iowa); Saline (Nebraska)
41	Knox (Indiana); Franklin (Kansas)	Jefferson (Nebraska)
42	Switzerland (Indiana)	Dodge (Nebraska)

Appendix A-3: Balance between Treated and Control Counties in the County-Year Matched Sample

Variable	***α:*** **Mean of control counties (robust std. error)**	***β:*** **Mean difference between treated and control counties (robust std. error)**
Variables used in the matching procedure		
Bankruptcy filings per capita (2006)	1.597 (0.073) ***	0.122 (0.102)
Bankruptcy filings per capita (2007)	2.119 (0.105) ***	-0.035 (0.145)
Bankruptcy filings per capita (2008)	2.469 (0.114) ***	0.002 (0.158)
Bankruptcy filings per capita (2009)	3.061 (0.131) ***	-0.061 (0.187)
Bankruptcy filings per capita (2010)	2.905 (0.129) ***	0.045 (0.184)
Unemployment rate (2006)	3.887 (0.084) ***	0.052 (0.117)
Unemployment rate (2007)	3.798 (0.084) ***	-0.001 (0.110)
Unemployment rate (2008)	4.272 (0.105) ***	0.048 (0.149)
Unemployment rate (2009)	6.162 (0.178) ***	0.285 (0.279)
Unemployment rate (2010)	5.853 (0.165) ***	0.956 (0.269) ***
Population (2006)	32.108 (5.435) ***	5.534 (8.378)
Variables *not* used in the matching procedure		
Average monthly earnings	2.649 (0.047) ***	-0.063 (0.065)
Median household income	44.279 (0.604) ***	-0.895 (0.966)
Number of employed individuals	13.361 (2.861) ***	0.757 (4.122)
% age 60 & above	23.716 (0.483) ***	-1.358 (0.750) *
% white	95.542 (0.513) ***	-2.252 (0.853) ***

Notes: *** p <0.01, ** p<0.05, * p<0.10. α and β derived from regressing the variable in the left-hand column on *Treated*, which equals 1 for treated counties and 0 otherwise. α is the intercept and β is the coefficient for *Treated*. Each regression was weighted by the CEM weights (see text).

The first section of Appendix A-4 shows that the procedure generated good balance on the matching variables. The only significant difference between the treated and control counties is for *unemployment rate* in 2010. We do not believe this harms our identification because we control for *unemployment rate* in the regressions directly.[14] We also examined the balance between

[14] We also created a new county-year matched sample, using stricter matching for the unemployment rate (2010) variable. In this sample, we achieved balance on all matching variables, although we have fewer

treated and control counties for variables that are not included in the matching procedure. For these regressions, we included all observations in the 2006-2010 pre-treatment period (n = 259 counties * 5 years = 1,295), clustered the standard errors by county, and used the CEM weights. The second section of Appendix A-4 shows that our matching procedure yields no statistically distinguishable differences between treated and control counties for *number of employed individuals*, *average monthly earnings*, and *median household income* and only small differences for *% age 60 & above* and *% white*. We control for each of these variables directly in our regressions, which accounts for any remaining imbalance in them.

Appendix A-4: Regressions Results for Bankruptcy Filings (Raw Count) for the County-Year Matched Sample

	Poisson regression		**Negative Binomial regression**	
Variable	**Baseline model**	**Full model**	**Baseline model**	**Full model**
Lending Club available	0.130 (0.031) ***	0.097 (0.026) ***	0.126 (0.016) ***	0.084 (0.016) ***
Population		0.003 (0.001) **		0.003 (0.001) ***
Number of employed individuals		-0.000 (0.002)		0.000 (0.001)
Average monthly earnings		-0.047 (0.088)		-0.112 (0.048) **
Unemployment rate		0.032 (0.014) **		0.035 (0.006) ***
Median household income		-0.006 (0.003) **		-0.006 (0.002) ***
% age 60 & above		0.015 (0.022)		0.030 (0.008) ***
% white		0.033 (0.024)		0.037 (0.008) ***
Year 2006	omitted baseline period			
Year 2007	0.290 (0.019) ***	0.306 (0.025) ***	0.271 (0.018) ***	0.298 (0.019) ***
Year 2008	0.467 (0.023) ***	0.474 (0.032) ***	0.447 (0.017) ***	0.467 (0.023) ***
Year 2009	0.666 (0.029) ***	0.576 (0.054) ***	0.661 (0.017) ***	0.580 (0.029) ***
Year 2010	0.674 (0.028) ***	0.577 (0.062) ***	0.654 (0.017) ***	0.576 (0.032) ***
Year 2011	0.476 (0.031) ***	0.407 (0.069) ***	0.457 (0.018) ***	0.417 (0.037) ***
Year 2012	0.318 (0.027) ***	0.272 (0.080) ***	0.274 (0.018) ***	0.260 (0.040) ***
Year 2013	0.195 (0.025) ***	0.173 (0.097) *	0.131 (0.020) ***	0.146 (0.046) ***
Year 2014	0.077 (0.028) ***	0.089 (0.116)	0.009 (0.020)	0.059 (0.052)
County fixed effects	✓	✓	✓	✓
n (counties)	259	259	259	259
n (observations)	2,331	2,331	2,331	2,331
Log likelihood	-7167.46	-7040.22	-6692.373	-6628.70
χ^2	1201.07	2063.86	3585.73	4050.59

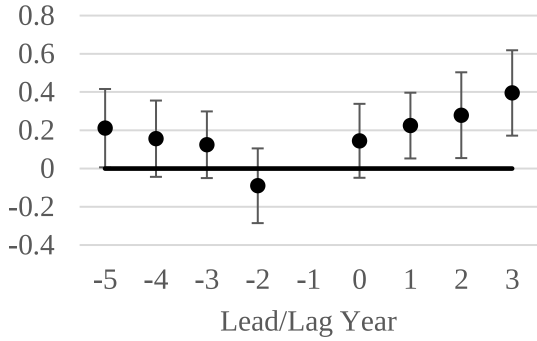
Notes: Regressions weighted using the CEM weights. Negative binominal models use OIM (Observed Information Matrix)-based standard errors. Poisson models use robust standard errors. *** p<0.01, **p<0.05, *p<0.1.

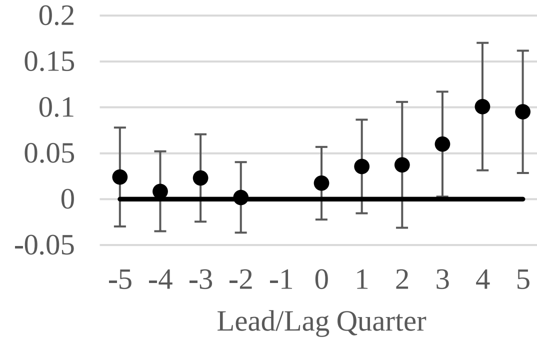
Appendix A-5: Plots of the Lead and Lag Coefficients from the Leads/Lags Models (Table 5)

Lead and Lag Coefficients (with 95% Confidence Intervals) from Leads/Lag Model: County-Year Matched Sample

Lead and Lag Coefficients (with 95% Confidence Intervals) from Leads/Lag Model: State-Quarter Sample

observations than in the focal sample. The results of the analysis using this sample are consistent with those from the focal sample.


0.8
0.6
0.4
0.2
0
-0.2
-0.4
-5
-4
-3
-2
-1
0
1
2
3
Lead/Lag Year


0.2
0.15
0.1
0.05
0
-0.05
-5
-4
-3
-2
-1
0
1
2
3
4
5
Lead/Lag Quarter

Appendix A-6: IV Regressions for Bankruptcy Filings Per Capita: First-Stage Results

Endogenous Variable Being Instrumented	Outstanding Loans per Capita	
LC availability / maturity	0.002 (0.000) ***	
Debt-to-income policy change		0.010 (0.000) ***
Population	0.032 (0.050)	0.036 (0.049)
Number of employed individuals	0.115 (0.076)	0.096 (0.079)
Average monthly earnings	-0.063 (0.038) *	-0.054 (0.040)
Unemployment rate	0.007 (0.011)	0.013 (0.009)
Median household income	0.035 (0.022)	0.037 (0.022) **
% age 60 & above	0.043 (0.060)	0.052 (0.062)
% white	-0.016 (0.027)	-0.014 (0.027)
% below high school attainment	0.094 (0.052) *	0.106 (0.056) *
% housing units rented	0.089 (0.047) *	0.117 (0.049) **
% housing units with mortgage	0.004 (0.046)	0.019 (0.048)
State fixed effects	✓	✓
Quarter fixed effects	✓	✓
n (states)	51	51
n (observations)	2,039	2,039
R^2	0.93	0.94
Partial R^2	0.37	0.49
F	70.59	449.25

Notes: Standard errors (in parentheses) are clustered by state. *** p<0.01, **p<0.05, *p<0.1.

Appendix A-7: Annual Individual Income of Lending Club Borrowers, Lending Club Borrowers who Default, and Residents of the 50 U.S. States plus District of Columbia

	Lending Club Borrowers	Lending Club Borrowers who Default	Residents of the 50 U.S. States and D.C.
75th percentile	90,000	80,000	28,629
50th percentile	65,000	58,560	25,631
25th percentile	45,000	42,000	23,602
Time period	2007-2015	2007-2015	2006-2015
n (observations)	887,440	99,276	2,040

Notes: Columns 1 and 2 are calculated based on Lending Club borrowers' self-reported income; approximately half of these values were verified by Lending Club. Column 3 is based on the *median individual income* of residents in each U.S. state, plus the District of Columbia, as measured by the U.S. Census. This figure is smaller than the *median household income*, which we use as a control variable in our regressions.

APPENDIX B. ADDITIONAL ANALYSES FOR CHAPTER 4

In Table B-1 we report results based on DID models within several subsamples. The focal specification is the treatment effect heterogeneity DID model, which is exactly the one used in Table 4-4. A subsample of the matched sample excluding California investors is used in Column 1 because California borrowers are treated. A subsample of the matched sample excluding all same-state pairs is used in Column 2 to remove "home bias" effect. Removing home bias weakens the average treatment effect, but the effect is still positive and the pattern of political distance is still consistent. A subsample of the full sample including listings that have a final funding percentage over 0.09 is used in Column 3 (to represent the top 25% listings in terms of funding percentage). A subsample of the full sample including listings that have a final funding percentage equaling to 1 is used in Column 4. The results are similar as our main findings: the investors holding similar political ideology increase bids after the treatment while the investors holding dissimilar political ideology increase fewer bids or even decrease bids after the treatment.

Table B-1: Additional Robustness Checks on DID Models				
Event/Treatment	2008 California Event	2008 California Event	2008 California Event	2008 California Event
Model Specification	No California Investors	No Same-State Dyads	Top 25% Funded Listings Only	Fully Funded Listings Only
Treated * >=20 Difference	0.003	0.003	-0.018	-0.045
	(0.035)	(0.035)	(0.098)	(0.150)
Treated * [10, 20) Difference	-0.006	-0.006	-0.030	-0.056
	(0.019)	(0.019)	(0.069)	(0.104)
Treated	0.021*	0.021*	0.103*	0.186**
	(0.012)	(0.012)	(0.058)	(0.089)
Treated * (-20, -10] Difference	-0.011	-0.011	-0.059	-0.115
	(0.016)	(0.016)	(0.065)	(0.099)
Treated * <=-20 Difference	-0.035***	-0.035***	-0.141**	-0.225**
	(0.013)	(0.013)	(0.063)	(0.094)
Time Fixed Effects	√	√	√	√
State-Listing Dyad Fixed Effects	√	√	√	√
# of Observations	102,557	102,557	42,350	20,650
# of Groups	14,651	14,651	6,050	2,950
Adjusted R^2	0.3457	0.4213	0.4134	0.4110

Notes: *** p<0.01; ** p<0.05; * p<0.1. Standard errors in parentheses are clustered at the state-listing level.

In Table B-2 we report results based on DID models with interaction terms of political distance to check the treatment effect heterogeneity approach (Table 4-4). "*Political Distance*" is defined as the absolute political distance between borrower state and investor state, so the term "*Treated * Political Distance*" reflects how political distance moderates the impact of treatment in the DID setting. "*Treated*" serves the baseline treatment effect, indicating the influence of same-sex marriage law passage on same state investors. Results in Columns 1 and 3 confirm that when investors have a dissimilar political ideology as the treated state, they tend to place less bids after the treatment than who have similar political ideology. Considering liberal investors and conservative investors seem to response to political distance differently (as suggested by Table 4-4 and Table 4-7), we further add "*Treated * Political Distance * Investors More Liberal*" into the specification used in Columns 1 and 3 as a formal check. The coefficient of "*Treated * Political Distance*" now denotes how political distance moderates the treatment effect when investors are more conservative than the treatment state. The sum of the coefficient of "*Treated * Political Distance*" and the coefficient of "*Treated * Political Distance * Investors More Liberal*" denotes how political distance moderates the treatment effect when investors are more liberal than the treatment state. Columns 2 and 4 confirm that political distance has a stronger impact when investors are more conservative than borrowers.

Table B-2: Additional Robustness Checks on The Impact of Political Distance

Event/Treatment	2008 California Event	2008 California Event	2008 Connecticut Event	2008 Connecticut Event
Treated	0.043*	0.043*	0.038***	0.040**
	(0.024)	(0.024)	(0.014)	(0.017)
Treated * Political Distance	-0.002[a]	-0.002*	-0.0006*	-0.0007*
	(0.001)	(0.001)	(0.0003)	(0.0004)
Treated * Political Distance * Investors More Liberal		0.0005		-0.015
		(0.0009)		(0.017)
Time Fixed Effects	√	√	√	√
State-Listing Dyad Fixed Effects	√	√	√	√
# of Observations	104,650	104,650	39,550	39,550
# of Groups	14,950	14,950	5,650	5,650
Adjusted R^2	0.4793	0.4793	0.0191	0.0190

Notes: *** p<0.01; ** p<0.05; * p<0.1. Standard errors in parentheses are clustered at the state-listing level.
a: p value = 0.159.

One empirical concern of gravity models is that each pair of investor state and borrower state has the same weight, although in reality each pair represents different numbers of investors and borrowers. As one robustness check (Columns 1 and 2 in Table B-3), we use the production of the number of investors and the number of listings as the weight to check whether the results are influenced by small states. As another robustness check, we require the investor states to have more than 100 investors and the borrower state to have more than 100 listings (Column3). We also remove same-state pairs to avoid the confounding effect of home bias (Column 4). Across all models, the coefficients of *Borrower Political Ideology* are always positive and significant, indicating an overarching preference on liberal borrowers. Political distance still exhibits a negative impact, making investors who hold dissimilar political ideology less likely to invest.

Table B-3: Additional Robustness Checks on Gravity Models

Model Specification	Population Weighted	Population Weighted	Excluding Inactive States	Excluding Same-state Pairs
Sample Year	2008	2008	2008	2008
Estimation Method	PPML	OLS	PPML	PPML
lnInvestors	1.104***	1.072***	1.062***	1.063***
	(0.018)	(0.011)	(0.010)	(0.009)
lnListings	0.918***	0.949***	0.987***	0.988***
	(0.027)	(0.015)	(0.011)	(0.012)
Average Credit Score	0.023***	0.025***	0.023***	0.023***
	(0.004)	(0.002)	(0.002)	(0.002)
Average DTI Ratio	1.048*	1.105***	0.822***	0.761***
	(0.557)	(0.201)	(0.178)	(0.179)
Average Monthly Payment	0.008***	0.007***	0.005***	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
lnMedianHouseholdIncome	0.050	0.053	0.049	0.016
	(0.106)	(0.078)	(0.067)	(0.067)
lnGeographicDistance	-0.024*	-0.023***	-0.013*	-0.006
	(0.014)	(0.009)	(0.007)	(0.010)
lnEconomicDistance	0.020	0.017**	0.007	0.012
	(0.013)	(0.007)	(0.006)	(0.008)
PoliticalDifference in Range (30, 100)	0.081 (0.067)	0.032 (0.045)	0.027 (0.040)	0.028 (0.041)
PoliticalDifference in Range (10, 30)	0.027	-0.029	-0.008	-0.006
	(0.051)	(0.032)	(0.028)	(0.027)
PoliticalDifference in Range (-10, 10)		Baseline		
PoliticalDifference in Range (-30, -10)	-0.057* (0.029)	-0.035 (0.023)	-0.027 (0.020)	-0.030 (0.020)
PoliticalDifference in Range (-100, -30)	-0.150*** (0.053)	-0.137*** (0.037)	-0.099*** (0.035)	-0.111*** (0.034)
Borrower Political Ideology	0.006***	0.003***	0.002**	0.002**
	(0.002)	(0.001)	(0.001)	(0.001)
# of Observations	2,450	2,450	2,205	2,401
Pseudo Log-likelihood	-87555.022		-74299.167	-73387.42
Adjusted R^2		0.9707		

Notes: *** $p<0.01$; ** $p<0.05$; * $p<0.1$. Heteroskedasticity robust standard errors in parentheses.