Essays on Agency Costs of Financial Intermediation

M.A. Economics, University of Virginia, 2017
B.A. Finance and Mathematics, Michigan State University, 2012

A Dissertation Presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of Doctor of Philosphy

Department of Economics

University of Virginia
July 2018

Professor Denis Nekipelov

Professor Andrew Kloosterman

Professor Zhaohui Chen

Professor William Wilhelm

Abstract

A financial intermediary is a delegated monitor that produces information and adds value to capital allocation between market agents. A borrower, an entrepreneur or a firm obtains capital investments, and in return, the lender or investor profits through interest payments or equity shares. These transactions are facilitated by or implemented under intermediaries, either commercial banks, investment banks, credit rating agencies, venture capitalists or platforms. As economic agents, financial intermediaries may not have their interest aligned with that of clients, which can result in inefficient capital allocation, market failure and financial instability. This dissertation studies conflicts of interest in different types of intermediation and from various aspects.

In the first chapter, I study how financial intermediaries balance between market share and reputation under competition, using unique datasets on loans originated and declined on peer-to-peer lending platforms. Using a platform entry event that intensifies market competition on borrowers and lenders, I document less prudent borrower screening, credit rating inflation and aggravated loan performance. In particular, post-entry borrowers are more likely to obtain financing, equally creditworthy borrowers receive better credit ratings and their average loan performance deteriorates significantly. It distorts platforms' incentive on truthfully reporting borrowers' risk to compete for market making. The incumbent platform lowers interest rates to encroach on creditworthy borrowers, indicating aggressive undercutting behavior. Raising interest rates on subprime borrowers maintains lenders' participation while accounting for competition-induced adverse selection.

I further document that, as disincentive for platforms' credit inflation, lenders exit

the platform upon their borrowers' underperformance. In particular, with vintage loan performance deterioration, the number of lenders on a newly originated loan decreases, credit crunches emerge and capital flows slow down. The magnitude of the 'punishments' mitigates significantly post-entry, arguably because the market size expands with the entry event, eliciting new and unfamiliar lenders to enter, which intensifies borrower competition and credit inflation.

By fuzzy matching borrowers' identities between the platforms, I identify the incumbent's first mover advantage, where incumbent-rejected borrowers get financing at the entrant but rarely vice versa. As a dominant player, the incumbent gets high-quality borrowers and induces severe adverse selection for the entrant. I contend that this first mover advantage is endogenized by the incumbent's active borrower screening beyond "hard information" and its capital provision for borrowers facing credit crunches.

The second chapter, coauthored with Zhaohui Chen, Alan Morrison and William Wilhelm, examines and identifies the underlying mechanism of the decline of investment bank-client relationships from 1960 — present. As investment banks know superior information about their clients in security underwriting, the internal agents, investment bankers, often face conflicts of interest and thus, have incentive to misuse the information against clients. Without contractibility, banks' internal governance and monitoring provide incentives for bankers to harness their relationship with clients by making agency problems costly.

The adoption of computing technology started in the 1960s has caused investment banks' internal governance to evolve. Advances in technology and novel financial economic theory make it profitable to be engaged in the trading and risk-taking business, which induces investment banks to get bigger in scale and more complex in financial innovation. The increasing internal liquidity dampens senior bankers' incentive to train and monitor younger partners by easing the mobility of their stake. These changes on the internal governance endogenize the breakdown of the investment bank-client relationship. We provide a causal econometric model to test how increasing bank complexity affects the propensity for their clients to switch underwriters in the succeeding deal. Measured by investment banks' capital, partnership size and an event study on investment banks' decision to go public, we find that the increasing complexity induces clients to switch out of their relationship bank.

In the last chapter, I study how a venture capitalist's information production induces an entrepreneur's effort. In particular, I design a contingent contracting mechanism where the principal's (a venture capitalist) private monitoring induces the agent's (an entrepreneur) effort and adds value to the project through the capital investment from the principal. Featuring double-sided moral hazard, the optimal contract subsumes a menu that entitles the principal to punish the agent upon negative information. Also, it is incentive compatible to prevent the principal from falsely punish to expropriate a bigger equity stake. Compared to the "second best" under "pay-for-performance" mechanism, this scheme grants the principal high ex ante equity stake. The project value and capital investments are commensurate with a higher marginal return on the investments. The optimal monitoring intensity increases with the value added by the agent's effort but decreases with the cost of monitoring.

JEL Classifications: L13, L14, L15, G23, G24, D86

Keywords: Financial Intermediation, Agency Cost, Competition, Contracts

Acknowledgements

To all of my advisors and their extremely generous support, I am immeasurably indebted. I benefited greatly from Professor Denis Nekipelov's mentorship and guidance in research directions and emphases on the big picture. I deeply thank Professor Andrew Kloosterman for his detailed facilitation on research technicalities and model constructions. I indefinitely owe Professor Zhaohui Chen for the help he offered when I needed it the most, mentally and academically, and for his continuous support throughout the graduate program. I very much thank Professor William Wilhelm for helping shape me to be a rigorous and critical thinker. I appreciate the ongoing help during my first 3 years from Professor Latchezar Popov. They all set perfect examples for me to be hardworking and excited about research and to keep challenging myself.

I benefited from the insightful comments from the research community at the University of Virginia and discussants and participants at numerous conferences. Particularly, I learned so much during conversations with Professors Eric Young, Karim Chalak, John Pepper, Gaurab Aryal, Yael Hochberg, Manju Puri, Stephen Karolyi, Nicholas Ryan, Mingfeng Lin, Alan Morrison, Andre Boik and Ruomeng Liu.

I thank the generous financial support from the Department of Economics at UVa, the Bankard Fund for Political Economy at UVa, the Dissertation Fellowship from the Kauffman Foundation and the Searle Center at Northwestern.

Lastly, I cannot be here without the unconditional love, faith, support and research assistance from my parents, Lixin Zhao and Hongsong Xue, and my wife, Nicole.

Chapter 1

Competition and Intermediaries' Incentives: Evidence from Peer-to-peer Lending Platforms

Abstract

How do financial intermediaries balance between market share and reputation under competition? Using data on the peer-to-peer lending market, I document less prudent borrower screening, credit rating inflation and aggravated loan performance induced by a platform entry event. Platforms sacrifice truthful information production to compete for market making. The decline of creditworthy borrowers' interest rates reveals aggressive undercutting. Increasing interest rates among subprime borrowers account for possible competition-induced adverse selection, while maintaining lender participation for market clearance. Vintage loan underperformance causes lenders to exit and credit crunches to emerge. This effect is significantly mitigated post-entry likely by the emergence of new lenders, and it provides additional incentive on credit rating inflation. Finally, cross-platform borrower matching shows the importance of being the dominant player. With a first mover advantage, the incumbent reduces credit risk and causes more severe adverse-selection problems for the entrant.

MARK BAUM: *Have you ever refused to rate any of these bonds upper-tranches AAA?...Can you name one time in the past year, where you checked the tape and you didn't give the banks the AAA-percentage they wanted?*

GEORGIA: *If we don't give them the ratings, they'll go to Moody's right down the block. If we don't work with them, they will go to the competitors. Not our fault. Simply the way the world works.*

—–The Big Short (2015): Mark Baum and Georgia from S&P

1 Introduction

Financial intermediaries specialize in capital allocation and market making by producing information to market agents (Campbel and Kracaw (1980)). However, as economic agents conflicted between reputation and immediate payoffs from market making, intermediaries may have incentive to falsify information that entices capital misallocation (Chen et al. (2014) and Hartman-Glaser (2017)). Competition largely affects intermediaries' prudent behavior and overall financial stability. Through numerous mechanisms, theory shows that competition reduces the average creditworthiness and can undermine the economy.[1] Empirical studies document a negative correlation between competition

[1]Through reputation and relationship banking, Sharpe (1990) argues that competition undermines relationship banking and forces banks to explore unfamiliar customers. Due to debtors' rate shopping behavior and investors' trusting nature, Bolton et al. (2012) argues that competition induces credit rating agencies (CRAs) to inflate credit ratings. Other mechanisms include intermediaries' moral hazard and adverse selection. Keeley (1990) argues that competition induces banks' risk taking and risk shifting by purchasing deposit insurance. Hauswald and Marquez (2006) show that competition reduces banks' incentive to acquire information on borrowers, allowing for more lemon problems. See other papers such as Marquez (2002) on information dispersion and adverse selection due to competition; Broecker (1990) on lower credit-worthiness caused by competition.
Although financial intermediation competition has its upside to the economy and is crucial to en-

and intermediaries' imprudent behavior but have yet to identify the proposed mechanisms.[2] Becker and Milbourn (2011) document that when Fitch enters to compete with S&P and Moody's on credit rating, the incumbents lower their rating quality by credit inflation, manifested by worse loan market yield and deteriorated default predictive power. Flynn and Ghent (2017) go one step further and identify that issuers' rate shopping induces credit inflation while credit rating agencies(CRAs) are conflicted between protecting their reputation and encroaching for market share.

I extend these studies by also examining the lender side of the market to better understand the channels through which intermediaries are conflicted. In particular, in addition to how competition affects intermediaries to encroach on issuers, I explore how it provides incentive for intermediaries to sustain lender participation and market clearance.

I use loan-level rejection and issuance data from peer-to-peer lending platforms, Lending Club and Prosper, and an entry event where a monopoly turns into a duopoly, to study the effect of competition on the market.[3] I show striking resemblant results to Becker and Milbourn (2011), where competition reduces borrower screening and induces credit inflation, and as a result, the ex post performance is aggravated. I go beyond this

trepreneurial activities, I do not focus on these aspects.Allen and Gale (1998) argue that perfect financial stability is socially undesirable and can lead to inefficient outcomes. Other theories suggest that an absence of competition leads to high interest rates, which induce entrepreneurs' risk-taking behavior.Boyd and De Nicolo (2005) argue that under a general equilibrium environment, entrepreneurs tend to choose risky projects when the banking industry is concentrated and capitals are expensive. Cetorelli and Strahan (2006) analyze data from U.S. local banks and find that potential entrepreneurs face greater difficulty gaining access to credit where banking is less competitive.

[2]Beck et al. (2013) use cross-country data and find that competition causes economic fragility, and it appears that the effect is more severe in countries with more generous deposit insurance and better credit information sharing.

[3]A peer-to-peer lending platform is a novel financial intermediary where borrowers and lenders match. A platform has the following responsibilities: collecting and aggregating information on borrowers from credit rating agencies, screening their loan applications, rating the borrowers and posting their interest rates. Lenders get to observe borrowers' credit history, the platform's recommended rating and the loan terms; they also make lending decisions. Its function highly resembles a credit rating agency, where it profits from loan origination fees from the borrowers/issuers. In addition, it sets interest rates to clear the market.

partial equilibrium to study how competition affects the market-clearing interest rates, market-clearing efficiency and market activities. In particular, how do lenders respond to the unobtrusive credit inflation and risk-adjusted interest rates, and how do they provide disincentives on the platforms to reduce potential borrower underperformance? Moreover, by bridging borrowers' identities between platforms, I further analyze the mechanism of competition and how that feeds back into the platforms' incentives.

Using Regression Discontinuity Design (SRD) and Average Treatment Effect (ATE) models, I estimate the effect of the entry on the propensity of borrower acceptance. I find that after the entry, the incumbent becomes less prudent and accepts less creditworthy applicants. Within the accepted borrowers, I use an Order Logistic model and find a significant inflation in credit rating policy resulted from the entry event, where borrowers of same creditworthiness get rewarded with higher ratings. As a consequence, I document that the post-entry loans originated yield significantly inferior performance. This deterioration is a direct result from platforms' imprudent behavior and credit inflation induced by competition, similar to findings by Becker and Milbourn (2011).

Post-entry borrowers' interest rate changes exhibit a highly heterogeneous pattern. I use a Propensity Score Matching model to compare the interest rates of borrowers with similar creditworthiness at different quantiles (or Quantile Treatment Effect model). I document that the most creditworthy borrowers start receiving cheaper loans, whereas risky borrowers obtain more expensive financing. While undercutting interest rates to encroach on creditworthy borrowers, the conflicted platform faces the market clearing constraint from the lender side. The increase in interest rates over risky borrowers may correspond to several coherent channels. First, by introducing more risky borrowers to the market, interest rates are raised to adjust for the risk and satiate market clearing. Second, differentiated information between the platforms may induce an unwanted adverse selection problem, which attenuates the undercutting incentive. (See Hauswald and Marquez (2006)). Third, price elasticity over lenders' participation may exhibit heterogeneity over borrowers' creditworthiness.

To make an identification, I analyze the market-clearing efficiency and lenders' credit supply measured by the number of lenders per loan, the time duration for a loan to issue and the platform's capital provision to resolve credit crunches. I discover that lower interest rates on creditworthy borrowers attract less lenders and slower funding flow but do not entice any credit crunch. The higher interest rates on risky borrowers receive more lender traction and largely mitigate previously existing credit crunches, and its funding flow commensurates with borrowers' interest rates. It shows that the lenders are inelastic to interest rate changes amongst the creditworthy borrowers but are quite responsive to those of risky borrowers.

I extend my study to the platforms' disincentive to upset the repeated lenders, by investigating lenders' responses to vintage loan underperformance. Without individual lenders' identities, I measure the lenders exposure to vintage loan defaults at the time of new loans' origination. I discover that lenders punish the platform by reducing investments and potentially leaving the market, evidenced by the decline in the number of lenders, diminishing funding flow and emergence of credit crunches. However, the magnitude of the post-entry "punishment" diminishes significantly. Entry very possibly brings in new and unfamiliar lenders, the effect of which dominates the upset repeated lenders' exit. It suggests additional incentives on credit inflation and market share competition.

The increase in potential market size induces the platform to take on the risk of being punished when the loans mature, while establishing its position as the dominant player.

I combine data from both platforms and delve into the competition mechanism to better understand the platforms' incentives. Following the algorithm used by Liu et al. (2013), I fuzzy-match borrowers' identities across the platforms using uniquely shared features, taking into account their institutional differences.[4] I document that borrowers who are rejected by the incumbent become financed on the entrant, but quite few vice versa. Among all loans originated, I identify market segmentation where the incumbent focuses on prime borrowers and the entrant serves a wider spectrum. Compared to the incumbent, a majority of the borrowers are charged at a premium by the entrant, but they do not appear to dominate in loan performance. Holding more than 3 quarters of the market share, the incumbent platform achieves a first mover advantage, which induces an unwanted adverse selection problem against the entrant. I attribute the "first mover advantage" to the following institutional differences. First, the incumbent actively screens borrowers using a strict standard and even information that is unobservable to lenders. In comparison, the entrant accepts any applicants and lets the lenders decide whom to finance. Therefore, the information asymmetry incentivizes good borrowers to self select into the incumbent. Second, the incumbent prices borrowers' interest rates, whereas the entrant uses an auction mechanism where lenders decide their reservation interest rates. Prone to adverse selection, lenders are bidding at much higher inter-

[4]Liu et al. (2013) use a fuzzy-matching procedure to identify the same apps between two platforms, but focus on "text" matching. The procedure is applicable to my research because it improves matching speed and accuracy. The matching process bears noise from several aspects. First, I am constrained by information discrepancies between the platforms. Second, borrowers censored by self selection, such as withdrawals and cancellations, cannot be captured. Moreover, the algorithm may yield multiple matches onto one borrower. However, I argue that despite the noise, my results delineate the status quo of platform competition on average.

est rates to compensate for unobservable risks. The high interest rates feed back into borrowers' decisions, and borrowers' rate-shopping behavior makes the incumbent the preferred platform. Third, the incumbent has established its reputation on its market-making competency by filling credit crunches and signaling the lenders with its "skin in the game".

The rest of the paper is organized as follows. First, I introduce the institutional background and business models of the two platforms and compare the peer-to-peer market with those of commercial loans and credit rating agencies. In the meantime, I discuss the history and the impact of the (re)-entry event. In the following sections, I describe the data and construction of key measurements. I further delineate the hypotheses, empirical strategies and economic interpretations. Using features from both platforms, I merge the data from both and yield identifications on the mechanisms.

2 Institutional Background

2.1 Institution Details of Peer-to-peer Lending in U.S.

2.1.1 Institutional Summary

Peer-to-peer (P2P) lending is a type of profit-seeking crowdfunding, serving as online platforms where lenders and borrowers match.[5] Since the introduction of this novel financing intermediation in the U.S. in 2005, the accrued loan volume issued has reached \$25 billion, leading the U.S. online alternative finance industry.[6] During the year of 2014

[5] Other types of crowdfunding include donation (Donorschoose), product reward (Kickstarter), and also equity venturing (AngelList).

[6] The Americas alternative finance benchmarking report, 2016.

alone, it accounts for $5.5 billion loans issued, and according to PwC's estimation, this figure could reach $150 billion annually.[7] Two major P2P lending platforms operating within U.S., Lending Club and Prosper, account for 98% of the peer-to-peer lending market. [8] More than 80% of the loans issued are claimed to be used for personal debt consolidation in 2015. [9]

2.1.2 Market Agents and Market Mechanism

Four types of agents participate in this market, borrowers, lenders, platforms and platform-partnered banks.[10] Any adult resident of the U.S. is entitled to apply for an unsecuritized loan less than $35,000 on a peer-to-peer lending platform. An application requires the borrower's Social Security Number, current employment and income verification, homeownership status, the intended term of the loan (3-year or 5-year) and loan size. Other self-reported information includes the loan usage and personal financial status.

The platform evaluates borrowers' creditworthiness by pulling their credit reports including their FICO scores, debt outstanding, previous delinquencies and default from credit agencies.[11] Along with verified information on borrowers' employment, income and homeownership, the platform decides whether to grant them the loans, and if so, the ratings of the borrowers, which map onto some interest rates.[12] The origination fee

[7]PwC 2015, Peer Pressure.

[8]According to an Economist article in 2014, "Peer-to-Peer Lending: Banking without Banks".

[9]However, the loans are not covenanted, and thus the usage of credit is not enforced. Based on Lending Club's description, the products include personal loans and small business loans (greater than $35,000). All loans exceeding $35K are secured (collateralized) but the data is not disclosed and thus will not be further discussed.

[10]Since their establishments, both Lending Club and Prosper operate alongside WebBank, an FDIC-insured Utah-chartered Industrial Bank. See Prosper Lender Registration Agreement and Lending Club Prospectus. Also, see "Where Peer-to-Peer Loans Are Born," Bloomberg.

[11]Lending Club claims to pull one or more credit reports from credit agencies such as Transunion, Experian and Equifax, whereas and Prosper pulls data from Experian.

[12]Both Lending Club and Prosper claim that they only accept prime borrowers. Lending Club strictly rejects any borrower with FICO score below 640 and for Prosper, 600. For Lending Club, the rating

on a loan ranges from 1% to 5% depending on its rating.[13] With the origination fee, the interest rate, the loan term and the loan size, it becomes a contract offer back to the borrower. The repayment structure on the loan contract amortizes monthly like a mortgage. Had the borrower accepted the loan, it will be posted on the platform's website along with her credit information to attract financing.

The listing period for a loan is typically up to 14 days.[14] Each loan is typically split into notes of $25. After observing the loan contracts and the borrowers' credit history, lenders make investment decisions.[15] A loan will be issued as soon as it's been filled by the lenders. If the lenders' pledged amount exceeds 60% of the requested at expiration, the borrower can either keep the funded amount under the same interest rate, or reject the loan and refund the lenders.[16] Although platforms keep their lenders' information opaque, it is reported that institutional investors chipped in more than 80% in peer-to-peer lending platforms.[17]

The loan issuances are jointly done by the platforms and their partner banks. A partner bank (or WebBank for Lending Club and Prosper) is the sole underwriter of the loan securities. WebBank, the Utah-based charter, is believed to be favored as the loan underwriter due to its freedom from the limitation by Glass-Steagall Act and favorable consumer finance code such as no caps on interest rates charged and exportation of

spectrum ranges over 35 categories. For Prosper, the rating is coarser, 7 categories.

[13]For example, for a loan of $1,000 with a 5% fee, the platform receives a $50 loan origination fee at the time of loan issuance. The borrower receives $950 of capital, but the principal on the loan stays at $1,000.

[14]Before December 2010, listing on Prosper was 7 days.

[15]Institutional Investors tend to invest in whole loans rather than a fraction. See 'Wall Street is hogging the peer-to-peer lending market', QUARTZ.

[16]It is 60% for LendingCLub and 70% for Prosper. Under any other cases, the loan will be dropped, and its lenders are refunded.

[17]Historically, Prosper disclosed information on lenders for each loan until late 2013. According to Lending Club, institutions can be banks, pension funds, asset management companies, etc. See 'The Evolving Nature Of P2P Lending Marketplaces', Techcrunch.

interest.[18]

Borrowers are expected to follow the payment schedule, and early payoffs are encouraged. If payments are delinquent more than 150 days, the loan is charged-off and sold to collection agencies for recovery. Defaults appear on borrowers' credit report and limit their future borrowing ability.

2.1.3 Comparison to Consumer Banking

Similar to commercial banking and credit rating agencies (CRA), peer-to-peer lending produces and aggregates information on borrowers, prices interest rates and makes a market between borrowers and lenders. However, unlike banks, a peer-to-peer lending platform typically does not take positions in securities for lenders and thus does not need capital requirements or deposit insurance.[19] Unlike CRAs, peer-to-peer lending platforms are responsible for market clearance between lenders and borrowers, in addition to information production.

Opinions are dispersed on the future of peer-to-peer lending.[20] Not being a perfect substitute for banking, it exists due to several comparative advantages, and is able to undercut banks on both borrowing and lending. Banks and borrowers benefit heavily from relationship formation, where borrowers have access to cheaper credit and banks in return get lower credit risk. Agarwal et al. (2009) show that 56% of accounts in their sample are "Relationship Accounts". Data provided by Prosper show that although once

[18]See "The future of finance, the rise of the new shadow bank" from Goldman Sach's Research.

[19]Due to this feature, liquidity shocks induced by agency costs between banks and lenders are not applied here. See Hellmann et al. (2000), Allen and Gale (2000) and Diamond and Dybvig (1983)

[20]Some suggest that peer to peer lending isn't a threat to the banking industry, while others claim it may be the future of banking and the credit market. See "Peer-to-peer lenders will never challenge the banks, says Deloitte", The Telegraph. See "Lending Club Can Be a Better Bank Than the Banks", Bloomberg.

peaked to 45% in 2011, the number stays around 20%-30%.[21] First time borrowers who are screened by banks get undercut by P2P platforms.

Second, the claimed purposes of most personal loans are for debt consolidation and refinancing. Facing higher interest rates from credit card debt, borrowers may receive lower interest on a peer-to-peer lending platform. According to the Fed, the average interest rate on an issued credit card is between 13-14% (APR) in 2014. On Lending Club alone, conditional on a FICO score 750 and above, the average interest rate is 8.4%. [22] P2P also caters to borrowers with lower creditworthiness, where the interest rate can go as high as 32%, 10% more than the highest legal rate among states with regulations.[23]

On the lending side, as deposit institutions, banks provide deposit insurance, and thus guarantee "risk-free" returns. First, P2P platforms cater to agents with heterogeneous risk preferences. The average deposit interest rate is less than 0.5% on a 3-year CD, whereas, the adjusted annual return for loans on Lending Club in 2016 averages between 4.9% and 8.3%. Second, institutions such as banks and funds account for more than 80% of the loan volume, which indicates that institutions are essentially undercutting each other using P2P lending platforms.[24] Banks compete locally due to geographical limitation and regulation. P2P lending provides the means for banks to undercut each other at the national level.

[21]'Is the Surge of Repeat Borrowers at Prosper Over?', Lendacademy

[22]Also, controlled for loan contract terms and borrower credit history, a loan with claimed purpose as 'debt consolidation' is 2.4% cheaper than "small business". Securitized loans from commercial bank such as auto-loan are generally cheaper (4% APR) than peer-to-peer lending, while incomparable since P2P lending does not require collateral.

[23]It is still debatable on what the credit rationing interest rate is. Some states are not regulated, such as ME, NH, NV, UT, SC and NM.

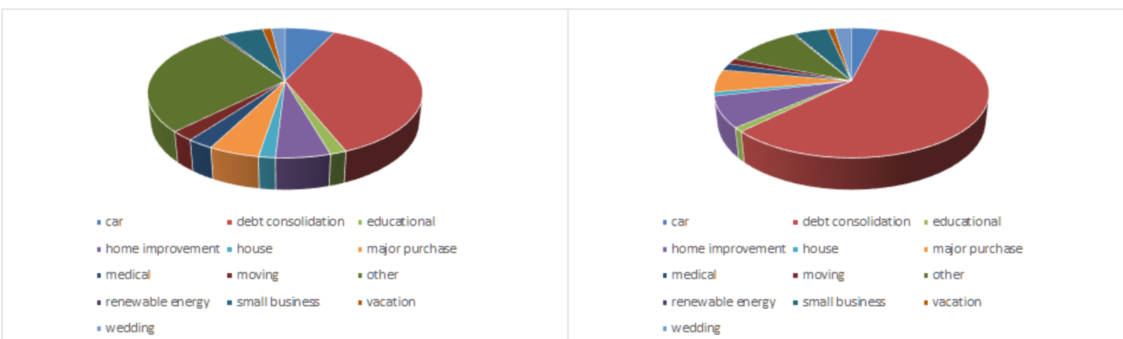
[24]Citigroup announces their partnership with Lending Club as an institutional investor. See "This Huge Bank Is Coming to Lending Club's Rescue." Fortune

2.2 Institutions: Lending Club vs. Prosper

Lending Club is currently the world's largest profit-seeking crowdfunding and the largest peer-to-peer lending platform, followed by its biggest competitor, Prosper.com.[25] By 2015, $22 billion in loans were issued between them.

2.2.1 Lending Club

Founded in 2007, Lending Club issued 646,389 among 5,317,010 loan applications by the end of 2014. With close to 40% applications intended for debt refinancing, the proportion is over 60% among the issued loans. Other major purposes include car financing, educational, housing and home improvement, purchases, medical, small business, vacation, wedding, etc. (see Figure 1). Without collateral, Lending Club targets 'prime' borrowers



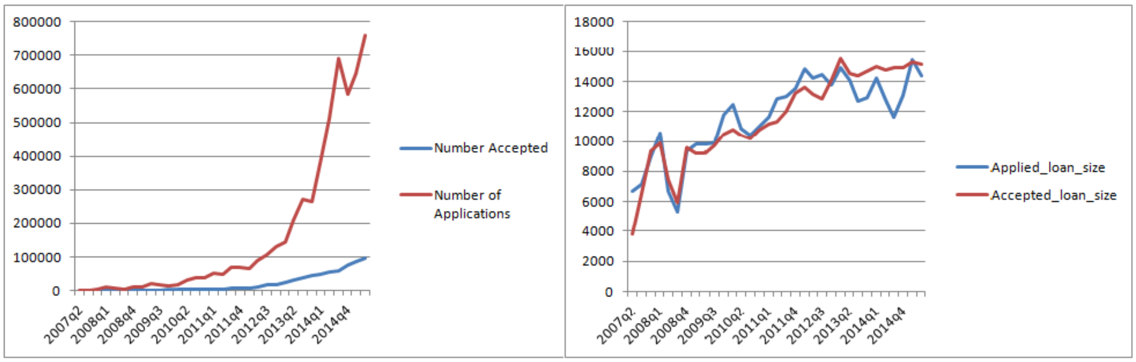
Reported loan purpose popularity ranks from debt consolidation with more than 40%, housing expenditure (housing, major purchase, home improvement) at 12% and car financing at 7%. With an average acceptance rate below 9%, loans originated for debt consolidation reach more than 60%.

Figure 1: Loan purposes among applications (left) and acceptances (right)

with FICO scores above 640. An algorithm-based screener inputs several categories of information. Some are self-reported such as age, address, income, employment length and homeownership, and may require verification by additional documents. Some are

[25]Lending Club currently accepts borrowers in all states but Iowa and West Virginia and lenders in all but Ohio, Pennsylvania, New Mexico, North Carolina and Hawaii. Other platforms such as Academic Capital Exchange, CapAlly, GreenNote, and so on have joined the market recently.

prompted for the intended loan such as loan size request, loan purposes and loan term. Others are pulled from a credit reporting agency that include FICO scores, total account, first credit line, revolving utilization and balance, total debt, credit history and public record such as delinquencies and defaults. Then, for the intended loan, it outputs a rating and an interest rate (or rejection).[26] The rating system consists of 35 grades



The number of loan applications on Lending Club grew since founded in 2007. Once it reached 100,000 per quarter in early 2012, the number jumps to 700,000 per quarter in 2014. However, the number of loans being accepted has been growing at a much steadier rate and reached about 100,000.
The average requested and originated loan sizes appear to grow in the same pattern to $15K.

Figure 2: Left: applications vs origination. Right: Loan size requested and issued (in $)

ordering from A1, A2,... A5, B1... G5 ascending with risk measurement, where A1 is the most creditworthy and G5 is the riskiest. Each category corresponds to an interest rate at a given time. Interest rates change over time. Table 1 shows the monthly interest rates in 2015 and loan origination fees corresponding to each loan grade. As aforementioned, loans are divided into notes of $25 and posted on Lending Club.com for up to 14 days. On a loan listing, an individual lender observes features including contract details, borrower credit history and instantaneous information on the funding status of the loan, which is captured by the current funded amount and the instantaneous number of lenders who pleaded on the loan. Contract description shows the borrower's requested loan size,

[26]A loan can either have a term of 3 or 5 years. Short-term loans are available on Prosper.com (see Figure 2)

Loan Grade	A	B	C	D	E	F	G
Monthly Interest Rate	0.44 ∼0.66%	0.68 ∼0.96%	1.02 ∼1.22%	1.30 ∼1.49%	1.52 ∼1.75%	1.83 ∼2.15%	2.23 ∼2.42%
Origination Fee	1∼4%	4∼5%	5%	5%	5%	5%	5%

Table 1: Monthly Interest Rates from APR and Origination Fees on Lending Club in 2015

The interest rates shown above for each rating are converted into monthly rates from APRs. The loan origination fees are profitted by the platform as soon as the loan is underwritten. They range from 1%-5% of the issued loan size, depending on borrower creditworthiness. For example, for an underwritten loan of $10K facing a 5% origination fee, the platform immediately profits $500, the borrower claims the residual,$9,500, and the loan principal stays at $10,000.

	A	B	C	D	E	F	G
2007	57	61	75	37	14	7	0
2008	295	507	438	222	75	21	4
2009	1,178	1,365	1,193	657	236	64	23
2010	2,709	3,284	2,293	1,472	663	200	72
2011	5,665	5,811	3,279	2,259	1,296	527	159
2012	7,667	12,010	7,799	4,766	1,953	831	174
2013	5,645	16,172	14,544	8,762	3,875	2,055	417
2014	4,107	8,580	10,310	7,389	3,988	1,506	448
Total	27,323	47,790	39,931	25,564	12,100	5,211	1297

Table 2: Numbers of Accepted Loans Across Year and Grades (2007-2014)

This table shows loans issued across ratings during 2007 - 2014. While B and C ratings cover 55% of all the issued loans, A and D are on par at 17%. E, F, G ratings are the highly risky loans at 10%.

the loan term, its listing expiration date, intended purpose, rating and interest rate.[27] Observable borrower characteristics include her 3-digit ZIP code, employment length, employment title and annual income. Borrower credit history features her debt-to-income ratio, recent FICO score range, delinquency record within the last two years, credit card revolver balance and utilization and default history. For loans issued before Sep, 2009 lenders observe some descriptions entailing the borrower's usage of the loan, her current financial situation and a Q&A between the borrower and other lenders.[28] Data show that absences of loan descriptions or descriptions with 10 characters or less largely emerge in the 4th quarter of 2009. Anecdotal research also shows this phenomenon (see Figure 3).[29] According to data and a former CEO, Lending Club has also partially funded some loans listed on its website.[30] The amount pledged by Lending Club on each loan was not disclosed to the lenders but observable to an econometrician.

For loans with vintage prior to 2010, the net annualized return stays between 5% to 7% across all the loan grades. Figure 4 illustrates the contractual vs. actual returns

[27] All loans that were issued before 2010 only had 3-year maturity.

[28] For example, one borrower elaborated "I am applying for this loan because I am trying to lower my credit card so I can start saving up some money. I graduated college two years ago and have had my current job for about a year and a half. I just moved home, (so no rent/bills- thanks mom and dad!) and I don't have a whole lot of other expenses. With some frugal months I could have this paid off, but I am getting married in about ten months and have been slammed with deposits, and a big dental bill for $3,000, virtually eliminating my savings. My parents are paying for the "big stuff" for the wedding, but I have been picking up the deposits. So, I am not in any way concerned with having to pay a few hundred dollars a month, I just would like to not be paying that high interest rate and would like to be saving some money on my end. No credit card debt with a steady amount at a lower interest rate is what I am hoping for. A monthly payment would be easily managed." See https://www.lendingclub.com/browse/loanDetail.action?loan_id=364451

[29] The reason for this event was not disclosed by Lending Club. I contend this incident is related to Prosper's re-entry. Data shows that the propensity of a borrower's comments or describe her loan went down to nearly 50% in 2010.

[30] This largely happened during the period when Lending Club was under evaluation by SEC in 2008. See 'A Look Back at the Lending Club and Prosper Quiet Periods,' Lending Academy. However, even before and after the 'quiet' period, Lending Club has also lent to its borrowers. Former CEO also told Wall Street Journal that Lending Club slowed down its activity in the 'quiet' period to use its own money to fund borrowers. "Peer-To-Peer Lenders Get Into Secondary Market," WSJ. This is one key institutional differences between Lending Club and Prosper, where Prosper has not played the role of a lender.

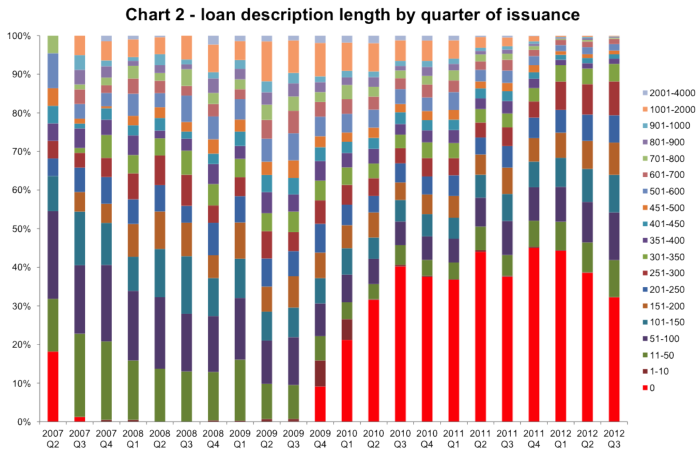


Figure 3: Loan Descriptions Length (in characteristics) with Vintage 2007-2012

Figure 3 is from an article describing research conducted by Sam Kramer on Lending Academy. http://www.lendacademy.com/lending-club-loan-descriptions-1/. While every loan originated from early 2008 to the 3rd quarter of 2009 is required to carry some description about the loan from the borrower, this requirement disappears in 2009. In late 2010, 40% of the loans do not contain any information directly from the borrower.

measured by the internal rate of return (IRR) for loans with vintage between 2008 and 2010 aggregated at monthly level.

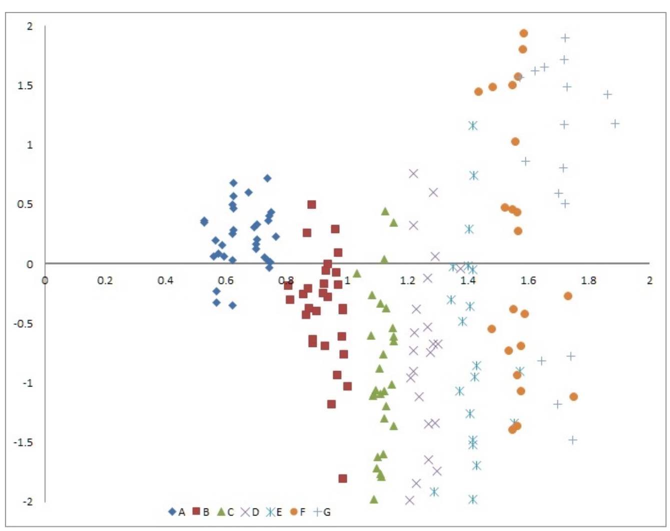


Figure 4: Lending Club: IRR for Loans Vintaged from 2008-2010

The X-axis: interest rates (IRR) of loans averaged within each rating within each origination month.
The Y-axis: their performance measured by IRR.
Note that, for loans with no or little repayment, the IRR are highly left skewed. Therefore, the average for loans with risky ratings can still be negative.

2.3 The Event: Prosper's (Re-)entry

In early 2008, SEC requested both Lending Club and Prosper (then Lending Club's sole competitor) for evaluation. On April 8, 2008, Lending Club underwent its SEC registration and entered into a 'quiet' period. It discontinued new lenders' registration, took a halt in advertising to borrowers, and funded many borrowers with its own capital. 6 months later, on October 14, 2008, Lending Club announced its immediate return, and all loans issued henceforth could be traded on a secondary market. On the second day, Prosper exited for its registration.[31] Since then, Lending Club monopolized the peer-to-peer lending market until July 13, 2009, when Prosper came back. Unlike Prosper 1.0, Prosper 2.0 set stricter guidelines to screen borrower credit background, where an eligible borrower must have a FICO score of more than 600 [32]

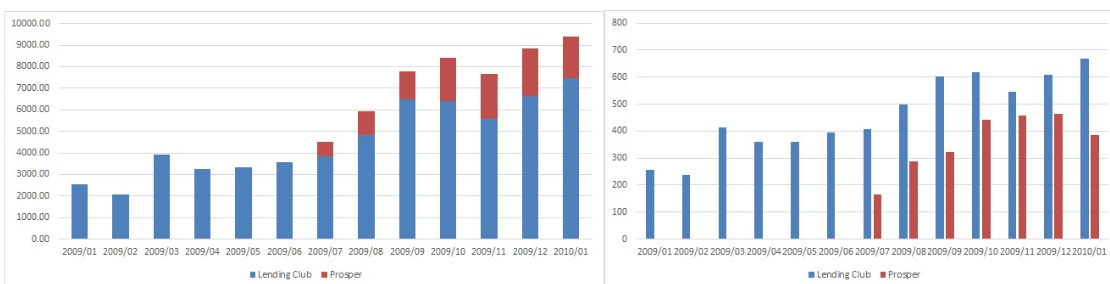


Figure 5: Left: Market Capital. Right: Loans Underwritten

Left: Note that the unit on the Y-axis is $1,000. During the six months preceding Prosper's re-entry, Lending Club monopolized the market and issued roughly $3 million per month. Joined by Prosper in the second half of 2009, the issuance grew to more than doubled. Between the two platforms, it reached around $9 millions monthly.
Right: the "incumbent," Lending Club, also dominates the market share measured by number of loans issued. Compared to the pre-entry period, the number of loans issued increased by 50% per month.

[31] See https://www.prosper.com/about-us/2008/10/15/prosper-filing-registration-statement-enters-quiet-period/

[32] Prosper 2.0 refers to Prosper after its re-entry. Proper 1.0 accepted borrowers with all credit background.

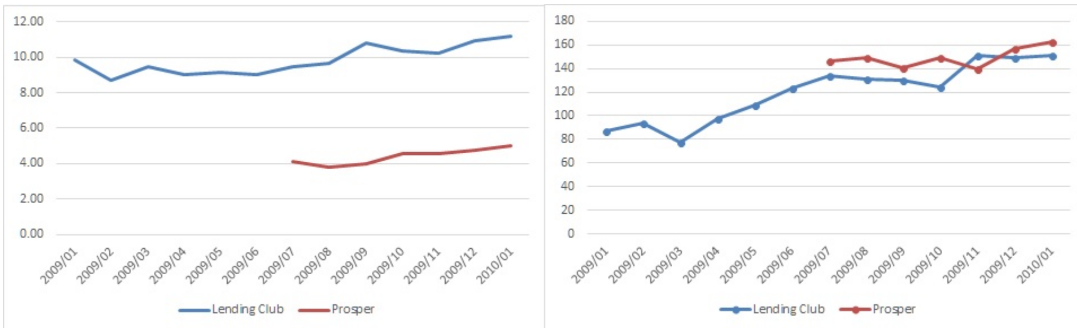


Figure 6: Left: Average Loan Size. Right: Average Lenders per Loan

Left: Note that the unit on the Y-axis is $1,000. Average loan size issued on Lending Club stays around $10K, compared to only $4K on Prosper.
Right: Number of lenders per loan between the platforms are mostly on par.

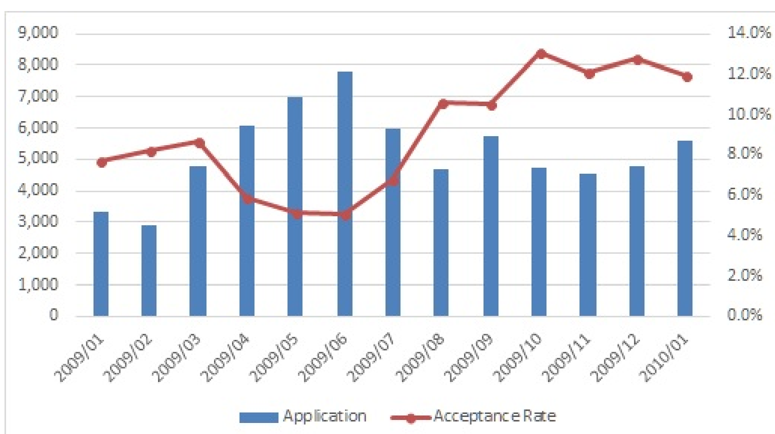


Figure 7: Lending Club's Acceptance before and after Prosper's Re-entry

Compared to the pre-entry period (before July 2009), the number of applications on Lending Club decreases by 2000. However, the average acceptance rate increases from 4% to above 10%.

Market Structure Timeline

2.3.1 Event Exogeneity and Treatment Time

I examine information asymmetry between market agents on the entry event. Prosper "re-entry" may have been well anticipated by Lending Club, and thus decisions that are unobservable to borrowers and lenders would have been made prior to the event. By not clocking the event discontinuity correctly, I may attenuate the result. To test if Prosper's re-entry is anticipated, I plot the incumbent's interest rate changes at the time of the event. (See Figure 8) The vertical lines pin down the actual timing of Prosper's entry on July 13, 2009 and the jump of the interest rates happens on August 1, 2009. First, I conclude that the incumbent, Lending Club, may or may not anticipate Prosper's re-entry, but the timing of the event is exogenous and unforeseen. Second, its interest rate changes did not happen until 15 days after Prosper's re-entry announcement. Consistent with the interest rate change, I set the time of the event at 8/1/2009. However, the result is fully robust to the event date set at 7/13/2009. (See Figure 12)

2.3.2 Prosper 2.0 vs Lending Club

Since I study how Prosper's entry affects loan origination and performance on Lending Club, I section most of Prosper's history and institutional details to the appendix. However, to identify the competition mechanism, I compare the similarities and differences in their business models, and how market participants respond to the entry event.

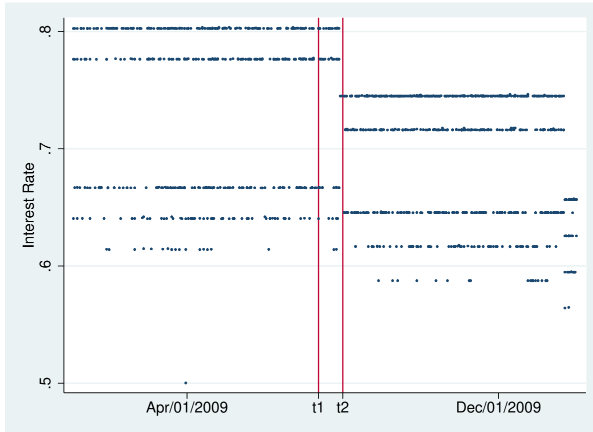


Figure 8: Entry and Interest Rates of "A" & "B"

This figure demonstrates the interest rate with rating "A" loans from early 2009 to early 2010. The first vertical line "t1" maps onto the date 7/13/2009 and the second, "t2", onto 8/1/2009. Remember that the entry announcement is on 7/13/2009, but the interest rate jump is on 8/1/2009.

While both Lending Club and Prosper screen borrowers using their credit reports, they have very different borrower selection and pricing mechanisms on the interest rates. Lending Club in general accepts borrower with higher FICO scores and lower debt-to-income ratios. After accepting the borrowers, Prosper does not price loan interest rates, but it uses an auction business model where borrowers provide reservation interest rates and lenders make bid offers.[33] Table 3 shows the institutional differences between the two platforms.

	Interest Rate Pricing	Credit Report Agency	Borrower Guideline	Fund Borrowers	Repeated Borrower Observable
Prosper 2.0	Auction (Before Dec, 2010), then Platform Pricing	Experian	FICO>600	No	Yes
Lending Club	Platform Pricing	Transunion	FICO>640	Yes	No

Table 3: Lending Club vs Prosper 2.0

This table provides some basic institutional differences between the two platforms.

[33] This business model was replaced with platform pricing in late 2010.

The funded rate on Prosper is much lower than that on Lending Club for two reasons. First, Lending Club did provide capital to some borrowers, making loans with credit crunches issued. Second, historically, Prosper was not able to maintain strong a reputation among lenders. Prior to Oct, 2008, Prosper 1.0 granted 28,936 loans with 18,480 fully paid off and 10,456 loans defaulted, consisting of total loan volume of $178K, $47K of which was written off, implying a loss rate of 26.1%. In order to maintain enough supply from the lending side, Prosper sets higher interest rates than Lending Club. Table 4, shows the unconditional average monthly interest rates across platform ratings issued from 2008-2010 on Lending Club and Prosper 2.0. Figure 9 shows an unconditional interest rates comparison across platforms over time.[34]

Lending Club	A	B	C	D	E	F	G
Interest Rate (IRR)	0.6323%	0.919%	1.1173%	1.263%	1.3993%	1.5469%	1.7039%
Prosper 2.0	AA	A	B	C	D	E	HR
Interest Rate (IRR)	0.7169%	0.9563%	1.3892%	1.7472%	2.1682%	2.6017%	2.6901%

Table 4: 2008-2010, Interest Rates Comparison

Similar to Figure 9, interest rates on Prosper are higher than Lending Club on average in the sample period. Note that here I only show the unconditional interest rates. Interest rates are measured using the internal rate of return to adjust for different loan terms.

[34]In a Reuters article on January 19, 2010, Renaud Laplanche, the former CEO of Lending Club, wrote "Lending Club approves 10% of the loan applications. That's an underwriting decision. These 10% most creditworthy loans are made available on the platform for investors to invest in, and all loan listings get fully funded. Currently, the platform is "demand constrained", meaning that I have more investors willing to invest in these loans than loans available." The article continues "Prosper's 10% is very different in nature: most loan applications received by Prosper get listed on their platform, and only 10% actually get funded, either because of insufficient supply of investors funds, or just because investors don't want to fund the other 90% of the loans. The question here is whether the 10% that get funded are "the right 10%?" See, "http://blogs.reuters.com/felix-salmon/2010/01/19/the-problem-with-peer-to-peer-lending/"

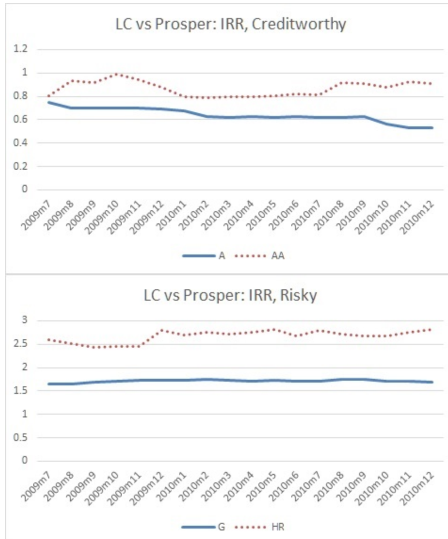


Figure 9: Interest Rate Comparison

Top: an interest rate comparison measured by IRR for loans with the most creditworthiness between the two platforms. Bottom: The same comparison for the least creditworthy borrowers. Lending Club appears to offer lower interest rates. Note that they are not conditional on borrowers' characteristics.

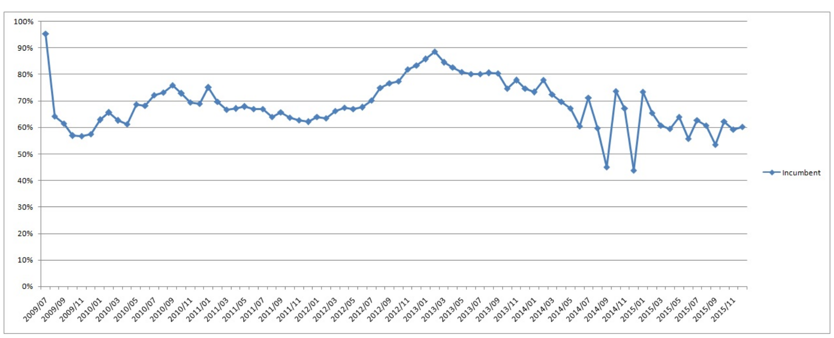


Figure 10: Lending Club's Market Share since Prosper's Entry

It shows the market share of the incumbent, Lending Club. The market share for the entrant, Prosper, is simply 1 minus the incumbent's market share. Lending Club has always been a dominant player in this market and once peaked at 90% in 2013.

3 Data and Measurements

3.1 Data Source, Lending Club

There are two pieces of unsecured loan data made available by Lending Club to the general public and its investors. One is a cross-sectional snapshot of all the loans originated from 2007 to 2015 on Lending Club. The other contains all the applications declined by the platform in the same time frame. The sample period for the majority of the analyses is around 2009, as it focuses on the entry event. I do not go beyond 2011, since loans originated after 2012 with maturity in 5 years may still under payments.

3.2 Loans Issued

Within the issued loan data, each observation uniquely identifies a loan and contains 3 categories of information, the loan's funding and contract, the borrower's characteristics and credit history, and the ex post payments.

A loan contract includes the amount requested by the borrower and the actual funded amount. The funded amount is divided into that by registered lenders and the platform. It further subsumes the loan term (36/60 month), interest rate measured in APR, monthly installment and payment schedule, loan rating assigned by the platform, loan origination date and borrower-reported loan purpose.

A borrower's attributes show her address at 3-digit zip-code and state, employment job title, length of the current employment, self-reported annual income, income verification dummy, homeownership category (rent, own or mortgage) and an optional self-reported full loan description. The credit history consists of her FICO score range at the time, debt-to-income ratio excluding mortgages or the loan from Lending Club, number of credit lines, revolving balance and utilization, number of delinquencies over the preceding two years and date of first credit line (in month/year).

A loan's payment information includes 3 key variables. The status of the loan indicates the status of the loan contract. In the sample period preceding 2012, loans are either fully paid off or charged off. I also observe the total payments on each loan, the sum towards the interest and the principal. Moreover, I observe the date of the last payment made.[35] Note that for some fully paid loans, the total payment amount can exceed or be short of the amount listed in the contract. Borrowers are able to prepay all the rest of the installments, and the interest payment will discounted accordingly. If the payments are delinquent, to fully payoff the loan the borrower has to pay more interest.

Information such as the number of lenders on each loan and loan application date are not provided. I obtain those datums from each loan's archived funding page on "Lending Club.com", and match them back to the main data using loan ID.

3.2.1 Measuring Unobservable Borrower Characteristics

The platform gets to observe the information of borrowers that neither an average lender nor an econometrician can see. The unobservable datum is likely used by the platform on loan pricing (see Blanchflower et al. (2003)), and may correlate with the observables. To measure the unobservable borrower characteristics, I combine the data uniqueness with several strands of literature. I observe the date of each borrower's first credit line. First credit lines are normally harder to obtain, since first time borrowers are riskier

[35] Recoveries and collection fees are also recorded but are not used in the paper and thus will not be mentioned.

due to information asymmetry (see Bertrand et al. (2010) and Sharpe (1990)) I define a borrower's experience as the number of months between the loan application date and that of her first credit line. This piece of information is observable by lenders but can be neglected. Moreover, based on the facts documented by Agarwal et al. (2017) where credit expansion can pass through to both borrowers and lenders, I obtain quarterly credit card delinquency data from the Fed to measure credit tightening, and pin down the 3-month trailing delinquency rate prior to the dates of borrowers' first credit lines. I argue that trailing delinquency rate measures average borrower ability to obtain credit under adverse conditions. If a borrower obtains her first credit when the market is prudent, her unobservable characteristics favor her. The first date of the credit line can be either exogenously determined by credit demand shock or endogenously by borrowers' quality. While using it to measure individual borrowers' unobservable quality induces noise, it still provides information on average. I denote this measure as "Delinquency Forbearance."[36].

3.3 Loans Rejected and Other Variables

For the loan rejection data, each observation stands for a loan application declined by the platform. The data only contains 7 variables, the application date, the requested loan size, its intended purpose, the borrower's vintage FICO score at the time of the application, the borrower's debt-to-income ratio, her zip-code and employment length.

I concatenate the loans rejected and originated by the same measures, and use a dummy variable, 'accept', to indicate if a loan is accepted and issued.[37]

[36]By including this measure in the regression analyses, I lose 546 data points out of a total of 5,972 during the "entry" sample period.

[37]Loan applications that were not funded to 70% or canceled by the borrowers are not observed in the datasets.

I add several control variables on the monthly macroeconomic environment: AAA bond yield, S&P return and mortgage rate, since the peer-to-peer lending market clearing depends both lender and borrower participation and their outside options. (See Freedman and Jin (2008))

3.3.1 Measuring Interest Rate: IRR

I do not use the interest rates provided by the platform (APR, or annual percentage rate), because loan terms can be 3 years or 5 years. I use the internal rate of return (IRR) to standardize the loan maturity. IRR is defined as the break-even discount rate where the net present value (NPV) of the investment is 0. $K = \sum_{t=1}^{T} \frac{P_t}{(1+\mathrm{IRR})^t}$, where, K is the funded amount, P_t is the payment of the loan at time t, and T is the loan term, 36 or 60 months. I compute the contractual interest rate using the cash flow out (funded amount) and cash flows in (monthly installments), where T is the loan term.

3.4 Loan Performance Measures

A loan's performance can be measured relative to its contract terms or in its absolute return. I propose two measures that are relative to contracts, Default and Percentage Nonpayment, and two others to measure the ex post state of the world, Internal Rate of Return (IRR) and Return on Investment (ROI).

3.4.1 Default

A loan's payment status can be fully paid or charged off. A dummy variable, 'Default', indicates if a loan is charged off. 'Default', the binary variable, is a straightforward measure on loan performance, but does not inform the severity of the underperformance.

3.4.2 Percentage Nonpayment

Percentage Nonpayment is defined to as percentage unpaid of the overall contractual payment: Percentage Nonpayment $= \frac{\text{Contractual Payment} - \text{Actual Payments}}{\text{Contractual Payment}}$, where 'Contractual Payment' is its monthly installment multiplied by the loan term. It measures how much it is short of the payments according to the contract, and is bounded below by 0. Since some loans are paid off early, its total payment may not coincide with the contractual payment.[38] Therefore, for loans that are fully paid, I let its 'Percentage Nonpayments' be 0. Nonetheless, this measure has a couple limitations. It ignores discounting since it does not take into account the timing of the payments.

3.4.3 Internal Rate of Return

Unlike computing for interest rate, I do not observe the actual cash flows of the loans. To compute the discount rate in IRR, I make the following assumptions. First, if a loan is paid off, regardless of possible delinquencies or early payments, its actual IRR is equal to its contractual value. Second, from total payments, loan origination dates and last payment dates, I compute IRR assuming evenly distributed payments, $P_t = \frac{\text{Actual Payments}}{\text{Number of Payments}}$. Lastly, if there is no payment ever made on a loan, I set its IRR to be -1.[39] One limitation of IRR to measure performance is that it is bounded above by the interest rates (max, 6.34%) and below by -100%. Outliers on the left spectrum cause a left skewing of the distribution, and the OLS estimates can be misinterpreted. For robustness, I introduce another measure, Return on Investment.

[38] A typical reason for a loan gets prepaid is to avoid further interest payment. There is no additional fee for a loan to be prepaid.

[39] The IRR goes to $-\infty$.

3.4.4 Return on Investment

Return on Investment is the ratio between the net profit and the investment, and in this context, ROI $= \frac{\text{Actual Payments-Funded Amount}}{\text{Funded Amount}}$. Similar to Percentage Nonpayment, ROI does not factor discounting.

3.5 Data on Prosper

Since this paper focuses on how the entry by Prosper 2.0 affects Lending Club, the data on Prosper is not directly needed. However, it provides a scope for us to identify the type of competition and mechanism.

Prosper also makes two pieces of data accessible to the general public on its website, loan listing data and loan performance data.[40] I match the sample period to Lending Club, and select loans listed or originated between July 2009 and 2011.[41]

The loan listing data includes observations of both issued and dropped loans that once were listed on Prosper. Identified by a listing number, a loan listed can end up with 4 different statuses, 'Expired,' 'Cancelled,' 'Withdrawn' or 'Completed.'[42] A loan issued if and only if it is marked 'Complete'. I merge the issued loan listing data with the performance data. [43]

Since the two platforms acquire similar information on borrowers, the structures of

[40]Note the institutional difference between the two platforms. Prosper does not show platform rejected borrowers, but loans that are failed to be cleared by the market are shown.

[41]I drop all samples earlier than 2009, since Prosper 1.0 and 2.0 are fundamentally different.

[42]For an 'Expired' listing, the funded amount does not reach a sufficient amount to issue (70%) by expiration date (7 days or 14 days after application). When Prosper changed its business model again in December 2010, it extended the listing period from 7 days to 14 days. A 'Canceled' listing means that the borrower's information is incomplete or cannot be verified. 'Withdrawn' indicates that the listing is dismissed by the borrower herself.

[43]The datasets do not have a common key to merge on. The loan performance data includes loan number as the identifier but does not contain listing number, and vice versa. Prosper API service provided by Prosper resolves the issue. Using the API, I could track down each 'Complete' listing's loan number and thus merged the two files.

their data are almost identical. Some major distinctions are separated by institutional differences. Prosper does not lend to the borrowers, and thus all loans are funded by lenders. Prosper's data indicates additional variables on repeated borrowers.[44] Borrowers' debt-to-income ratios include the current loan from Prosper.[45] Other minor differences spread among employment, address and homeownership.[46]

3.5.1 Data Concatenation across Platforms

This subsection is extended in the 'Identification' section, where I pool the data together between Lending Club and Prosper to determine the competition features and mechanisms. Here I address how I combine the two sources. I concatenate variables with the same nature such as borrower and loan characteristics including income, loan origination and application date, interest rate (IRR), performance measures (loan status, nonpayment percentage, actual IRR), funded and requested amount in dollars, loan purpose, employment length, address at state level and number of lenders.

For differentiated measures such as the FICO score, debt to income ratio and detailed addresses at county level, I make additional adjustments and assumptions. Since FICO reports from Experian (used by Prosper) and Transunion (used by Lending Club) are not mapped into each other, I assume they are highly correlated. In addition, FICO scores on Prosper are in coarse ranges. To be combined with Lending Club, I take the upper bounds of Prosper's FICO ranges. Using the variable income, debt-to-income ratio and loan size of Prosper's data, I compute debt-to-income ratios excluding Prosper's loans.

[44] Also borrowers' previous loan performance.

[45] Remember, Lending Club excludes mortgages and Lending Club loan in the calculation.

[46] In addition to employment length and occupation, prosper informs lenders on borrowers' employment type such as full-time/part-time. For borrowers' addresses, Prosper uses city name in comparison to ZIP-codes by Lending Club.

I look up the state and city names on 'www.zipinfo.com', and transform the addresses from city-level into county-level, or 3-digit zipcodes. With all the measures matched, I created an indicator "Platform" to differentiate loans from Lending Club and Prosper.

4 Empirical Strategy

I use data from the incumbent to study the effect of entry on borrower screening, credit rating, interest rate pricing, loan performance and lenders' responses.

4.1 Hypotheses Development and Sample Selection

Prosper's entry intensifies platform competition on both borrowers and lenders. Expecting similar results to Becker and Milbourn (2011), I hypothesize that the credit ratings inflate and the ex post loan performance deteriorates.[47] There is not a uniform theory on how interest rates change. The incumbent platform has an incentive to undercut the entrant to attract borrowers. On the other hand, two possible concerns arise about the argument that interest rates may rise. For one, since platforms' information may differentiate, winner's curse mitigates the undercutting behavior (See Marquez (2002)). The other is that an increase in interest rates would attract lenders. I form the following hypotheses based on these conjectures. After entry,

(i) the incumbent's incentive to screen borrowers mitigates, allowing more risky borrowers to the market and the ex post loan performance is aggravated.

[47] Flynn and Ghent (2017) find different results under an unique setting where the incumbent's reputation is low. They are able to study credit inflation on the entrant compared to the incumbent, since they credit-rate the same products, commercial loans. Here, I do not have any rationale to believe that the incumbent has low reputation. On the other hand, compared to the entrant, the incumbent may have been the preferred platform from the lender side (See institutional comparisons). Moreover, I cannot directly compare cross-platform rating schemes also due to their institutional differences.

(ii) controlling for borrower and loan characteristics, a borrower is more likely to receive better loan classification.

(iii) interest rate changes are ambiguous.

In conjunction, to better understand the mechanism, I study the lenders' responses and answer the questions:

1. How do lenders respond to the entry event and changes made by the incumbent platform?

2. What's the "disincentive" for the platform to inflate borrowers' credit?

To eliminate other possible exogenous changes in the credit market that I cannot control for, I restrict the sample period to a 1-year window around the time of the entry.[48] That is, I use 6-month data on loans rejected and issued from Lending Club both preceding and succeeding Prosper's re-entry. This sample gives me 61,441 total number of observations where an applicant has a FICO score. They include 5,972 accepted and issued loans, and 55,469 rejected by Lending Club. Separated by the entry event, 33,031 applicants appear before the entry, and 28,410 after. Of those applied before the entry event, 2431 obtained financing. Of those applied after the event, 3541 obtained loans. The these statistics along with variable summary statistics are listed in table 11 and 13.[49]

4.2 Borrower Screening

Borrower i with characteristics $\{X_i, \mu_i\}$ applies for a loan with size K_i and term T_i where K_i, T_i, X_i are observable to both econometricians and the platform, and μ_i is some latent

[48] An exogenous factor can affect the pool of the applicants' characteristics and the lenders' outside options.

[49] 14 are cross-platform statistics comparison for the last section.

variable that is only observable to the borrower. Whether a borrower is accepted is a binary choice, denoted by $\mathbb{1}_{\{\text{Accept}\}}$.[50] I use a dummy variable $\mathbb{1}_{\{\text{Entry}\}}$ to indicate post-entry loans, and its coefficient measures the Average Treatment Effect (ATE) of "Entry" on borrower acceptance propensity. I use both LPM (linear probability model) and Probit approaches to estimate ATE of Entry. That is:

$$\mathbb{1}_{\{\text{Accept}\}} = X\beta + \gamma \mathbb{1}_{\{\text{Entry}\}} + \varepsilon$$

$$\mathbb{E}\{\mathbb{1}_{\{\text{Accept}\}}\} = \Phi(X\beta + \gamma \mathbb{1}_{\{\text{Entry}\}})$$

where X subsumes FICO, Employment, Debt to Income, Requested Amount. I further control for time, address at state-level and loan purpose fixed effects. Monthly FE is perfectly collinear with the entry dummy, and thus is not added to any specifications. To show robustness, in other specifications, I shorten the window to 60 days (30 on each side of the entry point), and also experiment with more control variables: the first two statistical moments of FICO, debt-to-income ratio and requested loan size within each month.

In another setup, I apply a Sharp Regression Discontinuity Design (SRDD) to study the event as a quasi-experiment, see Lee and Lemieux (2010). I contend that SRDD is more clean and robust to ATE. The quality of the applicant pool may change over time which can contaminate the ATE. If applicant quality improve, the ATE overestimate the effect of entry. Otherwise, the previous estimate is attenuated. By focusing on borrower screening right around the event, I take advantage of the exogeneity of the event, where

[50] If the borrower's preliminary signal, $Y_i \subset X_i$, is good enough (above some latent cutoff, $\underline{x}$, chosen by the platform), she will be accepted. I implicitly assume μ_i follows i.i.d. Normal distribution.

I show an applicant quality does not change. For that reason, I am able to fully recover the effect of the event itself. First, I estimate the acceptance propensity separately before and after the entry event:

$$\mathbb{E}\{\mathbb{1}_{\{\text{Accept}\}}|Y_i, \text{Entry} = 0\} = \Phi(X\beta)_{Pre-entry}$$

$$\mathbb{E}\{\mathbb{1}_{\{\text{Accept}\}}|Y_i, \text{Entry} = 1\} = \Phi(X\beta)_{Post-entry}$$

Using the estimation above, I conduct an in-sample prediction on each applicant and fit two separate local Epanechnikov polynomials of degree 3 against time in days until the entry and after the entry. For example, for loans submitted 5 days prior to the entry, the value on the timeline is -5. For those submitted 10 days after the entry, the value is +10. To make it "local", I experiment 15-day, 20-day and 30-day windows on both sides of the entry event (See Figure 11).[51] Results show that even controlling for monthly FE, the increment of accepting propensity goes up by more than 4%, compared to 6.9% prior to the entry. The direction is consistent with Becker and Milbourn (2011), where the platform screens less and brings in riskier borrowers.

[Table 15 Here]

To verify the assumption that the event itself doesn't immediately induce higher borrower quality, in a similar discontinuity setup, I show that an average applicant's creditworthiness including her FICO score and debt-to-income ratios does not appear to improve within the same time frame. (See Figure 13)

[51] Other nonparametric kernels such as Gaussian do not change the result either.

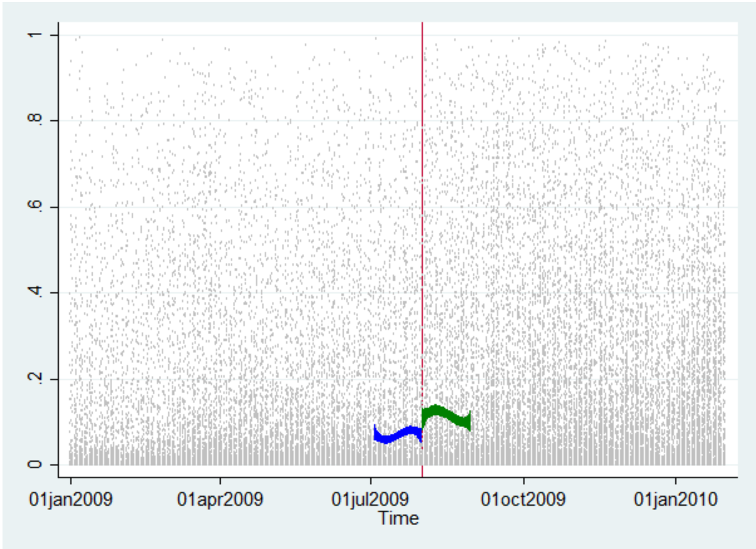


Figure 11: RDD with Confidence Interval: Loan Selection

Using the estimated equations from the two Probit above, one on the sample preceding the entry point (left of the vertical red line) and other succeeding it (right of the vertical red line), I respectively predict each applicant's propensity of being accepted. This is shown by the "gray" dots in the graph. The Epanechnikov local polynomials with 95% confidence intervals are separately fitted onto the predicted propensities, shown in blue and green, each with a 30-day window.

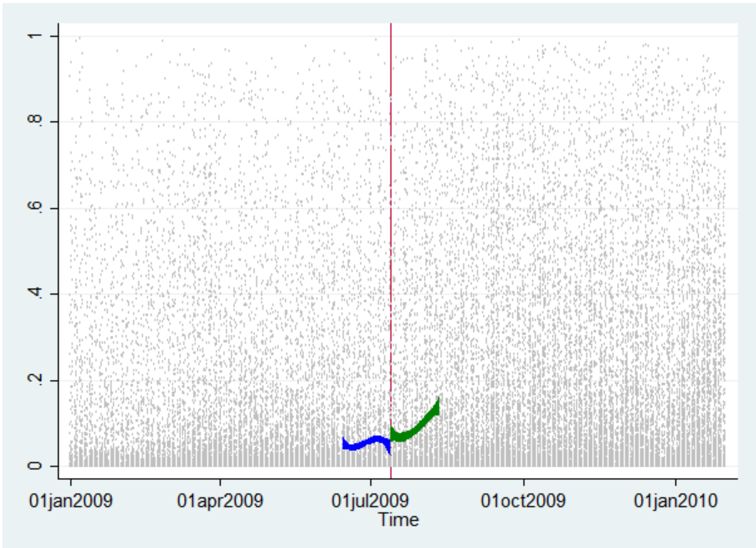


Figure 12: RDD with Confidence Interval: Loan Selection

Remember that the entry event announcement date is on July 13, 2009. In figure 11, the event date was set to August 01, 2009 to be consistent with the interest rate jump shown in Figure 8. This figure is a robustness check on the event date, by replicating Figure 11 with the event date on 07/13/2009.

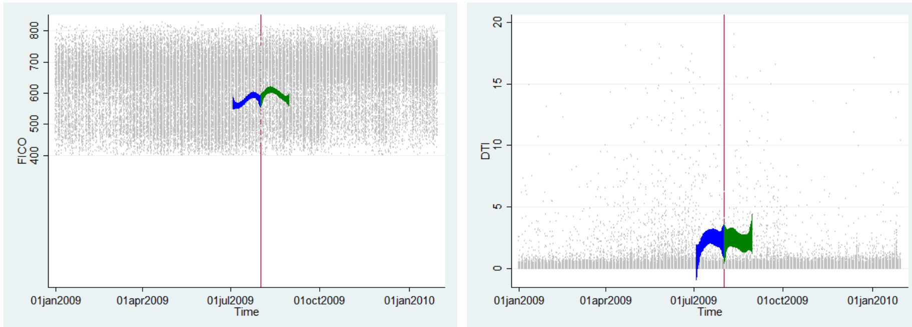


Figure 13: Robustness Check: Applicant FICO and DTI

Left: I show local polynomial fits of applicants' FICO scores against days until and after the entry event, with a 30-day window on each side. Right: Applicants' debt to income ratio. First, I do not observe an upward trend in the polynomials on either graphs. Second, the graphs do not show discontinuities at the time of the entry event. Both measures indicate that the applicant quality does not significantly improve.

4.3 Loan Classification

The platform classifies the accepted borrowers into 35 categories, A1, A2, ..., G5. To allow enough observations and statistical power, I reduce 35 categories into 7 ordered by creditworthiness from A to G. I adapt the methodology applied by most credit rating literature (see Becker and Milbourn (2011)). The credit rating is presumably monotonic in borrowers' creditworthiness, and thus, I use a linear approach by regressing the credit rating (Numerated from 1 to 7) onto borrower characteristics X:

$$\text{Rating} = X\beta + \gamma\mathbb{1}_{\{\text{Entry}\}} + \varepsilon \tag{1}$$

where X subsumes 2 categories of information, borrowers' attributes and loan contracts. A borrower's attributes are characterized by her revolving utilization percentage, credit open accounts, homeownership, annual income, income verification status, previous records on delinquency and default. A loan contract includes the term of the loan and the requested loan size. Note that interest rates are bijective to loan ratings and thus not included.

The information observed by the platform but not by the lenders may be used in credit rating.[52] I hypothesize that controlling for borrower characteristics, an average borrower is more likely to get better loan rating, i.e. $\gamma < 0$. I also use an Ordered Logistic Model to estimate credit rating, where the platform sets 7 latent cutoffs to classify borrowers' creditworthiness:

$$\text{Grade} = \begin{cases} A & \underline{x_A} \geq X\theta + \mu_i \\ B & \underline{x_B} \geq X\theta + \mu_i > \underline{x_A} \\ C & \underline{x_C} \geq X\theta + \mu_i > \underline{x_B} \\ D & \underline{x_D} \geq X\theta + \mu_i > \underline{x_C} \\ E & \underline{x_E} \geq X\theta + \mu_i > \underline{x_D} \\ F & \underline{x_F} \geq X\theta + \mu_i > \underline{x_E} \\ G & X\theta + Z\beta + \mu_i > \underline{x_F} \end{cases} \tag{2}$$

The OLS results in Table 16 show that on average, the borrower's rating goes up by 0.14, and if we take unobservable information into account, the magnitude goes up to 0.4. This result is arguably underestimated, because the average accepted borrower becomes riskier (see Table 7). The negative coefficient on "Delinquency Forbearance" shows that lenders who are able to obtain first credit under adverse environments are likely to get better ratings. The marginal effects derived from the Ordered Logistic Model yield that borrowers are 2.6% more likely to be rated into Grade A and 1.3% more into Grade B. (See Table 17)

Again, an SRDD specification would be robust to possible post-entry borrowers' characteristic improvement. I separately estimate an Ordered Logistic model over pre-entry

[52]The private information is measured by "Delinquency Forbearance".

and post-entry borrowers, and conduct in-sample predictions of the probabilities. [53] I take the sum of the predicted probabilities on Grade A and B, and denote it "Creditworthy Rating Propensity." Similarly, I denote the sum of the rest "Risky Rating Propensity." The goal here is to show if there is discontinuities on "Creditworthy Rating Propensity" between pre- and post-entry. Figure 14 shows that the relative "Creditworthy Rating Propensity" ("Risky Rating Propensity") appears to have a jump (dip) at the entry event. The magnitude is close to 10%. The results again show credit inflation, and indicate the incumbent's incentive to undercut the entrant to maintain borrower population and market share.

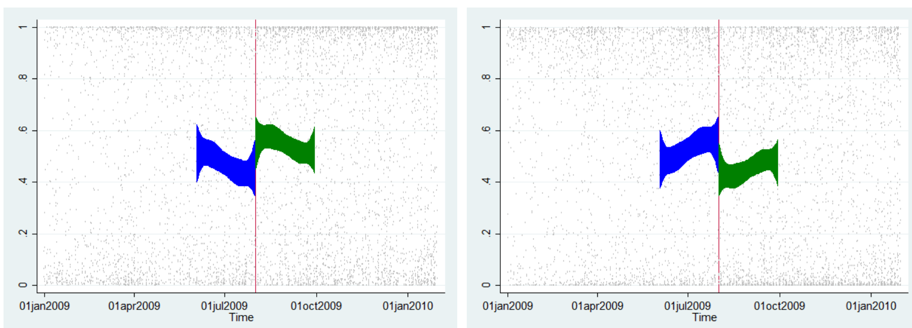


Figure 14: RDD on credit Rating

Left: I show local polynomial fits of "Creditworthy Rating Propensity" against days until and after the entry event, with a 60-day window on each side. Right: local polynomial fits of "Risky Rating Propensity" against days until and after the entry event, with a 60-day window on each side. Note that here I no longer use 30-day windows because observations on issued loans are fewer than the number of applicants. To observe a more significant effect, I extend the window size.

[Table 17 and 16 Here]

[53] Note that since the Ordered Logistic Model accounts for 7 categories, each borrower has 7 predicted probabilities, with sum equal to 1.

4.4 Entry and Interest Rate

In the previous section, I show credit inflation. However, borrowers are expected to be more elastic to their interest rates. As preliminary analyses, I plot interest rates within each loan rating category over the sample period (see Figure 15). It shows that creditworthy borrowers obtain cheaper financing whereas the subprime borrowers' interest rates increase. Since the accepted borrower pool is more heterogeneous after the entry,

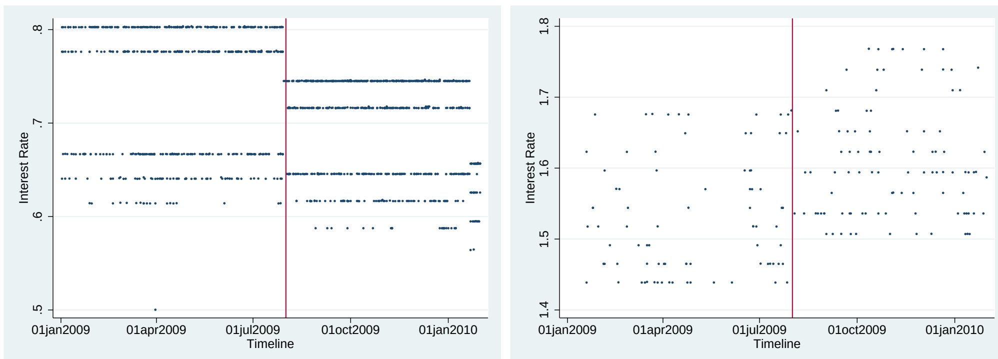


Figure 15: Entry and Interest Rate for Grade A & F,G

Left: I show the interest rates measured by internal rate of return for loans of rating A (most creditworthy borrowers). Right: the same measure for loans of rating F&G (riskiest borrowers). The vertical lines represents the time of the entry event. If only judged by ratings, creditworthy borrower obtain cheaper financing after the entry event, whereas the risk premium requirement over risky borrowers increases.

I control for observed borrower characteristics and loan classifications. First, using a semi-parametric approach from Firpo (2007), I estimate the quantile treatment effects (QTE) of entry on the interest rates. The intuition is similar to propensity score matching, where I group borrowers with same/similar characteristics from before and after the entry. Instead of being interested the average treatment effect, I estimate heterogeneous effects over creditworthiness evaluated at each decile.[54](See Table 5) It may be counter-intuitive to find that some borrowers bear higher interest rates. One effect comes from

[54]I also experiment with out-of-sample validation approach, where I use the estimated pre-entry mechanism to forecast out-of-sample interest rates on post-entry borrowers. (See Table 18)

Quantile	QTE	Z-score
0.001	−0.049	−14.72
0.1	−0.060	−10.83
0.2	−0.058	−7.86
0.3	−0.008	−1.18
0.4	−0.031	−5.35
0.5	0.000	−0.01
0.6	0.005	0.79
0.7	0.011	1.48
0.8	0.016	2.08
0.9	0.026	2.56
0.999	0.092	5.93

Table 5: Entry and Interest Rate: Quantile Treatment Effects

The table shows the estimated results on the treatment effects of entry on interest rates at different quantiles (or creditworthiness). I separate borrower creditworthiness by their interest rates at each decile. The quantile treatment effect model (QTE) matches borrowers' propensity score measured by their observable attributes such as FICO scores, debt to income ratios, requested loan size, income, etc, and compare their interest rates. The event study statisfies the exogenous treatment requirement by QTE. The result shows that the most creditworthy borrowers obtain 0.05% cheaper monthly interest rate, and the riskiest borrowers get loans almost 0.1% more expensive. Z-scores are measuring the statistical significance of the effects. Since more risky borrowers are introduced, this model cannot provide a precise comparison.

the fact that the platform screens less and brings in much more risky borrowers. To compensate the additional risk brought to the lenders, the platform needs to increase their interest rates for market clearing. The other possible effect stems from unwanted adverse selection induced by borrowers' rate shopping behavior. Specifically, the platforms' information upon a borrower is differentiated. Borrowers go to the platform that obtain better information about them. This unwanted adverse selection, or "Winner's Curse", entices them to raise interest rates due to information uncertainty from the other platform. To identify the underlying mechanism, it is important to examine the other side of the market.

4.5 Lenders' Response and Mechanism

In particular, I study the effect of interest rate changes on the platform's market-making through lender's participation and market clearing efficiency. Without observing lender identities or characteristics, I cannot analyze lender competition at the individual level. However, with the information on number of lenders per loan, the time duration between a loan application date and issuance date and the platform's capital provision, I examine lenders' participation and market clearance at the loan level.[55] [56]

4.5.1 Model Selection

Bolton et al. (2012) argues that lender's trusting nature is one cause for the platform to inflate ratings. To validate or disprove the assumption, in this section, I compare data fitness under two different models. On one hand, lenders ignore all the borrower characteristics, X, and make decisions only based on the loan ratings. On the other hand, lenders can perfectly evaluate a borrowers' risk profile without using the credit rating by the platform. We test the validity of the two models separately before and after the entry

---

[55]The less time a loan takes to issue, the more efficient the market is.

$$\text{Funding Duration} = \text{Loan Issuance Date} - \text{Loan Application Date}$$

[56]The less funding provided by the platform, the more efficient the market clears.

$$\begin{aligned}&\text{Pct Platform} = \frac{\text{Lending Amount by Lending Club}}{\text{Total Lending Amount}}\\&\text{Pct Lenders} = 1 - \text{Pct Platform}\end{aligned}$$

Note that, one does not observe how the platform selects to fund the borrowers. Using the ex post loan performance, I test the two opposing theories: the platform getting its "skin in the game" by making investments using own capital Vs. it signaling its market clearing competency. With a linear probability model, I estimate the equation below:

$$\mathbb{1}_{\{\text{Default}\}} = X_i \cdot \theta + \delta \text{Pct Platform} + \varepsilon \tag{3}$$

If $\delta > 0$, we can conclude that the correct theory is the latter one. (See Table 25) event. In the first case:

$$\text{Number of Lenders} = \boldsymbol{\beta}'\text{Grade} + \varepsilon^{57} \qquad \text{(Trusting)}$$

In the second case, if lenders ignore the rating, I have the following specification:

$$\text{Number of Lenders} = X\beta + \boldsymbol{\kappa}' \cdot I_i + \varepsilon \qquad \text{(Sophisticated)}$$

I compare their adjusted R-squared ($\bar{R}^2$), Bayesian information criteria (BIC) and Akaike information criteria (AIC). The three metrics in Table 6 show two implications. First, lenders are neither extremely sophisticated or trusting, but "Sophisticated" model has better explanatory power and goodness of fit. In addition, the post-entry fitness show improvement on "Sophisticated" model, but deterioration on "Trusting" model. Therefore, in the analyses below, I use "Sophisticated" model.

	Before Entry				After Entry			
	N	$\overline{R}^2$	AIC	BIC	N	$\overline{R}^2$	AIC	BIC
Trusting	2,431	0.1817	26225.19	26694.67	3,541	0.097	40162.19	40662.13
Sophisticated	2,431	0.2612	25867.16	26371.02	3,541	0.624	36911.52	37448.13

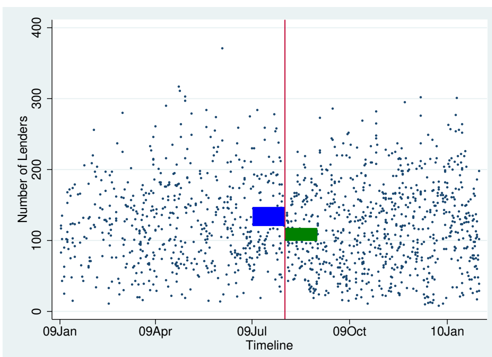
Table 6: Model Comparison

It makes the most sense to compare vertically, or the top with the bottom rows, since they have the same number of observations. Sophisticated model maintains a higher adjusted R-sqaured and lower AIC and BIC, all of which indicate that Sophisticated model contains more information and better data fit.

[57] Note that since the only variation of I_i within grade comes from the entry event, I_i is not identifiable in the equation above and thus is not included. Ideally, one would use an instrument or control function approach to separate the two effects. However, it is not applicable here, since the platform determines the co-movement of loan classification and interest rates.

4.5.2 Hypotheses and Estimation

Controlling for observable borrower characteristics, I expect that loans are funded more efficiently for those with higher interest rates. Also, with higher lender participation induced by higher interest rates, the capital provided by the platform decreases. As a preliminary pass, I separate the loan ratings into groups with different directions in their interest rate changes, and compare pre-entry with post-entry number of lenders within each group. (See Figure 16) Lenders' reaction to interest rate changes is rather intuitive, where I observe higher participation for those with higher interest rates and less traction for those with lower. Note that the jumps in figure 16 contain two effects. The direct effect



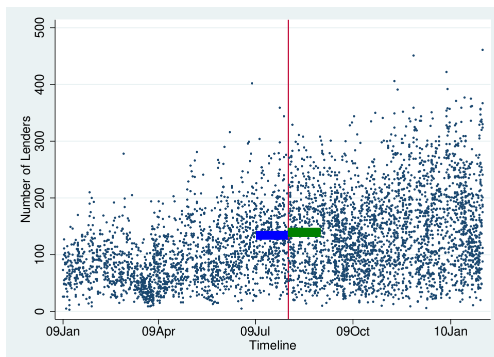


Figure 16: Entry and Number of Lenders: Left, Grade A | Right, Grade B & above

Left: this figure shows the number of lenders on loans with rating A, with local polynomial fits at 30-day windows. As expected, the number of lenders per loan drops because the credit borrowers gets cheaper interests. Right: the figure shows the same metric with loans of ratings B-G. However, the number of lenders does not seem to significantly improve. This might be attributed to the increasing risk among the borrowers.

comes from the interest rate changes induced by competition. Indirectly, the risk within each loan rating has also increase. By controlling for borrower observable characteristics, I recover the direct effect of entry on number of lenders:

$$\text{Number of Lenders} = X_i\theta + \boldsymbol{\gamma'}\mathbb{1}_{\{\text{Entry}\}} \times \text{Grade} + \varepsilon_i$$

$$\text{Number of Lenders} = X_i\theta + \boldsymbol{\kappa}\mathbb{1}_{\{\text{Entry}\}} \times I_i + \varepsilon_i$$

γ' compare the pre- and post-entry conditional number of lenders controlling for borrower and loan characteristics. κ, the coefficient on the interaction of the entry dummy with interest rates, measure the heterogeneity of average number of lenders responding to interest rates before and after the entry. Even though Grades and interest rates are one-to-one mapped onto each other, the interpretations of γ and κ differ. Credit Rating mechanism changes are not directly observable to lenders, or at least to some lenders, whereas the interest rate changes are one of the most important features lenders focus on.

I apply the estimation procedure above to the other measures. Figure 17 shows percentage of the loan size financed by the platform. The graph shows that the interest rate decrease does not affect the market clearing for creditworthy borrowers. However, lenders are much more elastic to the raise of interest rates over the risky borrowers. Market efficiency measured by Funding Duration in Figure 18, show that it takes longer for creditworthy loans to issue, but slightly less time for risky borrowers. Now, I further

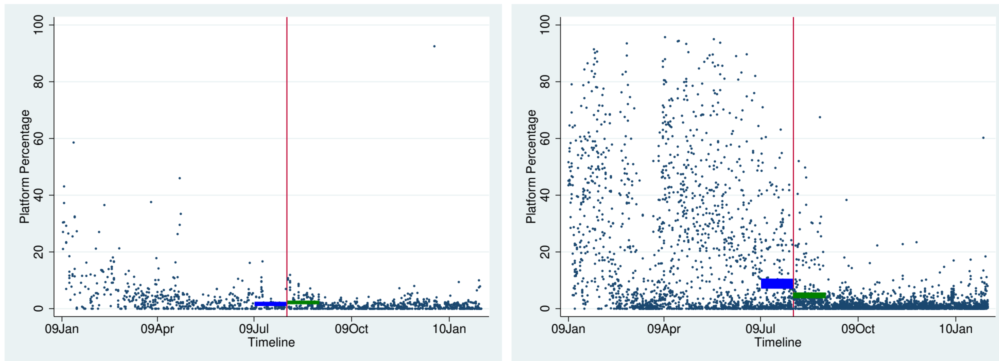


Figure 17: Entry and Pct Platform Capital: Left, Grade A | Right, Grade B & above

Left: this figure shows the percentage capital provided by the platform on loans with rating A, with local polynomial fits at 30-day windows. Credit crunches do not emerge following lowered interest rates. This implies that, to ensure market clearance, the platform doesn't not have incentive to overly undercut borrowers' interest rates. Right: the figure shows the same metric with loans of ratings B-G. From both the scatter plot as well as the polynomial fits, I observe credit crunches diminish significantly following the entry event. For one, rising interest rates induce more investments. For another, the event may induce new lenders to enter, which helps the loans to clear.

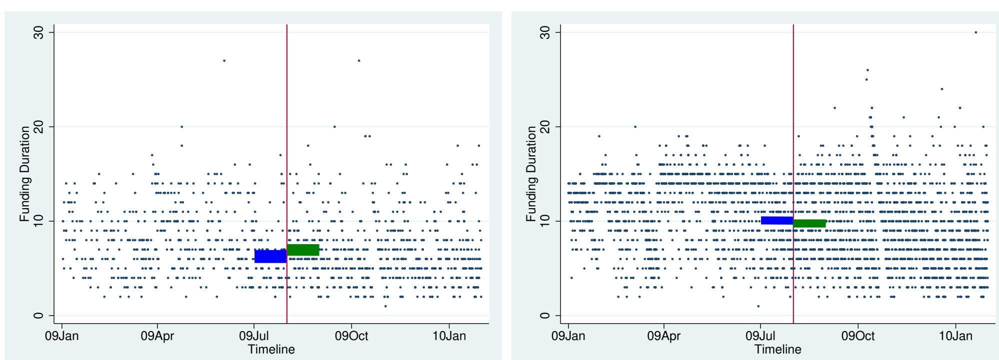


Figure 18: Entry and Funding Duration: Left, Grade A | Right, Grade B & above

Left: this figure shows the funding flow measured by funding duration in days on loans with rating A, with local polynomial fits at 30-day windows. As expected, with lower interest rates, it takes longer for loans to be cleared off the market. Right: the figure shows the same metric with loans of ratings B-G. I do not observe a significant increase in the funding flow induced by higher interest rates.

control for borrowers' characteristics and estimate the following equations:

$$\text{Pct Platform} = X_i\theta + \boldsymbol{\gamma}'\mathbb{1}_{\{\text{Entry}\}} \times \text{Grade} + \varepsilon_i$$

$$\text{Funding Duration} = X_i\theta + \boldsymbol{\gamma}'\mathbb{1}_{\{\text{Entry}\}} \times \text{Grade} + \varepsilon_i$$

$$\text{Pct Platform} = X_i\theta + \boldsymbol{\kappa}'\mathbb{1}_{\{\text{Entry}\}} \times I_i + \varepsilon_i$$

$$\text{Funding Duration} = X_i\theta + \boldsymbol{\kappa}'\mathbb{1}_{\{\text{Entry}\}} \times I_i + \varepsilon_i$$

Table 19 shows results that are consistent with the preliminary analyses, where creditworthy borrowers get less traction in terms of the number of lenders, but market clearing efficiency is inelastic to the decrease in interest rates. On the other hand, lenders' response to an increase in interest rates on risky borrowers is much more positive and significant. Column 1 shows that the most creditworthy loans receive 28 less lenders with the interest rate decrease. Among most of the loans whose interest rates go up, I observe a higher lender participation. However, in column 2 and 3, market efficiency estimated by funding duration and the platform's capital provision increase uniformly for all credit ratings. The last 3 columns are the estimation results for the linear approach. To better understand the intuition, based off the estimates, I predict the marginal effects to compare the heterogeneous responses of lenders facing different interest rates both before and after the entry (See Figure 19). I can infer that first, the entry event likely expands the market and induce lender participation. Second, lenders are more aware of the change in interest rates than creditworthiness, which is possibly due to their trusting nature.

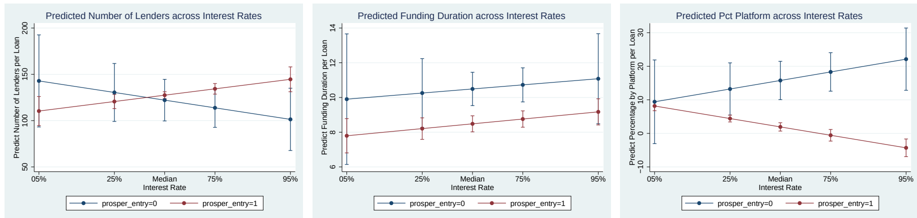


Figure 19: Model Predicted Lenders' Marginal Responses of Interest Rate

Left: Number of Lenders, Middle: Funding Duration, Right: Pct Platform. Horizontal Grid: 5th, 25th, 50th, 75th and 95th percentiles of in sample interest rates. See Left: Before entry, the number of lenders decreases with interest rate, but the direction reverses after the entry. The overall number of lenders appears to be higher. See Middle: Creditworthy loans take less time to be filled than risky loans. The efficiency increases uniformly after the entry. See Right: measured by platform's capital provision, the lenders' preference shows a more significant increase for risky loans than creditworthy ones.

4.6 Entry and Loan Performance

In addition to the effect of competition on loan screening and interest rates, it is the most important to understand its effect on the ex post loan performance. Loan performance determines lenders' payoffs and the reputation and continuation of the platform.

Based on results from Becker and Milbourn (2011) and prediction from Bolton et al. (2012) and Marquez (2002), I hypothesize that the relative loan performance (to the loan contract) of those originated after the entry is inferior to those from the pre-entry period. First, because the incumbent platform screens less prudently by accepting more risky borrowers and potentially lemons, the average loan performance deteriorates. Second, credit inflation makes the deterioration uniform across all ratings. It is ambiguous how loan returns measured by IRR or ROI change. The increase in interest rates can benefit the lenders by improving overall returns, or it can be dominated by the lower average borrower quality and hurts the lenders' payoff.

4.6.1 Default

Default is a binary variable, and indicates the status of loan payments. It however, does not capture the precise return on a loan. For borrower i with characteristics X_i, she defaults if her unobservable average monthly earnings R_i cannot cover her interest rate I_i, i.e. $\mathbb{1}_{\text{Default}} = \mathbb{1}_{R_i<I_i}$. Her monthly earning R_i is a function of her observable characteristics X_i, unobservables ε_i and ex post exogenous macroeconomic shocks, which I measure using the monthly fixed effect of her last payment date of the loan.

I use Probit to estimate the default propensity. To measure average treatment affect (ATE) of entry, I incorporate $\mathbb{1}_{\{\text{Entry}\}}$ in the equation of interest.

$$\mathbb{E}\{\mathbb{1}_{\{\text{Default}\}}|X_i\} = \Phi(X_i\theta + \gamma\mathbb{1}_{\{\text{Entry}\}})$$

$$\mathbb{E}\{\mathbb{1}_{\{\text{Default}\}}|X_i\} = \Phi(X_i\theta + \gamma\mathbb{1}_{\{\text{Entry}\}} \times Grade)$$

where X_i represents borrower attributes and loan characteristics, and I control for address fixed effects, last payment monthly fixed effects and macroeconomic environment at origination. Similar to the previous specifications, I again apply SRDD, and estimate the default propensity jump for loans originated around the entry point.

4.6.2 Percentage Nonpayment

Percentage Nonpayment measures how much the actual payments are short compared to that stipulated by contracts, without discounting. Remember that:

$$\text{Percentage Nonpayment} = \frac{\text{Contractual Payment} - \text{Actual Payment}}{\text{Contractual Payment}}$$

For a fully paid loan, its Percentage Nonpayment is 0. I estimate the equations below using OLS:

$$\text{Percentage Nonpayment} = X_i\theta + \gamma\mathbb{1}_{\{\text{Entry}\}} + \varepsilon$$

$$\text{Percentage Nonpayment} = X_i\theta + \gamma\mathbb{1}_{\{\text{Entry}\}} \times Grade + \varepsilon_i$$

Based on the hypothesis, loan performance relative to the contract should be worse after the entry, i.e. $\gamma > 0$. Note that the higher Percentage Nonpayment is, the worse off the lenders are. Similar to Default, this measure has its limitations, since it does not capture the absolute returns.

4.6.3 Absolute Performance: IRR & ROI

Finally, I estimate the effect of entry on loan performance captured by IRR and ROI:

$$\text{IRR}_i = X_i\theta + \gamma\mathbb{1}_{\{\text{Entry}\}} + \varepsilon_i$$

$$\text{ROI}_i = X_i\theta + \gamma\mathbb{1}_{\{\text{Entry}\}} + \varepsilon_i$$

$$\text{IRR}_i = X_i\theta + \boldsymbol{\gamma}'\mathbb{1}_{\{\text{Entry}\}} \times \text{Grade} + \varepsilon_i$$

$$\text{ROI}_i = X_i\theta + \boldsymbol{\gamma}'\mathbb{1}_{\{\text{Entry}\}} \times \text{Grade} + \varepsilon_i$$

Table 20 shows the estimated results for all specifications above. The post entry default propensity increases by 5.6% and the nonpayment on an average loan raises by 3.8%. Absolute performance measured by IRR and ROI also decreased significantly by 1.8% and 6.5%, indicating that the overall borrower quality deteriorates. Moreover, the performance decline appears in a similar magnitude and uniformly across all ratings.[58] I contend that as a result of credit inflation and possible adverse selection, the loan performance is aggravated.

Table 20 Here

4.7 Lenders' Punishment

One aspect literature has yet to document is lenders' responses to borrowers' poorer performance. Bolton et al. (2012) argue that the incentive for a CRA to be "truth-telling" is to avoid lenders' "punishment" and discontinuity of future payoffs. In this section, I examine how lenders punish the platform using measures on credit supply.

While individual lenders' identities are not observed, repeated lenders are documented to exist in this market (See Freedman and Jin (2008)). Upon a deteriorated vintage performance, I hypothesize that repeated lenders leave the market. I measure the observed performance by aggregating the realized vintage loan returns within each calendar month. Specifically, I tracks loans originated within the sample until their last payment dates. I group those vintage loans by their last payment months and ratings. Within each group, I compute the default rate and nonpayment percentage to measure realized platform underperformance. I map those measures back to the newly issued loans corresponding with their issuance months and loan ratings. To test the hypothesis, I examine if credit supply is affected by vintage loan performance. To account for other unobservable noise

[58] One except is that G rating appears to have better post-entry performance.

to the market, I add monthly fixed effects.[59]

$$\text{Number of Lenders}_i = X_i\theta + \eta\text{Realized Default}_{t-1} + \varepsilon_i$$

$$\text{Number of Lenders}_i = X_i\theta + \eta\text{Realized Nonpayment}_{t-1} + \varepsilon_i$$

$$\text{Platform Percentage}_i = X_i\theta + \lambda\text{Realized Default}_{t-1} + \varepsilon_i$$

$$\text{Platform Percentage}_i = X_i\theta + \lambda\text{Realized Nonpayment}_{t-1} + \varepsilon_i$$

where I also experiment Realized Default and Realized Nonpayment with both currently month and one-month-lagged values. Hypothetically, $\eta < 0$ and $\lambda > 0$, since underperformance should cause lenders to leave and shortages on credit supply. Also, I separately run regressions on loan origination prior to and also after the entry event. (See Table 21)

Estimated results show that first, facing loan underperformance, lenders leave and punish the platform. Specifically, with a 1% increase in nonpayment of the vintage loans, the newly originated loans during the pre-entry period lose 0.71 lenders on average (column 1). Or if the realized default rate jumps from 0 to 1, I expect an average loan loses 61 lenders (column 2). Second, the "punishment" is alleviated after the entry. The magnitudes are no longer significant after the entry. The percentage of capital contributed by the platform shows similar results, where a 1% increase in nonpayment induces 0.27% increase in platform capital provision during the pre-entry period.

Upon poor performance, lenders downsize their investment and may leave the platform. This is a disincentive for the platforms to bring in risky borrowers and inflate the credit. However, the post-entry lenders' response immediately becomes less elastic. This

[59]Note that by construction, the variation of the underperformance measure within each month comes from loan ratings, and thus monthly fixed effects do not wash away all the effects.

is likely due to the fact that as the market size expands post-entry, new and unfamiliar lenders join the platform, which mitigates the 'punishment' effect. This marks additional incentives for the platform to 'inflate' the credit rating to compete for borrowers.

5 Adverse Selection Identification

To better under stand the mechanism of competition and platform incentive, in this section, I combine data from both lending platforms, and make several key identifications on the underlying mechanism corresponding to credit inflation, interest rate changes and lenders' response. Perfect competition such as Bertrand and yield quite different equilibrium outcome from imperfect competition such as market segmentation. Moreover, since practically the platforms are differentiated, unobserved borrower/lender preference may yield asymmetric/non-simultaneous competition, such as "first-mover advantages"

Had I observe exact borrower identities, I can determine the exact type of competition by studying repeated borrowing and encroachment on borrowers. With the outcome of competition, I can identify the mechanisms of competition on credit inflation and lender competition. Without exact borrowers' identities, I use several key criteria to construct a fuzzy bilateral match metric and incorporate the difference between their business models.

5.1 Borrower Identification across Platforms

The basic idea of the fuzzy matching mechanism follows Liu, Nekipelov and Park (2017). They cross-platform match application identities between two online App stores, Itunes and Android.

Note that, in addition to borrowers who obtained loans, I also examines those that either were rejected by the platform or failed to be financed by lenders. Relative unique information on borrowers and loan characteristics that both platforms share include 3-digit zipcode borrowers addresses, application dates, FICO scores, opening month of first credit line, open accounts and revolving utilization.

Step 1: Sample Selection First, I assume that the reasons a borrower apply at both platforms can be either rate shopping or loan rejection either by the platform or not getting enough traction by the lenders. For either case, the timing of the loan applications is close to each other. I further separately consider two cases: borrower accepted or rejected. Given that a borrower (target) is accepted/financed by one platform, I limit the search pool of the target borrower to the rejected ones on the other within the preceding two-week period. Given that a target borrower is rejected by one platform, I restrict the search sample to a one-month-window around the time of its loan rejection (2 weeks on each side). The window length selection has two trade-offs between sample inclusiveness and noise. With a short window, I reduce the chance of a correct match, but also reduce the likelihood of a noisy match. [60]

Step 2: Sample Refinement Among the data features, I use several keys to refine the search pool and improve the search efficiency. I first match up the exact 3 digit zipcodes and opening date of the first credit line up to the monthly level between the target and the search sample. The platforms use different credit reporting agencies. Instead of matching the exact FICO scores, I restrict the sample to borrowers whose FICO score ranges overlap with the subject's.[61]

[60] I also have experimented a 10-day and 7-day window.

[61] If the target's FICO score is only a vintage score, I make it a range of ± 50 window. Although the platforms obtain credit reports from different agencies, I contend that their information is highly correlated.

For a target, the first two steps can generate no or many matches. For cases with no matches, I do not further explore. For borrowers with multiple matches, I go to step 3 to construct the metric.

Step 3. Fuzzy Matching Borrower and loan characteristics differ between the two platforms on FICO scores and loan application date. I construct the metric to measure the difference between the target and each candidate in the search sample. [62] The metric is a product of two differences. The first is the difference of the FICO scores between the target and searched candidates. The second is the difference between their application date. I choose the borrower with the highest matching metric.

This method of cross-platform matching has its limitations. The algorithm of matching is rather mechanical and dependent upon numerous assumptions and a market is defined. Moreover, features in data across platforms are unbalanced. Noise may be generated through several channels. Borrowers' addresses is at city level by Prosper, and manually mapped into a 3 digit zip-code level to match with Lending Club. Some key information among rejected borrowers is censored on Lending Club such as income and credit revolving rate. However, the noise is against matching from both sides. That is, matching rejected borrowers from Lending Club to Prosper does not produce more (or less) noise vice versa. Therefore, on average, had the competition between the platforms on par, I would observe similar number of borrowers matched from either direction.

However, I find that on average, borrowers rejected by the incumbent are more likely to be financed by lenders on the entrants, but not vice versa. (See Table 7) It shows evidence that the platform competition is rather imperfect, and the incumbent platform

[62] Note that information from borrowers' self report and gets later verified such as annual income is not used as a metric. It can serve as reference to weed out matches with large disparities. Also, observed information on one platform that is partially censored on the other is also not used as a first order match.

holds "first mover advantage" in the sense that some of its rejected borrowers are able to obtain financing on the entrant. Also, among the 40 borrowers who are accepted by Lending Club and not fully funded by lenders on Prosper, their average interest rate is 1.114% per month, and 33 of them fully paid off the loans. For the 697 borrowers who are rejected by Lending Club and receive loan contract on Prosper, their average monthly interest rate stay at 1.677%. Within 113 of them who fail to pay off, their monthly loss rate measured by IRR is -14%, compared to -6% on Lending Club. That is, the risky borrowers prefer the incumbent platform, and are likely to join the entrant only if they are rejected. To better observe the competition mechanism, I go to the market level.

	Prosper	2009/07		2009/08		2009/09		2009/10
Lending Club	Unfinanced	Financed	Unfinanced	Financed	Unfinanced	Financed	Unfinanced	Financed
Reject		34		89		80		97
Accept	2		6		2		13	
		2009/11		2009/12		2010/01		
Reject		125		138		134		
Accept	4		6		7			

Table 7: Borrowers' Identity Matched across Platforms

The table shows fuzzy-matched borrowers across the platforms separated by loan origination months. The first cell shows that 34 rejected borrowers by Lending Club were financed on Prosper in July, 2009. However only 2 who fail to obtain finance on Prosper got loans from Lending Club. During the sample period, the total instances that financed by Prosper and rejected by Lending Club reach 697, 6 times of those in the reverse direction.

5.2 Competition and Market Segmentation

I combine the datasets from both platforms using similar measures such as FICO scores, debt to income ratios, interest rates, loan terms, etc, and restrict the sample period to the following 6 month since the entry. I weed out rejected borrowers and only focus on those who obtain loans. This sample gives us a total of 6,341 loans, with more than 60% from the incumbent. The FICO scores for borrowers on the incumbent range from 664 to 824 with an average of 720 and a standard deviation of 36.

5.2.1 Data Sampling

In this section, I delve into their competition and differentiation at the market level. Figure 21 shows the distributions of borrowers' FICO scores between the platforms. While the entrant finances borrowers with a wider range of creditworthiness, competition among the prime borrowers appears to be intense. Figure 20 compares the interest rate distribution between the two platforms. Borrowers on average are more creditworthy and receive cheaper rates on the Incumbent. Figure 22 demonstrates the loan size distributions,

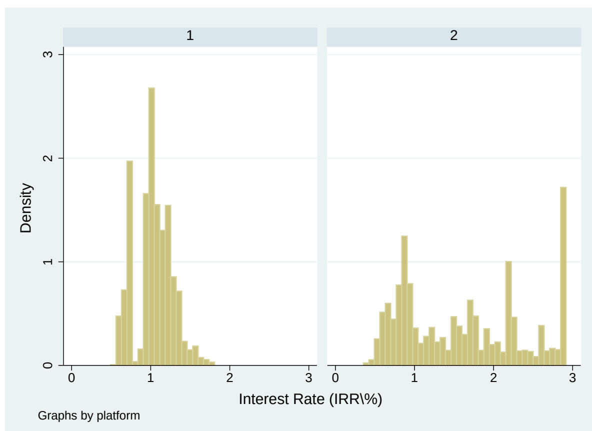


Figure 20: Interest Rate Comparison: Left, Lending Club; Right, Prosper

Left: In sample monthly interest rate distribution measured by IRR in percentage on Lending Club. Right: upper bounds of FICO score ranges of financed borrowers. Lending Club has monthly interest rates mostly below 1.5% and does not go beyond 1.8%. The lowest interest rate on Lending Club does not go below 0.5%. However, in comparison, Prosper's interest rates can go as high as 3% monthly and as low as 0.4% in IRR, and are more evenly distributed.

Since borrower distributions differ between the platforms, to study competition, I filter a subsample where the borrowers' traits overlap between the platforms. Using key features such as loan size, FICO, income and credit history, I only select borrowers from one platform that are not "outliers" of the other. Specifically, if a borrower's FICO score

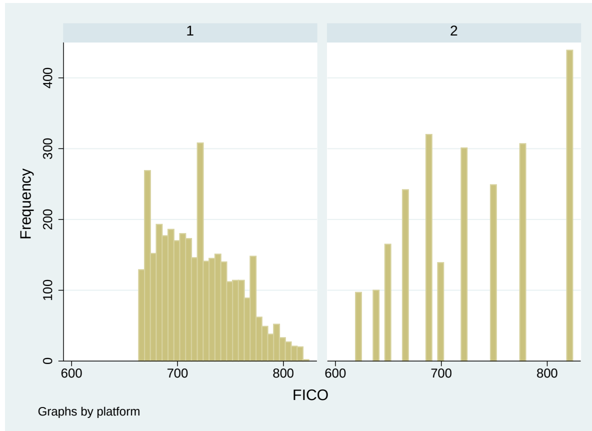


Figure 21: FICO Comparison: Left, Lending Club; Right, Prosper

Left: In sample monthly interest rate distribution measured by IRR in percentage on Lending Club. Right: upper bounds of FICO score ranges of financed borrowers. Lending Club cuts off borrowers with FICO scores below 660. FICO scores of borrowers on Lending Club are heavily distributed between 660 to 780. However, Prosper accepts borrowers with FICO scores above 600. Among those successfully financed, I observe lenders strongly prefer borrowers with high FICO scores.

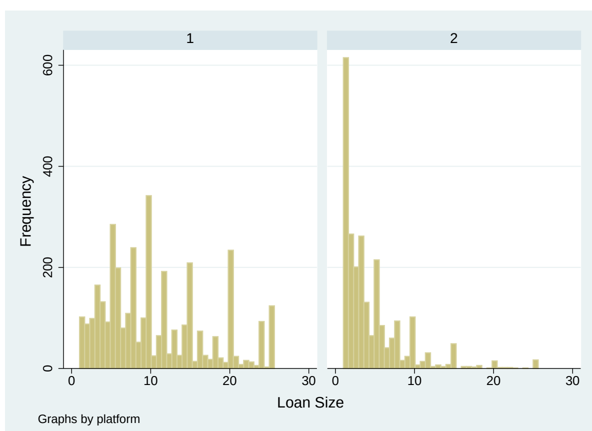


Figure 22: Loan Size Comparison: Left, Lending Club; Right, Prosper

Left: Distribution of in sample issued loan sizes on Lending Club. Right: the same metric on Prosper. Lending Club grants borrowers with loan size requests less than $25K. Among the issued loans, I still frequently observe cases where the borrowers receive $25K. On the other hand, Prosper's loans are unlikely to be issued at or beyond $10K.

(or other variables) on a platform is below the minimum or above the maximum, he is excluded from the sample. I end up with 2940 loans from the incumbent and 951 from the entrant (compared to 3541 and 2359).

5.2.2 Hypotheses and Empirical Strategy

I argue that if perfect competition exists, then borrower and platform's matching should be completely random, and no one can charge a markup or discount. To estimate borrowers' self-selection into platforms, I run a LPM model on borrower-platform matching:

$$1\{\text{Platform}\} = X\beta + \varepsilon \tag{4}$$

Note that in Table 22, the dependent variable 1{Platform} is equal to 1 for the entrant and 0 for the incumbent, and I report the standardized coefficients. Column (2) of Table 22 shows that, among the borrowers whose characteristics overlap, the platform selection is differentiated. Borrowers with higher FICO scores and self-reported annual income are more likely to appear on the entrant whereas they also appear to have more debt to income ratio, public record and open accounts. Also shown in Figure 22, successfully funded borrowers can hardly obtain loans that are above $10K. I further measure if either platform has pricing premium or discount over the borrowers' creditworthy spectrum. First, to allow for difference in platform pricing strategy and sensitivities on each variable, I interact all the observables with the entrant dummy. Second, to observe the pricing discrepancy at different creditworthiness, I use a quantile regression and estimate the equation below at 1th, 10th, 25th, median, 75th, 90th and 99th percentiles of interest rate.

$$\text{Interest Rate}_i = \alpha + \theta \mathbb{1}_{\{\text{Entrant}\}} \times X_i + \delta \mathbb{1}_{\{\text{Entrant}\}} + \mu_i \tag{5}$$

Here, the parameter of interest is θ, the premium (discount if negative) that the entrant charges. Table 23 shows that the entrant charges at a discount for the very creditworthy borrowers and at a premium for more than 3 quarters of the borrowers who obtain loans in the market.[63]

These results above show that the entrant does not have much ability to poach borrowers from the incumbent. Even with a discount on creditworthy borrowers, the entrant has to differentiate from the incumbent. Moreover, as an intermediary, the entrant has to take lenders' participation and market clearing into consideration. With cheaper interest rates, the incumbent is likely holding a first mover advantage, while inducing a winner's curse problems upon the entrant. Specifically, a lender with a "bad" signal on the incumbent can be viewed as "good" on the entrant, resulting an adverse selection problem. However, this problem is not as severe in the reverse direction (See Table 7).

5.2.3 Competition Mechanisms

I attribute the results and the incumbent's "first mover advantages" to several reasons. First, the incumbent actively screens borrowers using information observed by as well as censored to the lenders. However, in principle, the entrant accepts all borrowers and only lets the lenders decide who to finance. Therefore, with the information asymme-

[63]First, for precise interpretation of the difference, I need to delve into the marginal effects. Second, I do not use propensity score matching method because it requires that the borrowers select the platforms randomly, which is contradicted by Table 22.

try between the platforms, good borrowers self select into the incumbent. Second, the incumbent post interest rate on all loans and the entrant using an auction mechanism where lenders decide their reservation interest rates and amount to finance. Prone to adverse selection, lenders are posting much higher interest rates to compensate unobservable risks. The high interest rate feeds back into borrowers' rate shopping behavior, and thus the incumbent platform is preferred by the borrowers. Third, the incumbent established its reputation on its market-making competency by filling credit crunches, whereas the entrant does not provide such services.

Finally, I compare the borrower performance between the two platforms using borrowers' ex post internal rate of returns.

$$\mathrm{IRR}_i = \alpha + \theta X_i + \delta \mathbb{1}_{\{\mathrm{Entrant}\}} + \mu_i \tag{6}$$

Table 24 compares borrowers' default rate, monthly internal rate of returns and overall return on investments. Since it is perceived that the entrant covers more subprime borrowers, I find that the borrowers' default propensity is 12% higher among the platform "overlapped" sample and 15% in the whole sample. Meanwhile, borrowers' performance measured by return on investment is about 5-6% higher for the entrant. The average monthly discounted return shows no significant difference between the two platforms, compared to a 3% disparity between their interest rates at the median. This statistic confirms the theory that 1. borrower competition is segmented, and the incumbent platform has a first mover advantage over prime borrowers. 2. Adverse selection problem is exacerbated upon the entrant by the competitive advantages from the incumbent. Therefore, the interest rates on the entrant is much higher to adjust for the adverse risk.

6 Conclusion

In this paper, I use peer to peer lending data and an entry event to study how competition affects financial intermediaries' incentive on credit screening, credit rating and market clearing. Competition induce intermediaries' imprudent behavior through less rigorous screening and credit inflation. It is a result of an unbalanced trade-offs between defending its reputation by protecting lenders' interests and encroaching on borrowers for higher payoffs and more dominant market share. As the market size expands, new and unfamiliar lenders enter, which further dilutes the effect of upset repeated lenders and intensifies borrower competition. In the meantime, the dominant intermediary with first mover advantage further induces an unwanted adverse selection issue upon other players. I attribute the formation of the first mover advantages to its intermediation role: actively screening borrowers beyond using "hard" information and getting its skin in the game when borrowers encounter credit crunches.

References

Sumit Agarwal, Souphala Chomsisengphet, Chunlin Liu, and Nicholas S Souleles. Benefits of relationship banking: Evidence from consumer credit markets. 2009.

Sumit Agarwal, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel. Do banks pass through credit expansions to consumers who want to borrow? The Quarterly Journal of Economics, 133(1):129–190, 2017.

Franklin Allen and Douglas Gale. Optimal financial crises. The Journal of Finance, 53 (4):1245–1284, 1998.

Franklin Allen and Douglas Gale. Financial contagion. Journal of political economy, 108 (1):1–33, 2000.

Thorsten Beck, Olivier De Jonghe, and Glenn Schepens. Bank competition and stability: cross-country heterogeneity. Journal of financial Intermediation, 22(2):218–244, 2013.

Bo Becker and Todd Milbourn. How did increased competition affect credit ratings? Journal of Financial Economics, 101(3):493–514, 2011.

Marianne Bertrand, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman. What's advertising content worth? evidence from a consumer credit marketing field experiment. The Quarterly Journal of Economics, 125(1):263–306, 2010.

David G Blanchflower, Phillip B Levine, and David J Zimmerman. Discrimination in the small-business credit market. Review of Economics and Statistics, 85(4):930–943, 2003.

Patrick Bolton, Xavier Freixas, and Joel Shapiro. The credit ratings game. The Journal of Finance, 67(1):85–111, 2012.

John H Boyd and Gianni De Nicolo. The theory of bank risk taking and competition revisited. The Journal of finance, 60(3):1329–1343, 2005.

Thorsten Broecker. Credit-worthiness tests and interbank competition. Econometrica: Journal of the Econometric Society, pages 429–452, 1990.

Tim S Campbel and William A Kracaw. Information production, market signalling, and the theory of financial intermediation. The Journal of Finance, 35(4):863–882, 1980.

Nicola Cetorelli and Philip E Strahan. Finance as a barrier to entry: Bank competition and industry structure in local us markets. The Journal of Finance, 61(1):437–461, 2006.

Zhaohui Chen, Alan D Morrison, and William J Wilhelm Jr. Traders vs. relationship managers: Reputational conflicts in full-service investment banks. The Review of Financial Studies, 28(4):1153–1198, 2014.

Douglas W Diamond and Philip H Dybvig. Bank runs, deposit insurance, and liquidity. The journal of political economy, pages 401–419, 1983.

Sergio Firpo. Efficient semiparametric estimation of quantile treatment effects. Econometrica, 75(1):259–276, 2007.

Sean Flynn and Andra Ghent. Competition and credit ratings after the fall. Management Science, 64(4):1672–1692, 2017.

Seth Freedman and Ginger Zhe Jin. Do social networks solve information problems for peer-to-peer lending? evidence from prosper.com. 2008.

Barney Hartman-Glaser. Reputation and signaling in asset sales. Journal of Financial Economics, 125(2):245–265, 2017.

Robert Hauswald and Robert Marquez. Competition and strategic information acquisition in credit markets. Review of Financial Studies, 19(3):967–1000, 2006.

Thomas F Hellmann, Kevin C Murdock, and Joseph E Stiglitz. Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough? American economic review, pages 147–165, 2000.

Michael C Keeley. Deposit insurance, risk, and market power in banking. The American Economic Review, pages 1183–1200, 1990.

David S Lee and Thomas Lemieux. Regression discontinuity designs in economics. Journal of economic literature, 48(2):281–355, 2010.

Yongdong Liu, Denis Nekipelov, and Minjung Park. Timely versus quality innovation: The case of mobile applications on itunes and google play. working paper, 2013.

Robert Marquez. Competition, adverse selection, and information dispersion in the banking industry. Review of financial Studies, 15(3):901–926, 2002.

Steven A Sharpe. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. The journal of finance, 45(4):1069–1087, 1990.

Zaiyan Wei and Mingfeng Lin. Market mechanisms in online peer-to-peer lending. Management Science, 2016.

Appendix A Institutional Detail, Prosper

Founded two years before Lending Club in 2005, Prosper debuts peer-to-peer lending in U.S. under an auction model, while accepting borrowers with any credit background. Prior to Oct, 2008, Prosper had granted 28,936 loans with 18,480 fully paid off and 10,456 loans defaulted, consisting of total loan volume of \$178K, \$47K of which was written off, implying a loss rate of 26.1%. The auction mechanism works as follows. Borrowers put down reservation interest rates. Prosper acquired the borrower's credit reports and posted them online for a 7-day open-bid multi-unit uniform-price auction with reservation price. Lenders (bidders) specified the amount and the interest rate bids. Lending position are ranked in a descending order by their interest rate bids. Once the pledged amount exceeds the requested loan size, the lowest winning bid is the ongoing interest rate for the loan. If the loan is not fully funded by expiration, the ongoing interest rate is the borrower's reservation price. .[64] Prosper 2.0 announced that it would only accept borrowers with

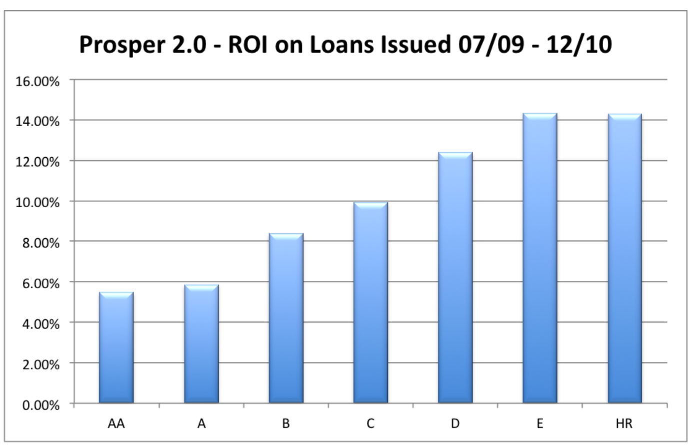


Figure 23: Return on Investments Prosper 2.0

FICO above 600 and started classifying borrowers into different risk ratings ranging from

[64]Prosper 2.0 is much improved compared to Prosper 1.0. (See figure 23, from A Look Back at Prosper 1.0 ? How Relevant are the Numbers? Lending Academy)

AA to HR by its evaluation of borrowers' creditworthiness, so that lenders can better understanding the default risk.[65] Similar to Lending Club, debt-consolidation is the main reason for loan request. Other purposes including home improvement and small business are also quite popular. (Figure 24) At the time of its re-entry, Prosper and Lending Club

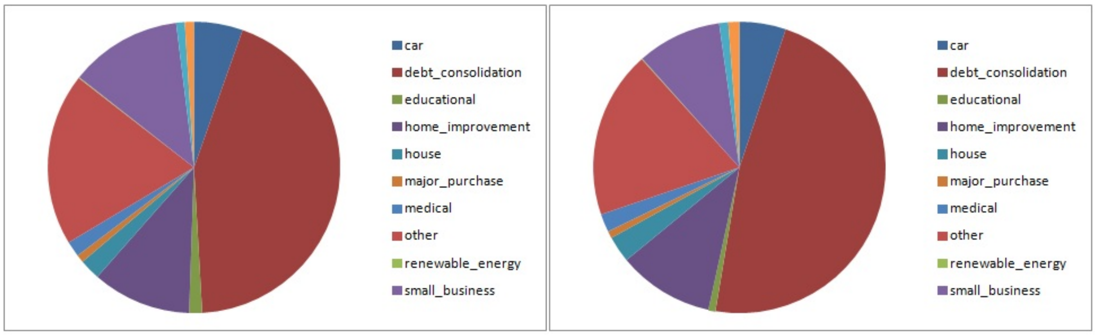


Figure 24: Loan Purpose for Prosper 2.0, Left: application Right: issued

Loan Grade	AA	A	B	C	D	E	HR
Interest Rate	0.57%	0.77%	1.02%	1.36%	1.80%	2.22%	2.57%
Origination Fee	0.5%	4%	5%	5%	5%	5%	5%

Table 8: Monthly Interest Rates and Origination Fees on Prosper in 2015

were almost identical except several key discrepancies as follows. Foremost, still under the auction business model, lenders on Prosper placed bids on the interest rates, and thus did not observe the final interest rate until the loan was issued.[66] Both borrowers and lenders were and had always been price takers on Lending Club. Second, Prosper's address information was at city level whereas Lending Club was at county level. Third, as aforementioned, the FICO scores on Prosper and Lending Club came from different agencies. More than what was observable on Lending Club, a lender can observe if the borrower was a repeated borrower on Prosper.[67] Regardless of the differences between

[65]'P2P lender Prosper is back and better than ever', AOL Finance

[66]In Dec, 2010, Prosper got rid of the auction business model and switched to the posting-interest-rate business model as Lending Club, and the listing expiration for a loan increased from 7 days to 14 days. This event was studied by Wei and Lin (2016).

[67]Later on, Lending Club also added this feature.

them, preliminary results show that Prosper's market re-entry tightens the competition with Lending Club.

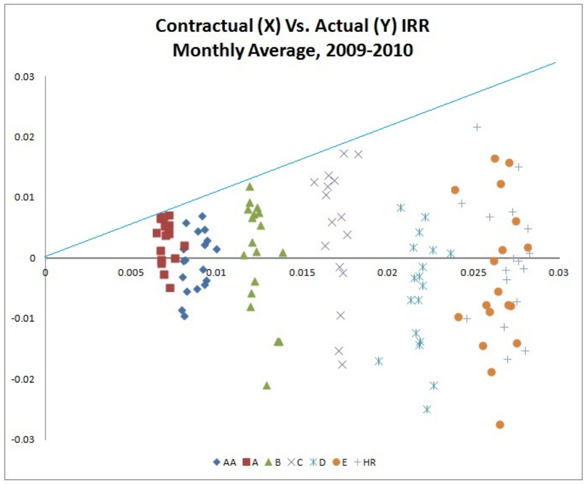


Figure 25: Prosper 2.0: IRR for Loans Vintaged from 2009-2010

Appendix B Tables

Table 9: Variable Definition

Variable Name	Variable Definition
Loan Contract and Funding Information	
Interest Rate IRR %	The interest rate of the loan measured by IRR, in %.
Loan Size (1K)	The principal borrowed and underwritten in the loan contract, in $1K.
Funding Duration	The number of days gapped between the loan application date and loan issuance date.
Platform Pct %	The percentage of Loan Size funded by the platform
Number of Lenders	The number of lenders that fund the loan
Mortgage Rate	The mortgage rate in the month when the loan originates
AAA Yield	AAA bond yield in the month when the loan originates
S&P Return	Monthly S&P return in the month when the loan originates
Borrower Attributes and Borrower Credit History	
Requested Size (1K)	The principal (Loan Size) the borrower intends to borrow
DTI	The borrower's debt-to-income ratio. The debt excludes the current loan and any mortgages.
Employment Length	The length of the borrower's current employment in years.
FICO	The borrower's low FICO score at the time of application
Revolving Utilization	The borrower's revolving utilization in percentage.
Revolving Balance (1K)	The borrower's revolving balance in $1K.
Open Account	The borrower's number of open accounts.
Total Account	The borrower's total accounts since the first credit line.
Public Record	The borrower's number of public records
Annual Income (1K)	The borrower's annual income in $1K.
Income Verified	A dummy variable to indicate if the borrower's income is verified by the employer.
Delinquency (Preceding 2yr)	The number of delinquencies the borrower has in the past 2 years
1(Borrower's Description)	A dummy variable to indicate if the borrower has put up description to the lenders.
Borrower Experience (Month)	The number of month gapped between the loan application date and the borrower's first credit line
Delinquency Forbearance	The market delinquency rate at borrower's first credit line
Performance	
Nonpayment%	The ratio of the loan's nonpayment and its contractual payment (installment*loan term) in %
IRR in Return%	The IRR % of the loan's payment stream.
Default	A dummy variable to indicate if the loan is charged off.
ROI %	Total Payment on the loan/Loan Size -1, in %

Date	Mean(Lenders)	Sd(Lenders)	Mean(Loan Size)	Sd(Loan Size)
2009m1	87.58	41.08	6030.95	3088.14
2009m2	93.83	47.51	6246.40	3406.48
2009m3	77.14	47.08	8411.57	4617.92
2009m4	97.81	56.46	6117.07	3463.34
2009m5	109.30	60.02	6614.41	3631.93
2009m6	123.03	59.61	7310.30	4097.97
2009m7	133.75	61.02	8343.60	4395.08
2009m8	131.38	67.04	9094.19	5155.95
2009m9	129.74	63.83	10602.45	6081.95
2009m10	123.86	63.99	10238.30	6483.34
2009m11	150.67	77.62	10053.49	6216.13
2009m12	149.59	81.54	10722.28	6819.21
2010m1	150.91	76.70	11018.30	6428.73

Table 10: Pre- vs Post-Entry Lender Participation and Loan Size on Lending Club

This table presents the means and standard deviations for Number of Lenders and Loan Size grouped by loans originated within a month. Note that the number of lender per loan does not significantly improve due to the entry event, whereas its standard deviation trends upward. Average loan size increases more than 50% along with the standard deviation.

Pre-Entry	N	Mean	SD	Min	Max	Median
FICO	33031	580.41	172.51	0.00	828.00	627.00
DTI	37872	3.48	17.52	0.00	100.00	0.13
Requested Size ($1K)	37872	9.82	7.12	1.00	35.00	8.00
Employment Length	37872	3.16	3.41	0.00	10.00	2.00
Post-Entry						
FICO	28410	610.20	174.19	0.00	825.00	663.00
DTI	30006	3.76	18.29	0.00	100.00	0.18
Requested Size ($1K)	30006	11.32	8.11	1.00	30.00	10.00
Employment Length	30006	3.62	3.53	0.00	10.00	2.00

Table 11: Summary Statistics: Pre- vs Post-Entry Applicant Borrower Characteristics

In this table, I show summary statistics on observable applicant characteristics before and after the entry. Note that, not every applicant has a FICO score. Second, the number of applicant drop from 37K to 30K, possibly due to platform competition.

Pre-Entry	N	Mean	SD	Min	Max	Median
FICO	4920	575.19	174.22	0.00	824.00	620.00
DTI	5685	2.20	13.53	-0.01	100.00	0.14
Requested Amount ($1K)	5685	9.83	7.37	1.00	25.00	8.00
Employment Length	5685	3.24	3.45	0.00	10.00	2.00
Post-Entry						
FICO	4304	596.19	168.53	0.00	820.00	646.00
DTI	4700	2.26	13.78	-0.01	100.00	0.15
Requested Amount ($1K)	4700	9.48	6.56	1.00	25.00	8.00
Employment Length	4700	3.62	3.51	0.00	10.00	2.00

Table 12: Summary Statistics 60-day Window: Pre- vs Post-Entry Applicant Borrower Characteristics

In this table, I show the summary statistics for loan applicants within 30-day windows of the entry event, compared to the 6-month windows in the previous table.

Variable	N	Mean	SD	Min	Max	Median	N	Mean	SD	Min	Max	Median
Loan Contract and Funding Information												
Interest Rate IRR %	2431	1.04	0.21	0.50	1.68	1.04	3541	1.03	0.23	0.50	1.77	1.04
Loan Size (1K)	2431	9.25	5.51	1.00	25.00	8.00	3541	10.56	6.46	1.00	25.00	9.60
Funding Duration	2431	10.52	3.82	1.00	27.00	11.00	3541	8.55	4.06	1.00	30.00	8.00
Platform Pct	2431	16.17	22.21	0.00	95.71	4.25	3541	1.57	3.92	0.00	92.50	0.57
Number of Lenders	2431	104.61	57.91	3.00	402.00	95.00	3541	139.57	73.10	6.00	461.00	132.00
Mortgage Rate	2431	4.97	0.16	4.80	5.31	5.00	3541	4.41	0.14	4.27	4.68	4.33
AAA Yield	2431	5.42	0.16	5.05	5.61	5.41	3541	5.21	0.06	5.13	5.26	5.26
S&P Return	2431	0.04	0.06	-0.11	0.09	0.05	3541	0.01	0.03	-0.04	0.06	0.03
Borrower Attributes and Borrower Credit History												
Requested Size (1K)	2431	9.30	5.53	1.00	25.00	8.00	3541	10.56	6.46	1.00	25.00	9.60
DTI	2431	0.12	0.07	0.00	0.25	0.12	3541	0.13	0.07	0.00	0.25	0.13
Employment Length	2431	3.96	3.46	0.00	10.00	3.00	3541	4.17	3.57	0.00	10.00	3.00
FICO	2431	713.12	35.44	660.00	820.00	710.00	3541	716.59	36.75	660.00	820.00	710.00
Revolving Utilization	2431	0.45	0.29	0.00	1.00	0.44	3541	0.46	0.29	0.00	1.00	0.45
Revolving Balance (1K)	2431	16.13	33.87	0.00	95.20	8.76	3541	16.95	30.02	0.00	60.25	8.69
Open Account	2431	9.35	4.31	3.00	35.00	9.00	3541	9.19	4.64	2.00	44.00	8.00
Total Account	2431	20.98	11.33	4.00	69.00	19.00	3541	21.67	11.97	3.00	73.00	20.00
Public Record	2431	0.05	0.24	0.00	3.00	0.00	3541	0.05	0.22	0.00	3.00	0.00
Annual Income (1K)	2431	65.26	46.95	4.20	700.00	55.00	3541	72.43	69.22	4.00	1440.00	60.00
Income Verified	2431	1.55	0.50	1.00	2.00	2.00	3541	1.24	0.43	1.00	2.00	1.00
Credit Policy Meet	2431	0.88	0.33	0.00	1.00	1.00	3541	0.90	0.30	0.00	1.00	1.00
Delinquency (Preceding 2yr)	2431	0.13	0.44	0.00	5.00	0.00	3541	0.14	0.46	0.00	5.00	0.00
1(Borrower's Description)	2431	1.00	0.02	0.00	1.00	1.00	3541	0.93	0.25	0.00	1.00	1.00
Borrower Experience (Month)	2431	149.25	81.62	17.00	507.00	133.00	3541	158.86	81.60	36.00	547.00	143.00
Delinquency Forbearance	2222	3.43	0.34	2.69	4.21	3.52	3204	3.43	0.34	2.69	4.21	3.53
Performance												
Nonpayment%	2431	7.83	21.89	-2.78	100.00	0.00	3541	7.23	21.30	-11.88	100.00	0.00
IRR in Return%	2431	-1.15	10.84	-100.00	6.34	1.02	3533	-1.01	10.84	-100.00	6.21	1.01
Default	2431	0.14	0.35	0.00	1.00	0.00	3541	0.13	0.34	0.00	1.00	0.00
ROI %	2431	8.36	25.93	-100.00	38.17	16.355	3541	8.57	25.25	-100.00	42.63	15.83

Table 13: Summary Statistics: Pre- vs Post-Entry Accepted Borrower Characteristics

Note that borrowers' state address dummies and homeownership dummies are not included. Homeownership includes 4 categories: own, rent, mortgage and others.

Table 14: Summary Statistics: Financed Borrower Characteristics across Platforms

	Incumbent Platform						Entrant Platform					
	N	Mean	SD	Min	Max	Median	N	Mean	SD	Min	Max	Median
Number of Funders	3541	139.57	73.10	6.00	461.00	132.00	2359	149.54	154.13	1.00	1189	95
Interest Rate IRR %	3541	1.03	0.23	0.50	1.77	1.04	2359	1.62	0.77	0.35	2.92	1.58
Loan Size (1K)	3541	10.56	6.46	1.00	25.00	9.60	2359	4.51	4.18	1.00	25.00	3.00
Annual Income (1K)	3541	72.43	69.22	4.00	1440.00	60.00	2359	61.77	39.22	0.00	425.00	54.00
(Requested) Loan Size (1K)	3541	10.56	6.46	1.00	25.00	9.60	2359	4.51	4.18	1.00	25.00	3.00
DTI	3541	0.13	0.07	0.00	0.25	0.13	2359	9.96	29.60	0.01	100.00	0.23
Employment Length	3533	4.17	3.57	0.00	10.00	3.00	2359	5.07	3.62	0.00	10.00	4.00
FICO Score (High)	3541	720.59	36.75	664.00	824.00	714.00	2359	725.64	62.09	619.00	820.00	723.00
Total Account	3541	21.67	11.97	3.00	73.00	20.00	2359	6.64	4.38	0.00	31.00	6.00
Public Record	3541	0.05	0.22	0.00	3.00	0.00	2359	0.19	0.57	0.00	10.00	0.00
Open Account	3541	9.19	4.64	2.00	44.00	8.00	2359	8.79	4.81	0.00	33.00	8.00
Revolving Utilization	3534	0.46	0.29	0.00	1.00	0.45	2359	0.50	0.31	0.00	1.00	0.49
Default	3541	0.00	0.00	0.00	0.00	0.00	2359	0.16	0.37	0.00	1.00	0.00
IRR in Return %	3533	-1.01	10.84	-100.00	6.21	1.01	2358	-0.81	11.83	-100.00	8.96	1.33
ROI %	3541	8.57	25.25	-100.00	42.63	15.83	2359	15.70	31.61	-100.00	143.26	16.93

	LPM			Probit	RD	
	(1)	(2)	(3) 60-day	(4)	(5) Pre-Entry	(6) Post-Entry
main						
Entry	0.0384***	0.0251***	0.0119**	0.2592***		
	(0.0062)	(0.0075)	(0.0054)	(0.0581)		
FICO	0.0004***	0.0004***	0.0003***	0.0134***	0.0112***	0.0118***
	(0.0000)	(0.0000)	(0.0000)	(0.0002)	(0.0002)	(0.0002)
DTI	−0.0010***	−0.0010***	−0.0009***	−3.3619***	−2.0131***	−3.0960***
	(0.0001)	(0.0001)	(0.0002)	(0.0714)	(0.0690)	(0.0726)
Requested Size (1K)	−0.0035***	−0.0035***	−0.0035***	−0.0427***	−0.0346***	−0.0273***
	(0.0001)	(0.0001)	(0.0004)	(0.0015)	(0.0020)	(0.0015)
Employment Length	0.0031***	0.0030***	0.0018**	0.0277***	0.0195***	0.0322***
	(0.0003)	(0.0003)	(0.0008)	(0.0028)	(0.0036)	(0.0032)
Mortgage Rate	0.0692***	−0.0345*		0.3427***		
	(0.0085)	(0.0188)		(0.0752)		
AAA Bond Yield	0.0285**	−0.0621**		0.0806		
	(0.0111)	(0.0244)		(0.1071)		
S&P Return	0.0321	0.1658***		0.1107		
	(0.0277)	(0.0386)		(0.2442)		
FICO_M1		−0.0021***				
		(0.0004)				
FICO_M2		−0.0006***				
		(0.0002)				
DTI_M1		0.3643**				
		(0.1588)				
DTI_M2		0.0025**				
		(0.0011)				
LOANSIZE_M1		0.0000***				
		(0.0000)				
LOANSIZE_M2		−0.0000***				
		(0.0000)				
Observations	61441	61441	9224	61439	33031	28408
R^2	0.178	0.180	0.290			
Time FE	−	−	−	−	*Yes*	*Yes*
Monthly Controls	*Yes*	*Yes*	*Yes*	*Yes*	−	−

Standard errors in parentheses
Suppressed Variable: loan purpose FE and state FE
Standard Errors are Robust
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Entry and Borrower Screening

This table shows estimation results corresponding to section "Borrower Screening". Entry is a dummy variable to indicate if a loan application happens prior to or after the entry. In the Linear Probability Model specification, I regress the dummy variable "accept" on all the observable controls and the "Entry" dummy. Column (1) includes borrower observables and monthly macroeconomic controls using 6-month windows on each side of the entry event. Column (2) subsumes additional controls including first two moments of the applicants' characteristics. Column (3) reduces the window size to 30-days on each side, and thus excludes monthly controls and fixed effects. Column (4) uses a Probit approach to test the models' nonlinear robustness. Column (5) and (6) correspond to the models for the Sharp Regression Discontinuity design, where separate regressions were estimated with monthly fixed effects.

	OLS		Ordered Logit			
	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled	Pooled	Pre-Entry	Post-Entry	Pooled	Pooled
main						
Entry	-0.1238^{**}	-0.5439^{**}	-0.3676^{**}	-1.3513^{**}		
	(0.0600)	(0.2358)	(0.1504)	(0.6015)		
Requested Size (1K)	0.0615^{***}	0.0624^{***}	0.1968^{***}	0.2017^{***}	0.2072^{***}	0.2131^{***}
	(0.0019)	(0.0020)	(0.0057)	(0.0060)	(0.0106)	(0.0083)
DTI	-1.6896^{***}	-1.6620^{***}	-3.4470^{***}	-3.5128^{***}	-3.7677^{***}	-3.9315^{***}
	(0.1916)	(0.2007)	(0.4855)	(0.5149)	(0.8060)	(0.6850)
FICO	-0.0275^{***}	-0.0279^{***}	-0.0870^{***}	-0.0896^{***}	-0.0997^{***}	-0.0927^{***}
	(0.0004)	(0.0005)	(0.0019)	(0.0021)	(0.0035)	(0.0028)
Employment Length	-0.0040	-0.0001	-0.0121	-0.0040	0.0031	-0.0041
	(0.0031)	(0.0033)	(0.0079)	(0.0085)	(0.0139)	(0.0112)
Revolving Utilization	0.1397^{**}	0.1546^{***}	0.4445^{***}	0.5005^{***}	0.0665	0.8564^{***}
	(0.0550)	(0.0578)	(0.1287)	(0.1382)	(0.2128)	(0.1886)
Revolving Balance (1K)	-0.0000	-0.0001	-0.0002	-0.0000	0.0015	-0.0012
	(0.0004)	(0.0005)	(0.0010)	(0.0011)	(0.0019)	(0.0015)
Annual Income (1K)	0.0002	0.0003	0.0000	0.0003	-0.0002	0.0006
	(0.0002)	(0.0002)	(0.0006)	(0.0006)	(0.0013)	(0.0008)
Delinquency (Preceding 2yr)	0.0044	0.0070	-0.0181	0.0084	-0.1353	0.1167
	(0.0234)	(0.0252)	(0.0463)	(0.0524)	(0.0847)	(0.0734)
Borrower Experience	-0.0001	-0.0009^{***}	-0.0009^{**}	-0.0034^{***}	-0.0037^{***}	-0.0028^{***}
	(0.0002)	(0.0002)	(0.0004)	(0.0006)	(0.0010)	(0.0008)
Delinquency Forbearance		-0.2819^{***}		-0.7280^{***}	-0.5635^{***}	-0.5047^{***}
		(0.0523)		(0.1338)	(0.1281)	(0.1121)
Entry=1 × Delinquency Forbearance		0.1334^{**}		0.3152^{*}		
		(0.0660)		(0.1682)		
Open Accounts	-0.0072^{*}	-0.0151^{***}	-0.0454^{***}	-0.0620^{***}	-0.0846^{***}	-0.0525^{***}
	(0.0044)	(0.0044)	(0.0105)	(0.0110)	(0.0166)	(0.0154)
Total Accounts	-0.0037^{***}	-0.0009	-0.0099^{***}	-0.0024	-0.0001	-0.0039
	(0.0013)	(0.0014)	(0.0034)	(0.0037)	(0.0058)	(0.0047)
Public Record	-0.0020	0.0085	0.0186	0.0730	0.1213	-0.0369
	(0.0481)	(0.0506)	(0.1021)	(0.1111)	(0.1761)	(0.1613)
Observations	5972	5426	5972	5426	2222	3204
R^2	0.655	0.662				
Time FE	–	–	–	–	Yes	Yes
Monthly Controls	Yes	Yes	Yes	Yes	–	–

Standard errors in parentheses
Suppressed Variable: monthly controls, loan purpose FE, homeownership FE, state FE
Standard errors are clustered at Loan Grade
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 16: Entry and Loan Classfication

This table presents estimation results associated with section "Loan Classification" (Credit Rating). Entry is a dummy variable to indicate if a loan application happens after the entry. I regress loan ratings (1-7) on observable borrower characteristics, credit history, loan contract terms and "Delinquency Forbearance" that measures borrowers' quality unobservable to econometricians and lenders. The first 2 columns are under the OLS specification. The last 4 are estimated by the Order Logistic Regression. For specifications where I include Delinquency Forbearance, (2) and (4), I lose 546 observations due to missing data. Column (5) and (6) represent regressions separated by the entry event to yield the result for regression discontinuity.

Grade	Pre-Entry Probability	SE	Z-Score	Post-Entry Probability	SE	Z-Score	Difference	Z
A	0.212	0.007	30.610	0.238	0.005	44.830	−0.026	−2.979
B	0.274	0.005	50.620	0.287	0.005	55.030	−0.013	−1.738
C	0.257	0.005	54.660	0.253	0.005	55.010	0.003	0.515
D	0.160	0.005	29.680	0.144	0.005	31.410	0.017	2.336
E	0.063	0.004	16.000	0.052	0.003	17.770	0.011	2.284
F	0.022	0.002	10.050	0.017	0.002	10.860	0.005	1.790
G	0.012	0.002	7.510	0.009	0.001	7.740	0.003	1.566

Table 17: Ordered Logit: Predicted Marginal Effects

The marginal effect estimates corresponds to the specification in column (5) in Table 16. First, it estimates the conditional probabilities for a borrower to be classified into each rating prior to and after the event. Then, using T-tests, I compare those probabilities within each rating. Borrowers are 2.6% and 1.3% relatively more likely to rated into the creditworthy categories: A and B.

Grade	Post-Entry Interest Rate	$\widehat{\text{Pre-Entry}}$ Interest Rate	T-Test	Observations
A	0.695	0.805	−28.370	670
B	0.980	0.996	−4.998	1,008
C	1.125	1.142	−5.658	796
D	1.268	1.210	11.514	484
E	1.408	1.257	18.789	160
F	1.563	1.304	16.671	60
G	1.710	1.336	18.674	26

Table 18: Risk Adjusted Average Interest Rate Comparison

Using a out-of-sample validation approach, I estimate a linear model on interest rates with borrowers before the entry point, and conduct out-of-sample predictions on those after the entry, denoted by "pseudo interest rate". That is, the "pseudo interest rate" is the forecasted interest rates under the pre-entry pricing mechanism. I pairwise compare post-entry borrowers' interest rates and "pseudo interest rates". This pairwise comparison approach does not take into account of the discrepancy of the credit rating mechanisms or borrowers' characteristic distributions in the forecast and estimation samples. Therefore, the magnitude of the estimates should not be as precisely interpreted as the QTE result. (See Table 5)

Discussion: It shows that had a post-entry A-rated borrower obtained her loan during the pre-entry period, her monthly interest rate (IRR) would increase from 0.695% to 0.802%. However, for a G-rated borrower, there is a 0.4% interest rate increase.

	Loan Ratings			Interest Rates		
	(1) Number of Lenders	(2) Fuding Duration	(3) Pct Platform	(4) Number of Lenders	(5) Fuding Duration	(6) Pct Platform
A × Entry=1	−27.0023***	−1.0096***	−3.0279*			
	(2.9582)	(0.1519)	(1.4927)			
B × Entry=1	5.0439	−3.0667***	−14.9770***			
	(5.1279)	(0.1411)	(1.7585)			
C × Entry=1	25.1195***	−2.5499***	−20.4363***			
	(3.8611)	(0.1279)	(1.5226)			
D × Entry=1	21.6779***	−1.5896***	−20.2459***			
	(4.6301)	(0.1639)	(1.5305)			
E × Entry=1	25.6994***	−1.0438***	−18.7650***			
	(4.5762)	(0.1860)	(1.3011)			
F × Entry=1	30.7652***	−0.0067	−19.3227***			
	(3.5548)	(0.2074)	(1.4248)			
G × Entry=1	−5.6355	1.1505***	−15.9756***			
	(7.4381)	(0.1628)	(2.2927)			
Entry				−70.4080**	−2.2085	11.3410
				(25.9171)	(2.5346)	(8.2627)
Interest Rate (IRR%)				−41.5161	1.1767	12.7099
				(29.4392)	(2.5199)	(7.6731)
Entry=1 × Interest Rate (IRR%)				75.7666**	0.2002	−25.1803**
				(21.3736)	(2.1967)	(7.1428)
Requested Size (1K)	7.0397***	0.2917***	0.7494***	6.6764***	0.3011***	0.8134***
	(1.0412)	(0.0116)	(0.1226)	(0.9266)	(0.0139)	(0.1170)
DTI	−45.1795*	2.9256**	2.5404	−42.5564	3.0441**	2.4804
	(20.9139)	(0.9878)	(4.3124)	(22.2333)	(1.0044)	(4.7092)
Employment Length	0.0783	−0.0472**	−0.1858***	0.0890	−0.0476**	−0.1876***
	(0.3595)	(0.0156)	(0.0366)	(0.3694)	(0.0156)	(0.0402)
FICO	0.1181	−0.0091*	−0.0220	0.2953***	−0.0140**	−0.0492***
	(0.0656)	(0.0041)	(0.0126)	(0.0457)	(0.0055)	(0.0117)
Revolving Utilization	8.3429***	−0.0884	−0.0019	6.1693	0.0328	0.4909
	(2.2314)	(0.2105)	(0.7473)	(3.3756)	(0.1768)	(0.8207)
Revolving Balance (1K)	−0.0170	0.0033*	0.0053	0.0013	0.0022	0.0010
	(0.0328)	(0.0017)	(0.0033)	(0.0328)	(0.0019)	(0.0036)
Open Accounts	0.1496	0.0043	0.0457	0.2281	0.0009	0.0338
	(0.0995)	(0.0089)	(0.0604)	(0.1213)	(0.0106)	(0.0567)
Total Accounts	−0.0133	−0.0123***	0.0007	−0.0106	−0.0113**	0.0013
	(0.0888)	(0.0025)	(0.0197)	(0.0868)	(0.0031)	(0.0203)
Public Record	−1.3844	−0.1529	0.8098	−0.5873	−0.1682	0.7811
	(2.2550)	(0.1548)	(0.5064)	(2.4123)	(0.1610)	(0.5008)
Not Verified × Annual Income (1K)	−0.0565**	−0.0015	−0.0033	−0.0533**	−0.0017	−0.0039
	(0.0172)	(0.0014)	(0.0042)	(0.0164)	(0.0014)	(0.0042)
Verified × Annual Income (1K)	−0.0563***	0.0002	0.0018	−0.0522**	−0.0000	0.0016
	(0.0122)	(0.0022)	(0.0093)	(0.0148)	(0.0021)	(0.0090)
Delinquency (Preceding 2yr)	−6.0259**	0.2593*	1.4161***	−6.1250***	0.2769	1.4396***
	(1.6320)	(0.1331)	(0.2937)	(1.4887)	(0.1434)	(0.2882)
1(Borrower's Description)	22.4255**	0.6674	−0.9570	22.7387**	0.6691	−1.0262
	(8.4652)	(0.4437)	(0.5308)	(8.6005)	(0.4406)	(0.6078)
Observations	5972	5972	5972	5972	5972	5972
R^2	0.532	0.352	0.406	0.510	0.326	0.386
Monthly Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Loan Grade FE	*Yes*	*Yes*	*Yes*	–	–	–

Standard errors in parentheses are clustered at Loan Grade

Suppressed Variable: loan purpose, homeownership, state FE, monthly controls

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 19: Entry and Lenders' Responses

In this table, I show regression results corresponding to the section "Lenders' Response". I measure lenders' response with 3 variables, Number of Lenders, (1) & (4), Funding Duration (2) & (5) and Percentage of loan size lent by the Platform, (3) & (6). In the first three columns, the results show lenders' responses with respect to entry in within each loan rating. Within each specification, I interact Entry dummy with loan grades, including loan grades as independent variables along with all the observable borrower characteristics. Note that, here I do not include "Delinquency Forbearance", since it is not observable to lenders. In the last three columns, I interact interest rates with Entry dummy to observe the heterogeneity of responses at different interest rates.

	Probit	OLS				OLS Interact Grade			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Default	Default	% Nonpayment	Return (IRR%)	ROI%	Default	% Nonpayment	Return (IRR%)	ROI%
main									
Entry	0.3061***	0.0582***	3.9551***	−2.2444***	−6.7307***				
	(0.0331)	(0.0123)	(0.6038)	(0.3845)	(0.8941)				
Entry=1 × A						0.0554***	4.2425***	−2.2241***	−6.8655***
						(0.0100)	(0.6174)	(0.3382)	(0.8411)
Entry=1 × B						0.0568***	4.2880***	−2.3872***	−6.8418***
						(0.0116)	(0.6423)	(0.3070)	(0.8551)
Entry=1 × C						0.0649***	3.3971***	−1.5382***	−6.2873***
						(0.0142)	(0.7459)	(0.3655)	(0.9116)
Entry=1 × D						0.0695***	4.7260***	−3.0916***	−7.8274***
						(0.0128)	(0.7904)	(0.5226)	(0.9947)
Entry=1 × E						0.0605***	4.0834***	−3.0519***	−7.5977***
						(0.0145)	(0.9442)	(0.6022)	(1.1042)
Entry=1 × F						0.0423	0.4768	−0.5121	−3.0182
						(0.0234)	(1.8131)	(1.1821)	(2.0810)
Entry=1 × G						−0.0920**	−3.3782	−2.0208	2.7803
						(0.0278)	(1.9231)	(1.1827)	(2.2164)
Interest Rate (IRR%)	0.8199***	0.1407**	5.9570*	−0.3281	9.7880***	0.0647	6.3022	−0.5850	10.6959**
	(0.2669)	(0.0485)	(2.6179)	(1.2678)	(2.5943)	(0.0897)	(5.0076)	(1.8347)	(4.2210)
Loan Size (1K)	0.0137**	0.0020*	0.1369**	−0.0329	−0.1515**	0.0022**	0.1404**	−0.0226	−0.1599**
	(0.0069)	(0.0009)	(0.0423)	(0.0337)	(0.0429)	(0.0009)	(0.0453)	(0.0348)	(0.0455)
DTI	0.8229**	0.1408**	7.3990**	−2.3420	−8.2757**	0.1401**	7.1261**	−2.1872	−8.0084*
	(0.3230)	(0.0512)	(2.6366)	(2.0688)	(3.2025)	(0.0474)	(2.6907)	(2.1125)	(3.2942)
Revolving Utilization	0.2271	0.0501*	4.2882**	−1.2301	−4.1346**	0.0530*	4.3737**	−1.1797	−4.2583**
	(0.1396)	(0.0237)	(1.2400)	(0.6717)	(1.3581)	(0.0235)	(1.2248)	(0.6951)	(1.3392)
Revolving Balance (1K)	0.0006	0.0000	0.0043	−0.0038	−0.0047	−0.0000	0.0026	−0.0043	−0.0024
	(0.0011)	(0.0002)	(0.0116)	(0.0063)	(0.0126)	(0.0002)	(0.0122)	(0.0061)	(0.0134)
Not Verified × Annual Income (1K)	−0.0021**	−0.0002**	−0.0127*	0.0003	0.0127*	−0.0002**	−0.0128*	0.0002	0.0130*
	(0.0010)	(0.0001)	(0.0054)	(0.0029)	(0.0054)	(0.0001)	(0.0054)	(0.0029)	(0.0054)
Verified × Annual Income (1K)	−0.0021*	−0.0003	−0.0188	0.0021	0.0209	−0.0003*	−0.0189	0.0017	0.0211*
	(0.0011)	(0.0001)	(0.0099)	(0.0068)	(0.0108)	(0.0001)	(0.0098)	(0.0068)	(0.0106)
Total Accounts	−0.0169***	−0.0030**	−0.1890***	0.0580***	0.1989***	−0.0031**	−0.1893***	0.0581***	0.1996***
	(0.0039)	(0.0009)	(0.0457)	(0.0132)	(0.0489)	(0.0009)	(0.0454)	(0.0134)	(0.0480)
Public Record	0.1867**	0.0361	1.1848	0.4508	−1.4448	0.0355	1.1359	0.4578	−1.3731
	(0.0887)	(0.0201)	(0.8455)	(0.2540)	(0.9500)	(0.0196)	(0.8330)	(0.2526)	(0.9304)
Open Accounts	0.0183	0.0039	0.3051*	−0.0842**	−0.3261*	0.0040	0.3088*	−0.0854**	−0.3303*
	(0.0150)	(0.0027)	(0.1480)	(0.0331)	(0.1601)	(0.0026)	(0.1464)	(0.0332)	(0.1580)
Employment Length	0.0077	0.0011	0.0137	0.0450	−0.0199	0.0011	0.0149	0.0430	−0.0212
	(0.0084)	(0.0016)	(0.0880)	(0.0468)	(0.1013)	(0.0016)	(0.0882)	(0.0471)	(0.1017)
FICO	−0.0017	−0.0001	−0.0103	0.0028	0.0092	−0.0002	−0.0130	0.0007	0.0143
	(0.0028)	(0.0005)	(0.0231)	(0.0070)	(0.0238)	(0.0005)	(0.0252)	(0.0075)	(0.0257)
Delinquency Forbearance	0.0742	0.0181	1.5792	−0.5469*	−1.6969	0.0177	1.5691	−0.5451*	−1.6729
	(0.1026)	(0.0171)	(0.9397)	(0.2497)	(1.0951)	(0.0171)	(0.9340)	(0.2466)	(1.0843)
Borrower Experience	0.0020***	0.0003***	0.0165**	−0.0025	−0.0167**	0.0003***	0.0167**	−0.0024	−0.0169**
	(0.0004)	(0.0001)	(0.0045)	(0.0016)	(0.0048)	(0.0001)	(0.0045)	(0.0015)	(0.0048)
Observations	5186	5405	5405	5405	5405	5405	5405	5405	5405
R^2		0.199	0.276	0.281	0.415	0.201	0.277	0.283	0.415
Loan Grade FE	–	–	–	–	–	*Yes*	*Yes*	*Yes*	*Yes*

Standard errors in parentheses
Suppressed Variable: loan purpose, homeownership, Grade FE, state FE, ex post time FE
Standard errors are clustered at Loan Grade
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 20: Entry and Loan Performance

This table shows the estimation results on the effect of Entry on loan performance. Loan performance is measured by Default, Percentage Nonpayment, ROI and IRR. In the first 5 columns, I regress the loan performance measure on the entry dummy, observable borrower characteristics, loan contract terms and "Delinquency Forbearance" (unobservable to lenders). Column 1 is a Probit specification on Default dummy. Column 2 — 5 are OLS estimates. In the last 4 columns, I interact Entry dummy with loan grades, in order to discover the heterogeneity of loan performance changes for different ratings.

	Number of Lenders		Platform's Percentage	
	(1) Nonpayment	(2) Default	(3) Nonpayment	(4) Default
Entry=0 × Realized Default		−61.4133***		21.4929***
		(10.5207)		(1.4755)
Entry=1 × Realized Default		4.2930		−2.8838
		(7.6402)		(1.8006)
Entry=0 × Realized Nonpayment	−0.7123***		0.2793***	
	(0.1719)		(0.0190)	
Entry=1 × Realized Nonpayment	0.2341		−0.0625	
	(0.1562)		(0.0346)	
Requested Size (1K)	6.9206***	6.9248***	0.7984***	0.7939***
	(0.8639)	(0.8680)	(0.1126)	(0.1156)
DTI	−50.9515*	−50.3739*	3.3555	3.0309
	(22.5920)	(22.6854)	(3.0871)	(3.0746)
FICO	0.1936***	0.1934***	−0.0478***	−0.0482***
	(0.0137)	(0.0144)	(0.0086)	(0.0096)
Revolving Utilization	7.5374**	7.5240**	−0.2005	−0.1866
	(2.7771)	(2.7874)	(0.7264)	(0.7637)
Revolving Balance (1K)	−0.0117	−0.0129	0.0060*	0.0063*
	(0.0342)	(0.0349)	(0.0028)	(0.0031)
Delinquency (Preceding 2yr)	−6.4083**	−6.3249**	1.2623***	1.2511***
	(1.8162)	(1.8335)	(0.2229)	(0.2119)
Interest Rate (IRR%)	108.2534	103.4825	−7.2503**	−8.7522***
	(73.1969)	(72.5751)	(2.7204)	(2.3564)
Observations	5962	5962	5962	5962
R^2	0.552	0.554	0.448	0.451
Time FE	Yes	Yes	Yes	Yes
Loan Grade FE	Yes	Yes	Yes	Yes

Standard errors in parentheses
Suppressed Variable: loan purpose, homeownership, time FE, state FE
Standard errors are clustered at Loan Grade
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 21: Entry and Lender's Punishment

The estimation results in this table are corresponding to the section on "Lenders' Punishment". I use two dependent variables to measure lenders' reaction, number of lenders per loan and percentage of the loan size lent by the platform. Remember, by definition, for a loan at origination, its Realized Default measures the number of default occurrence over the number of loans maturing during the current month within its rating. Its Realized Nonpayment measures the average percentage nonpayment during the current month within its rating. This construction gives me variations within a month and within a rating, and thus, both Monthly FE and Grade FE at loan origination are included.

	Linear Probability Model		
	(1)	(2)	(3)
	All	Overlapped	Non-overlapped
FICO Score	0.144***	0.142***	0.101***
	(0.0002)	(0.0001)	(0.0001)
Requested Size (1K)	−0.230***	−0.301***	−0.227***
	(0.0017)	(0.0051)	(0.0012)
Annual Income(1K)	0.060***	0.214***	−0.026
	(0.0001)	(0.0005)	(0.0001)
DTI	0.110**	0.145***	0.089*
	(0.0011)	(0.1773)	(0.0006)
Total Accounts	−0.684***	−0.613***	−0.706***
	(0.0042)	(0.0045)	(0.0023)
Public Record	0.067**	0.076***	0.035
	(0.0272)	(0.0197)	(0.0171)
Open Accounts	0.377***	0.225***	0.458***
	(0.0019)	(0.0032)	(0.0040)
Revolving Utilization	0.062**	−0.010	0.064*
	(0.0405)	(0.0363)	(0.0472)
Observations	5900	3701	2199
R^2	0.632	0.538	0.728

Standardized beta coefficients; Standard errors in parentheses

Suppressed Variable: loan purpose FE, monthly FE, state FE

Standard Errors are clusterd at platform-month level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 22: Borrower Platform Choice

Corresponding to section "Competition and Market Segmentation", this table shows results that verify market segmantation between the platforms, where I regress borrowers' platform choices on their observable characteristics. The dependent variable is a dummy variable to indicate if the platform is the entrant. Remember, I rule out outliers in borrower characteristics across platforms by indicaing if a borrower's characteristics on one platform can be found on the other, i.e. overlap. Using Linear Probability Models, in column (1), I look at the whole sample without filtering borrower characteristics. In column (2), I focus on the subsample where overlap is 1; and in (3), I examine the complement sample where borrower characteristics do not overlap.

	Quantile Regression						
	(1) 5%	(2) 10%	(3) 25%	(4) 50%	(5) 75%	(6) 90%	(7) 95%
main							
Entrant	0.5166	1.4428	3.5545***	5.1735***	6.7262***	7.1115***	3.4909**
	(0.4909)	(0.9692)	(0.4697)	(0.5071)	(0.7577)	(1.2479)	(1.4845)
Entrant=0 × FICO Score	−0.0062***	−0.0061***	−0.0056***	−0.0051***	−0.0048***	−0.0047***	−0.0048***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)
Entrant=1 × FICO Score	−0.0069***	−0.0081***	−0.0102***	−0.0116***	−0.0129***	−0.0127***	−0.0077***
	(0.0007)	(0.0014)	(0.0006)	(0.0007)	(0.0010)	(0.0018)	(0.0022)
Entrant=0 × Requested Size (1K)	0.0136***	0.0134***	0.0121***	0.0112***	0.0108***	0.0114***	0.0088***
	(0.0006)	(0.0004)	(0.0004)	(0.0004)	(0.0006)	(0.0011)	(0.0014)
Entrant=1 × Requested Size (1K)	0.0338***	0.0374***	0.0400***	0.0552***	0.0604***	0.0930***	0.0300
	(0.0112)	(0.0085)	(0.0068)	(0.0110)	(0.0145)	(0.0193)	(0.0235)
Entrant=0 × Annual Income(1K)	−0.0001	−0.0001	−0.0000	−0.0000	0.0000	−0.0001	−0.0003
	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0002)	(0.0003)
Entrant=1 × Annual Income(1K)	−0.0009*	−0.0011	−0.0007	−0.0024***	−0.0013	−0.0026	0.0009
	(0.0006)	(0.0007)	(0.0008)	(0.0008)	(0.0013)	(0.0018)	(0.0018)
Entrant=0 × DTI	−0.1053**	−0.1371***	−0.1510***	−0.1499***	−0.1643***	−0.2723**	−0.2861*
	(0.0536)	(0.0438)	(0.0330)	(0.0282)	(0.0444)	(0.1255)	(0.1462)
Entrant=1 × DTI	−0.2154	−0.1279	0.1637	0.0005	0.1930	−0.2108	0.9583
	(0.3111)	(0.2974)	(0.5241)	(0.5223)	(0.4504)	(0.7976)	(1.0487)
Entrant=0 × Employment Length	−0.0007	−0.0008	−0.0005	−0.0005	−0.0010	−0.0016	−0.0007
	(0.0008)	(0.0006)	(0.0005)	(0.0006)	(0.0007)	(0.0012)	(0.0022)
Entrant=1 × Employment Length	0.0047	0.0055	0.0066	0.0072	0.0077	0.0117	0.0015
	(0.0074)	(0.0058)	(0.0054)	(0.0062)	(0.0114)	(0.0162)	(0.0161)
Entrant=0 × Total Accounts	−0.0006**	−0.0005***	−0.0004	−0.0006***	−0.0009**	−0.0008	−0.0013
	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0004)	(0.0005)	(0.0008)
Entrant=1 × Total Accounts	−0.0180	−0.0203*	−0.0133	−0.0288**	−0.0317*	−0.0582	−0.0141
	(0.0115)	(0.0110)	(0.0114)	(0.0139)	(0.0164)	(0.0400)	(0.0611)
Entrant=0 × Public Record	0.0030	0.0023	−0.0055	−0.0024	0.0097	0.0030	0.0299
	(0.0142)	(0.0062)	(0.0052)	(0.0067)	(0.0167)	(0.0200)	(0.0350)
Entrant=1 × Public Record	0.1151***	0.1062**	0.1260**	0.1366***	0.1762**	0.1549*	0.1988
	(0.0416)	(0.0494)	(0.0522)	(0.0376)	(0.0726)	(0.0855)	(0.1245)
Entrant=0 × Open Accounts	−0.0002	−0.0004	−0.0017***	−0.0022**	−0.0017	0.0003	0.0045
	(0.0008)	(0.0006)	(0.0005)	(0.0009)	(0.0011)	(0.0017)	(0.0031)
Entrant=1 × Open Accounts	0.0162	0.0151	−0.0054	0.0066	−0.0108	0.0077	−0.0195
	(0.0117)	(0.0100)	(0.0135)	(0.0161)	(0.0176)	(0.0338)	(0.0538)
Entrant=0 × Revolving Utilization	0.0326***	0.0273**	0.0285***	0.0324***	0.0384***	0.0331*	0.0269
	(0.0098)	(0.0110)	(0.0090)	(0.0082)	(0.0145)	(0.0180)	(0.0339)
Entrant=1 × Revolving Utilization	−0.0015	−0.0111	0.0219	0.0752	0.0836	−0.0571	−0.0592
	(0.0682)	(0.1149)	(0.1178)	(0.1376)	(0.1640)	(0.1723)	(0.1861)
Observations	3701	3701	3701	3701	3701	3701	3701

Standard errors in parentheses
Suppressed Variable: loan purpose FE, monthly FE, state FE
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 23: Platform Pricing Premium

This table shows the estimation results that compare interest rates between the two platforms. I use a Quantile Regression Model to evaluate pricing premium or discount between the platforms at different borrower creditworthiness. To account for the fact that pricing mechanisms may differ between the two, I interact the indicator Entrant with all the observable borrower characteristics. The point estimates on indicator Entrant indicate the premium (or discount) the entrant charges.
Note that, in this table, I only use the subsample where the borrower characteristics overlap between the platforms, i.e. overlap is 1. I estimate the equation at 5%, 10%, 25%, 50%, 75%, 90% and 95% percentile of the borrower interests within the subsample.

	Overlapped			All		
	(1) Default	(2) IRR	(3) ROI	(4) Default	(5) IRR	(6) ROI
Entrant	0.1402***	0.7145	7.0743***	0.1694***	0.5497	7.4123***
	(0.0130)	(0.5933)	(1.8093)	(0.0085)	(0.4053)	(0.8884)
FICO Score	−0.0003	0.0087	−0.0573**	−0.0007***	0.0090*	−0.0563***
	(0.0002)	(0.0061)	(0.0202)	(0.0002)	(0.0047)	(0.0134)
Loan Size(1K)	0.0004	−0.0304	−0.0812	0.0007	−0.0389	0.0257
	(0.0004)	(0.0444)	(0.1423)	(0.0006)	(0.0423)	(0.1029)
Annual Income(1K)	−0.0001	0.0108**	0.0358*	−0.0001	−0.0022	0.0050
	(0.0001)	(0.0046)	(0.0182)	(0.0000)	(0.0035)	(0.0046)
DTI	0.0211	5.2672	0.7133	−0.0000	−0.0096	−0.0069
	(0.0350)	(4.1755)	(11.2322)	(0.0001)	(0.0092)	(0.0158)
Employment Length	0.0023	0.0010	−0.0171	0.0028*	0.0045	0.0090
	(0.0014)	(0.0520)	(0.1435)	(0.0015)	(0.0528)	(0.0931)
Total Accounts	0.0003	0.0089	−0.0338	0.0004	0.0190	−0.0151
	(0.0002)	(0.0197)	(0.0493)	(0.0002)	(0.0155)	(0.0299)
Public Record	0.0162	0.9715	1.8884	−0.0033	0.8284*	2.5029***
	(0.0169)	(0.5422)	(1.0835)	(0.0135)	(0.4419)	(0.7673)
Open Accounts	−0.0014	0.0022	0.0743	−0.0019	0.0362	0.1287
	(0.0008)	(0.0412)	(0.0768)	(0.0013)	(0.0510)	(0.1110)
Revolving Utilization	−0.0185	−1.1250	−1.5973	−0.0138	0.1284	2.2287
	(0.0206)	(0.8676)	(2.2332)	(0.0118)	(0.6670)	(1.9564)
Observations	3701	3701	3701	5900	5900	5900
R^2	0.141	0.031	0.045	0.131	0.018	0.045

Standard errors in parentheses
Suppressed Variable: loan purpose FE, monthly FE, state FE
Standard Errors are clusterd at platform-month level
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 24: Platform Performance Premium

In this table, I compare loan performance between the two platforms, using 3 measures Default, IRR and ROI. The point estimates on Entrant dummy indicate the average underperformance or outperformance from the entrant in the measures. I estimate the 3 regressions using OLS under 2 different samples, for one, borrower with overlapped characteristics between the platform and for another, the whole sample.

	(1) Default	(2) Nonpayment	(3) ROI	(4) IRR
Entry=0 × Platform Pct%	0.0017***	0.1225***	−0.1746***	−0.0381**
	(0.0002)	(0.0149)	(0.0152)	(0.0120)
Entry=1 × Platform Pct%	0.0095*	0.6565*	−0.8157**	−0.2714*
	(0.0045)	(0.2747)	(0.3206)	(0.1143)
Requested Size (1K)	0.0027*	0.1482**	0.0539	−0.0399
	(0.0012)	(0.0571)	(0.0514)	(0.0291)
DTI	−0.0187	−4.4621*	4.2308	3.2316
	(0.0244)	(1.8303)	(3.0579)	(2.4566)
Employment Length	0.0023	0.1068	−0.1589	0.0050
	(0.0023)	(0.1457)	(0.1934)	(0.0787)
FICO	−0.0010**	−0.0555*	−0.0181	0.0157
	(0.0004)	(0.0227)	(0.0211)	(0.0088)
Observations	5426	5426	5426	5426
R^2	0.052	0.048	0.038	0.026

Standard errors in parentheses
Suppressed Variable: loan purpose, homeownership, state, monthly FE. Other borrower and loan characteristics.
Standard errors are clustered at Loan Grade
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 25: Test Platform's Objective

The objective of this table is to show the platform's objective of providing capital to borrowers. For one hypothesis, the platform is profit-seeking, and lends to high quality borrowers for return. For another, the platform protects its reputation by providing capital to those with credit crunches and clearing the market. To test the hypotheses, I examine the correlation between the platform's capital provision and borrowers' ex post loan performance, controlling for the observables. In addition, to account for any mechanism change induced by the entry event, I interact the key independent variable Platform Percentage with Entry dummy.
Discussion: It shows that for loans with higher platform's capital provision, the loan performance is inferior. This confirms the hypothesis that the platform has an objective to clear the market and protect its reputation. Second, the magnitude of this incentive is significantly stronger in the post-entry period. This can be attributed to several factors. Borrower performance worsens and the capital provided by the platform significantly declines post-entry. These effects jointly strengthen the magnitude of the estimates.

Appendix C Theoretical Section

C.1 Model Preview

I construct a model to explain the empirics I document. Borrowing the basic setup from Bolton et al. (2012), I add new components and relax some assumptions to endogenize some of the most important parameters. Bolton et al. (2012) shed light upon the conflict of interest between clientèles and intermediaries (Credit Rating Agencies) emerged from CRA competition. Competition distorts intermediaries' incentive on its information production since it yields trade-offs between a short-run gain from poaching debtors and a long-run reputation cost from deceiving creditors. In this model setting, I first introduce interest rates between a borrower and a lender to explain the asset price changes due to competition. Second, I relax the assumption where the lender can be trusting ex ante, which in Bolton et al. (2012), serves as a main mechanism for the CRA to inflate credit. While only observing the market maker's message, the lender makes sequentially rational decisions with the market maker in the Perfect Bayesian Equilibrium. Third, I introduce adverse selection as it was identified in the empirical section.

Appendix D Monopoly

In a one period economy, there are 3 types of risk-neutral agents, one borrower, one market maker and a lender.[68] The borrower can be either of the two types, 'good' and 'bad', in short "g" or "b", denoted by $\omega \in \{g, b\}$. A g type defaults with probability 0 and a b type with probability p. The borrower demands 2 units of investments for her

[68]Note that by definition risk-neutrality requires linear utility functions. Later in this section, I introduce increasing marginal reservation utility that may have a "risk-aversion" flavor.

project. The first unit of investment generates a gross return $\overline{R}$ and the second unit, a marginal gross return of $\underline{R}$, given that the project does not fail. If the project fails, the total return is 0 and the borrower defaults without liability. As marginal valuations, $\underline{R}$ and $\overline{R}$ pin down the credit demand, regardless of the borrower's type. If the gross interest rate is greater than $\overline{R}$, the borrower does not borrow. If the interest rate is between $\underline{R}$ and $\overline{R}$, the borrower demands 1 unit. Otherwise, the borrower demands 2 units.

The credit demand, denoted by $D(R)$, is as follows:

$$D(R) = \begin{cases} 2 & R \leq \underline{R} \\ 1 & R \in (\underline{R}, \overline{R}]) \\ 0 & o.w. \end{cases}$$

Assumption 1. *First best outcome: $\overline{R} > \underline{R} > 1$, i.e. a 'good' type borrower deserves 2 units investments. $\overline{R}(1-p) > 1$ and $\underline{R}(1-p) < 1$. That is, a 'bad' type borrower deserves 1 and only 1 unit of investment.*

The borrower's type is not publicly observable, and all other agents hold a prior belief that the borrower is 'good' with probability 1/2. The market maker has a private monitoring technology, and can obtain an imperfect signal $\theta \in \{g, b\}$. The signal is only observable to the borrower and the market maker, and has precision e, with e>1/2. That is, $Pr(\theta = g|\omega = g) = Pr(\theta = b|\omega = b) = e$. After receiving the signal, the market maker produces a credit report and an interest rate to the lender, $m \in \{(G, R_G), (B, R_B)\}$.

With deep pocket, the lender can either invest 1 unit or 2 units on the borrower through the market maker, conditional on that the borrower accepts the interest rate agreement. The lender's marginal reservation utilities for the investment are increasing. Specifically, if the ex ante gross return on the first unit (marginal) is greater than **u**=1, the investors purchase at least 1 unit. If the ex ante gross return on the second unit is greater than $\mathbf{U} > \mathbf{u}$, the investors purchases 2 units. **U and u** are called the marginal reservation utilities. For every unit of the loan issued, the market maker profits some exogenous fee, denoted by ϕ from the lenders. Here, I assume ϕ exogenous and charged on lenders for simplicity and tractability.[69]

Assumption 2. *I assume without the market maker, the public prior belief on the borrower's types just yields at most 1 unit of investments,* $\frac{U}{(1-p/2)} = \underline{R} + \varepsilon$*, where* ε *is an arbitrarily small positive real number.*[70]

With an arbitrarily small ε, the signal can be (in)precise enough but still guarantees (on average) more investment to a "g" type borrower than a "b" type. Therefore, the market making activities can improve efficiency compared to a bilateral trade.

I do not follow the assumption imposed by Bolton et al. (2012), where lenders ex post can infer if the market maker has truthfully reported. Instead, ex ante, the lender rationally forms some beliefs on if the market maker was truth-telling. Ex post, the lender also punishes the market maker, only if the market maker reported 'G' and the borrower defaults. In practice, the lender can only hurt the market maker by exiting the market. If she does so, the market maker loses a future value ρ.[71]

[69] The lender's problem becomes — if the ex ante return on the first unit is greater than u+ϕ, the lenders purchase 1 unit. For the second unit, their reservation utilities is U+ϕ. Since both U, u, ϕ are exogenous, WLOG, I let $u := u + \phi$ and $U := U + \phi$.

[70] We do not require this assumption to be true for the following properties.

[71] First, note that it is important to assume that the lender is nonstrategic. In a static game, we cannot endogenize lender's "punishment". By allowing punishment, we change the payoff structure of the static game, and don't want to end up in an equilibrium where punishment is a dominant strategy. Second, note that it is WLOG to assume the lender punishes with probability 1 given 'G' message and a default outcome.

D.1 Lender's Problem and Credit Supply

I denote the probability of inflating the rating form "b" to "G" as μ_i and the probability of deflating the rating from "g" to "B" as μ_d. Depending on parameter values, in equilibrium μ_i and μ_d can be 0 or 1.

$$P(\text{"}G\text{"}|\text{"}g\text{"}) = 1 - \mu_d \quad P(\text{"}B\text{"}|\text{"}g\text{"}) = \mu_d$$

$$P(\text{"}G\text{"}|\text{"}b\text{"}) = \mu_i \quad P(\text{"}B\text{"}|\text{"}b\text{"}) = 1 - \mu_i$$

By Bayes Rule, the lender's belief on the probability of default given "G" message is then:

$$\begin{aligned} P(\text{default}|G) =& \frac{P(\text{default}|g)P(G|g)P(g) + P(\text{default}|b)P(G|b)P(b)}{P(G|g)P(g) + P(G|b)P(b)} \\ =& \frac{(1-e)(1-\mu_d)p + e\mu_i p}{(1-\mu_d) + \mu_i} \end{aligned} \tag{7}$$

For the lender, the ex ante marginal return must exceed U for 2 units and 1 for 1 unit:

$$(1 - P(\text{default}|G))R_G \geq U \tag{G-2 Units}$$

$$U > (1 - P(\text{default}|G))R_G \geq 1 \tag{G-1 Unit}$$

Similarly, The lender's belief on the probability of default given "B" becomes:

$$\begin{aligned} P(\text{default}|B) =& \frac{P(\text{default}|g)P(B|g)P(g) + P(\text{default}|b)P(B|b)P(b)}{P(B|g)P(g) + P(B|b)P(b)} \\ =& \frac{(1-e)(\mu_d)p + e(1-\mu_i)p}{(\mu_d) + (1-\mu_i)} \end{aligned} \tag{8}$$

The investment decision stems out from the following inequalities

$$(1 - P(\text{default}|B))R_B \geq U \tag{B-2 Units}$$

$$U > (1 - P(\text{default}|B))R_B \geq 1 \tag{B-1 Unit}$$

By the payoffs, the ex ante credit supplies are:

$$S(R_G, G) = \begin{cases} 2 & R_G \geq \frac{U}{(1-P(\text{default}|G))} \\ 1 & R_G \in [\frac{1}{1-P(\text{default}|G)}, \frac{U}{1-P(\text{default}|G)}) \\ 0 & R_G \leq \frac{1}{1-P(\text{default}|G)} \end{cases}$$

$$S(R_B, B) = \begin{cases} 2 & R_B \geq \frac{U}{(1-P(\text{default}|B))} \\ 1 & R_B \in [\frac{1}{1-P(\text{default}|B)}, \frac{U}{1-P(\text{default}|B)}) \\ 0 & R_B \leq \frac{1}{1-P(\text{default}|B)} \end{cases}$$

Proposition 1. *By reporting 'G', the market maker is able to issue 2 units. By 'B', he can at most issue 1 unit.*

Proof. Since the signal is informative, i.e. $e > 1/2$, we know that the least marginal return for a 'g' signal: $\frac{U}{(1-(1-e)p)} < \frac{U}{(1-p/2)} = \underline{R} + \varepsilon$. Since e and p are given parameters and ε can be arbitrarily small, $\frac{U}{(1-(1-e)p)} < \underline{R}$. Therefore, the market clears at 2 units when the interest rate is set $R_G \in [\frac{U}{(1-(1-e)p)}, \underline{R}]$.

For a 'b' signal, its least required marginal return to the lenders is bigger than that of the second unit of investment, $\frac{U}{(1-ep)} > \frac{U}{(1-p/2)} > \underline{R}$. The market can at most clear at 1 unit. □

D.2 Market Maker's Problem

Given a "g" signal, if the market maker is truth-telling by reporting 'G' and some market clearing interest rate R_G, there is some nontrivial probability that he cannot capture the future value ρ. This probability is $\mathbb{P}(\text{Default}|\theta = g) = (1-e)p$. If the market maker deflates the reporting to "B", he will capture ρ with probability 1.

$$\Pi(R_G, G|g) = \max_{R_G} \min\{S(R_G, G), D(R_G)\}\phi + (1-(1-e)p)\rho \qquad \text{(Truthful-g)}$$

$$\Pi(R_B, B|g) = \max_{R_B} \min\{S(R_B, B), D(R_B)\}\phi + \rho \qquad \text{(Deflate-g)}$$

Given a "b" signal, if the market maker reports "B" is reported, he won't be punished ex post. If the market maker inflates it "G", he can only capture ρ if the project happens to be safe. The probability that the project is safe given $\theta = b$ is $e(1-p)+(1-e) = 1-ep$.

$$\Pi(R_B, B|b) = \max_{R_B} \min\{S(R_B, B), D(R_B)\}\phi + \rho \qquad \text{(Truthful-b)}$$

$$\Pi(R_G, G|b) = \max_{R_G} \min\{S(R_G, G), D(R_G)\}\phi + (1-ep)\rho \qquad \text{(Inflate-b)}$$

Note that the credit supply $S(R_G, G)$ and $S(R_B, B)$ incorporate lender's belief on the default probability, given the message "G" or "B".

Proposition 2. *1. A pooling equilibrium does not exist, i.e.* μ_i *or* μ_d *cannot equal to 1 ; 2. In a partially separating equilibrium, if* $\mu_i > 0$ *then* $\mu_d = 0$ *and if* $\mu_d > 0$*, then* $\mu_i = 0$

Proof. 1. To support a pooling equilibrium, the market maker either always chooses to report "G" (or equivalently "B"). Given that the message the market maker reports does not contain any information, the lender can bypass the market maker and engages in a bilateral trade with the borrower for 1 unit of investment (See assumption 2).

2. The only case to rule out is both $\mu_i > 0$ and $\mu_d > 0$. By compare the payoffs at different signals "g" or "b", Deflate-g = Truthful-b and Truthful-g>Inflate-b. If the lender uses a pure strategy on either 1 unit or 2 units of investment upon messages, then we are done, since the market maker just maximizes his payoff without the need to mix. If the lender offers a mixed strategy q, between 1 unit and 2 units to make the market maker indifferent between both Deflate-g & Truthful-g and Truthful-b & Inflate-b, due to the fact that Deflate-g = Truthful-b and Truthful-g>Inflate-b, at most one equality can be satisfied. Either Deflate-g = Truthful-g or Truthful-b = Inflate-b, but not both. □

I focus on the equilibrium where $\mu_d = 0$ and $\mu_i \geq 0$. The credit demand gives the least upper bound(s) where $R_G \leq \underline{R}$ to satisfy 2 units and $R_B \leq \overline{R}$ to satisfy 1. For the market maker, under all μ_i, he weakly prefers setting $R_G = \underline{R}$ to any other interest rates. Since the interest rate does not go into the market maker's problem directly, by choosing a higher interest rate without affecting the market clearing, the market maker gets more degree of freedom to inflate the credit. Given 'g' signal, either reporting "G" or "B", he gets different payoffs if the lender chooses different units:

$$\Pi(R_G, G|g) = 2\phi + (1 - (1 - e)p)\rho \tag{G|g-2 units}$$

$$\Pi(R_G, G|g) = \phi + (1 - (1 - e)p)\rho \tag{G|g-1 unit}$$

$$\Pi(R_B, B|g) = \phi + \rho \tag{B|g-1 unit}$$

Since $R_B > \underline{R}$, the credit demand stays at 1, and thus 2 units for "B" message is not considered. Similarly, given "b" signal,

$$\Pi(R_G, G|b) = 2\phi + (1 - ep)\rho \qquad \text{(G|b-2 units)}$$

$$\Pi(R_G, G|b) = \phi + (1 - ep)\rho \qquad \text{(G|b-1 unit)}$$

$$\Pi(R_B, B|b) = \phi + \rho \qquad \text{(B|b-1 unit)}$$

D.3 Equilibrium Strategy

Definition 1. *A Perfect Bayesian Equilibrium of the game is a strategy profile for the market maker and the lender, $\sigma_m^*(\theta), \sigma_l^*(M)$, contingent on the "signal" the market maker obtains (Type, $\theta = \{g, b\}$) and the "message" the lender observes ($M = \{G, B\}$), and the posterior belief $P(Default|M)$ such that:*

- *for $\theta \in \{g, b\}$, the market maker maximizes his payoff by choosing $\sigma_m^*(\theta) = \{\mu_i^*|b, \mu_d^*|g, R_G, R_B\}$, given the strategy profile of the lender $\sigma_l^*(M)$*
- *for $M \in \{G, B\}$, the lender maximizes her payoff by choosing $\sigma_l^*(M) \in \{1, 2\}$, given $\sigma_m^*(\theta)$*
- *The $P(Default|M)$ follows Bayes rule given $\sigma_m^*(\theta)$.*

As derived in the past section, $R_G = \underline{R}$. Since the market maker does not deflate "g" to B, a "B" message would indicate that the borrower has a "b" signal. $P(\text{Default}|B) = P(\text{Default}|b) = ep$. To satisfy one unit of investment, $\frac{1}{1-ep} \leq R_B \leq \overline{R}$, and the market maker is indifferent for any R_B within the set. Depending the parameter values of ϕ, p, e, they may end up in a fully separating equilibrium or partially separating equilibrium. Since it is not interesting to pursue an equilibrium where the market maker report messages with no information (pooling equilibrium in Proposition 2), the pure equilibrium sets where it is a dominant strategy for the market maker to report the truth and the lender invests 2 units. In the mixed strategy equilibrium, the lender chooses a probability q^* to invest 2 units (vs 1 unit) upon G to make the market maker indifferent between inflating to "G" and telling the truth by reporting "B" given that he obtains a "b" signal. At the same time, q^* guarantees that the market maker tells the truth under "g". If $(1 - ep)\rho \leq \phi \leq ep\rho$, then it is a dominant strategy for the lender to investment 2 units upon "G" message. The punishment is too high and the market maker would rather settle for 1 unit than taking the risk of being punished.

If the fee is large enough where $\phi > ep\rho$, then the lender knows that the market maker has incentive to inflate the rating. The lender chooses a probability q^* to make the market maker indifferent between inflating the rating and telling the truth upon "b" signal, $q^* = \frac{ep\rho}{\phi}$. The market maker chooses $\mu_d^* = 0$ and $\mu_i^* \geq 0$ to make the lender indifferent between offering 1 unit or 2 units upon "G" message.

$$\mu_i^* = \frac{\underline{R}(1 - (1 - e)p) - U}{U - (1 - pe)\underline{R}} \tag{9}$$

Remember, by assumption 2, $U = \underline{R}(1 - \frac{p}{2}) + \varepsilon$, then the incentive to inflate is almost 1, $\mu_i^* = 1 - \varepsilon$.

Figure 26: Equilibrium for Possible Values of ϕ

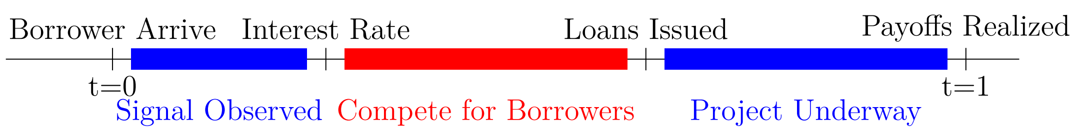
Bottom (Blue) brackets represent parameter region of ϕ for truthfully reporting with probability 1 or inflating with a high probability under monopoly.

Appendix E Duopoly Setup

Let's consider two identical market makers (denoted by player 1 and 2) competing for both the borrower and lender. Both platforms obtain identically precise signals. However, even with the same precision, signal draws can differ. That is, one platform may obtain "g" and another may obtain "g" or "b". In particular, if player 1 observes a 'g' signal, player 1 acknowledges that with probability $\mathbb{P}(\theta_2 = \text{"g"}|\text{"g"}) = e^2 + (1-e)^2$, player 2 observes a 'g' signal, and with probability $\mathbb{P}(\theta_2 = \text{"b"}|\text{"g"}) = 2e(1-e)$, a 'b' signal. Similarly, conditional on a 'b' signal, i.e. $\mathbb{P}(\theta_2 = \text{"b"}|\text{"b"}) = e^2 + (1-e)^2$ and $\mathbb{P}(\theta_2 = \text{"g"}|\text{"b"}) = 2e(1-e)$. It is symmetric from player 2's perspective.

E.0.1 The Game

Timeline



The timing of the game with market maker competition is as follows:

- The borrower visits both market makers simultaneously
- Each market maker observes respective signal and offers a rating and an interest rate to the borrower
- The borrower chooses the interest rate offer to maximize his payoff:

  Specifically, if player 1 offers interest rate R and player 2 R'. If $R = R'$, the borrower chooses each with probability 1/2. If $R < R'$, the borrower goes to player 1. Otherwise, the borrower goes to player 2.

- The lender follows the borrower to clear the market

In addition to the description above, I make an additional assumption:

Assumption 3.
- *upon observing the interest rates, the borrower can perfectly foresee the lender's credit supply, and strictly prefers two units to one being cleared in equilibrium.*
- *the future value for either market maker becomes less, $\rho/2$, since in equilibrium, on average the market maker can only get the borrower half of the times.*

This assumption is not needed in the monopoly case, since it involves overly undercutting. It states that a market maker cannot undercut the other too much so that the credit demanded exceeds credit supplied. I apply several other restrictions. First, the borrower cannot obtain a loan from both market makers. In practice, loan underwriters typically observe the debt outstanding and loan application history of the borrowers. Second, I do not incorporate competition on the lender's side. In practice, lender competition and lenders' loyalty are non-negligible to market making activities. By assuming that lenders are fully mobile between the market makers, we implicitly assume that lender population is large enough with deep pockets, and they do not have a clear preference between the market makers.

E.1 Market Makers' Problem

Borrowers' marginal reservation interest rate does not change due to competition. However, the situation is a bit tricky for the lenders. Because market makers may obtain differentiated signals and they only observe their own, the competition may result in a winner's curse problem. Specifically, a borrower with 'g' signal on one platform may have 'b' on the other. Although the interest rate doesn't directly go into the market maker's payoff, it affects the market clearing unit through the borrower and the lender's decisions. The previous results on the interest rates from monopoly no longer sustain in equilibrium:

Proposition 3. *Due to the existence of adverse selection, undercutting the interest rates to the lower bounds of the monopoly case, i.e.* $R_G = \frac{U}{(1-(1-e)p)}$ *for 2 units of investment and* $R_B = \underline{R}$ *for 1 unit is not a Bertrand Equilibrium.*

Proof. I show it by contradiction. I assume that the equilibrium holds. The lender fully trust the borrower's message, m. If the lender observes $m = G$, she believes that the signal is $\theta = g$. She can also deduce that if the borrower accepted player 1's interest rate offer at $R_G = \frac{U}{(1-(1-e)p)}$, the borrower must hold signals (g,g) or (g,b). In equilibrium, a (g,g) borrower accepts $R_G = \frac{U}{(1-(1-e)p)}$ with 1/2 probability and a (g,b) borrower accepts $R_G = \frac{U}{(1-(1-e)p)}$ with probability 1. Note that the conditional probability distribution of the other player's signal is $\mathbb{P}(\theta_2 = g|\theta_1 = g) = e^2 + (1-e)^2$ and $\mathbb{P}(\theta_2 = b|\theta_1 = g) = 2e(1-e)$. Therefore, given the borrower accepted the offer, the conditional probability $\mathbb{P}(g, g|\text{Accept}) = \frac{e^2+(1-e)^2}{1+2e(1-e)}$ and $\mathbb{P}(g, b|\text{Accept}) = \frac{4e(1-e)}{1+2e(1-e)}$.

Therefore, the ex ante required return to the lender on the second unit is

$$(1-\mathbb{P}(\omega = b|\theta = (g, g))p)\frac{e^2 + (1-e)^2}{1 + 2e(1-e)}+(1-\mathbb{P}(\omega = b|\theta = (g, b))p)\frac{4e(1-e)}{1 + 2e(1-e)} \geq U/R_G$$

Plugging R_G into the equation above, we have

$$\begin{aligned}&(1-\mathbb{P}(\omega=b|\theta=(g,g))p)\frac{e^2+(1-e)^2}{1+2e(1-e)}+(1-\mathbb{P}(\omega=b|\theta=(g,b))p)\frac{4e(1-e)}{1+2e(1-e)}\\ \geq&(1-\mathbb{P}(\omega=b|\theta=g)p)\end{aligned}$$

This is equivalent to

$$\begin{aligned}&(\mathbb{P}(\omega=b|\theta=(g,g)))\frac{e^2+(1-e)^2}{1+2e(1-e)}+(\mathbb{P}(\omega=b|\theta=(g,b)))\frac{4e(1-e)}{1+2e(1-e)}\\ \leq&(\mathbb{P}(\omega=b|\theta=g))\\ =&\mathbb{P}(\omega=b|\theta=(g,g))\mathbb{P}(\omega=g|\theta=g)+\mathbb{P}(\omega=b|\theta=(g,b))\mathbb{P}(\omega=b|\theta=g)\\ =&\mathbb{P}(\omega=b|\theta=(g,g))(e^2+(1-e)^2)+\mathbb{P}(\omega=b|\theta=(g,b))(2e(1-e))\end{aligned}$$

Note that $\frac{e^2+(1-e)^2}{1+2e(1-e)}+\frac{4e(1-e)}{1+2e(1-e)}=1$ and $(e^2+(1-e)^2)+(2e(1-e))=1$. We are essentially compare a weighted average of $\{\mathbb{P}(\omega=b|\theta=(g,g)),\mathbb{P}(\omega=b|\theta=(g,b))\}$ with two sets of weights. By Bayes Rule, $\mathbb{P}(\omega=b|\theta=(g,g))\ll\mathbb{P}(\omega=b|\theta=(g,b))$. In order for the in equality above to hold, we require a lower weight on $\mathbb{P}(\omega=b|\theta=(g,g))$, i.e. $\frac{e^2+(1-e)^2}{1+2e(1-e)}>e^2+(1-e)^2$. It is thus a contradiction. $\square$

Following the proposition above, a symmetric pure strategy equilibrium on the interest rates can no longer be sustained. I denote in equilibrium R_G follows $\mathcal{F}_G(\cdot)$ and R_B follows $\mathcal{F}_B(\cdot)$, where R_G and R_B are separating. That is, the support of R_G and R_B do not overlap. Also, the distributions of R_G and R_B depend on the market makers' decision to inflate or deflate the reporting. At each node, the other player may choose to inflate, deflate or tell the truth. (See Figure 27) Using those conditional probabilities, I derive

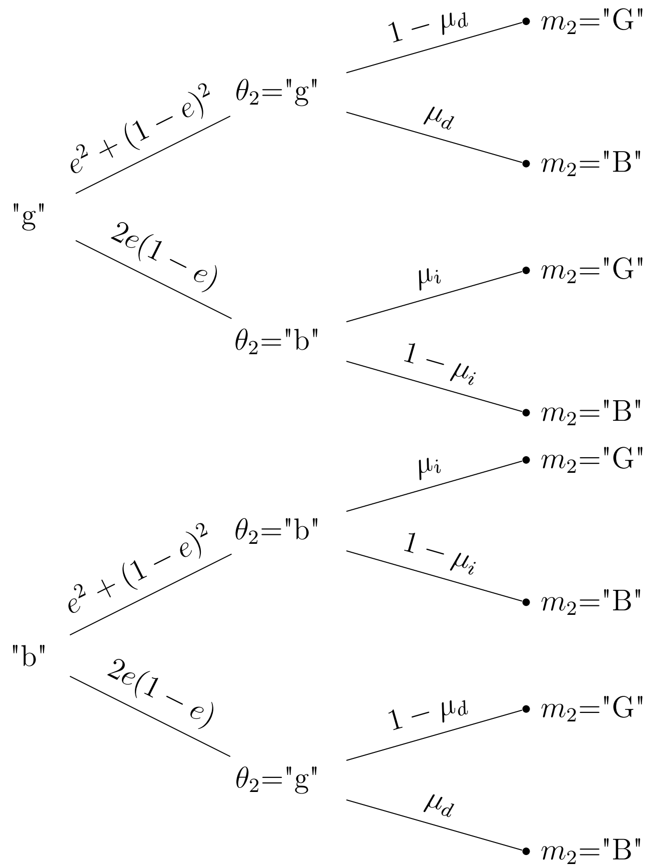


Figure 27: Conditional Probabilities

the payoffs for player 1. With "g" signal, if player 1 reports "G" and draws an interest rate R from R_G, the probability that the borrower takes his bid is:

$$\begin{aligned}
P(\text{winning}|\text{"g","G"}) &= P(R_G < R') \\
&= \underbrace{(1-\mu_d)\mathbb{P}(\theta_2 = \text{"g"}|\text{"g"})(1-\mathcal{F}_G(R_G))}_{\text{competition}} + \underbrace{\mu_d \mathbb{P}(\theta_2 = \text{"g"}|\text{"g"})}_{\text{player 2's deflation}} \\
&+ \underbrace{(1-\mu_i)\mathbb{P}(\theta_2 = \text{"b"}|\text{"g"}) + \mu_i \mathbb{P}(\theta_2 = \text{"b"}|\text{"g"})(1-\mathcal{F}_G(R_G))}_{\text{winner's curse}}
\end{aligned}$$

I derive the other probabilities of winning: $P(\text{winning}|\text{"g","B"})$, $P(\text{winning}|\text{"b","G"})$ and $P(\text{winning}|\text{"b","B"})$. Use those probabilities, I derive player 1's ex ante payoffs under each action contingent on his own signals. Given signals and actions, player 1's payoffs are

$$\begin{aligned}
\Pi_1(R_G|\text{"g"}) =& P(\text{"g"}|\text{"g"})(\mu_d + (1-\mu_d)(1-\mathcal{F}_G(R_G)))(\min\{S(R_G), D(R_G)\}\phi - P(\text{default}|\text{"g","g"})\frac{\rho}{2}) \\
&+ P(\text{"b"}|\text{"g"})(1-\mu_i + \mu_i(1-\mathcal{F}_G(R_G)))(\min\{S(R_G), D(R_G)\}\phi - P(\text{default}|\text{"g","b"})\frac{\rho}{2}) + \frac{\rho}{2}
\end{aligned} \tag{g-truth-telling}$$

$$\Pi_1(R_B|\text{"g"}) = (P(\text{"g"}|\text{"g"})(1-\mathcal{F}_B(R_B))\mu_d + P(\text{"b"}|\text{"g"})(1-\mu_i)(1-\mathcal{F}_B(R_B)))(\min\{S(R_B), D(R_B)\}\phi) + \frac{\rho}{2} \tag{deflate}$$

$$\begin{aligned}
\Pi_1(R_G|\text{"b"}) =& P(\text{"b"}|\text{"b"})(1-\mu_i + \mu_i(1-\mathcal{F}_G(R_G)))(\min\{S(R_G), D(R_G)\}\phi - P(\text{default}|\text{"b","b"})\frac{\rho}{2}) \\
&+ P(\text{"g"}|\text{"b"})(\mu_d + (1-\mu_d)(1-\mathcal{F}_G(R_G)))(\min\{S(R_G), D(R_G)\}\phi - P(\text{default}|\text{"g","b"})\frac{\rho}{2}) + \frac{\rho}{2}
\end{aligned} \tag{inflate}$$

$$\Pi_1(R_B|\text{"b"}) = (P(\text{"b"}|\text{"b"})(1-\mu_i)(1-\mathcal{F}_B(R_B)) + P(\text{"g"}|\text{"b"})\mu_d(1-\mathcal{F}_B(R_B)))(\min\{S(R_B), D(R_B)\}\phi) + \frac{\rho}{2} \tag{b-truth-telling}$$

E.1.1 Lender's Problem

In the monopoly case, I derive that by reporting "G" and R_G, the market maker can issue at most 2 units. With "B" reported, the market can only clear at 1 unit. In duopoly, the adverse selection can change lender's belief on the accepted borrower's type and subsequently his credit supply. With the probability of winning, the lender can also derive her belief on the probability of default:

$$\begin{aligned} P(\text{default}|\text{winning,"G"}) =& \frac{\sum_{\theta_1,\theta_2} P(\text{winning,"G"}|\theta_1,\theta_2)P(\theta_1,\theta_2)P(\text{default}|\theta_1,\theta_2)}{\sum_{\theta_1,\theta_2} P(\text{winning,"G"}|\theta_1,\theta_2)P(\theta_1,\theta_2)} \\ =& \frac{\sum_{\theta_1,\theta_2} P(\text{winning,"G"}|\theta_1,\theta_2)P(\text{default}|\theta_1,\theta_2)}{(1-\mu_d+\mu_i)^2(1-\mathcal{F}_G(R_G)) + (1-\mu_d+\mu_i)(1-\mu_i+\mu_d)} \end{aligned} \tag{10}$$

, where $P(\text{default}|\text{"g","g"}) = \frac{(1-e)^2 p}{e^2+(1-e)^2}$, $P(\text{default}|\text{"g","b"}) = \frac{p}{2}$ and $P(\text{default}|\text{"b","b"}) = \frac{e^2 p}{e^2+(1-e)^2}$. Moreover, the winner probabilities are

$$P(\text{winning,"G"}|g,g) = (1-\mu_d)^2(1-\mathcal{F}_G(R_G)) + (1-\mu_d)\mu_d$$

$$P(\text{winning,"G"}|g,b) = (1-\mu_d)\mu_i(1-\mathcal{F}_G(R_G)) + (1-\mu_d)(1-\mu_i)$$

$$P(\text{winning,"G"}|b,g) = (1-\mu_d)\mu_i(1-\mathcal{F}_G(R_G)) + \mu_d\mu_i$$

$$P(\text{winning,"G"}|b,b) = \mu_i^2(1-\mathcal{F}_G(R_G)) + (1-\mu_i)\mu_i$$

For 2 units of investments to clear at equilibrium, the ex ante marginal payoff for a lender's second unit is at least U. That is,

$$(1 - P(\text{default}|\text{winning,"G"}))R_G \geq U$$

Note that in some equilibrium if μ_i is large enough, 2 units of investments cannot be sustained. To compare to the monopoly case, I derive the conditions on the parameter values where the market maker is just indifferent between credit inflation and telling the truth when facing "g". Therefore, the 2 units of investment in equilibrium is not fully compromised. Similarly, for R_B:

$$P(\text{default}|\text{winning,"B"}) = \frac{\sum_{\theta_1,\theta_2} P(\text{winning,"B"}|\theta_1,\theta_2)P(\theta_1,\theta_2)P(\text{default}|\theta_1,\theta_2)}{\sum_{\theta_1,\theta_2} P(\text{winning,"B"}|\theta_1,\theta_2)P(\theta_1,\theta_2)}$$

, where the probabilities of default given the signals do not change, but those of winning become

$$P(\text{winning,"B"}|g,g) = (\mu_d)^2(1 - \mathcal{F}_B(R_B))$$

$$P(\text{winning,"B"}|g,b) = (\mu_d)(1-\mu_i)(1 - \mathcal{F}_B(R_B))$$

$$P(\text{winning,"B"}|b,g) = (\mu_d)(1-\mu_i)(1 - \mathcal{F}_B(R_B))$$

$$P(\text{winning,"B"}|b,b) = (1-\mu_i)^2(1 - \mathcal{F}_B(R_B))$$

Similarly to R_B, the indifference condition for R_B guarantees 1 unit. Again, for some parameter values, we may end up with 2 units of investment if μ_d is high enough. To make it comparable to monopoly, I do not consider the case.

$$(1 - P(\text{default}|\text{winning,"B"}))R_B \geq 1$$

E.2 Equilibrium Strategy

Proposition 2 still holds in the duopoly, since in a pooling equilibrium where both market makers map all signal into one message, "G" (or "B"), there can be only one unit issued.[72] I consider a partially(or fully) separating equilibrium where $\mu_d^* = 0$ and $\mu_i^* \geq 0$.

Since $\mu_d^* = 0$, then with a "B" message, the lender knows that the borrower obtained two "b" signals and is indifferent between the market makers. Therefore, the interest rate, R_B, is priced such that it just covers the probability of default and satisfies one unit of investment. With two "b" signals, the likelihood of default becomes $\frac{e^2 p}{e^2+(1-e)^2}$, and in equilibrium

$$R_B = \frac{e^2 + (1-e)^2}{e^2(1-p) + (1-e)^2} \tag{11}$$

By comparing the following payoffs, I derive the parameter values where it is a dominant strategy for players to be truth-telling

$$\begin{aligned}\Pi_1(R_G|\text{"b"}) =& P(\text{"b"}|\text{"b"})(1-\mu_i+\mu_i(1-\mathcal{F}_G(R_G)))(\min\{S(R_G), D(R_G)\}\phi - P(\text{default}|\text{"b","b"})\frac{\rho}{2})\\ &+P(\text{"g"}|\text{"b"})(\mu_d+(1-\mu_d)(1-\mathcal{F}_G(R_G)))(\min\{S(R_G), D(R_G)\}\phi - P(\text{default}|\text{"g","b"})\frac{\rho}{2}) + \frac{\rho}{2}\end{aligned} \tag{inflate}$$

$$\Pi_1(R_B|\text{"b"}) = (P(\text{"b"}|\text{"b"})(1-\mu_i)(1-\mathcal{F}_B(R_B)) + P(\text{"g"}|\text{"b"})\mu_d(1-\mathcal{F}_B(R_B)))(\min\{S(R_B), D(R_B)\}\phi) + \frac{\rho}{2} \tag{b-truth-telling}$$

[72] In this case, the market makers do not add any value to a bilateral trade between the lender and the borrower.

If it is a dominant strategy for the players to tell the truth, i.e. $\Pi_1(R_G|"b") \leq \Pi_1(R_B|"b")$, for all $\mu_i > 0$, then the players end up in a separating equilibrium. To derive the condition above, ex ante, each player has an average chance of winning equal to 1/2, or in other words, $\int(1 - \mathcal{F}_G(R_G))dR_G = \int(1 - \mathcal{F}_B(R_B))dR_B = 1/2$. From the inequality above, a separating equilibrium exists if $\phi \leq \frac{e^2}{2((1-e)^2+e^2)}p\rho$, i.e. $\mu_i = 0$. With $\mu_i = 0$ and the probability of default, I derive a closed form solution for $\mathcal{F}_G(R_G)$:

$$\mathcal{F}_G(R_G) = \frac{R_G - U - R_G p(1-e)}{(e^2 + (1-e)^2)(R_G - U) - (1-e)^2 R_G p} \tag{12}$$

The support of R_G is bounded below by $\frac{U}{(1-(1-e)p)}$ and above by $\frac{U}{(1-p/2)} = \underline{R} + \varepsilon$.[73]

If $\phi > \frac{e^2}{2((1-e)^2+e^2)}p\rho$, the players end up in a partially separating equilibrium strategy. $R_B = \frac{e^2+(1-e)^2}{e^2(1-p)+(1-e)^2}$, since $\mu_d = 0$. Here, there is additional degree of freedom between R_G and μ_i, since R_G is drawn from $\mathcal{F}_G(R_G)$, which depends on the value of μ_i. In the meantime, μ_i is endogenized by the value of R_G. To bypass this issue, I denote the following rule. Instead of randomly drawing R_G from $\mathcal{F}_G(R_G)$, players agrees on a pre-drawn R^*_G from $\mathcal{F}_G(R_G)$. If any player deviates, the lender would treat the message as B, since it is on an off-equilibrium path. This rule prevents players to deviate. Since the goal of the theory section is to show that it is more difficult for players to be truth-telling under competition, I do not solve the equilibrium to a closed form, but characterize it as follows.

Definition 2. *A symmetric Perfect Bayesian Equilibrium under* $\phi > \frac{e^2}{2((1-e)^2+e^2)}p\rho$ *is a strategy profile for the market makers and the lender,* $\sigma^*_{m,1}(\theta), \sigma^*_{m,2}(\theta), \sigma^*_l(M)$*, contingent on the "signal" the market makers obtain* ($\theta = \{g, b\}$) *and the "message" the lender*

[73] See Assumption 1

observes $(M = \{G, B\})$*, and the posterior belief* $P(Default|M)$ *such that:*

- *for* $\theta \in \{g, b\}$*, a market maker maximizes his payoff by choosing* $\sigma_m^*(\theta) = \{\mu_i^*|b, \mu_d^* = 0, R_G^*, R_B^*\}$*, given the strategy profile of the lender* $\sigma_l^*(M)$ *and the other player's strategy profile* $\sigma_l^*(\theta')$
- *for* $M \in \{G, B\}$*, the lender maximizes her payoff by choosing* $\sigma_l^*(M) \in \{1, 2\}$*, given* $\sigma_m^*(\theta)$
- *The* $P(Default|M)$ *follows Bayes rule given* $\sigma_m^*(\theta)$*.*
- *Any off equilibrium path* $R \notin \{R^*G, R^*B\}$ *will be treated as* R_B *by the lender.*

E.2.1 Equilibrium Comparison

In the past section, I show that under competition, it is more difficult to sustain a truth-telling equilibrium according to the parameter value of ϕ:

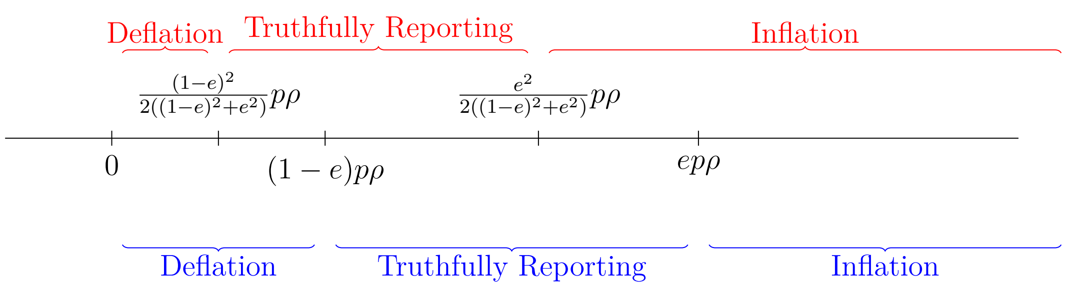


Figure 28: Equilibrium for Possible Values of ϕ

Top (Red) brackets denote the values of ϕ to support Deflation, Truthfully Reporting or Inflation under duopoly. Bottom (Blue) represent that for monopoly.

Chapter 2

Investment Bank Governance and Client Relationships

Investment Bank Governance and Client Relationships*

Zhaohui Chen,
McIntire School of Commerce,
University of Virginia

Alan D. Morrison,
Saïd Business School,
University of Oxford

William J. Wilhelm, Jr.,
McIntire School of Commerce,
University of Virginia

Xin Xue,
University of Virginia

July 13, 2018

Abstract

The relational contract at the heart of an investment banking relationship is valuable because it engenders and requires mutual trust in a setting where conflicts of interest are significant and are not easily resolved through formal contract. However, a bank's ability to commit to the relational contract depends on internal governance mechanisms that align the interests of individual bankers with those of the bank. We argue that increasing complexity in investment banks weakens internal governance and estimate a causal model that indicates that the likelihood of a relationship being broken is increasing in bank complexity.

*Wilhelm received generous support from the Janemarie D. and Donald A. King Faculty Fellowship in Real Estate. We thank Mike Gallmeyer, Steve Karolyi, and Michael Roberts for helpful comments..

Investment banking is not what it used to be. Investment banks were once partnerships whose employees formed close-knit social communities (Pak 2013). Partners had long tenure, seldom moved between banks, and formed long-lived relationships with their clients; they appeared to be more concerned with their reputational than their financial capital. In contrast, modern investment banking is dominated by very large, complex, publicly-owned firms that increasingly struggle to address internal conflicts of interest. Labor mobility is high among today's senior bankers, and bank-client relationships have weakened steadily for almost a half century.[1] Many observers have expressed concerns that behavioral standards have declined in financial firms. The spirit of these concerns was captured in a 2013 speech by William Dudley, the president of the Federal Reserve Bank of New York. President Dudley identified "deep-seated cultural and ethical failures" in the banking sector, as well as an "apparent lack of respect for law, regulation, and public trust." But he also noted that it is hard to determine whether these failures are a consequence of "size and complexity, bad incentives or some other issues".[2]

This paper examines the effect that the increasing scale and complexity of investment banks has had on their relationships with securities issuers. We claim that close relationships both engender and require mutual trust. Trust is valuable because investment bankers are better informed than their clients about market conditions and the transactions on which they advise and often face conflicts of interest stemming from their intermediary role.[3] Formal contract is insufficient to prevent bankers from abusing their superior knowledge. A trust-based relationship can therefore be a rational response to agency problems within banks. But both parties to a relationship bear opportunity costs. Issuers forgo competitive bidding for their underwriting mandates; banks may decline business that poses a threat to their client relationships; and each party to the relationship may sacrifice an opportunity to match with a counterparty that is more complementary to its requirements

[1]See Morrison and Wilhelm (2008) on banks' incentives to go public and Morrison, Thegeya, Schenone, and Wilhelm (2018) on long-term patterns in banker tenure and relationship exclusivity.

[2]"Ending Too Big to Fail," Remarks at the Global Economic Policy Forum, New York City, November 7, 2013. https://www.newyorkfed.org/newsevents/speeches/2013/dud131107.html

[3]Kang and Lowery (2014), Reuter (2006), Nimalendran, Ritter, and Zhang (2007) study conflicts between banks and securities issuers that stem from institutional brokerage relationships. Asker and Ljungqvist (2010) provide evidence that issuers avoid banks that may be conflicted by serving multiple clients within a product market. Bodnaruk, Massa, and Simonov (2009), Griffin, Shu, and Topaloglu (2012), and Jegadeesh and Tang (2010) provide evidence of banks' ability to exploit information gained from advising M&A clients. Mehran and Stulz (2007) for a broad review of the literature on conflicts of interest in financial institutions.

or capabilities (Fernando, Gatchev, and Spindt 2005).

In the theory that we present in Section 1, we consider the costs and benefits of relationships and how internal governance of an investment bank affects the strength of its client relationships. We argue that, if a bank's internal governance weakens, it becomes less able to control employee opportunism. This makes it more costly for the bank to keep the promises that underpin its long-term relationships, and from which those relationships derive their value. It follows that anything that weakens an investment bank's internal governance should increase the likelihood that its clients switch to a different bank. In particular, we argue that it is harder to govern a more complex bank, so that increases to a bank's complexity should weaken its client relationships.

We test this hypothesis using a sample of debt and equity issues that were brought to market between January 1960 and December 1998 by issuers who had a prior relationship with one of 30 sample banks. We estimate models in which securities issuers condition their decision to break or to continue a relationship on three measures of bank complexity that are intended to proxy for the unobservable underlying agency problem: bank capital, the number bank partners or comparable senior officers, and discrete changes in organizational structure. It is worth noting that banks that are more complex judged by these measures often are able to offer a wider range of services and can draw from a larger pool of human resources for their delivery. As a consequence, finding evidence in favor of our hypothesis will indicate that any benefits associated with greater scale and scope of operations are dominated by the costs of complexity.

Our analysis is complicated by the fact that decisions regarding a bank's organizational structure are not exogenous. We address the endogeneity problem by constructing an instrument that links banks' organizational complexity to advances in technology and their incentives to adopt new technology. Our sample period witnessed unprecedented technological and organizational change in investment banking. Morrison and Wilhelm (2008) argue that much of this change was driven by technological advances in both information technology and financial economic theory that increased the efficient scale of investment banks and contributed to increasing conflicts of interest within investment banks.[4]

[4]Philippon and Reshef (2012) provide evidence that technological change and deregulation placed a premium on highly skilled workers during the second half of our sample period. See Chen, Morrison, and Wilhelm (2014) for a

Our instrument is motivated by this connection between technological change and bank complexity. Its construction reflects theoretical work on the ways that early career experience affects human capital formation (for example Jovanovic (1979) and Greenwood and Jovanovic (1999)) and empirical work that links early life or career experience to future behavior.[5] Finally, it rests on our ability to identify on an annual basis the relationship bank's partners and place them in cohort years according to their first year of service in that capacity.[6]

The instrument is an annual index for each bank that reflects a weighted average of each partner cohort's exposure to the state of technology during a window immediately preceding their first year of partnership. This it varies in the cross-section of banks with differences in their partner cohort structures in a given year. We assume that banks with cohort structures dominated by bankers from cohort years when information technology is relatively primitive will be less inclined to adopt new technologies because those technologies are less likely to complement their human capital, and may even undermine it. Alternatively, banks with cohort structures dominated by partners of more recent vintage will be closer to the the cutting edge of new technology. Individual banks' cohort structures vary through time as current partners leave the firm and new ones join. These changes interact with the technology state variable for each cohort to increase or decrease receptivity to technology adoption.

First-stage regressions are consistent with our theory: bank complexity increases with greater receptivity to technology adoption, as measured by our instrument. In the second-stage regressions, switching propensity is increasing in our proxies for bank complexity and the marginal effect is statistically significant in both linear probability and probit model specifications. We also show that bank complexity weakens the tendency for issuers with strong existing relationships to continue

model of agency problems that stem from individual bankers facing strong incentives to build their personal reputation at the expense of their clients and their bank's reputation. Chen, Morrison, and Wilhelm (2015) present a model in which client trust is undermined by conflicts of interest between divisions of full-service investment banks.

[5]Bertrand and Schoar (2003) show that older CEOs are more conservative, Oyer (2008) demonstrates that career outcomes for MBAs entering investment banking are influenced by the state of the stock market at the outset of their careers, Malmendier and Nagel (2011) demonstrate a lower willingness to assume financial risk among people who have experienced lower stock-market returns, Malmendier, Tate, and Yan (2011) present evidence that CEOs who grew up during the Great Depression are averse to debt and lean excessively on internal finance, and Schoar and Zuo (2017) finds that managers who began their careers during recessions are more conservative.

[6]Our use of the term "partner" includes analogous titles in publicly listed banks. Details are provided in Section 2.2.

their relationship. At high levels of complexity, the presence of a strong banking relationship does little to deter an issuer from breaking its relationship. Finally, we show that relationships involving poorly "matched" banks and issuers are more likely to be broken. But when we interact our measure of the mismatch between the issuer and its relationship bank with proxies for bank complexity, the apparent preference for positive assortative matching is amplified. This result is consistent with the existence of the tradeoff between the costs and benefits of a relationship described above.

Our work contributes to the broad literature on the securities issuance process and, specifically, how investment-banking relationships influence the assignment of underwriting mandates.[7] The switching model is similar in spirit to those used by Krigman, Shaw, and Womack (2001), Fernando, Gatchev, and Spindt (2005), and Ljungqvist and Wilhelm (2005) to examine why firms switch banks between their initial public offering of equity (IPO) and first subsequent equity offering. However, we do not restrict our attention to IPOs. Our work is most closely related to work by Yasuda (2005, 2007), Ljungqvist, Marston, and Wilhelm (2006, 2009), and Asker and Ljungqvist (2010). Just as we do, these papers find that issuers favor banks with whom they have a relationship and that the effect is increasing in the degree of relationship exclusivity.[8] Morrison, Thegeya, Schenone, and Wilhelm (2018) show that many relationships were exclusive, or nearly so, prior to the 1970s but that relationship exclusivity and the influence of the state of the relationship on issuer decisions weakened substantially thereafter. Our evidence points to increasing organizational complexity as a likely contributor to this change. Our theory suggests that to the extent that strong relationships sustain trust between banks and their clients, greater organizational complexity in modern investment banks contributes to a decline in trust.

1. Theoretical Framework and Identification

Kahn and Whited (2018, p. 3) argue that identification is always based on a verbal or mathematical theory and, hence, that identification depends upon the plausibility of the assumptions

[7] See Ljungqvist (2007) and Eckbo, Masulis, and Norli (2005) for reviews of the literature on equity offerings.

[8] Also see Schenone (2004) for the benefits to IPO issuers that select a bank with which they have a lending relationship.

underlying the theory. We therefore begin by presenting our theory and the assumptions upon which it rests. We then discuss our identification strategy.

First, note that most firms access the capital markets infrequently. We therefore assume that investment banks are better-informed than their clients about market conditions and about the best way to meet their clients' needs.[9] Indeed, this knowledge is one of the most important things that the investment bank has to sell to its clients. But the knowledge is complex and nuanced and, hence, the quality of the advice tendered by an investment bank is seldom verifiable. This problem gives rise to conflicts of interest, because an investment banker has incentives to sell advice or products that are sub-optimal from the perspective of its clients, but that generate benefits for the banker. Those benefits could be earned by favoring clients' counterparties or competitors or by bundling high-margin products with advice; more subtly, they could be earned if the banker succumbs to the temptation to build her reputation at the expense of her clients by performing excessively complex deals. Ely and Välimäki (2003) demonstrate that this type of "bad reputation" concern arises whenever technically able advisers are better informed than their clients.[10] Securities issuance is only possible if a solution can be found to this agency problem. Much of traditional investment banking relies upon promises and tacit understandings that are hard-to-codify and probably impossible-to-verify and, hence, investment bank clients cannot usually rely upon contract alone.

Our second assumption is that bank-client relationships evolved as another way to address agency problems between bank and client. The relationship underpins a tacit, extra-legal, promise by the bank to work in its clients' best interests. Investment bank relationships are therefore necessary precisely because it is impossible to write formal contracts to govern the quality of investment bank advice. Relationships are sustained because, on the one hand, clients are willing to pay a

[9] This assumption is in contrast to the relationship lending literature, which usually stresses the relative informational advantage that borrowers have over their banks. Skilled lenders address the associated agency problems by screening their borrowers ex ante, and monitoring their performance after loans have been extended. See, for example, Boot (2000). The type of knowledge studied in this literature concerns the nature of the borrower, rather than the congruence between the borrower and the products it receives from its investment bank. Like commercial banks, investment banks are better able to check client quality than other investments and, hence, can have a certification role, as in studies by Booth and Smith (1986), Titman and Trueman (1986), Carter and Manaster (1990), and Chemmanur and Fulghieri (1994).

[10] See also Morris (2001), Ely, Fudenberg, and Levine (2008), Bolton, Freixas, and Shapiro (2007) and Chen, Morrison, and Wilhelm (2014), (2015).

premium for as long as they receive reliable advice and, on the other hand, because bankers will provide that advice so as to prevent damage to the relationship and, with it, to their relationship rents. The relational contract over service quality can be sustained provided the bank is sufficiently patient, so that it is more concerned for its long-term rents than for the short-run gains of breaking the contract, and provided also that there is a reasonable chance that the client will find out if the bank breaks its promise.[11]

Last, we assume that, while the relationships of the previous paragraph may maximize the investment bank's profits, they need not maximize an individual banker's utility. It follows immediately that relational contracts between investment banks and their clients are only possible if corporate governance systems in the investment bank incentivize individual bankers to maintain the bank's relationships. That is, bank-client relationships are effective, and therefore survive, only if the bank's internal governance systems are effective.

An issuer's decision to maintain a bank relationship reflects a trade-off. On the one hand, a strong bank relationship addresses bank-client agency problems. On the other hand, maintaining an existing relationship prevents the issuer from realizing any benefits from seeking competitive bids for its underwriting mandate. As we argue below, those benefits could include lower issuance fees and a closer match between bank skill and issuer requirements, but our analysis hinges upon the existence of a tradeoff, and not upon the specific nature of the benefits of relationship breaking.

Our analysis suggests that any change in the internal governance of investment banks that weakens their ability to control internal agency problems will weaken their ability to commit to relationships and, hence, undermine their value. Clients should therefore respond to weaker internal bank governance by breaking their banking relationships. The final building block of our theory is the claim that it is harder to govern a more complex institution. We therefore estimate equations of the following form:

$$\mathbb{P}\left[\text{Relationship breaks}\right] = \beta_1 \mathit{Complexity} + \mathbf{x}\beta. \tag{1}$$

[11] See, for example, Levin (2003, Theorem 6). Eccles and Crane (1988) provide evidence of investment banks exercising such patience during the early part of our sample period. Banks routinely provided advisory services in the expectation of future compensation from underwriting mandates.

We experiment with three proxies for *Complexity*: total bank capitalization (*Capital*), number of partners or senior officers (*Partners*), and a discrete variable that measures whether or not the bank has raised external equity (*Public*). The first two proxies are direct measures of bank scale and the primary focus of our analysis. The last is motivated by Morrison and Wilhelm's (2008) observation that, when investment banks choose to expand aggressively, they use external equity funding to do so. We describe the data used to measure each proxy in Section 2.

Our theory predicts $\beta_1 > 0$. Note that our proxies for bank complexity are likely to associate greater complexity with banks that offer a wider range of services and can draw from a larger pool of human resources for their delivery. A common refrain during the latter part of our sample period, especially among among elite mergers and acqusition bankers, was that conflicts of interest across divisions of large, full-service investment banks posed a serious threat to client relationships. This is precisely the sort of tension that we intend to capture in these proxies. If our prediction is borne out by the data, then any benefits associated with greater scale and scope of operations will have been outweighed by the costs of complexity identified by our theory.[12]

Equation (1) does not identify the effect of agency problems within investment banks on client decisions to continue a banking relationship. Our proxies for *Complexity* are subject to measurement error. Moreover, we cannot rule out the possibility of omitted variables or simultaneity in a client's decision to continue a banking relationship and bank decisions regarding organizational complexity. We address the endogeneity problem by constructing an instrument that seeks to measure incentives for technology adoption and, hence, the complexity of the adopting bank, but has no direct effect upon the issuer's decision to continue its investment banking relationship.

We assume that incentives for technology adoption vary in the cross-section of banks to the extent that their senior bankers, or partners, have different preferences. One reason to believe that preferences may be heterogenous is advanced by Greenwood and Jovanovic (1999), who suggest that agents may resist adoption of technologies that would devalue human capital formed during an earlier technological regime. It follows that, if decision-taking powers in a bank are mostly held

[12]Chen, Morrison, and Wilhelm (2015) explicitly model this tension and demonstrate how it motivates elite bankers to leave full-service banks to found narrow "boutique" advisory banks. The first prominent example occurred in 1988 when Bruce Wasserstein and Joseph Parella left First Boston to start Wassserstein Parella.

by partners whose human capital is of less recent vintage, then that bank is less likely to adopt new technology.

Our instrument, *Technology Exposure*, reflects this reasoning. We begin with an annual measure of the natural log of the *minimum* cost to date per million computations per second (in 2006 constant dollars) based on data compiled by Nordhaus (2007).[13] We compute a technology state variable by averaging this measure over the three years prior to every year in our sample.[14] Each partner is assigned the value of the state variable for the cohort year that he was admitted to the partnership. We then calculate the average partner state variable, which is equivalent to taking a cohort size-weighted average of the annual value of each partner's technology state variable. Finally, for ease of interpretation, we define *Technology Exposure* by reversing the sign of this quantity, so that higher levels correspond to lower computational costs and banks with relatively young partnership cohort structures. Assuming that new technology is more complementary to human capital of recent vintage, we predict that our measures of bank complexity will be increasing in *Technology Exposure*.

Figure 1 shows the annual equally-weighted average of bank *Technology Exposure* from 1960 to 1998. The time trend is dominated by the declining cost of computation (for which we reversed the sign). The one-standard-deviation bands around the average show that cross-sectional variation in the instrument is declining over time. In large part, this is a consequence of a declining number of banks in the sample for reasons described in Section 2.2. If we remove the cost of computation from the construction of the instrument, we are left with each bank's partner cohort structure in a given year. This is easily translated into the average tenure of a bank's partners in a given year. The annual average across banks is shown in Figure 2 which provides a clear picture of the source of cross-sectional variation in the instrument. Average partner tenure declines until the early 1970s

[13]Figure A.1 in the appendix shows the decline in this measure of the cost of computation over the sample period. The raw data underlying the series are summarized in Figure 3 in Nordhaus (2007) and were downloaded from http://www.econ.yale.edu/ nordhaus/Computers/Appendix.xls where the data series appears as "Cost per million computations (2006 $)" in the "Data" page of Appendix.xls. It is worth noting that over the 1966-2006 period our time series has a correlation of 0.92 with the natural log of the Bureau of Economic Analysis' chain-type quantity index of the net capital stock of mainframes and PCs held by firms in the Securities, Commodity Contracts, and Investments sector (BEA Code 5230).

[14]Our results are insensitive to alternative specifications of the measurement window. **Summarize what we've experimented with.**

and then begins a steady increase. As was clear in Figure 1, variation across banks generally declined over time.

Technology Exposure is highly correlated with both log(*Capital*) (0.76) and *Partners* (0.57). In Section 3.1 we show that partial correlations between *Technology Exposure* and proxies for bank complexity in the first-stage regressions are statistically different from zero at the 1% level. Moreover, Cragg-Donald F-statistics for the first-stage regressions range from 21.02 to 68.25. Thus we can reject, by a fairly wide margin, the hypothesis that *Technology Exposure* is a weak instrument judged by the relevant Stock and Yogo (2005) criterion. In summary, *Technology Exposure* appears to be a relatively strong instrument for our proxies for bank complexity.

Satisfying the exclusion restriction for instrument validity requires that *Technology Exposure* be uncorrelated with the error term in regressions taking the form of Equation 1. We include time fixed effects in our model specifications to absorb any time trend that would otherwise appear in the error term and correlate with the time series behavior in *Technology Exposure* shown in figure 1. Thus our primary concern should be whether there is a source of variation in the error term that is correlated with the cross-sectional and residual time series variation in average partner tenure summarized in figure 2.

Our theory provides a strong foundation for our proxies for bank complexity but does not provide any guidance regarding the precise nature or magnitude of measurement error. Similarly, to the extent that the model suffers from simultaneity bias, we have no insight regarding its potential influence on the error term. Our theory does suggest primary concern lies with our inability to measure the influence of the state of the formal contracting environment on issuer decisions. Out theory makes clear that a banking relationship is especially valuable when formal contract is weak. In general, we would expect the formal contracting environment to improve with advances in information technology and thus, potentially diminish the value of a banking relationship. However, while technological change almost surely improved performance measurement and verification in bank's brokerage and risk-taking functions, these are the functions that were likely the greatest source of governance problems. Our proxies for bank complexity are intended to capture these effects.

On the other hand, we do not control for the possibility that an improvement in the contracting environment would improve formal contracting between issuers and their banks and thereby diminish the relative value of the relational contract. This does not strike us as a serious issue because the nature of advisory work has change relatively little over time. Specifically, it remains difficult to measure the quality of advice and service and even more difficult to verify that the bank was not acting in the best interest of the issuer. With that said, this is perhaps the greatest potential challenge to the exclusion restriction to which our theory can speak.

2. Data and Variable Construction

Our unit of observation is a securities transaction for which the issuer engaged one or more of 30 banks described below to manage its *previous transaction*. We refer to the bank(s) that managed the issuer's last transaction as its "relationship bank(s)" and estimate models of the issuer's propensity for switching away from the relationship bank for the present transaction. In the remainder of this section we describe the bank and transaction sample and describe our measures of bank complexity, the "state" of bank-client relationships, and a battery of control variables.

2.1. Transaction Sample

The transaction sample includes public and private underwritten common equity and debt offerings by U.S. issuers between January 1960 and December 1998. Additionally, we draw on transaction data from 1933-1959 to construct several variables described below. Transaction data from 1970 forward are collected from the Thomson Reuters SDC database. Pre-1970 data are collected from Issuer Summaries (1933-1949) prepared by counsel for several defendants in *United States v. Henry S. Morgan, et al* and Investment Dealers' Digest, Corporate Financing (1950-1960 and 1960-1969).[15]

The 1975-2003 sample period studied by Asker and Ljungqvist (2010) is the closet comparable to ours. For the sake of comparison, we follow their lead in screening out financial and governmen-

[15]See Morrison, Thegeya, Schenone, and Wilhelm (2018) for details.

tal issues. Panel 1 in Table I shows that this screening criterion yields 52,883 transactions between 1960 and 1998 that raised \$5.1 trillion in proceeds (all dollar values are in 1996 GDP-deflated constant dollars).[16] All annual market share measures that we use in the paper are calculated relative to this "full sample". By comparison, Asker and Ljungqvist (2010) report 50,128 deals over the 1975-2003 period with proceeds of about \$4.7 trillion. Equity offerings account for 44% of transactions and 23% of proceeds in our sample versus 39% of transactions and 26% of proceeds in the Asker and Ljungqvist (2010) sample.

Panel 2 reports characteristics of the sample used to estimate the switching model. For a transactions to be included in the "estimation sample" the issuer must have had at least one prior transaction after January 1, 1930 and its *last transaction* must have been led by at least one of the 30 sample banks described below. Compustat coverage is less comprehensive during the early part of our sample period thus limiting our ability to consistently measure issuer characteristics. Rather than exclude issuers for lack of Compustat coverage we impose the weaker requirement that the issuer's 2-digit SIC code be available. We then use an industry fixed effect to complement observable characteristics of the issuer's transaction to control for issuer characteristics.

These requirements yield an estimation sample of 16,280 transactions that raised \$2.2 trillion in proceeds. Public and private equity issues account for 28% of transactions and 19%, or about \$415 billion, of proceeds. Among the equity transactions, 206 were initial public equity offerings that accounted for 1% of total proceeds. Public and private debt (including preferred equity) issues raised about \$1.8 trillion. Private offerings (both debt and equity) account for 24% of transactions and 15% of proceeds. The mean (median) number of transactions per year was 426 (371). The maximum of 948 occurred in 1986 and only 155 transactions took place in 1998. The reason for the small number of sample transactions from 1998 will be clear when we describe the bank sample in the next section.

Finally, on average 45% of issuers switched away from their relationship bank in a given year with a minimum switching frequency of 21% in 1970 and a maximum of 60% in 1988. Figure 3 shows that the 3-year moving average of switching frequency increased over the sample period,

[16]There are some instances in which the issuer carries out more than one transaction on the same day. In such cases, we treat the bundle of transactions as a single transaction.

was less than 40% prior to 1973 and remained above 49% from 1985 forward. Fernando, Gatchev, and Spindt (2005) find a similar pattern from 1970-2000 in issuers' first seasoned equity offering following their IPO. The increase in switching also is consistent with the long-run decline in relationship exclusivity documented by Morrison, Thegeya, Schenone, and Wilhelm (2018, Figure 1).

2.2. Bank Sample

Table II reports the 30 sample banks ranked by the number of deals they led in the sample of 16,280 transactions reported in Table I, panel 2. There are 344 transactions for which two sample banks served as lead underwriter and 3 transactions for which three sample banks were identified as a lead underwriter. Thus the transactions collectively led by these banks yield 16,630 observations.

With one exception, all of the sample banks were New York Stock Exchange (NYSE) member firms prior to 1970, the year during which the NYSE lifted its prohibition on member firms being publicly listed.[17] The sample is representative of a broad cross-section of large, full-service banks with relatively large retail brokerage operations (e.g., Merrill Lynch), large, full-service banks with a predominantly institutional focus (e.g., Goldman Sachs), and smaller, more specialized banks dealing with both large and middle market clients (e.g., Lazard, William Blair).[18]

The mean (median) bank led 554 (279) transactions worth \$76 (\$31) billion. Both the mean and median client-switching frequency is 44%. Among the ten banks that remained in the sample for at least 35 years, the average switching frequency is slightly higher at 45%. Among these ten banks, William Blair had the lowest switching frequency (31%) while Salomon Brothers had the highest (59%).

The number of years that a bank appears in the sample from 1960 through 1998 ranges from 7 to 39 with a mean (median) of 26 (29) years. Variation across banks occurs for several reasons.

[17] First Boston had long been publicly listed and became a member firm in 1971.

[18] Although both U.S. and non-domestic universal banks were active in securities underwriting during the 1990s, they do not appear in our sample because we could find no source that would enable us to track the career histories of senior officers with status similar to that of invetment-banking partners. It is worth noting, however, that U.S. commercial banks played only a modest role prior to 1992. U.S. commercial banks collectively accounted for less than 10% of debt proceeds through 1994 and less than 17% in 1997. As late as 1997, they accounted for less than 4% of equity issues.

The first year in which banks appear in the sample is determined by the availability of information required to form the partner cohort structure for the instrument described in Section 1. In 23 cases, we were able to identify each bank's partners in 1960 and determine their first year of service as a partner.[19]

Banks leave the sample before 1998 for two reasons. Some were acquired by another sample bank. For example, Merrill Lynch acquired Goodbody in 1970 and White Weld in 1978.[20] Alternatively, banks leave the sample because we are unable to maintain consistency in identification of the bank's partners or senior officers. As long as they operated as partnerships, bank reporting standards were consistent from year to year. After 1970, most of the sample banks either went public (e.g., Merrill Lynch), continued to operate under the same name following acquisition by a publicly-listed entity (e.g., Salomon Brothers), or acquired a substantial private equity infusion (e.g., Goldman Sachs).[21] These are the criteria for coding $Public = 1$, reported in the last column of Table III, that we use as a proxy for complexity in Section 4. But going public or being acquired by a publicly-listed firm also typically led, at some point, to a change in reporting standards for the bank's senior officers. At one extreme, Merrill Lynch went public in 1971 but maintained relatively consistent reporting standards through 1988 (its last year in the sample). In contrast, Morgan Stanley began identifying a much smaller number of senior bankers immediately following its public offering in 1986 (its last year in the bank sample).

Six banks did not meet one of the criteria for coding $Public = 1$. William Blair, Cowen, Goodbody, and Lazard remained private partnerships and did not raise substantial external equity throughout their time of inclusion in the sample. Hayden Stone was dropped from the sample before being merged into Shearson in 1975. First Boston merged with Credit Suisse in 1988, but

[19]Data on bank partners or their post-partner analog was collected from the New York Stock Exchange's annual Member Firm Directories, annual issues of Standard & Poors' Securities Dealers of North America, and through historical news searches. The earliest partner cohorts date to the first decade of the 20th century. The longest standing partner in our sample at 1960 was B.H. Griswold, Jr., who joined the Alex. Brown partnership in 1904.

[20]In such cases, the acquired bank's partner (and client) histories are merged into those of the acquiring sample bank.

[21]Goldman accepted a $500 million private equity investment from Sumitomo Bank and raised additional private capital from a Hawaiian education trust and a group of private insurers at a time when the partners' capital was around $1 billion (See Endlich (2013, pp. 9-15)). Also see Ljungqvist, Marston, and Wilhelm (2006, p. 310, Figure 1), Morrison and Wilhelm (2007, Figure 1, p. 298), and Morrison and Wilhelm (2008, Table I, p. 327) for further details of the timing and nature of organizational change among investment banks.

each bank was already quite large, heavily capitalized, and offered a wide range of services to its clients. Thus we do not believe that this organizational change was comparable to others in the sample and exclude First Boston from the estimation sample used in Section 4.

2.3. Explanatory Variables

Table III provides summary statistics for variables included in the model of client switching propensity as well as for the instruments described in Section 1. Measures of "bank" or "bank-client" characteristics refer to the relationship bank for the issuer whose transaction appears in the estimation sample described in Table I. For each transaction, the relationship bank is defined as any bank that played a management role in the issuer's preceding transaction.

2.3.1. Bank Characteristics

The first panel of Table III provides summary statistics for characteristics of each issuer's relationship bank. *Capital* is the relationship bank's equity plus long-term debt in 1996 constant dollars measured during the year of the current transaction.[22] The wide spread between the mean ($3,063.10m) and median ($702.88m) capitalization reflects substantial cross-sectional variation among the sample banks as well as the large, general increase in the scale and scope of bank operations over the long sample period.

Partners is the number of partners or senior officers reported by the issuer's relationship bank during the year of the transaction at hand. Although the difference in the mean (161.52) and median (127) is not as large as for *Capital*, the wide variation evidenced by the difference between the minimum and maximum values again is a reflection of the long sample period as well as cross-sectional variation among banks.

For each transaction, we also measure the relationship bank's debt and equity market share during the year preceding the issuer's transaction. The mean market share for debt and equity

[22] Capital data were collected from annual capitalization rankings published by Finance magazine (prior to 1978) and the Securities Industry Association (from 1978 forward). It is worth noting that the NYSE imposed new net capital rules on member firms effective August 1, 1971 requiring at least $1 of capital for each $10 (as opposed to $15) of aggregate debt. For details see Finance, March 1972 and Seligman (1982, p.458).

is 7% as is the median. Market share is often interpreted as a measure of a bank's market-wide reputation. To the extent that these measures are correlated with a bank's bilateral reputation within a client relationship, we expect them to be negatively correlated with switching propensity.[23]

2.3.2. Bank-Client Characteristics

The second panel in Table III reports summary statistics for two measures of the state of bank-client relationships. *RelationshipStrength* is the bank's share (scaled 0-1) of the dollar value of a client's securities issued during the preceding 7 years.[24] As such, it is a measure of the client's history with the bank at the point of its decision whether or not to break the relationship. In the event that one sample bank is acquired by another, the surviving bank inherits the relationships of the acquired bank.[25] The sample mean (median) value of *RelationshipStrength* is 0.48 (0.43).[26] The extant literature in which this measure is used suggests a strong existing relationship will be a moderating force against any incentive a client may have to break its banking relationship. In some specifications of the switching model we also interact *RelationshipStrength* with *Capital* or *Partners* to estimate the extent to which this moderating force is undermined by increasing complexity within the relationship bank.

SIC Share measures the client's share of total proceeds (inclusive of the client's proceeds) raised by the relationship bank for firms in the client's 2-digit SIC code industry during the preceding seven years. The variable (scaled from 0-1) has a mean (median) value of 0.23 (0.08). *SIC Share* is intended to control for the tension between potential benefits from a bank having industry expertise (Morrison, Thegeya, Schenone, and Wilhelm 2018) and concern for conflicts of interest stemming from strong ties to the client's competitors (Asker and Ljungqvist 2010). In light

[23]See Krigman, Shaw, and Womack (2001) for evidence that issuers prefer more prestigious banks.

[24]When multiple banks manage a transaction, each bank is assigned full credit for measurement purposes. **Also report alternative lags with which we have experimented.**

[25]See Ljungqvist, Marston, and Wilhelm (2006, 2009), Asker and Ljungqvist (2010), and Morrison, Thegeya, Schenone, and Wilhelm (2018) for details.

[26]About 25% (4,155) of the 16,630 observations in the estimation sample are cases where the issuer did not do a deal with its "relationship bank" during the preceding 7 years but did at least one deal with the bank after January 1, 1930. In these cases, *RelationshipStrength* is set to its minimum value of zero. Among these cases, 43% (1,790/4,155) switched from the bank that underwrote its last transactions. There is no obvious time pattern in switching frequency among these transactions.

of this tradeoff, the expected net effect on switching propensity is ambiguous.

2.3.3. Transaction and Client Characteristics

A final set of variables control for transaction and client characteristics. *Proceeds* is the dollar amount of securities sold in the transaction measured in 1996 constant dollars. The mean value ($137m) of *Proceeds* is much larger than the median amount ($75m). Again, this reflects both the wide range of issuers in our sample as well as their increasing scale through time. If bank-client relationships involve an element of "quality" matching (Fernando, Gatchev, and Spindt 2005) and large issuers rank higher on the quality dimension, we would expect them to have more options among banks and greater bargaining power, each of which might cut against an existing relationship at the margin.

Last Deal measures the number of years since the client's last transaction. Although the median value of *Last Deal* is one, it ranges from 0 to 40 years, where the former indicates instances in which the issuer carried out two (or more) transactions in the same year. This variable is intended to control for two things. First, to the extent that the relational contract between the client and the bank that managed its last transaction is sustained by expectations of a rent stream that motivates the bank to place the client's interests before its own, we expect clients that dealt more recently with the bank to have greater confidence in the bank's commitment to the relationship. Moreover, although *RelationshipStrength* is intended to measure the state of a bank-client relationship, this measure is likely to be less meaningful in cases where the client's last experience with the bank is more distant. Regardless of the interpretion, we expect switching propensity to be increasing in *Last Deal*.[27]

Client Deal Experience measures the number of client transactions (managed by any bank) from 1930 to the present transaction. It is intended to control for the possibility that more experienced issuers will be less dependent on their bank. It is also likely to be correlated with firm age, which we are unable to measure for a number of the issuers in our sample. The number of prior transactions ranges from 1 to 157, with mean (median) of 17.07 (9.00).

[27] See Fernando, Gatchev, and Spindt (2005) for related evidence.

Finally, *Equity*, *Public Offering*, and *IPO* are binary variables intended to control for differences between types of transactions. In each case, the transaction type identified by the variable name is likely to be subject to more severe informational friction than its alternative. It is conceivable that preserving a relationship in such cases would improve certification but also plausible, especially in the case of IPOs, that the issuer will require certification from a bank that has a stronger reputation in the marketplace than does its relationship bank.[28]

3. The Effects of Bank Complexity on Client Relationships

3.1. Linear Probability Model First-Stage Regressions

We begin with a description of the first-stage regression model in which either log(*Capital*) or *Partners* is the (endogenous) proxy for bank complexity. In columns 1 and 4 we report results for second-stage specifications in which there is no interaction between the proxy for complexity and *RelationshipStrength*. We also estimate second-stage specifications in which the proxy for complexity is interacted with *RelationshipStrength*. In such cases, we must estimate two first-stage regressions. The first regresses the endogenous variable on *Technology Exposure* and the interaction of *Technology Exposure* with *RelationshipStrength*. The second regresses the interaction of the endogenous variable and *RelationshipStrength* on *Technology Exposure* and *Technology Exposure* interacted with *RelationshipStrength*.[29] These results are reported in columns 2 and 3 for the second-stage specification in which log(*Capital*) is the endogenous proxy for bank complexity and in columns 5 and 6 where *Partners* is the second-stage proxy.

The coefficients estimated for *Technology Exposure* in columns 1, 2, 4 and 5 have the predicted positive sign, are insensitive to whether or not there is an interaction with *RelationshipStrength*, and statistically significant. We interpret the positive coefficient estimates for *Technology Exposure* as an indication that bank complexity is increasing in the propensity for technology adoption embod-

[28] One might argue for estimating separate models for debt and equity transactions under the assumption that there exists a degree of independence among business units within the bank. However, our motivation for bank-client relationships is that agency problems within the bank at large undermine its ability to commit to a client relationship.

[29] Failing to instrument for this second-stage interaction is equivalent to incorrectly assuming that the linear projection of the interaction is equivalent to the interaction of the linear projections of each variable. See Wooldridge (2010, pp. 236-7) for discussion of this "forbidden regression" problem.

ied in the bank's partner cohort structure. For both log(*Capital*) and *Partners*, partial correlations with the instrument(s) are statistically different from zero at the 1% level. The Cragg-Donald F=statistics suggest that *Technology Exposure* is a relatively strong instrument.

3.2. Second-Stage Switching Regressions

Table V reports coefficients and robust standard errors (in parentheses) for ordinary least squares (OLS) and second-stage linear probability models (LPMs) of client switching in which either log(*Capital*) or *Partners* is the endogenous proxy for bank complexity. Each regression includes bank, year, and 2-digit SIC fixed effects.

Coefficients for the control variables are uniformly statistically signficant, with the exception of the lagged value of the bank's debt market share and *Last Deal*, and they generally conform with expectations based on prior work. Banks with large (lagged) equity and debt market share, often interpreted as a proxy for a strong market-wide reputation (Megginson and Weiss 1991), are less exposed to client switching. Issuers are less likely to break their relationship when their business accounts for a large share of the relationship bank's business in their 2-digit SIC industry (*SIC Share*). This result is consistent with clients being disinclined to share a bank with its primary competitors (Asker and Ljungqvist 2010). Issuers also are less likely to break their relationship when undertaking (non-IPO) equity offerings (*Equity* $= 1$) as opposed to debt offerings and when their offering is public as opposed to private (*Public Offering* $= 1$). Assuming that equity and public offerings are more susceptible to asymmetric information, these results are consistent with clients preserving a relationship in the interest of more credible certification of their quality. Greater propensity for switching in large transactions (*Proceeds*) is consistent with our conjecture that large issuers rank higher on the quality dimension and thereby have more options among banks and greater bargaining power. Higher switching propensity among IPOs is consistent with demand for certification exceeding the capacity of the client's relationship bank. Finally, more active participants in the capital markets (*Public Offering*) and those for whom there has been a longer period of time since their last transaction (*Last Deal*) are more likely to switch. The signs, magnitudes, and statistical significance of these control variables are generally insentive to model specification throughout Table V.

The OLS results for log(*Capital*) in columns 1 and 2 provide a point of comparison for the 2SLS results. In each OLS specification, the coefficient for log(*Capital*) is negative but, at most, only marginally statistically different from zero. The negative sign on *RelationshipStrength* is consistent with the existing literature which finds that issuers are less likely to switch away from or more likely to select a bank with which it has a strong relationship. The positive and statistically significant coefficient associated with the interaction between *RelationshipStrength* and log(*Capital*) suggests that bank complexity undermines the value of an existing relationship. Again, the coefficient for *RelationshipStrength* is negative and of similar magnitude to the corresponding OLS estimate in column 1.

Column 3 reports second-stage results for log(*Capital*). The first noteworthy result is that the estimated coefficient for log(*Capital*) (0.4155) is large relative to the OLS coefficient (-0.0034) and is statistically signficant at the 1% level. The positive sign is consistent with our theoretical framework's prediction that technological change undermines bank governance and therefore client relationships: a 1% increase in capital corresponds with an average increase in switching propensity of 0.42%. The coefficient estimate for *RelationshipStrength* is virtually identical to the OLS analog and it remains statistically significant at the 1% level.

In column 4 we interact log(*Capital*) with *RelationshipStrength*. At the 0.43 median level of *RelationshipStrength*, the marginal effect of log(*Capital*), 0.3121 (i.e., 0.2844 + [0.0637 * 0.43]), is now somewhat smaller than in the absence of the interaction. At the median level of log(*Capital*), the marginal effect of *RelationshipStrength* is -0.1632 (-0.5808 + (0.0637 * 6.56), or nearly identical to the marginal effect in the absence of the interaction. The positive coefficient for the interaction term indicates that greater organizational complexity diminishes the moderating effect of an existing relationship on client switching propensity. Moving to the 75th percentile level of log(*Capital*), the marginal effect of *RelationshipStrength* declines by roughly 60% in absolute value to -0.0631. At the 95th percentile level of the instrumented value of log(*Capital*), the marginal effect of *RelationshipStrength* on switching propensity is only marginally statistically different from zero (p = 0.102). In other words, at high levels of bank complexity, the strength of an existing relationship has little effect on the issuer's decision whether to break the relationship.

The results in columns 5-8 show that using *Partners* as the proxy for bank complexity yields qualitatively similar results. Moving immediately to column 8 where the model includes the interaction between *Partners* and *RelationshipStrength*, the marginal effect of *Partners* (at the median level of *RelationshipStrength*) is 0.0025, indicating that an additional partner corresponds with an average increase in switching propensity of 0.25%. The marginal effect of *RelationshipStrength* (at the median level of 127 bank partners) is -0.1996, or somewhat larger in absolute value than when using log(*Capital*) as the proxy for bank complexity. However, it more sensitive to an increase in complexity as it declines in absolute value by about 70% at the 75th percentile level of 242 partners. Once again, at the 95th perecentile level of *Partners* the marginal effect of *RelationshipStrength* is not statistically different from zero.

We noted in the discussion of Table II that some of the explanatory variables have relatively wide value ranges. In such cases, there is a greater likelihood that predicted probabilities from the LPM specifications will lie outside the unit interval and, potentially, lead to poor estimates of marginal effects averaged across the distribution of the explanatory variables (Wooldridge 2010, p. 563). Table VI shows that the results using either log(*Capital*) or *Partners* as the proxy for bank complexity are not sensitive to the LPM specification. For ease of comparison, columns 1 and 3 in Table VI repeats the results from columns 4 and 8 in Table V. Columns 2 and 4 report *marginal effects* from maximum likelihood estimation of instrumental variables (IV) probit specifications.[30] The marginal effects from the probit models are not meaningfully different from those obtained with the LPM specifications. It is also worth noting that maximum likelihood estimation of the IV probit model enables a Wald test of the exogeneity of log(*Capital*) or *Partners*. The *p*-values for these tests reported in Table VI provide strong evidence against treating log(*Capital*) or *Partners* as exogenous.

In summary, the 2SLS results reported in Tables V and VI are relatively insensitive to alternative model specifications and consistent with our theoretical framework and prior research. Switching propensity is increasing in both proxies for bank complexity. Consistent with existing

[30]See Wooldridge (2010, p. 591) for details. In the interest of brevity, we do not report first-stage estimation results for the probit specifications but they are qualitatively similar to those reported for the LPM specifications.

research on investment-banking relationships, issuers with relatively strong relationships are less likely to switch banks. However, increasing bank complexity undermines this effect. At high levels of bank complexity strong relationships have little moderating effect on the issuer's propensity to break its banking relationship.

3.3. Bank Complexity and Positive Assortative Matching

Our theoretical framework predicts that investment banking relationships are valuable because they underpin the formation of the trust that facilitates securities underwriting. But the need to establish trust is only one of the factors that influences the issuer's decision to maintain or to break a banking relationship. As we emphasize in Section 1, issuers trade off the value of the existing relationship against any benefits that could be realized by breaking it. For example, ceteris paribus, the demands of an issuer's transaction might better fit the capabilities of a different bank. In that case, the issuer would rationally choose to break its existing relationship if the cost of destroying the trust inherent in that relationship was outweighed by the benefit from switching to a bank with capabilities that better complement the characteristics of its transaction. In this Section, we examine whether such positive assortative matching influences an issuer's decision to break an existing banking relationship.

Although they do not consider the tradeoff we have just described, Fernando, Gatchev, and Spindt (2005) use data from from 1970-2000 to examine whether switching propensity in initial seasoned equity offerings (SEOs) is related to a measure of the "quality" mismatch between the issuer and the underwriter of its initial public offering at the time of the SEO. In any year t, Fernando *et al.* measure the degree of mismatch by ranking underwriters by the total proceeds of lead-underwritten deals in years $t-2$, $t-1$, and t, and issuers by the total proceeds of issues in year t. The absolute difference in issuer and underwriter percentile ranks can be thought of as a measure of quality mismatch. Fernando *et al.* show that switching propensity is increasing in this measure of mismatch and they interpret it as evidence of positive assortative matching.

Our key hypothesis is that internal agency problems in banks are harder to manage when banks become more complex. It follows immediately that issuer's should be more inclined to switch to achive a batter match at higher levels of bank complexity. We test this hypothesis by constructing a measure, *Mismatch*, that is computed identically to their mismatch measure using data from our estimation sample. Table VII shows results from 2SLS estimation of linear probability models of switching propensity that include *Mismatch* and its interaction with either log(*Capital*) or *Partners*. For comparison purposes, the first and fourth columns are taken from columns 1 and 3 in Table VI.

Note first in columns 2 and 5 that introducing *Mismatch* into the model has little impact on the coefficients for log(*Capital*) and *Partners*. Morever, there is little change in the estimated coefficients for other variables with the exception of log(*Proceeds*). The coefficient for *Mismatch* in the model using log(*Capital*) as the proxy for bank complexity (0.1762) is positive and statistically significant. The coefficient for *Mismatch* in the model using *Partners* as the proxy for bank complexity (0.1820) is quite similar. Thus, like Fernando *et al.*, we find evidence of positive assortative matching in that switching propensity is increasing in the degree of mismatch between the issuer and its relationship bank.

In columns 3 and 6, we interact *Mismatch* with either log(*Capital*) or *Partners*. In the log(*Capital*) model, the marginal effect of *Mismatch* at the median level of log(*Capital*) is 0.1721 (0.0499 + [-0.1533 * 6.6]) and statistically signficant at the 1% level. The 0.1542 marginal effect of *Mismatch* at the median level of *Partners* (127) is similar in magnitude and also statistically significant at the 1% level. In each case, the marginal effect at the 95th percentile is roughly double the median level. In summary, these results are consistent with our hypothesis that the tradeoff between the trust inherent in an investment banking relationship and the benefits from switching to a bank that is a better match with the issuer favors the latter at higher levels of complexity in the relationship bank.

Insert discussion of whether switchers move to more or less complex banks.

4. Robustness to Alternative Model Specifications

4.1. Switching Conditional on Discrete Change in Organizational Structure

As we noted in Section 2.2, the NYSE lifted its prohibition on member firms being publicly listed in 1970. Over the remainder of the sample period, most of the sample banks went public, were acquired by (or merged with) publicly-listed firms, or were otherwise subject to a discrete shock in their capital structure. We code the binary variable *Public* = 1 in the year of the shock and every year thereafter for each bank that was subject to a discrete shock in its capital structure (see table III). Morrison and Wilhelm (2008) develop a model in which discrete organizational change is an optimal response to technological advances that favor greater operating scale. In their model, banking partnerships face a tradeoff between scale efficiencies and the bank's ability to maintain a reputation for mentoring junior bankers in tacit functions such as the preservation of client relationships. At sufficiently high levels of technological development, banking partnerships sacrifice human capital development in favor of further investment in physical capital by going public.

It is tempting to identify these events as exogenous shocks, but they are not. Like our continuous proxies for bank complexity, a bank's decision to go public or raise a large amount of equity from outside the partnership cannot be interpreted as exogenous to the state of its client relationships. Moreover, we cannot rule out the possibility that clients foresee the potential for organizational change in their relationship bank and select in or out of a "quasi-experiment" comparing switching decisions before and after the change. With that said, we can address these problems with our instrumental variable.

To that end, we carry out an exercise in "IPO time" that includes only transactions involving relationship bank(s) for which *Public* = 1 during the sample period. We define the year in which the organizational shock occurred as year t. For each bank, we then construct a 4-year, event-time sample of transactions from years $t-2$ through $t+1$. This sampling procedure yields $2,212$ observations, or roughly 14% of the estimation sample used in the preceding sections.[31]

[31] The sampling procedure for Blyth and Eastman Dillon is complicated by the fact that Blyth was acquired by I.N.A Corporation in January 1970. I.N.A. then acquired Eastman Dillon in 1972 and merged its operations with

As before, we address the endogeneity of *Public* using *Technology Exposure* as an instrumental variable in a first-stage regression. We control for sample-selection bias by estimating a MLE probit selection equation using transactions from time $t-10$ through $t+1$. The inverse Mills ratio obtained from the selection model is then included as a regressor in both the first- and second-stage LPM regressions.[32] Estimation of second-stage regression specifications identical to those reported in Table V yields qualitatively identical results.

4.2. Bank Choice Specifications

The switching models estimated in Section 3 are attractive for their simplicity but they assume that issuers do not condition the assignment of their underwriting mandate on characteristics of banks other than those we define as their relationship bank(s). We have addressed this concern by estimating LPMs in which the issuer selects one or more banks from the full set of banks in our sample at the time of their transaction. This approach brings more information to bear on the issuer's decision, including any history (embodied in *RelationshipStrength*) the issuer had with banks other than the underwriter(s) of itspreceding transaction, as well as concurrent information related to each bank's complexity, market share, and industry expertise (*SIC Share*). It also increases the transaction sample size because it does not exclude transactions for which the issuer had no previous history with a sample bank. On the other hand, most of the banks in the choice set are probably not plausible candidates for any given transaction. Again, we do not report results other than to note that they are qualitatively similar to those obtained from the switching models: issuers are less likely to select organizationally complex banks to underwrite their securities offerings.

Blyth's to form Blyth, Eastman Dillon. For Blyth, we define 1970 as year t and sample transactions from 1968-1971 for which it was the relationship bank. For Eastman Dillon, year t is 1972 and we include transactions from 1970-1973 for which it was the relationship bank.

[32] See Wooldridge (2010, p. 809, 939) for details. The second-stage coefficient for the inverse Mills ratio is not statistically different from zero suggesting that selection bias is not a serious problem.

4.3. Commercial Bank Entry to Securities Underwriting

Although it is a small part of our sample period, it is worth considering whether our results or their interpretation are sensitive to commercial bank entry to securities underwriting following incremental relaxation of Glass-Steagall restrictions beginning in 1987. Two domestic commercial banks, Citicorp and J.P. Morgan appeared among the top 25 debt underwriters in 1987 but accounted for only 0.55% of the dollar value public debt offerings reported by SDC. By 1997, still there were only 5 domestic commercial banks among the top 25, accounting for about 16% of the market. By 1997 only 2 domestic commercial banks were among the top 25 equity underwriters with market share of about 3.5%. Thus, during our sample period, the *direct* effect of domestic commercial bank entry on investment-banking relationships was probably relatively modest. Consistent with this observation, if, for example, we simply end our estimation sample at 1986, 1989, or 1994, there is no meaningful change in our results.

Of course this does not imply that commercial banks had no impact on investment-banking relationships. For example, Chen, Morrison, and Wilhelm (2015, pp. 1181-4) argue that competitive pressure from commercial banks on investment banks' risk-taking functions may have amplified internal conflicts of interest that, in turn, *indirectly* undermined trust with their investment-banking clients. But these risk-taking functions expanded rapidly beginning in the 1980s, in no small part, as a consequence of the same technological changes that we identify at the root of agency problems within investment banks.

5. Conclusion

We examine the effect that an investment bank's internal governance has upon the strength of its client relationships. Our analysis rests upon theories that identify a strong bank-client relationship as an important foundation for mutual trust. Trust is important in investment banking, where information asymmetries are rife and there are few formal devices for addressing them. But an investment bank's ability to sustain trust, and so to earn the resultant relationship rents, is only as good as the governance systems it uses to control opportunistic behavior by its investment bankers.

We therefore hypothesize that, if an investment bank's corporate governance systems weaken, then clients for whom they underwrite securities offerings are more likely to break their relationship with the bank.

We test this hypothesis using three alternative measures of bank complexity and a variety of alternative specifications of the empirical model. In every instance, we find that relationships involving more complex banks are more likely to be broken. Even highly exclusive relationships are unlikely to be preserved at high levels of complexity. We also provide evidence that parties to a relationship trade off the benefits of the relationship against its opportunity costs. Specifically, we show that issuers are more likely to break a relationship in search of a better (positive assortative) match when the relationship bank is more complex.

Investment bankers appear to have experienced a crisis of trust in the last decade: their clients appear no longer to believe that banks can be relied upon to look out for the clients' best interests and, in line with this observation and our theoretical framework, investment banking relationships are weaker and less exclusive than at any time in the past (Morrison, Thegeya, Schenone, and Wilhelm 2018). Our results suggest that regulatory concern that bank complexity contributes to poor governance and loss of trust is well-placed.

References

Asker, J., and A. Ljungqvist, 2010, "Competition and the Structure of Vertical Relationships in Capital Markets," *Journal of Political Economy*, 118, 599–647.

Bertrand, M., and A. Schoar, 2003, "Managing with Style: The Effect of Managers on Firm Policies," *Quarterly Journal of Economics*, 118, 1169–1208.

Bodnaruk, A., M. Massa, and A. Simonov, 2009, "Investment Banks as Insiders and the Market for Corporate Control," *The Review of Financial Studies*, 22, 4989–5026.

Bolton, P., X. Freixas, and J. Shapiro, 2007, "Conflicts of Interest, Information Provision, and Competition in the Financial Services Industry," *Journal of Financial Economics*, 85, 297–330.

Boot, A. W. A., 2000, "Relationship Banking: What Do We Know?," *Journal of Financial Intermediation*, 9, 7–25.

Booth, J. R., and R. Smith, 1986, "Capital Raising, Underwriting and the Certification Hypothesis," *Journal of Financial Economics*, 15, 261–281.

Carter, R. B., and S. Manaster, 1990, "Initial Public Offerings and Underwriter Reputation," *Journal of Finance*, 45, 1045–1067.

Chemmanur, T. J., and P. Fulghieri, 1994, "Investment Bank Reputation, Information Production, and Financial Intermediation," *Journal of Finance*, 49, 57–79.

Chen, Z., A. D. Morrison, and W. J. Wilhelm, Jr., 2014, "Investment Bank Reputation and "Star" Cultures," *The Review of Corporate Finance Studies*, 2, 129–153.

Chen, Z., A. D. Morrison, and W. J. Wilhelm, Jr., 2015, "Traders vs. Relationship Managers: Reputational Conflicts in Full-Service Investment Banks," *Review of Financial Studies*, 28, 1153–1198.

Eccles, R. G., and D. B. Crane, 1988, *Doing Deals: Investment Banks at Work*, Harvard Business Review Press, Boston, Mass.

Eckbo, B. E., R. W. Masulis, and O. Norli, 2005, "Security Offerings," SSRN Scholarly Paper ID 863664, Social Science Research Network, Rochester, NY.

Ely, J., D. Fudenberg, and D. K. Levine, 2008, "When is reputation bad?," *Games and Economic Behavior*, 63, 498–526.

Ely, J. C., and J. Välimäki, 2003, "Bad Reputation," *Quarterly Journal of Economics*, 118, 785–814.

Endlich, L. J., 2013, *Goldman Sachs: The Culture of Success*, Knopf, 1 edition edn.

Fernando, C. S., V. A. Gatchev, and P. A. Spindt, 2005, "Wanna Dance? How Firms and Underwriters Choose Each Other," *Journal of Finance*, 60, 2437–2469.

Greenwood, J., and B. Jovanovic, 1999, "The Information-Technology Revolution and the Stock Market," *American Economic Review*, 89, 116–122.

Griffin, J. M., T. Shu, and S. Topaloglu, 2012, "Examining the Dark Side of Financial Markets: Do Institutions Trade on Information from Investment Bank Connections?," *Review of Financial Studies*, 25, 2155–2188.

Jegadeesh, N., and Y. Tang, 2010, "Institutional Trades Around Takeover Announcements: Skill vs. Inside Information," Working Paper, Emory University.

Jovanovic, B., 1979, "Firm-Specific Capital and Turnover," *Journal of Political Economy*, 87, 1246–1260.

Kahn, R., and T. M. Whited, 2018, "Identification Is Not Causality, and Vice Versa," *Review of Corporate Finance Studies*, 7, 1–21.

Kang, A., and R. Lowery, 2014, "The Pricing of IPO Services and Issues: Theory and Estimation," *The Review of Corporate Finance Studies*, 2, 188–234.

Krigman, L., W. H. Shaw, and K. Womack, 2001, "Why Do Firms Switch Underwriters?," *Journal of Financial Economics*, 60, 245–284.

Levin, J., 2003, "Relational Incentive Contracts," *The American Economic Review*, 93, 835–857.

Ljungqvist, A., 2007, "IPO Underpricing," in B. Espen Eckbo (ed.), *Handbook of Empirical Corporate Finance*Handbooks in Finance, . pp. 375–422, Elsevier, San Diego.

Ljungqvist, A., F. Marston, and W. J. Wilhelm, 2006, "Competing for Securities Underwriting Mandates: Banking Relationships and Analyst Recommendations," *Journal of Finance*, 61, 301–340.

Ljungqvist, A., F. Marston, and W. J. Wilhelm, 2009, "Scaling the Hierarchy: How and Why Investment Banks Compete for Syndicate Co-Management Appointments," *Review of Financial Studies*, 22, 3977–4007.

Ljungqvist, A., and W. J. Wilhelm, 2005, "Does Prospect Theory Explain IPO Market Behavior?," *The Journal of Finance*, 60, 1759–1790.

Malmendier, U., and S. Nagel, 2011, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *Quarterly Journal of Economics*, 126, 373–416.

Malmendier, U., G. Tate, and J. Yan, 2011, "Overconfidence and Early-Life Experiences: The Effect of Managerial Traits on Corporate Financial Policies," *The Journal of Finance*, 66, 1687–1733.

Megginson, W. L., and K. A. Weiss, 1991, "Venture Capitalist Certification in Initial Public Offerings," *Journal of Finance*, 46, 879–903.

Mehran, H., and R. Stulz, 2007, "The Economics of Conflicts of Interest in Financial Institutions," *Journal of Financial Economics*, 85, 267–296.

Morris, S., 2001, "Political Correctness," *Journal of Political Economy*, 109, 231–265.

Morrison, A. D., A. Thegeya, C. Schenone, and W. J. Wilhelm, Jr., 2018, "Investment-Banking Relationships: 1933-2007," *Review of Corporate Finance Studies*, Forthcoming.

Morrison, A. D., and W. J. Wilhelm, Jr., 2007, *Investment Banking: Institutions, Politics, and Law*, Oxford University Press, Oxford.

Morrison, A. D., and W. J. Wilhelm, Jr., 2008, "The Demise of Investment Banking Partnerships: Theory and Evidence," *Journal of Finance*, 63, 311–350.

Nimalendran, M., J. R. Ritter, and D. Zhang, 2007, "Do Today's Trades Affect Tomorrow's IPO Allocations?," *Journal of Financial Economics*, 84, 87–109.

Nordhaus, W. D., 2007, “Two Centuries of Productivity Growth in Computing,” *Journal of Economic History*, 67, 128–159.

Oyer, P., 2008, “The Making of an Investment Banker: Stock Market Shocks, Career Choice, and Lifetime Income,” *Journal of Finance*, 63, 2601–2628.

Pak, S. J., 2013, *Gentlemen Bankers: The World of J. P. Morgan*, Harvard University Press.

Philippon, T., and A. Reshef, 2012, “Wages and Human Capital in the U.S. Finance Industry: 1909–2006,” *Quarterly Journal of Economics*, 127, 1551–1609.

Reuter, J., 2006, “Are IPO Allocations for Sale? Evidence from Mutual Funds,” *Journal of Finance*, 61, 2289–2324.

Schenone, C., 2004, “The Effect of Banking Relationships on the Firm’s IPO Underpricing,” *Journal of Finance*, 59, 2903–2958.

Schoar, A., and L. Zuo, 2017, “Shaped by Booms and Busts: How the Economy Impacts CEO Careers and Management Styles,” *Review of Financial Studies*, 30, 1425–1456.

Seligman, J., 1982, *The Transformation of Wall Street: A History of the Securities and Exchange Commission and Modern Corporate Finance*, Houghton Mifflin Company, Boston, Mass.

Stock, J. H., and M. Yogo, 2005, “Testing for Weak Instruments in Linear IV Regression,” in Donald W. K. Andrews, and James H. Stock (ed.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* . pp. 80–108, Cambridge University Press, New York, NY.

Titman, S., and B. Trueman, 1986, “Information Quality and the Valuation of New Issues,” *Journal of Accounting and Economics*, 8, 159–172.

Wooldridge, J. M., 2010, *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Mass.

Yasuda, A., 2005, “Do Bank Relationships Affect the Firm’s Underwriter Choice in the Corporate-Bond Underwriting Market?,” *Journal of Finance*, 60, 1259–1292.

Yasuda, A., 2007, “Bank Relationships and Underwriter Competition: Evidence from Japan,” *Journal of Financial Economics*, 86, 369–404.

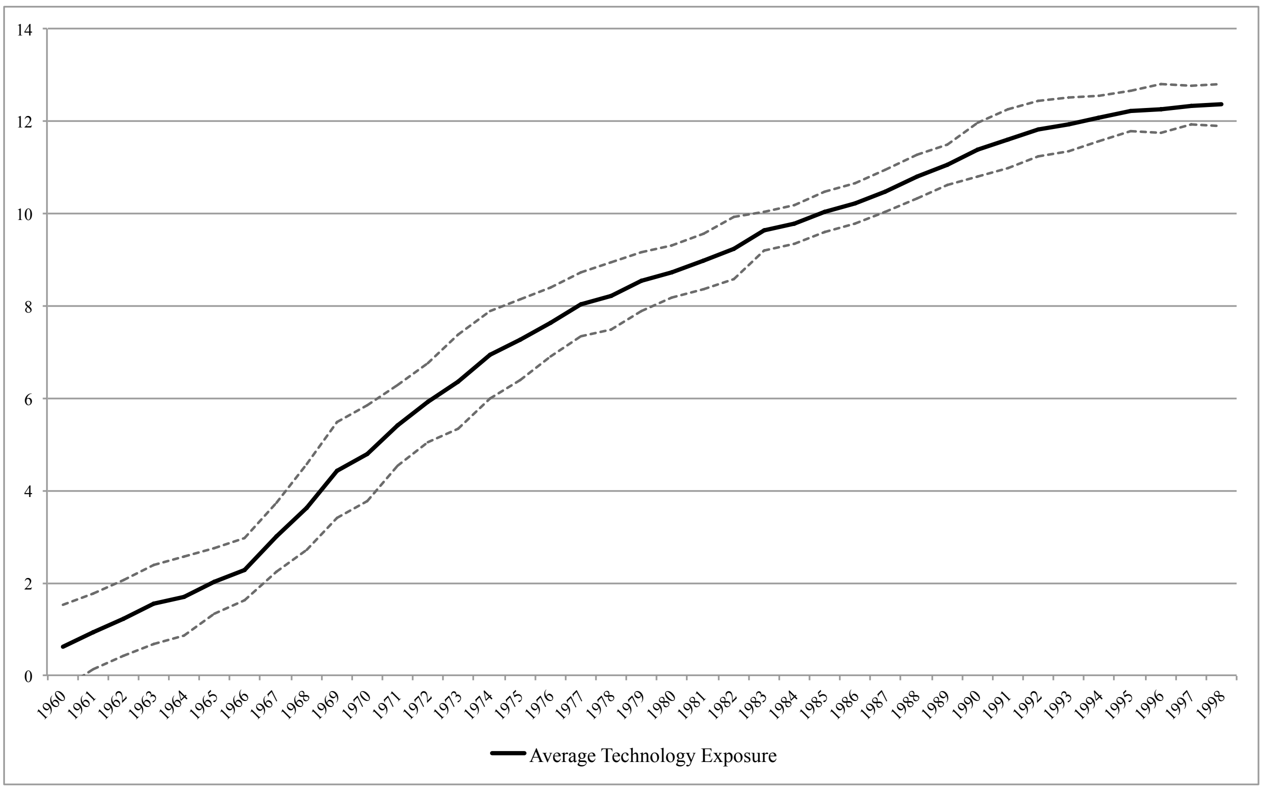


Figure 1: Average Technology Exposure ± one standard deviation.

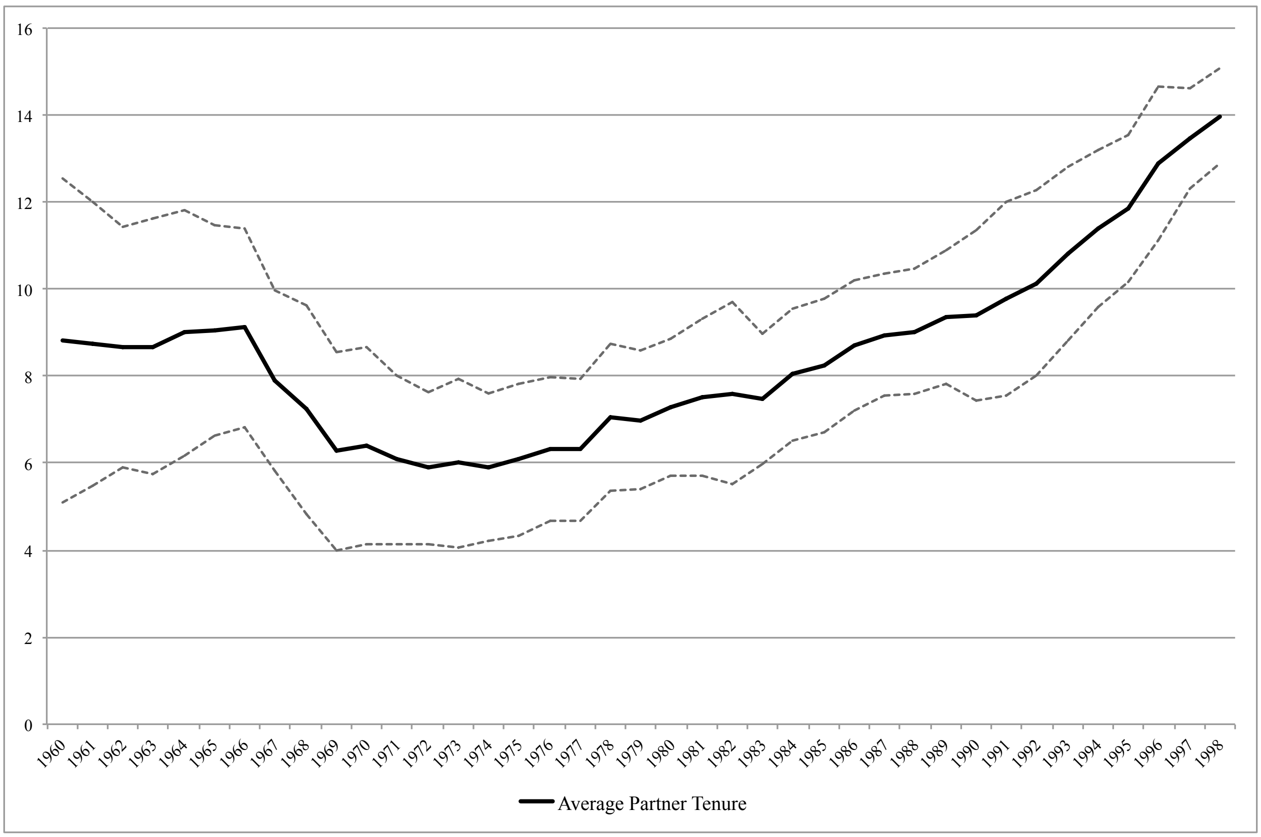


Figure 2: Average Partner Tenure ± one standard deviation.

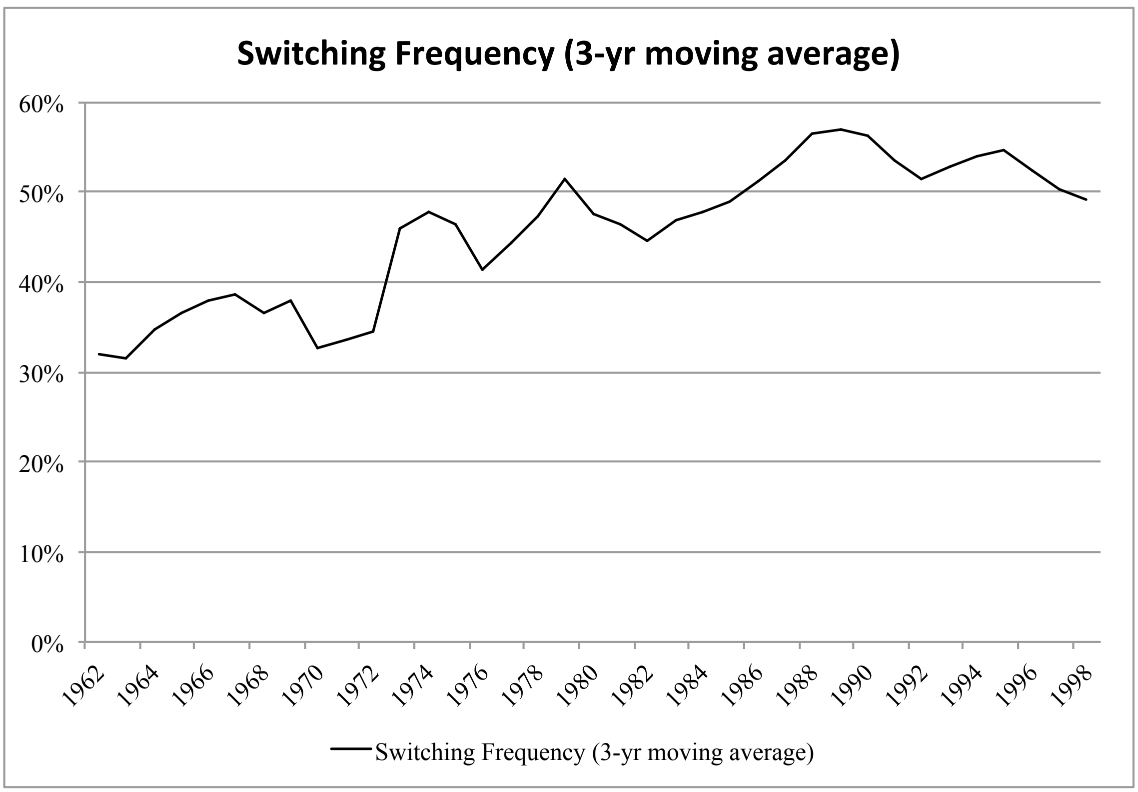


Figure 3: Switching Frequency.

Table I

Transaction Sample Summary Statistics

The table reports summary statistics for underwritten capital-raising transactions by nonfinancial and nongovernmental U.S. issuers between January 1960 and December 1998. Pre-1970 data are collected from *Investment Dealers' Digest, Corporate Financing*. Data from 1970 forward are collected from the *Thomson Reuters SDC* database. The estimation sample includes nonfinancial and nongovernmental issues underwritten by at least one of 30 banks for which we have identified the first year of partnership for each of the bank's partners in 1960. Proceeds are converted to constant 1996 dollars using the annual GDP Deflator.

	Number of Deals	% of Total Deals	Proceeds Raised ($m)	% of Total Proceeds
Panel 1: Full Sample				
Equity Offerings	23,490	44	1,152,730	23
Debt Offerings	29,393	56	3,960,963	77
Total	52,883	100	5,113,693	100
Panel 2: Switching Model Estimation Sample for Transactions led by at least one of 30 Sample Banks				
Equity Offerings	4,604	28%	414,978	19%
Public	4,383	27%	390,497	18%
IPO	206	1%	16,788	1%
Private	221	1%	24,481	1%
Debt Offerings	11,676	72%	1,804,291	81%
Public	6,748	41%	1,364,580	61%
Private	3,495	21%	280,620	13%
Public Preferred	1,151	7%	145,043	7%
Private Preferred	282	2%	14,048	1%
Total	16,280		2,219,269	
	Mean	Median	Minimum	Maximum
Transactions per Year	426	371	155	948
% Issuer Switched Banks	45%	48%	21%	60%

Table II

Sample Banks and Summary Statistics

The table reports summary statistics for 30 banks between January 1960 and December 1998. The total number of *transactions* (16,280) corresponds with panel 2 of Table I. The total number of *observations* (16,630) reflects the presence of 347 transactions with multiple bookrunners (344 with 2 bookrunners, and 3 with 3 bookrunners) from among the 30 sample banks, each of which is given full credit for the transaction. Proceeds are converted to constant 1996 dollars using the annual GDP Deflator.

Banks	Number of Observations	% of Total Observations	Proceeds Raised ($m)	% of Total Proceeds	Switching Frequency	First year in Sample	Last Year in Sample	Years in Sample	First year *Public* = 1
Goldman Sachs	2,143	13	381,682	17	49.0%	1960	1998	39	1986
Salomon Brothers	2,017	12	324,551	14	59.2%	1960	1997	38	1981
First Boston/CSFB	2,006	12	323,993	14	50.4%	1960	1995	36	--
Kidder Peabody	1,381	8	117,650	5	42.5%	1960	1994	35	1986
Merrill Lynch	1,362	8	168,862	7	50.2%	1960	1988	29	1971
Lehman Brothers	1,086	7	141,594	6	43.9%	1960	1991	32	1984
Morgan Stanley	1,045	6	227,232	10	41.1%	1960	1986	27	1986
Paine Webber	953	6	92,265	4	52.6%	1960	1997	38	1972
Smith Barney	549	3	59,896	3	43.7%	1960	1993	34	1987
Dillon Read	498	3	75,342	3	32.4%	1960	1997	38	1986
Dean Witter (Reynolds)	404	2	32,773	1	41.7%	1960	1989	30	1972
Donaldson Lufkin & Jenrette	384	2	65,851	3	44.6%	1969	1998	30	1970
White Weld	366	2	36,419	2	43.7%	1960	1978	19	1978
Alex. Brown	326	2	21,117	1	39.4%	1961	1997	37	1997
Bear Stearns	284	2	27,675	1	53.6%	1960	1997	38	1985
EF Hutton	273	2	21,014	1	50.4%	1960	1987	28	1972
Blyth	269	2	34,585	2	45.0%	1960	1971	12	1970
Blyth, Eastman Dillon	259	2	37,268	2	57.5%	1972	1978	7	1972
Eastman Dillon	203	1	13,269	1	37.3%	1960	1971	12	1972
Lazard	147	1	29,995	1	50.0%	1960	1997	38	--
Kuhn, Loeb	146	1	24,238	1	32.0%	1960	1977	18	1978
Shearson Hammill	135	1	7,021	<1	46.5%	1960	1984	25	1979
William Blair	104	1	4,221	<1	32.0%	1961	1997	37	--
Dupont	80	<1	3,342	<1	62.3%	1960	1972	13	1971
Hornblower	55	<1	2,537	<1	31.9%	1964	1977	14	1977
Loeb Rhoades	52	<1	4,070	<1	29.8%	1960	1979	20	1979
Hayden Stone	46	<1	1,968	<1	25.0%	1960	1972	13	--
Reynolds Securities	34	<1	1,589	<1	30.0%	1960	1977	18	1971
Cowen	13	<1	729	<1	46.2%	1987	1996	10	--
Goodbody	10	<1	290	<1	50.0%	1961	1970	10	--
Mean	554		76,101		44%			26	
Median	279		31,384		44%			29	
Total Number of Observations	16,630		2,283,038						
Total Number of Transactions	16,280		2,219,269						

Table III
Explanatory Variable Summary Statistics

Capital equity plus long-term debt in 1996 dollars reported by the relationship bank in the year (*t*) of the client's transaction. *Partners* is the number of partners or senior officers reported by the relationship bank in year *t*. *Market Share* is a bank's share of the total dollar value of equity or debt in year *t-1*. *Relationship Strength* is a bank's share of the dollar value of a client's securities issued during the seven years preceding year (*t*). *SIC Share* is the client's share of total proceeds (inclusive of the client's proceeds) raised by the bank for firms in the client's 2-digit SIC code industry during the seven years preceding year *t*. *Proceeds* is the dollar value of securities issued in 1996 dollars. *Last Deal* is the number of years since the client's last transaction. Client Deal Experience is the number of deals by the client from 1930 to year *t*. *Equity* = 1 for equity issues. *Public Offering* = 1 for public debt and equity issues. *IPO* = 1 for initial public offerings of equity. *Technology Exposure* is a bank's annual partner-cohort-weighted measure of exposure to an annual index of -log(cost per million computations per second). Default *Exposure* is a bank's annual partner-cohort-weighted measure of exposure to Moody's annual default rate for speculative grade borrowers.

	Obs.	Mean	Median	SD	Min.	Max.
Bank Characteristics						
Capital ($m)	16,630	3063.10	702.88	5287.96	4.67	27162.58
Partners	16,630	161.52	127.00	116.68	4.00	494.00
Market Share Debt(t-1)	16,630	0.07	0.07	0.06	0.00	0.27
Market Share Equity(t-1)	16,630	0.07	0.05	0.06	0.00	0.32
Bank-Client Characteristics						
Relationship Strength	16,630	0.48	0.43	0.41	0.00	1.00
SIC Share	16,630	0.23	0.08	0.30	0.00	1.00
Transaction and Client Characteristics						
Proceeds ($m)	16,630	137.29	75.17	232.39	0.10	5951.00
Last Deal	16,630	1.53	1.00	2.17	0.00	40.00
Client Deal Experience	16,630	17.07	9.00	20.85	1.00	157.00
Equity	16,630	0.28	0.00	0.45	0.00	1.00
Public Offering	16,630	0.69	1.00	0.46	0.00	1.00
IPO	16,630	0.01	0.00	0.12	0.00	1.00
Instruments						
Technology Exposure	16,630	8.56	9.51	3.38	-0.65	12.90
Default Exposure	16,630	1.75	1.71	0.87	0.16	3.72

Table IV

First-Stage Regressions

The table reports first-stage regressions for the client switching model in which *log*(*Capital*) or *Partners* is the (endogenous) dependent variable and *Technology Exposure* is the instrument. For each set of regressions, the first column corresponds with a second-stage regression in which either *log*(*Capital*) or *Partners* is *not* interacted with *Relationship Strength* (*RelStr*). The next two columns report the set of first-stage regressions that correspond with the second-stage model in which *log*(*Capital*) is interacted with *Relationship Strength*. Each regression includes year, bank, and (client) 2-digit SIC fixed effects. Robust standard errors (clustered at the deal level) are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

	log(Capital)			*Partners*		
	(1)	(2)	(3)	(4)	(5)	(6)
Instruments						
Technology Exposure	0.0567***	0.0645***	-0.1203***	7.2466***	7.6596***	-3.9130***
	(0.0092)	(0.0094)	(0.0099)	(0.8771)	(0.8926)	(0.6862)
Bank-Client Characteristics						
Relationship Strength	0.0073	0.1157***	3.2773***	1.3228	7.0498**	-0.8965
	(0.0077)	(0.0229)	(0.0273)	(0.8766)	(2.8293)	(1.9517)
Relationship Strength x Technology Exposure		-0.0127***	0.3843***		-0.6701**	18.8556***
		(0.0024)	(0.0034)		(0.2961)	(0.2469)
SIC Share	0.0483***	0.0499***	0.0034	-3.6893**	-3.6055**	-1.526
	(0.0156)	(0.0156)	(0.0199)	(1.6677)	(1.6691)	(1.6226)
Bank Characteristics						
Market Share Equity(t-1)	1.3598***	1.3582***	0.3298***	-10.1906	-10.2758	2.842
	(0.0659)	(0.0658)	(0.0913)	(7.8774)	(7.8735)	(8.8012)
Market Share Debt(t-1)	4.3629***	4.3422***	2.1601***	22.6757**	21.5772*	-13.5628
	(0.1143)	(0.1142)	(0.1330)	(11.1898)	(11.1668)	(10.2892)
Client and Transaction Characteristics						
log(Proceeds)	-0.0143***	-0.0135***	-0.0110***	1.0548***	1.0922***	-0.6500*
	(0.0030)	(0.0031)	(0.0039)	(0.3458)	(0.3460)	(0.3359)
Last Deal	-0.0026*	-0.0019	0.0042**	-0.2952*	-0.254	-0.3555**
	(0.0015)	(0.0015)	(0.0021)	(0.1678)	(0.1683)	(0.1681)
Client Deal Experience	0.0008***	0.0008***	-0.0007***	-0.5322	-0.4744	-0.6898
	(0.0002)	(0.0002)	(0.0002)	(0.8551)	(0.8548)	(0.8758)
Equity	-0.0059	-0.0048	-0.0034	1.2632	1.19	0.6044
	(0.0075)	(0.0075)	(0.0099)	(0.9723)	(0.9728)	(0.9652)
Public Offering	0.0518***	0.0504***	0.0225**	-5.7141**	-5.6862**	-8.7532**
	(0.0085)	(0.0085)	(0.0107)	(2.6017)	(2.5896)	(3.5469)
IPO	0.0127	0.0133	0.0194	0.0393**	0.0380*	-0.0320*
	(0.0278)	(0.0278)	(0.0428)	(0.0198)	(0.0199)	(0.0178)
Observations	16630	16630	16630	16630	16630	16630
R^2	0.956	0.956	0.969	0.867	0.867	0.798
Cragg-Donald F-statistic	37.84	21.02		68.25	35.56	

Table V

Second-Stage Regressions

The table reports second-stage regressions for the linear probability model of client switching behavior in which the instrumented value of either *log*(*Capital*) or *Partners* from the first-stage regression is the explanatory variable of interest. Results are provided for four sets of models. Each set includes two independent regressions where the second is distinguished by the interaction between *Relationship Strength* and the instrumented endogenous variable. OLS results are provided as a benchmark. Each regression includes year, bank, and (client) 2-digit SIC fixed effects. Robust standard errors (clustered at the deal level) are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

	OLS		2SLS		OLS		2SLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Capital)	-0.0034	-0.0186*	0.4155***	0.2844**				
	(0.0097)	(0.0098)	(0.1583)	(0.1419)				
Partners					-0.0001	-0.0003***	0.0033***	0.0020*
					(0.0001)	(0.0001)	(0.0012)	(0.0011)
Bank-Client Characteristics								
Relationship Strength	-0.1610***	-0.4561***	-0.1635***	-0.5808***	-0.1609***	-0.2201***	-0.1647***	-0.3520***
	(0.0096)	(0.0351)	(0.0101)	(0.0555)	(0.0096)	(0.0155)	(0.0101)	(0.0257)
Relationship Strength x log(Capital)		0.0450***		0.0637***				
		(0.0053)		(0.0084)				
Relationship Strength x Partners						0.0004***		0.0012***
						(0.0001)		(0.0002)
SIC Share	-0.2190***	-0.2206***	-0.2394***	-0.2363***	-0.2195***	-0.2204***	-0.2073***	-0.2129***
	(0.0179)	(0.0179)	(0.0201)	(0.0196)	(0.0178)	(0.0179)	(0.0192)	(0.0189)
Bank Characteristics								
Market Share Equity(t-1)	-0.3893***	-0.3857***	-0.9832***	-0.8229***	-0.3944***	-0.3979***	-0.3850***	-0.3983***
	(0.0945)	(0.0946)	(0.2458)	(0.2249)	(0.0934)	(0.0935)	(0.0968)	(0.0963)
Market Share Debt(t-1)	-0.0023	-0.001	-1.8096***	-1.3352**	-0.0156	0.0062	-0.0703	0.0105
	(0.1395)	(0.1391)	(0.6960)	(0.6317)	(0.1334)	(0.1334)	(0.1397)	(0.1380)
Client and Transaction Characteristics								
log(Proceeds)	0.0198***	0.0191***	0.0258***	0.0233***	0.0200***	0.0200***	0.0165***	0.0173***
	(0.0039)	(0.0038)	(0.0047)	(0.0044)	(0.0039)	(0.0039)	(0.0042)	(0.0041)
Last Deal	-0.0012	-0.0025	-0.0001	-0.0022	-0.0012	-0.0015	-0.0002	-0.0015
	(0.0018)	(0.0018)	(0.0020)	(0.0019)	(0.0018)	(0.0018)	(0.0020)	(0.0019)
Client Deal Experience	0.0027***	0.0028***	0.0024***	0.0026***	0.0027***	0.0028***	0.0026***	0.0027***
	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Equity	-0.0660***	-0.0675***	-0.0637***	-0.0664***	-0.0660***	-0.0665***	-0.0644***	-0.0662***
	(0.0094)	(0.0094)	(0.0099)	(0.0096)	(0.0094)	(0.0094)	(0.0098)	(0.0096)
Public Offering	-0.1031***	-0.1014***	-0.1249***	-0.1169***	-0.1031***	-0.1024***	-0.1075***	-0.1042***
	(0.0105)	(0.0104)	(0.0137)	(0.0130)	(0.0105)	(0.0104)	(0.0110)	(0.0108)
IPO	0.0957***	0.0944***	0.0906**	0.0901**	0.0952***	0.0972***	0.1145***	0.1167***
	(0.0339)	(0.0335)	(0.0371)	(0.0353)	(0.0340)	(0.0338)	(0.0354)	(0.0347)
Observations	16630	16630	16630	16630	16630	16630	16630	16630
R^2	0.128	0.131	0.001	0.051	0.128	0.129	0.02	0.047

Table VI

Second-Stage Marginal Effects for Linear Probability and Probit Models

The table reports second-stage marginal effects for both linear probability and probit specifications of the client switching model. The endogenous explanatory variable of interest is either *log*(*Capital*) or *Partners.* First-stage regressions use both *Technology Exposure* and *Default Exposure* as instruments. Each regression includes year, bank, and (client) 2-digit SIC fixed effects. Robust standard errors (clustered at the deal level) are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

	Capital LPM	Capital Probit	Partners LPM	Partners Probit
	(1)	(2)	(3)	(4)
log(Capital)	0.2844**	0.2963**		
	(0.1419)	(0.1180)		
Partners			0.0020*	0.0021**
			(0.0011)	(0.0010)
Bank-Client Characteristics				
Relationship Strength	-0.5808***	-0.5833***	-0.3520***	-0.3447***
	(0.0555)	(0.0501)	(0.0257)	(0.0295)
Relationship Strength x log(Capital)	0.0637***	0.0666***	0.0012***	0.0012***
	(0.0084)	(0.0071)	(0.0002)	(0.0001)
SIC Share	-0.2363***	-0.2217***	-0.2129***	-0.1969***
	(0.0196)	(0.0204)	(0.0189)	(0.0249)
Bank Characteristics				
Market Share Equity(t-1)	-0.8229***	-0.7930***	-0.3983***	-0.3498***
	(0.2249)	(0.1662)	(0.0963)	(0.0926)
Market Share Debt(t-1)	-1.3352**	-1.4288***	0.0105	-0.0278
	(0.6317)	(0.5177)	(0.1380)	(0.1254)
Client and Transaction Characteristics				
log(Proceeds)	0.0233***	0.0212***	0.0173***	0.0150***
	(0.0044)	(0.0036)	(0.0041)	(0.0042)
Last Deal	-0.0022	-0.0015	-0.0015	-0.0008
	(0.0019)	(0.0018)	(0.0019)	(0.0017)
Client Deal Experience	0.0026***	0.0023***	0.0027***	0.0024***
	(0.0003)	(0.0004)	(0.0002)	(0.0003)
Equity	-0.0664***	-0.0563***	-0.0662***	-0.0557***
	(0.0096)	(0.0102)	(0.0096)	(0.0098)
Public Offering	-0.1169***	-0.1072***	-0.1042***	-0.0939***
	(0.0130)	(0.0098)	(0.0108)	(0.0110)
IPO	0.0901**	0.0777**	0.1167***	0.1052***
	(0.0353)	(0.0314)	(0.0347)	(0.0289)
Observations	16630	16630	16630	16630
R^2	0.001		0.047	
Log Likelihood		-27567.13		-180172.49
Percent correctly predicted		65.81		65.77
Wald Exogeneity Test (*p*-value)		0.0001		0.0001

Table VII
Bank Complexity and Positive Assortative Matching

The table reports second-stage linear probability models in which the endogenous explanatory variable of interest is either *log(Capital)* or *Partners*. *Mismatch* is an absolute measure of the "quality" difference between an issuer and its relationship bank. Each regression includes year, bank, and (client) 2-digit SIC fixed effects. Robust standard errors (clustered at the deal level) are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

	(1)	(2)	(3)	(4)	(5)	(6)
log(Capital)	0.2581**	0.2314**	0.2185**			
	(0.1030)	(0.1024)	(0.1011)			
Partners				0.0006	0.0005	0.0002
				(0.0006)	(0.0006)	(0.0006)
Bank-Client Characteristics						
Mismatch		0.1762***	-0.1553		0.1820***	0.0527
		(0.0237)	(0.1141)		(0.0232)	(0.0565)
Mismatch x log(Capital)			0.0499***			
			(0.0166)			
Mismatch x Partners						0.0008**
						(0.0003)
Relationship Strength	-0.5798***	-0.5810***	-0.5751***	-0.3099***	-0.3132***	-0.3092***
	(0.0486)	(0.0483)	(0.0480)	(0.0233)	(0.0232)	(0.0232)
Relationship Strength x log(Capital)	0.0636***	0.0638***	0.0630***			
	(0.0073)	(0.0073)	(0.0072)			
Relationship Strength x Partners				0.0009***	0.0009***	0.0009***
				(0.0001)	(0.0001)	(0.0001)
SIC Share	-0.2350***	-0.2468***	-0.2464***	-0.2177***	-0.2317***	-0.2313***
	(0.0189)	(0.0187)	(0.0187)	(0.0181)	(0.0180)	(0.0180)
Bank Characteristics						
Market Share Equity(t-1)	-0.7854***	-0.7821***	-0.7895***	-0.4001***	-0.4362***	-0.4570***
	(0.1749)	(0.1736)	(0.1753)	(0.0940)	(0.0938)	(0.0943)
Market Share Debt(t-1)	-1.2212***	-1.1767**	-1.1766**	0.0200	-0.0494	-0.0482
	(0.4695)	(0.4659)	(0.4681)	(0.1345)	(0.1344)	(0.1343)
Client and Transaction Characteristics						
log(Proceeds)	0.0229***	0.0373***	0.0399***	0.0189***	0.0343***	0.0369***
	(0.0042)	(0.0045)	(0.0047)	(0.0039)	(0.0043)	(0.0044)
Last Deal	-0.0023	-0.0026	-0.0025	-0.0017	-0.0021	-0.0021
	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0018)	(0.0018)
Client Deal Experience	0.0026***	0.0026***	0.0026***	0.0028***	0.0028***	0.0028***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Equity	-0.0665***	-0.0666***	-0.0655***	-0.0666***	-0.0667***	-0.0667***
	(0.0096)	(0.0095)	(0.0095)	(0.0094)	(0.0094)	(0.0094)
Public Offering	-0.1155***	-0.1121***	-0.1073***	-0.1028***	-0.1005***	-0.0965***
	(0.0119)	(0.0119)	(0.0117)	(0.0105)	(0.0105)	(0.0106)
IPO	0.0904***	0.0893***	0.0898***	0.1068***	0.1047***	0.1026***
	(0.0351)	(0.0346)	(0.0346)	(0.0338)	(0.0336)	(0.0335)
Observations	16630	16630	16630	16630	16630	16630
R^2	0.06	0.072	0.073	0.092	0.098	0.099

Chapter 3

Optimal Equity Financing Contracts and Private Monitoring

Optimal Equity Financing Contracts and Private Monitoring

Xin Xue

University of Virginia

July 30, 2018

Abstract

I design a contingent contracting mechanism where the principal's (a venture capitalist) private monitoring induces the agent's (an entrepreneur) effort and adds value to the project through the capital investment from the principal. Featuring double-sided moral hazard, the optimal contract subsumes a menu that entitles the principal to punish the agent upon negative information, and is incentive compatible to avoid the principal to falsely punish to expropriate a bigger equity stake. Compared to the "second best" under 'pay-for-performance', this scheme grants the principal high ex ante equity stake. The project value and capital investments commensurate with a higher marginal return on the investments. The optimal monitoring intensity increases with the value added by the agent's effort but decreases with cost of monitoring.

1 Introduction

Venture capital or VC is a financial intermediation that incubates start-up companies and entrepreneurs. According to PwC (2011), in 1980, total VC investments in the US were $610 million. By 1990, this figure had increased to 2.3 billion dollars, once peaked at 100 billion during the "dot-com bubble" in 2000, and stayed around 30 billion (See Da Rin et al. 2011). VC plays an important role in both the product and labor market.[1] VC-backed firms also significantly contribute to the capital and asset market.[2]

Distinct from other types of financing intermediation, VC provides incentives from various aspects.[3] This paper studies how VC monitoring provides incentives to entrepreneurs' effort devotion and VC's capital investments.[4] Empirically, Gompers (1995) uses the number of rounds of financing to measure monitoring intensity, and discovers monitoring increases with corporate intangible assets, such as companies intense in R&D and innovation. Tian et al. (2011) find a positive correlation between a lack of VC monitoring and corporate fraud which induces costly IPOs and thus, disparages VCs'

[1]Puri and Zarutskie (2012), using US Census data, find that only 0.11% of new companies created over a 25 year sample period from 1981-2005 are funded by VC, yet these companies account for 4% to 5.5% of total employment. Popov and Roosenboom (2012) use a panel of industries across several European countries, and discover that higher levels of VC investment are associated with more entry, especially in high-R&D (and also high-entry) industries. Kortum and Lerner (2000) claim that a dollar of venture capital appears to be about three times more potent at stimulating patents than a dollar of traditional corporate R&D.

[2]Ritter (2011) reports that between 1980 and 2010, 35% of all US Initial Public Offerings were VC backed.

[3]Banks screen new borrowers by requiring large amount of collateral (see Ueda (2004), Stiglitz and Weiss (1981)), which most start-up firms cannot afford. Without collaterals as an insurance, VC has more incentive to care for and add value to the projects (see De Bettignies and Brander (2007)).In contrast to angel investors, VCs actively manage their portfolio and provide incentives through advising and exerting effort(See Chemmanur and Chen (2014)).

Incentives can be in the forms of offering contingent securities, cash flow rights and control rights (Hellmann (2006), Marx (1998)), contract renegotiating (Aghion and Bolton (1992) and Hart and Moore (1991) Hart and Moore (1997)), and monitoring and advising the entrepreneurs (Winton and Yerramilli (2008)).

[4]Banks also monitor, or acquire information on their borrowers. However, the information banks acquire is borrowers' accounting information or credit history (see Diamond (1984)). In contrast, VCs monitor entrepreneurs' action, and here specifically their effort.

reputation. Bernstein et al. (2015) use an exogenous variations of introduction of direct flights between VCs and entrepreneurs that change monitoring cost, and find that monitoring helps entrepreneurs' innovations and their likelihood of IPOs.

The theoretical literature on VC financing contracts argues that monitoring reduces information asymmetry and induces agents' effort. General theories on contracts and costly state verification (Harris and Raviv (1979), Hölmstrom (1979) and Townsend (1979)) argue that contract with monitoring (on either agents' action or outcome) would be Pareto-superior than that where the agent is unsupervised. However, they did not provide mechanisms on how monitoring can induce Pareto-superior outcome. Moreover, they do not incorporate principal's incentives such as the capital investments into the contractual setting. Other papers on moral hazard has examined optimal investment allocation under information asymmetry, but does not shed light on monitoring.[5]

I design a complete contracting mechanism where the principal's private monitoring can induce the agents' effort. This mechanism allows the principal to punish the agent upon negative information, while it prevents moral hazard from the principal's side, which is to 'punish' the agent even upon positive information. In the contract, I also incorporate the principal's optimal investment decision along with their monitoring decision.

Compared to a contract simply under 'pay-for-performance' scheme, the equilibrium outcome of the contract with monitoring improves the principal's payoff by two ways. First, the principal is able to provide a lower share to induce the same amount of effort from the agent. Second, with a higher share of the pie to the principal, the optimal investments increase and so does the value of the project.

[5]See Clementi and Hopenhayn (2006), Fulghieri and Sevilir (2009), Inderst et al. (2007), and Casamatta (2003)

Moreover, a stochastic monitoring decreases the cost to the principal without demolishing the agent's incentives. The optimal monitoring intensity increases with the value added by the agent's effort and decreases with the monitoring cost. These factors characterize the necessary conditions for monitoring to take place.

I arrange the rest of the paper in the following way. In section II, I layout the setups of the model. In section III, I derive 3 contracts, 1) a contract under 'pay-for-performance' scheme, 2) a contract with monitoring w.p.1, 3) a contract with optimal monitoring probability p. In the following sections, I compare the ex ante outcomes of the contracts.

2 Model

2.1 Framework

In a two period economy? there are two risk neutral agents, an entrepreneur (agent) and a venture capitalist (principal).[6] The entrepreneur, endowed with no capital, has a production technology, denoted by $f(\cdot)$, and attempts to get the project financed by the deep pocketed VC. The production function $f(\cdot,\cdot)$ requires two period capital investments, K_1, K_2, which are publicly observable and contractible at time 0. We assume $f(K_1, K_2)$ increasing, smooth and decreasing return to scale in both arguments, and $f(0,\cdot) = f(\cdot, 0) = 0$.

The production requires effort, e_t, provided only by the entrepreneur. Effort can be either 'high', 'low' or 0, i.e. $e_t \in \{h, l, 0\}$. His effort generates a productivity random variable, Z, normalized to support $[0, 1]$. Z follows distribution function $G(z|e_1, e_2)$,

[6]The risk averse case is also solved. See the appendix

with parameters e_1, e_2. I assume the monotone likelihood ratio property on the density functions of Z, $g(z|e_1, e_2)$.[7] 0 effort in any period will result in failure, $G(0|0, e_2) = G(0|e_1, 0) = 1$. Effort is costly. Let $c_h, c_l, 0$ be the costs corresponding to the effort $h, l, 0$, where $c_h > c_l > 0$.

The value of the project is realized at the end of the late stage, denoted by R. It is observable to both parties and assumed a product of the productivity shock and the production function.

$$R = Zf(K_1, K_2) \tag{1}$$

2.2 Monitoring

The effort is the entrepreneur's private information. The VC can perfectly learn e_1 and e_2 at a cost, and we call this learning procedure monitoring.[8] That is, the signal on entrepreneur's effort does not have noise, and it's the VC's private information. Monitoring a binary choice in each period, denoted by E_t, and it is costly. If $E_t = 1$, the VC learns the entrepreneur's effort at time t, and pays a cost of C. I assume monitoring is committable. That is, the VC can commit to her monitoring choice prior to the contract.[9]

[7] See Milgrom (1981). $\forall z_0, z_1, \in [0, 1]$, where $z_1 > z_0$, $\frac{g(z_1|h,h)}{g(z_1|h,l)} \geq \frac{g(z_0|h,h)}{g(z_0|h,l)}$ and $\frac{g(z_1|h,l)}{g(z_1|l,l)} \geq \frac{g(z_0|h,l)}{g(z_0|l,l)}$. Moreover we assume $g(z|h,l) = g(z|l,h)$. (This is not a necessary assumption. I do not require the perfect substitutability between the early and late stage effort. However, it simplifies mathematical derivation. A more general expression is that $\mathbb{E}|(Z|h,l) - (Z|l,h)| \leq \delta$ for some positive δ.) Instantly, we can conclude that high effort First Order Stochastic Dominates low effort, $G(z|h,h) \succ G(z|h,l) = G(z|l,h) \succ G(z|l,l)$.

[8] In the literature, the monitor is a perfect private monitor.

[9] For the noncommittable case, I discuss in the appendix.

2.3 Incentive Mechanisms, Contract and Timeline

The VC offers the entrepreneur a take-it-or-leave-it contract and commits on her monitoring technology.[10] The entrepreneur participates for non-negative ex ante payoff.[11] The objective for the contract is to maximize the VC's (principal's) payoff while inducing the entrepreneur's effort.[12] I assume the effort is not contractible whereas the capital investments and the sharing rule are.[13] Since I assumed the entrepreneur has no fund endowment, a lump sum transfer doesn't need to be considered.[14] I assume S is the equity share to the entrepreneur. I propose and compare two mechanisms that both induce high effort in the following.

2.3.1 Benchmark: Contingent Equity Mechanism

Assume no monitoring, to elicit his effort, we let his equity share be contingent on the outcome, i.e. $S := S(z)$.[15] In principal-agent literature, the scheme is similar to 'performance pay', see Holmstrom and Milgrom (1991) and Gibbons (1998). As his share expands with the outcome, he is willing to make the pie bigger.

At time 0, the contract pins down capital investments for each period, (K_1, K_2), and the sharing rule at the end of period 2, $S(z)$. The revenue gets disclosed and distributed

[10]I assume the contract is non-renegotiatable. Contract renegotiation may improve both parties' payoffs at certain game nodes, but it makes the contract no longer enforceable.

[11]The agent is risk neutral and we assume his reservation utility is 0.

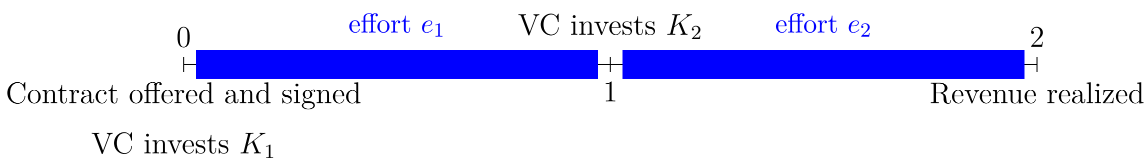
[12]I assume it favors her interest to induce his high effort at a cost. Later, we denote it as the hiring assumption.

[13]The entrepreneur's effort is not observed to any third party, and thus cannot be legally enforced.

[14]A lump sum transfer replicates the mechanism of debt financing. Any promised lump sum transfer from the principal to the agent does not induce effort. A lump sum transfer in the reversed direction is not feasible. If the project value is less than the promised transfer, the defaulting rule must be clearly defined. Moreover, a convertible bond or an option can be fully replicated by contingent equity sharing.

[15]For every observed fixed pair of capital investments, R and Z are bijective. The sharing rule that is contingent on R is equivalent to that on Z. WLOG, let $S(z)$ be the promised equity share assigned to the entrepreneur when $Z = z$.

according to the contract at $t = 2$.



2.3.2 Monitoring and Two-stage Contracting Mechanism

Adding the dimension of information acquisition for the principal, how does monitoring affect the ex ante contract to provide incentives? Since I have assumed VC's signal through monitoring is her private information, any punishment cannot be contracted or committable. I propose a contingent contracting mechanism.

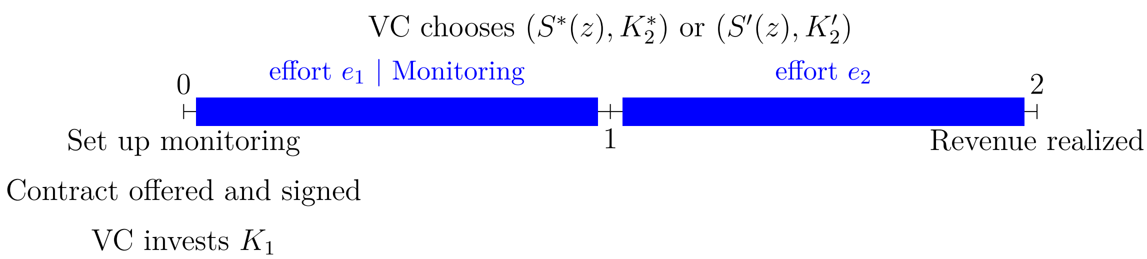
A menu contract settles investment K_1, and prompts two bundled choices of late stage investments and the sharing rule, $\{(S^*(z), K_2^*), (S'(z), K_2')\}$. The contract entitles the VC to choose one bundle after she monitors. The idea of the menu is that it provides the principal the choice to punish the agent. The menu incorporates the equilibrium path $(S^*(z), K_2^*)$, and a credible threat $(S'(z), K_2')$, reserved as a punishment.

Note that this contract must elimiate the potential double-sided moral hazard problem. Moral hazard from the agent is the propensity to exert low effort to reduce the private cost. Therefore, the punishment must be credible. Moral hazard from the principal is the propensity to punish and expropriate when she observes high effort. Neither moral hazard provides incentives for effort exertion.

To induce effort, the design of $(S^*(z), K_2^*)$ and $(S'(z), K_2')$ must satisfy the credibility for the principal (i) to not 'punish' if she observes 'h' effort, (ii) to 'punish' if she observes 'l' effort. Also, 'h' effort exertion is ex ante preferred to 'l' for the entrepreneur.

Corollary 1. *The cardinality of the menu is at least 2.*[16]

Notice that, late stage monitoring cannot induce agent's effort under any mechanism. Hence, monitoring only occurs at the early stage.



2.4 Ex Ante Payoffs

VC's payoff is her equity value deducting capital investments and monitoring cost:

$$\mathbb{E}(\Pi(K_1, K_2, S(z), E; e_1, e_2)) = \int_0^1 zf(K_1, K_2)(1 - S(z))dG(z|e_1, e_2) - K_1 - K_2 - C(E) \tag{2}$$

Entrepreneur's payoff is his share claimed subtracting his effort expense:

$$\mathbb{E}(\pi(e_1, e_2; K_1, K_2, S(z))) = \int_0^1 zf(K_1, K_2)S(z)dG(z|e_1, e_2) - c(e_1) - c(e_2) \tag{3}$$

2.5 Equilibrium Concept

The equilibrium is in the concept of a pure strategy Sub-game Perfect Equilibrium.[17]

Definition 1. *A Sub-game Perfect Equilibrium consists of*

[16]The corollary is a direct result by proposition in the appendix. Also, refer to the appendix for the general case where $T \geq 2$.

[17]Here pure strategy means that players' moves under each game node are not random variables. Monitoring technology can be a random variable for pure strategy equilibria. Here signals from monitoring are perfectly informative. Therefore, a Perfect Baysian Nash Equilibrium is equivalent to a SPE.

(i) Given the monitoring technology, the contract between the VC and the entrepreneur.

(ii) Effort the entrepreneur exerts in both periods.

(iii) The choice by the VC at $t = 1$, conditional on that the VC monitors.

3 Contracts

In this section, I derive and compare the contracts with and without monitoring.

3.1 Benchmark Contract, No Monitoring

3.1.1 Incentive and Participation, ICs-1

Offered with the incentive contract specifying capital investments and the sharing, the entrepreneur is induced with high effort, and also his payoff weakly dominates the outside option. That is, given $\{K_1, K_2, S(z)\}$

$$\mathbb{E}(\pi(h,h)) \geq \max\{\mathbb{E}(\pi(h,l)), \mathbb{E}(\pi(l,h)), \mathbb{E}(\pi(l,l)), 0\} \tag{IC/IR-agent-1}$$

The VC chooses the sharing rule and capital investments to maximize her expected payoff, subject to his incentive constraints, and $S(z)$ is bounded by $[0, 1]$.

$$\max_{(S(z),K_1,K_2)} \mathbb{E}(\Pi(K_1, K_2, S(z)|h,h)) \tag{4}$$

$$\text{s.t. (IC/IR-agent-1)}, \; 0 \leq S(z) \leq 1$$

3.1.2 The Contract

By solving the VC's objective, the optimal contract is as follows:

$$S^*(z) = \mathbb{1}\{z \geq \min\{x_1, x_2\}\} \tag{5}$$

$$K_1^* = argmax_{K_1} \int_0^{\min\{x_1,x_2\}} zdG(z|h,h)f(K_1,K_2) - K_1 - K_2 \tag{6}$$

$$K_2^* = argmax_{K_2} \int_0^{\min\{x_1,x_2\}} zdG(z|h,h)f(K_1,K_2) - K_1 - K_2 \tag{7}$$

where $\int_{x_1}^1 z(g(z|h,h) - g(z|l,l))dz = \frac{2c_h - 2c_l}{f(K_1,K_2)}$ and $\int_{x_2}^1 z(g(z|h,h) - g(z|h,l))dz = \frac{c_h - c_l}{f(K_1,K_2)}$.[18]

3.1.3 Existence condition

The hiring condition suffices the existence condition of this equilibrium. The hiring condition states that as long as it is in the best interest for the VC to induce high effort, the optimal contract exists.

3.2 Menu Contract, Monitoring w.p. 1

The VC commits to monitoring with probability 1, pays a cost C, and offers contract $\{K_1, \{(S^*(z), K_2^*), (S'(z), K_2')\}\}$, entitling herself to select one bundle at $t = 1$.

3.2.1 Equilibrium

If the entrepreneur exert 'l' effort, the VC has incentive to punish him $(S'(z), K_2')$. Her best-response to low effort is $(S'(z), K_2')$ rather than $(S^*(z), K_2^*)$.[19] If the VC recognizes high effort, her best response is $(S^*(z), K_2^*)$ rather than $(S'(z), K_2')$.[20] $(S^*(z), K_2^*)$ also

[18]See the appendix for the steps.

[19]This refers to the moral hazard from the entrepreneur's side. Its existence condition is provided in the appendix.

[20]This associates to the moral hazard from the VC's side. Its existence condition is provided in the appendix.

elicits late stage effort. Both bundles are constructed incentive compatible to both parties.

3.2.2 Incentive Constraints (ICs-2)

1. $S^*(z)$ induces him to work hard again given that he worked hard before

$$\mathbb{E}(\pi(K_1, K_2^*, S^*(z)|h,h)) \geq \mathbb{E}(\pi(K_1, K_2^*, S^*(z)|h,l)) \tag{IC-agent-1}$$

2. Existence of the VC's moral hazard

$$\mathbb{E}(\Pi(K_1, K_2', S'(z)|h,h)) \geq \mathbb{E}(\Pi(K_1, K_2^*, S^*(z)|h,h)) \tag{MH-principal}$$

Entrepreneur's incentive to credibly punish VC with 'l' effort

$$\mathbb{E}(\pi(K_1, K_2', S'(z)|h,l)) \geq \mathbb{E}(\pi(K_1, K_2', S'(z)|h,h)) \tag{Agent's Threat}$$

$$\mathbb{E}(\Pi(K_1, K_2^*, S^*(z)|h,h)) \geq \mathbb{E}(\Pi(K_1, K_2', S'(z)|h,l)) \tag{Agent Punishment}$$

3. Existence of the Entrepreneur's moral hazard

$$\mathbb{E}(\pi(K_1, K_2^*, S^*(z)|l,l)) \geq \mathbb{E}(\pi(K_1, K_2^*, S^*(z)|h,h)) \tag{MH-Agent}$$

If the entrepreneur shirks, high effort can no longer be induced, since

$$\mathbb{E}(\pi(K_1, K_2^*, S^*(z)|h,h)) \geq \mathbb{E}(\pi(K_1, K_2^*, S^*(z)|h,l))$$

$$\mathbb{E}(\pi(K_1, K_2^*, S^*(z)|l,h)) = \mathbb{E}(\pi(K_1, K_2^*, S^*(z)|h,l,)),$$

Therefore the VC punishes low effort:

$$\mathbb{E}(\Pi(K_1, K_2', S'(z)|l,l)) \geq \mathbb{E}(\Pi(K_1, K_2^*, S^*(z)|l,l)) \tag{Principal's Punishment}$$

4. Exerting high effort is weakly preferred.

$$\mathbb{E}(\pi(K_1, K_2^*, S^*(z)|h,h)) \geq \max\{\mathbb{E}(\pi(K_1, K_2', S'(z)|l,l)), 0\} \tag{IC/IR-agent-2}$$

3.2.3 VC's objective

The VC chooses sharing rules and capital investments to maximize her expected payoff.[21]

$$\max_{(S(z),S'(z),K_1,K_2,K_2')} \mathbb{E}(\Pi(K_1, K_2, S(z)|h,h)) \tag{8}$$

$$\text{s.t. ICs-2, } 0 \leq S(z) \leq 1 \text{ , } 0 \leq S'(z) \leq 1$$

3.2.4 The Contract

By solving the VC's objective, first the equilibrium path of the menu contract is:

$$S^*(z) = \mathbb{1}\{z \geq x_2\} \tag{9}$$

$$K_1^* = argmax_{K_1} \int_0^{x_2} z dG(z|h,h) f(K_1, K_2) - K_1 - K_2 - C \tag{10}$$

$$K_2^* = argmax_{K_2} \int_0^{x_2} z dG(z|h,h) f(K_1, K_2) - K_1 - K_2 - C \tag{11}$$

In the appendix, we show the existence and non-uniqueness of $S'(z)$ and K_2'. Note that $S'(z)$ and K_2' satisfy the condition: (1) $S'(z)$ and K_2' are bad enough for the entrepreneur so that the punishment is credible. (2) The VC is incentive compatible not to choose $S'(z)$ and K_2' on the equilibrium path.

In the special case I show in the appendix, where I force $K_2' = K_2^*$, the lower bound on $S'(z)$ is:

$$\underline{S'(z)} = \mathbb{1}\{z \geq x_2'\}, \tag{12}$$

[21] Note that $S'(z)$ and K_2' are in the constraints.

where $\int_0^{x_2'} zg(z|h,l)dz = \int_0^{x_2} zg(z|h,h)dz$.

3.3 Random Monitoring

How would the equilibrium be affected, if the VC commits to random monitoring? Let monitoring technology be a Bernoulli random variable, with monitoring probability p, i.e. $\mathbb{P}(E=1) = p \in (0,1)$. The associated cost is $p \cdot C$.

3.3.1 Monitoring Intensity, p and Equilibrium

There are two potential outcomes for the VC's information structure at time 1. With probability p, she observes e_1, and with probability $1-p$, she doesn't obtain any signal.

If p is small, the VC likely won't observe any information. A pure strategy equilibrium does not exist. If the VC always punishes with $(S'(z), K_2')$, the entrepreneur's effort is not induced. If she always chooses $(S^*(z), K_2^*)$, the entrepreneur can always shirk and receives a better payoff. This case is similar to where monitoring is not committable ex ante. It will be further discussed in the appendix.

We assume the monitoring cost is small enough for the pure strategy to exist. Monitoring intensity p must be large enough so that his effort is highly likely to be revealed. He is better off to endeavor than to shirk. Given the committed monitoring intensity p, high effort are induced and $(S^*(z), K_2^*)$ is chosen on the equilibrium path. The optimal monitoring intensity p^* is at where the entrepreneur is indifferent between devoting high effort and low effort in the first period.

3.3.2 Additional Incentive Constraint, ICs-3

The monitoring intensity, p is large enough so that high effort is induced.

$$\mathbb{E}(\pi(K_1, K_2, S(z)|h, h)) \geq$$

$$p\mathbb{E}(\pi(K_1, K_2', S'(z)|l, l)) + (1 - p)\mathbb{E}(\pi(K_1, K_2, S(z)|l, l)) \quad \text{(IC-agent-3)}$$

The optimal p^* occurs at the equality, and moreover, it is an implicit function of all the other arguments, $p^* = p^*(K_1, K_2, K_2', S(z), S'(z))$

3.3.3 VC's objective

The VC chooses monitoring intensity, sharing rules and capital investments to maximize her expected payoff.[22]

$$\max_{(p,S(z),S'(z),K_1,K_2,K_2')} \mathbb{E}(\Pi(K_1, K_2, S(z)|h, h)) \quad (13)$$

$$\text{s.t. (ICs-2\&3)}, \ 0 \leq p \leq 1 \ , 0 \leq S(z) \leq 1 \ , 0 \leq S'(z) \leq 1$$

3.3.4 The Contract

The pure strategy equilibrium path stays with the random monitoring. The sharing rule and the capital investments still need to provide the least incentive to induce effort in period 2. The form of $S^*(z)$ stay the same as those in the deterministic monitoring case. However, since the objective contains an implicit function of $p^*(K_1, K_2, K_2', S(z), S'(z))$, K_1^*, K_2^* differ from those in the deterministic monitoring case. Therefore, the position of

[22]Note that this objective is different from the previous ones. $p^*(K_1, K_2, K_2', S(z), S'(z))$ also contains K_1 and K_2. Moreover, $S'(z)$ and K_2' are in the constraints.

x_2 also differs.

$$S^*(z) = \mathbb{1}\{z \geq x_2\} \tag{14}$$

$$K_1^* = argmax_{K_1} \int_0^{x_2} zdG(z|h,h)f(K_1,K_2) - K_1 - K_2 - p^*C \tag{15}$$

$$K_2^* = argmax_{K_2} \int_0^{x_2} zdG(z|h,h)f(K_1,K_2) - K_1 - K_2 - p^*C \tag{16}$$

Remember, $\int_{x_2}^{1} z(g(z|h,h) - g(z|h,l))dz = \frac{c_h - c_l}{f(K_1,K_2)}$. However, in order to minimize the monitoring cost or its intensity, the punishment must be the harshest credible punishment that satisfies the conditions: (1) $S'(z)$ and K_2' are the worst (among all possible) for the entrepreneur so that the punishment is credible. (2) The VC is incentive compatible not to choose $S'(z)$ and K_2' on the equilibrium path.

The incentive for the VC to prevent her moral hazard. IC-principal becomes:

$$\int_0^{x_2} zg(z|h,h)dzf(K_1,K_2) - K_2 = \int_0^{x_2'} zg(z|h,l)dzf(K_1,K_2') - K_2'$$

The optimal monitoring intensity p^* by ICs-3 is:

$$p^*(K_1,K_2,K_2',S(z),S'(z)) = \frac{\int_{x_2}^{1} z(g(z|l,l) - g(z|h,h))dzf(K_1,K_2) + 2(c_h - c_l)}{\int_{x_2}^{1} zg(z|l,l)dzf(K_1,K_2) - \int_{x_2'}^{1} zg(z|l,l)dzf(K_1,K_2')}$$

4 Equilibrium Analysis

4.1 Necessary Condition 1

The venture capitalist has incentive to provide the contract with monitoring if her ex ante payoff is higher than that without. Her ex ante payoff differs between the contracts by three factors, **the sharing rule, return on investments and monitoring cost**.

Now I compare the principal's (and the agent's) payoffs among different contracts by comparing those factors.

4.1.1 The sharing rule

$$\text{Benchmark contract, no monitoring: } S_N^*(z) = \mathbb{1}\{z \geq \min(x_1, x_2)\}$$

$$\text{(Deterministic/Random) Monitoring: } S_M^*(z) = \mathbb{1}\{z \geq x_2\}$$

where $\int_{x_1}^{1} z(g(z|h,h) - g(z|l,l))dz = \frac{2(c_h - c_l)}{f(K_1, K_2)}$ and $\int_{x_2}^{1} z(g(z|h,h) - g(z|h,l))dz = \frac{c_h - c_l}{f(K_1, K_2)}$

Case 1 $x_1 \geq x_2$

If $x_1 \geq x_2$, $S_N^*(z) = S_M^*(z)$. With the same sharing rule, the principal's objectives are maximized at the same capital investments, $K_{1,M}^* = K_{1,N}^*$, and $K_{2,M}^* = K_{2,N}^*$. Monitoring is costly, and it does not affect the underlying contract. Therefore, when $x_1 \geq x_2$, the VC will not monitor.

Case 2 $x_1 < x_2$**, A Necessary Condition for Monitoring**

$x_1 < x_2$ has a direct implication: $S_N^*(z) > S_M^*(z)$ equity share the VC receives is higher for the contract with monitoring. **Necessary condition 1** for monitoring: $x_1 < x_2$, where $\int_{x_1}^{1} z(g(z|h,h) - g(z|l,l))dz = \frac{2(c_h - c_l)}{f(K_1, K_2)}$ and $\int_{x_2}^{1} z(g(z|h,h) - g(z|h,l))dz = \frac{c_h - c_l}{f(K_1, K_2)}$. This condition has two indirect implications, given by the proposition below:

Proposition 1. *Implication 1: If $x_1 < x_2$, efforts are increasing return to scale, i.e. 'h' effort in two stages can synergize.*

Implication 2: this condition is equivalent to the existence of the agent's Moral Hazard:

$$\mathbb{E}(\pi(K_1^*, K_2^*, S^*(z)|l,l)) \geq \mathbb{E}(\pi(K_1^*, K_2^*, S^*(z)|h,h) \Leftrightarrow x_1 \leq x_2$$

In other words, if at least 'h' effort in one stage is preferable by the agent, WLOG the early stage, effort can be induced in the late stage when monitoring is redundant nonetheless.

4.2 Optimal Monitoring Intensity p^*

4.3 Necessary Condition 2, Monitoring Cost C

This subsection should be jointly read with the next one and the section of numerical approach. Another **necessary condition** for the principal to monitor is that the monitoring cost C cannot be large. A large monitoring cost destroys the principal's incentive to offer the contract with monitoring.

4.4 Payoffs and Welfare

From now on, we assume $x_1 < x_2$, and thus, $S_N^*(z) > S_M^*(z)$. Therefore, the VC extracts more share in the contract where she monitors. Denote the optimal investments in the benchmark contract as K_1^N, K_2^N, and those in the contract with monitoring (deterministic/random) is K_1^M, K_2^M.

$$K_1^N = argmax_{K_1} \int_0^{x_1(K_1,K_2)} zdG(z|h,h)f(K_1,K_2) - K_1 - K_2 \tag{17}$$

$$K_2^N = argmax_{K_2} \int_0^{x_1(K_1,K_2)} zdG(z|h,h)f(K_1,K_2) - K_1 - K_2 \tag{18}$$

In the contract where VC monitors (Deterministically, where $p^* \equiv 1$/Randomly),

$$K_1^M = argmax_{K_1} \int_0^{x_2(K_1,K_2)} zdG(z|h,h)f(K_1,K_2) - K_1 - K_2 - p^*C \tag{19}$$

$$K_2^M = argmax_{K_2} \int_0^{x_2(K_1,K_2)} zdG(z|h,h)f(K_1,K_2) - K_1 - K_2 - p^*C \tag{20}$$

Proposition 2. *Compared to the benchmark, the contract with monitoring has higher* *promised capital investments per round. Moreover, the ex ante project value is improved.*[23]

That is $K_1^N < K_1^M$ and $K_2^N < K_2^M$. Intuitively, as the VC possesses a higher share of the project, her marginal return on her capital investments increases, which incentivizes her to devote more investments. Therefore, total project value increases. (See Figure 2) Therefore, the VC's payoff is improved in two ways by monitoring. First, she compensates the entrepreneur with less share of the pie. In addition, the size of pie expands as the marginal return on her capital investments increases.

5 Conclusion

In this paper, I design a contingent contracting mechanism where private monitoring by the venture capitalist induces the entrepreneur's effort and adds value to the project through the capital investment. Featuring double-sided moral hazard, the optimal contract subsumes a menu that entitles the principal to punish the agent upon negative information, and is incentive compatible to avoid the principal to falsely punish to expropriate a bigger equity stake. Compared to the "second best" under 'pay-for-performance', this scheme grants the principal high ex ante equity stake. The project value and capital investments commensurate with a higher marginal return on the investments, approaching to the "first best." The optimal monitoring intensity increases with the value added by the agent's effort but decreases with cost of monitoring.

[23] The most efficient outcome of investments achieves at the first best, i.e. $\{K_1^*, K_2^*\} = argmax_{K_1,K_2} \int_0^1 zdG(z|h,h)f(K_1,K_2) - K_1 - K_2$. More efficiency means the outcome is closer to the first best.

References

Aghion, P. and Bolton, P. (1992). An incomplete contracts approach to financial contracting. The review of economic Studies, 59(3):473–494.

Bernstein, S., Giroud, X., and Townsend, R. R. (2015). The impact of venture capital monitoring. The Journal of Finance.

Casamatta, C. (2003). Financing and advising: optimal financial contracts with venture capitalists. The Journal of Finance, 58(5):2059–2086.

Chemmanur, T. J. and Chen, Z. (2014). Venture capitalists versus angels: The dynamics of private firm financing contracts. Review of Corporate Finance Studies, 3(1-2):39–86.

Clementi, G. L. and Hopenhayn, H. A. (2006). A theory of financing constraints and firm dynamics. The Quarterly Journal of Economics, pages 229–265.

Da Rin, M., Hellmann, T. F., and Puri, M. (2011). A survey of venture capital research. Technical report, National Bureau of Economic Research.

De Bettignies, J.-E. and Brander, J. A. (2007). Financing entrepreneurship: Bank finance versus venture capital. Journal of Business Venturing, 22(6):808–832.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. The Review of Economic Studies, 51(3):393–414.

Fulghieri, P. and Sevilir, M. (2009). Size and focus of a venture capitalist's portfolio. Review of Financial Studies, 22(11):4643–4680.

Gompers, P. A. (1995). Optimal investment, monitoring, and the staging of venture capital. The journal of finance, 50(5):1461–1489.

Gompers, P. A. (1996). Grandstanding in the venture capital industry. Journal of Financial economics, 42(1):133–156.

Harris, M. and Raviv, A. (1979). Optimal incentive contracts with imperfect information. Journal of economic theory, 20(2):231–259.

Hart, O. and Moore, J. (1991). A theory of debt based on the inalienability of human capital. Technical report, National Bureau of Economic Research.

Hart, O. and Moore, J. (1997). Default and renegotiation: A dynamic model of debt. Technical report, National Bureau of Economic Research.

Hellmann, T. (2006). Ipos, acquisitions, and the use of convertible securities in venture capital. Journal of Financial Economics, 81(3):649–679.

Hölmstrom, B. (1979). Moral hazard and observability. The Bell journal of economics, pages 74–91.

Inderst, R., Mueller, H. M., and Münnich, F. (2007). Financing a portfolio of projects. Review of Financial Studies, 20(4):1289–1325.

Kortum, S. and Lerner, J. (2000). Assessing the contribution of venture capital to innovation. RAND journal of Economics, pages 674–692.

Marx, L. M. (1998). Efficient venture capital financing combining debt and equity. Review of Economic Design, 3(4):371–387.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. The Bell Journal of Economics, pages 380–391.

Popov, A. and Roosenboom, P. (2012). Venture capital and patented innovation: evidence from europe. Economic Policy, 27(71):447–482.

Puri, M. and Zarutskie, R. (2012). On the life cycle dynamics of venture-capital-and non-venture-capital-financed firms. The Journal of Finance, 67(6):2247–2293.

PwC (2011). Pricewaterhousecoopers moneytree. NVCA, PwC, and Thomson Reuters.

Ritter, J. R. (2011). Initial public offerings: Tables updated through 2010. Unpublished working paper.

Stiglitz, J. E. and Weiss, A. (1981). Credit rationing in markets with imperfect information. The American economic review, 71(3):393–410.

Tian, X., Udell, G., and Yu, X. (2011). Disciplining delegated monitors: the consequences of failing to prevent fraud. Unpublished working paper.

Townsend, R. M. (1979). Optimal contracts and competitive markets with costly state verification. Journal of Economic theory, 21(2):265–293.

Ueda, M. (2004). Banks versus venture capital: Project evaluation, screening, and expropriation. The Journal of Finance, 59(2):601–621.

Winton, A. and Yerramilli, V. (2008). Entrepreneurial finance: Banks versus venture capital. Journal of Financial Economics, 88(1):51–79.

Appendix A Hiring Assumption

A.1 The Assumption

We assume high effort is sufficient of low cost, and enough value adding such that it is optimal for the VC to provide incentives to induce high effort.

Mathematically, Hiring Assumption is delineated as follows: $\exists\, S(z) \in [0,1]$ such that she better provides incentives to induce hard work:

$$\begin{aligned} &\max_{K_1,K_2} \mathbb{E}(\Pi(K_1,K_2,S(z))|h,h) \geq \max\{\max_{K_1,K_2} \mathbb{E}(\Pi(K_1,K_2,S(z))|h,l), \\ &\qquad \max_{K_1,K_2} \mathbb{E}(\Pi(K_1,K_2,S(z))|l,h), \max_{K_1,K_2} \mathbb{E}(\Pi(K_1,K_2,S(z))|l,l), 0\} \qquad \text{(IC/IR-principal)} \end{aligned}$$

She also provides the capital investments for her best interest, denoted by

$$K_1^*(S(z)|e_1,e_2) = argmax_{K_1}\mathbb{E}(\Pi(K_1,K_2,S(z))|e_1,e_2)$$

$$K_2^*(S(z)|e_1,e_2) = argmax_{K_2}\mathbb{E}(\Pi(K_1,K_2,S(z))|e_1,e_2)$$

For the same sharing rule $S(z)$ and capital investments, the entrepreneur is better off to work than to shirk:

$$\begin{aligned} &\mathbb{E}(\pi(K_1^*(S(z)),K_2^*(S(z)),S(z))|h,h) \geq \max\{\mathbb{E}(\pi(K_1^*(S(z)),K_2^*(S(z)),S(z))|h,l), \\ &\qquad \mathbb{E}(\pi(K_1^*(S(z)),K_2^*(S(z)),S(z))|l,h), \mathbb{E}(\pi(K_1^*(S(z)),K_2^*(S(z)),S(z))|l,l), 0\} \\ &\qquad\qquad \text{(IC/IR-agent)} \end{aligned}$$

The two inequalities above guarantee the existence of the optimal sharing rule, $S^*(R)$, such that the equilibrium path reaches optimal 'game node'. Since effort costs are private to the agent and lump sum transfer is infeasible, the first best cannot be achieved. The hiring assumption guarantees that $S(z)$, the effort level (h,h) and the corresponding capital investment (K_1^*, K_2^*) are the second best.

A.2 Necessary conditions for the assumption

A.2.1 Bottomline: 'h' effort not induced

In the bottemline case, the principal offers a contract where it just provides incentive for the entrepreneur to work l. That is, his rationality is satisfied, but not incentive for high effort. This contract has a sharing rule $S^\dagger(z) = \mathbb{1}(z \geq x^\dagger)$, where $\int_{x^\dagger}^1 zg(z|l,l)dz = \frac{2c_l}{f(K_1^\dagger, K_2^\dagger)}$. Moreover $K_1^\dagger, K_2^\dagger = argmax_{K_1,K_2}\Pi(K_1, K_2, S^\dagger(z), l, l)$.

$$
\begin{aligned}
\pi^\dagger =& 0 \\
\Pi^\dagger =& \int_0^{x^\dagger} zg(z|l,l)dzf(K_1^\dagger, K_2^\dagger) - K_1^\dagger - K_2^\dagger \\
=& \mathbb{E}(Z|l,l)f(K_1^\dagger, K_2^\dagger) - 2c_l - K_1^\dagger - K_2^\dagger
\end{aligned}
$$

A.2.2 'h' effort induced

Now let the contract be such that 'h' effort is induced.

$$
\begin{aligned}
S^*(z) =& \mathbb{1}\{z \geq \min\{x_1, x_2\}\} \\
K_1^* =& argmax_{K_1} \int_0^{\min\{x_1,x_2\}} zdG(z|h,h)f(K_1, K_2) - K_1 - K_2 \\
K_2^* =& argmax_{K_2} \int_0^{\min\{x_1,x_2\}} zdG(z|h,h)f(K_1, K_2) - K_1 - K_2
\end{aligned}
$$

where $\int_{x_1}^{1} z(g(z|h,h)-g(z|l,l))dz = \dfrac{2c_h - 2c_l}{f(K_1,K_2)}$ and $\int_{x_2}^{1} z(g(z|h,h)-g(z|h,l))dz = \dfrac{c_h - c_l}{f(K_1,K_2)}$.

If $x_1 < x_2$:

$$\begin{aligned}
\pi^* &= \int_{x_1}^{1} zg(z|h,h)dzf(K_1^*,K_2^*) - 2c_h = \int_{x_1}^{1} zg(z|l,l)dzf(K_1^*,K_2^*) - 2c_l \\
\Pi^* &= \int_{0}^{x_1} zg(z|h,h)dzf(K_1^*,K_2^*) - K_1^* - K_2^* \\
&= \mathbb{E}(Z|h,h)f(K_1^*,K_2^*) - \pi^* - K_1^* - K_2^* - 2c_h
\end{aligned}$$

If $x_1 \geq x_2$:

$$\begin{aligned}
\pi^* &= \int_{x_2}^{1} zg(z|h,h)dzf(K_1^*,K_2^*) - 2c_h = \int_{x_2}^{1} zg(z|h,l)dzf(K_1^*,K_2^*) - c_l - c_h \\
\Pi^* &= \int_{0}^{x_2} zg(z|h,h)dzf(K_1^*,K_2^*) - K_1^* - K_2^* \\
&= \mathbb{E}(Z|h,h)f(K_1^*,K_2^*) - \pi^* - K_1^* - K_2^* - 2c_h
\end{aligned}$$

A.2.3 Analysis

When is the principal willing to induce effort and the agent to participate? That is, when are $\Pi^* \geq \Pi^\dagger$ and $\pi^* \geq \pi^\dagger = 0$ true?

$$\begin{aligned}
\mathbb{E}(Z|h,h)f(K_1^*,K_2^*) - \pi^* - K_1^* - K_2^* - 2c_h &\geq \mathbb{E}(Z|l,l)f(K_1^\dagger,K_2^\dagger) - 2c_l - K_1^\dagger - K_2^\dagger \\
\int_{x_1}^{1} zg(z|l,l)dzf(K_1^*,K_2^*) &\geq \int_{x^\dagger}^{1} zg(z|l,l)dzf(K_1^\dagger,K_2^\dagger) \quad \text{if } x_1 < x_2 \\
\int_{x_2}^{1} zg(z|h,l)dzf(K_1^*,K_2^*) - c_h &\geq \int_{x^\dagger}^{1} zg(z|l,l)dzf(K_1^\dagger,K_2^\dagger) - c_l \quad \text{if } x_1 \geq x_2
\end{aligned}$$

The first inequality above says that the principal's incentive is satisfied as long as the net project value (deducting the labor costs and the agent's rent) of high effort exceeds that of low effort. Note that since I assume that $f(\cdot,\cdot)$ is concave and increasing in both arguments, by theorem of maximum, $\Pi^* > \Pi^\dagger$ implies $f(K_1^*,K_2^*) > f(K_1^\dagger,K_2^\dagger)$.

Therefore, the difference in the cost of effort cannot too large so that it's less than the increment of the project value.

The latter two inequalities mean that the agent's payoff under high effort must exceeds 0, his participation or the least incentive for low effort.

Appendix B Optimal Contract

B.1 Benchmark: No monitoring

By (IC/IR-agent-1), the optimal sharing rule is at when the entrepreneur is indifferent between working hard and shirking. Apply Euler-Lagrange equation and functional derivatives, the optimal sharing rule is a step function:

$$S(z) = \mathbb{1}\{z \geq \min\{x_1, x_2\}\}, \tag{21}$$

where $\int_{x_2}^{1} z(g(z|h,h)-g(z|h,l))dz = \frac{c_h - c_l}{f(K_1, K_2)}$ and $\int_{x_1}^{1} z(g(z|h,h)-g(z|l,l))dz = \frac{2c_h - 2c_l}{f(K_1, K_2)}$.

Now we find the first order conditions on capital investments

$$\int_{0}^{\min\{x_1,x_2\}} zf(K_1, K_2)\ d\ G(z|h,h) - K_1 - K_2 \tag{22}$$

B.1.1 Case 1: $x_1 \leq x_2$

F.O.C on K_1:

$$x_1 g(x_1|h,h)\frac{\partial x_1}{\partial K_1} f(K_1, K_2) + \int_{0}^{x_1} z(g(z|h,h))dz f_1(K_1, K_2) - 1 = 0 \tag{23}$$

Since we know $\int_{x_1}^{1} z(g(z|h,h) - g(z|l,l))dz = \frac{2c_h - 2c_l}{f(K_1, K_2)}$, by Implicit Function Thereon

$$\frac{\partial x_1(K_1, K_2)}{\partial K_1} = \frac{2(c_h - c_l) f_1(K_1, K_2)}{(f(K_1, K_2))^2 \cdot x_1(g(x_1|h,h) - g(x_1|l,l))} \tag{24}$$

Eventually we have:

$$f_1(K_1, K_2)(\frac{g(x_1|h,h)}{g(x_1|h,h) - g(x_1|l,l)} \int_{x_1}^{1} z(g(z|h,h) - g(z|l,l))dz + \int_{0}^{x_1} zg(z|h,h)dz) = 1$$
$$f_2(K_1, K_2)(\frac{g(x_1|h,h)}{g(x_1|h,h) - g(x_1|l,l)} \int_{x_1}^{1} z(g(z|h,h) - g(z|l,l))dz + \int_{0}^{x_1} zg(z|h,h)dz) = 1$$

where $f_1(\cdot,\cdot)$ and $f_2(\cdot,\cdot)$ are partial derivatives of $f(\cdot,\cdot)$ on each argument.

B.1.2 Case 2: $x_1 > x_2$

F.O.C on K_1:

$$x_2 g(x_2|h,h) \frac{\partial x_2}{\partial K_1} f(K_1, K_2) + \int_{0}^{x_2} z(g(z|h,h))dz f_1(K_1, K_2) - 1 = 0 \tag{25}$$

Since we know $\int_{x_2}^{1} z(g(z|h,h) - g(z|h,l))dz = \frac{c_h - c_l}{f(K_1, K_2)}$, by Implicit Function Thereon

$$\frac{\partial x_2(K_1, K_2)}{\partial K_1} = \frac{(c_h - c_l) f_1(K_1, K_2)}{(f(K_1, K_2))^2 \cdot x_2(g(x_2|h,h) - g(x_2|h,l))} \tag{26}$$

Eventually we have:

$$f_1(K_1, K_2)(\frac{g(x_2|h,h)}{g(x_2|h,h) - g(x_2|l,l)} \int_{x_2}^{1} z(g(z|h,h) - g(z|h,l))dz + \int_{0}^{x_2} zg(z|h,h)dz) = 1$$
$$f_2(K_1, K_2)(\frac{g(x_2|h,h)}{g(x_2|h,h) - g(x_2|l,l)} \int_{x_2}^{1} z(g(z|h,h) - g(z|h,l))dz + \int_{0}^{x_2} zg(z|h,h)dz) = 1$$

where $f_1(\cdot,\cdot)$ and $f_2(\cdot,\cdot)$ are partial derivatives of $f(\cdot,\cdot)$ on each argument.

B.2 Deterministic monitoring

By (ICs-2), the optimal sharing rule and optimal capital investments are

$$S(z) = \mathbb{1}\{z \geq x_2\}, \tag{27}$$

where $\int_{x_2}^{1} z(g(z|h,h) - g(z|h,l))dz = \dfrac{c_h - c_l}{f(K_1, K_2)}$.

By Implicit Function Theorem

$$\frac{\partial x_2(K_1, K_2)}{\partial K_1} = \frac{(c_h - c_l) f_1(K_1, K_2)}{(f(K_1, K_2))^2 \cdot x_2(g(x_2|h,h) - g(x_2|h,l))} \tag{28}$$

The first order condition on capital investments are as follows:

$$f_1(K_1, K_2)(\frac{g(x_2|h,h)}{g(x_2|h,h) - g(x_2|h,l)} \int_{x_2}^{1} z(g(z|h,h) - g(z|h,l))dz + \int_0^{x_2} zg(z|h,h)dz) = 1$$
$$f_2(K_1, K_2)(\frac{g(x_2|h,h)}{g(x_2|h,h) - g(x_2|h,l)} \int_{x_2}^{1} z(g(z|h,h) - g(z|h,l))dz + \int_0^{x_2} zg(z|h,h)dz) = 1$$

B.2.1 K_2' and $S'(z)$

K_2' and $S'(z)$ depends on the functional form of $F(\cdot,\cdot)$ and $G(\cdot)$. A **special case** where I let $K_2 \equiv K_2'$, $S'(z)$ satisfies

$$\int_0^1 z(1 - S'(z))g(z|h,h)dz > \int_0^1 z(1 - S^*(z))g(z|h,h)dz > \int_0^1 z(1 - S'(z))g(z|h,l)dz \tag{29}$$
$$\int_0^1 z(S'(z) - S^*(z))(g(z|h,h) - g(z|h,l))dz < 0 \tag{30}$$
$$x_1 < x_2 \tag{31}$$
$$\int_0^1 zS'(z)g(z|l,l)dz < \int_0^1 zS^*(z)(2g(z|h,l) - g(z|h,h))dz \tag{32}$$

where $\int_{x_1}^{1} z(g(z|h,h) - g(z|l,l))dz = \dfrac{2c_h - 2c_l}{f(K_1, K_2)}$ and $\int_{x_2}^{1} z(g(z|h,h) - g(z|h,l))dz = \dfrac{c_h - c_l}{f(K_1, K_2)}$

Note that $S'(z)$ is not unique. Now we solve the boundary conditions for $S'(z)$, and thus prove its existence. As we don't acknowledge the functional form of $f(\cdot,\cdot)$, we assume $K_2' = K_2$. The upper bound for $S'(z)$ is just $S^*(z)$, because $S^*(z)$ is the least amount of share to induce high effort. Now, we solve the lower bound, and by showing the entrepreneur's punishment against the VC's moral hazard, we prove the existence of $S'(z)$. The intuition to establish the lower bound is that the VC's worst potential punishment option must not be expropriating, which does not induce first stage effort. We have $S'(z) \geq \underline{S'(z)}$, where

$$\underline{S'(z)} = \mathbb{1}\{z \geq x_2'\}, \tag{33}$$

where $\int_0^{x_2'} zg(z|h,l)dz = \int_0^{x_2} zg(z|h,h)dz$.

B.3 Random monitoring

The selection on $S'(z)$ is no longer up to a bound. To get the least monitoring intensity to induce high effort, let $S'(z)$ be the harshest possible credible punishment. That is, the best choice on the alternative sharing rule reaches its lower bound.

The incentive for the VC to prevent her moral hazard:

$$\int_0^1 z(1-S^*(z))g(z|h,h)dzf(K_1^*,K_2^*) - K_2^* \geq \int_0^1 z(1-S'(z))g(z|h,l)dzf(K_1^*,K_2') - K_2' \tag{IC-principal}$$

The incentive for the entrepreneur's effort ex ante, at the threshold:

$$\begin{aligned} &\mathbb{E}(\pi(K_1^*,K_2^*,S^*(z)|h,h)) = \\ &p\mathbb{E}(\pi(K_1^*,K_2',S'(z)|l,l)) + (1-p)\mathbb{E}(\pi(K_1^*,K_2^*,S^*(z)|l,l)) \end{aligned} \tag{34}$$

From the condition above, the monitoring intensity $p(S'(z), K_2')$ is as follows

$$p^*(K_1, K_2, K_2', S(z), S'(z)) = \frac{\int_{x_2}^{1} z(g(z|l,l) - g(z|h,h))dz f(K_1, K_2) + 2(c_h - c_l)}{\int_{x_2}^{1} zg(z|l,l)dz f(K_1, K_2) - \int_{x_2'}^{1} zg(z|l,l)dz f(K_1, K_2')} \tag{35}$$

Since the solution to the general case requires the functional form of $f(\cdot,\cdot)$, here I show the special case when $K_2' \equiv K_2$. Remember $\int_{x_1}^{1} z(g(z|h,h) - g(z|l,l))dz = \frac{2c_h - 2c_l}{f(K_1, K_2)}$,

$$\begin{aligned} p^*(K_1, K_2, K_2', S(z), S'(z)) =& \frac{\int_{x_2}^{1} z(g(z|l,l) - g(z|h,h))dz + \frac{2(c_h - c_l)}{f(K_1, K_2)}}{\int_{x_2'}^{x_2} zg(z|l,l)dz} \\ =& \frac{\int_{x_1}^{x_2} z(g(z|h,h) - g(z|l,l))dz}{\int_{x_2'}^{x_2} zg(z|l,l)dz} \in (0,1) \end{aligned}$$

Therefore, proposition 1 still holds. That is, $x_1 < x_2$ is a necessary condition to monitor.

We already know that:

$$\begin{aligned} \frac{\partial x_1(K_1, K_2)}{\partial K_1} &= \frac{2(c_h - c_l) f_1(K_1, K_2)}{(f(K_1, K_2))^2 \cdot x_1(g(x_1|h,h) - g(x_1|l,l))} \\ \frac{\partial x_2(K_1, K_2)}{\partial K_1} &= \frac{(c_h - c_l) f_1(K_1, K_2)}{(f(K_1, K_2))^2 \cdot x_2(g(x_2|h,h) - g(x_2|h,l))} \end{aligned}$$

$S'(z)$ reaches the lower bound derived in B.2.1.

$$S'(z) = \mathbb{1}\{z \geq x_2'\},$$

where $\int_{0}^{x_2'} zg(z|h,l)dz = \int_{0}^{x_2} zg(z|h,h)dz$

Appendix C Proofs of Propositions

Proof of Proposition 1, implication 1. The premises of the proof satiate: the Hiring Assumption and $x_1 < x_2$. The proof of the proposition is facilitated by the graphs below.

The graphs show 3 pdfs of $g(z|l,l)$, $g(z|h,l)$ and $g(z|h,h)$ that follow monotone likelihood ratio property. By MLRP, $g(z|l,l)$, $g(z|h,l)$ and $g(z|h,h)$ intersect each other once and only once on (0,1). (1. If any of them intercept more than once, MLRP is violated. 2. Their intersections are on (0,1) because they are pdfs with [0,1] support.) Denote the largest among all the three intersection to be $\underline{x}$.

Corollary 2. $\min(x_1, x_2) > \underline{x}$, *where* $\int_{x_2}^{1} z(g(z|h,h) - g(z|h,l))dz = \frac{c_h - c_l}{f(K_1, K_2)}$ *and* $\int_{x_1}^{1} z(g(z|h,h) - g(z|l,l))dz = \frac{2c_h - 2c_l}{f(K_1, K_2)}$.

Proof of the corollary. The corollary is simply a result of the Hiring Assumption. In the hiring assumption, we assumed that the principal has incentive to induce the agent's effort. If the principal gets even more from low effort, i.e. $\int_{0}^{\underline{x}} z(g(z|h,h) - g(z|h,l))dz < 0$, the Hiring Assumption is not satisfied. □

Therefore, $x_1 > \underline{x}$ and $x_2 > \underline{x}$. Remember, for monitoring to provide incentive, $x_1 < x_2$, where $\int_{x_2}^{1} z(g(z|h,h) - g(z|h,l))dz = \frac{c_h - c_l}{f(K_1, K_2)}$ and $\int_{x_1}^{1} z(g(z|h,h) - g(z|l,l))dz = \frac{2c_h - 2c_l}{f(K_1, K_2)}$.

$\int_{x_1}^{1} z(g(z|h,h) - g(z|l,l))dz = 2\int_{x_2}^{1} z(g(z|h,h) - g(z|h,l))dz$. Graphically, x_1 pins down the definite integral which marks the area, A_1, of gray and purple. x_2 pins down the area, A_2, of the gray only. In the threshold condition, where $x_1 = x_2$, $A_1 = 2A_2$. In other words, effort is constant return to scale. (See figure 1)

Since the premises are $x_1 < x_2$, $A_1 = 2A_2$, it must be the case showing in figure 2.

In this case, the efforts are increasing return to scale. □

Proof of Proposition 1, implication 2. The agents' moral hazard states that there's in-

Figure 1: g(z)

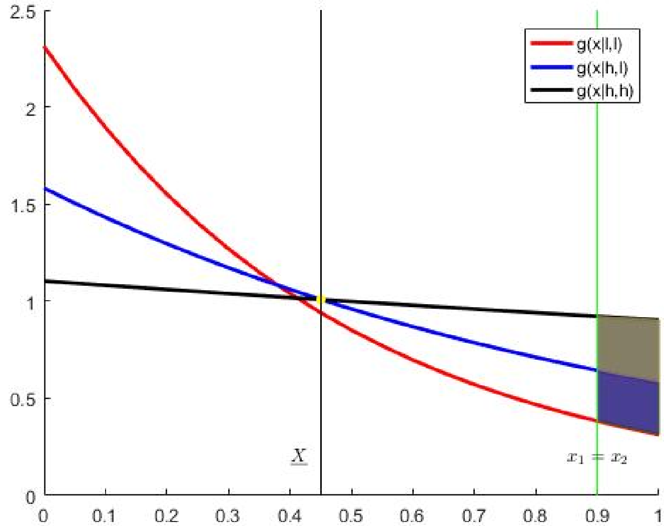
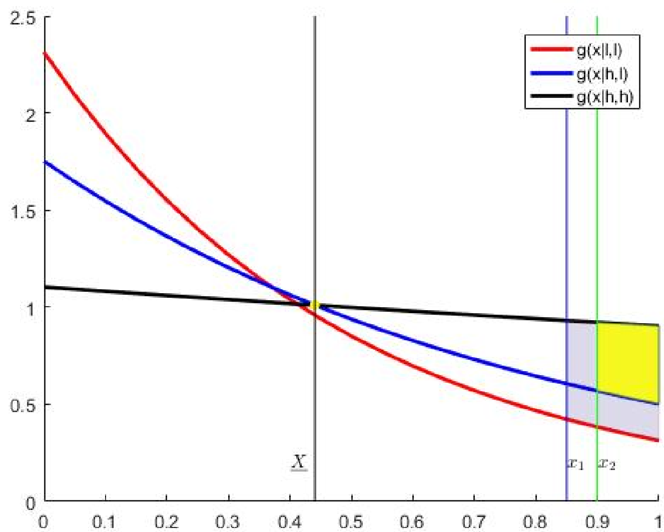


Figure 2: g(z)



centive to devote 'l' effort, had him been rewarded with the higher share $\{S^*(z), K_2^*\}$.

$$\mathbb{E}(\pi(K_1^*, K_2^*, S^*(z)|l,l)) \geq \mathbb{E}(\pi(K_1^*, K_2^*, S^*(z)|h,h)$$

Note that

$$\mathbb{E}(\pi(K_1^*, K_2^*, S^*(z)|l,l)) = \int_{x_2}^{1} zg(z|l,l)dzf(K_1^*, K_2^*) - 2c_l$$

$$\mathbb{E}(\pi(K_1^*, K_2^*, S^*(z)|h,h)) = \int_{x_2}^{1} zg(z|h,h)dzf(K_1^*, K_2^*) - 2c_h$$

Therefore

$$\mathbb{E}(\pi(K_1^*, K_2^*, S^*(z)|l,l)) - \mathbb{E}(\pi(K_1^*, K_2^*, S^*(z)|h,h)$$

$$= \int_{x_2}^{1} z(g(z|l,l) - g(z|h,h))dzf(K_1^*, K_2^*) + 2(c_h - c_l) \geq 0$$

We have

$$\int_{x_2}^{1} z(g(z|h,h) - g(z|l,l))dz \leq \frac{2(c_h - c_l)}{f(K_1^*, K_2^*)}$$

x_2 is the least amount of share to the entrepreneur that provokes his moral hazard. Remember $\int_{x_1}^{1} z(g(z|h,h) - g(z|l,l))dz = \frac{2(c_h - c_l)}{f(K_1^*, K_2^*)}$. Therefore,

$$\int_{x_2}^{1} z(g(z|h,h) - g(z|l,l))dz \leq \int_{x_1}^{1} z(g(z|h,h) - g(z|l,l))dz$$

Then we have $x_1 \leq x_2$. □

Proof of proposition 2. First, let's see the first order conditions on the capital investments. WLOG, I only compare the first order conditions on K_1. Denote the optimal investment of K_1 in the benchmark contract as K_1^N, and that in the contract with monitoring (deterministic/random) is K_1^M.

Remember that our objective is to maximize $\Pi(K_1, K_2, S(z)|h, h)$, and we assumed $x_1 < x_2$. In the benchmark contract:

$$\Pi(K_1, K_2, S(z)|h, h) = \int_0^{x_1} zg(z|h, h)dzf(K_1, K_2) - K_1 - K_2 \tag{36}$$

In the contract with monitoring

$$\Pi(K_1, K_2, S(z)|h, h) = \int_0^{x_2} zg(z|h, h)dzf(K_1, K_2) - K_1 - K_2 - p(S'(z), K_2') \tag{37}$$

Note that x_1 and x_2 are implicit functions of K_1, K_2: $\int_{x_2}^1 z(g(z|h, h) - g(z|h, l))dz \equiv \frac{c_h - c_l}{f(K_1, K_2)}$ and $\int_{x_1}^1 z(g(z|h, h) - g(z|l, l))dz \equiv \frac{2c_h - 2c_l}{f(K_1, K_2)}$.
F.O.C on K_1 for benchmark

$$f_1(K_1^N, K_2^N)(\frac{g(x_1|h, h)}{g(x_1|h, h) - g(x_1|l, l)} \int_{x_1}^1 z(g(z|h, h) - g(z|l, l))dz + \int_0^{x_1} zg(z|h, h)dz) = 1$$

and F.O.C on K_1 for the contract with monitoring, and $\frac{dp^*}{K_2} = 0$

$$f_1(K_1^M, K_2^M)(\frac{g(x_2|h, h)}{g(x_2|h, h) - g(x_2|h, l)} \int_{x_2}^1 z(g(z|h, h) - g(z|h, l))dz + \int_0^{x_2} zg(z|h, h)dz) = 1$$

First we multiply the production on both side of the equations:

$$f(K_1^N, K_2^N)(\frac{g(x_1|h, h)}{g(x_1|h, h) - g(x_1|l, l)} \int_{x_1}^1 z(g(z|h, h) - g(z|l, l))dz + \int_0^{x_1} zg(z|h, h)dz) = \frac{f(K_1^N, K_2^N)}{f_1(K_1^N, K_2^N)}$$

and

$$f(K_1^M, K_2^M)(\frac{g(x_2|h, h)}{g(x_2|h, h) - g(x_2|h, l)} \int_{x_2}^1 z(g(z|h, h) - g(z|h, l))dz + \int_0^{x_2} zg(z|h, h)dz) = \frac{f(K_1^M, K_2^M)}{f_1(K_1^M, K_2^M)}$$

which yield

$$\frac{2g(x_1|h, h)(c_h - c_l)}{g(x_1|h, h) - g(x_1|l, l)} + \mathbb{E}(\Pi(h, h, K_1^N, K_2^N, S_N(z))) = \frac{f(K_1^N, K_2^N)}{f_1(K_1^N, K_2^N)}$$

and

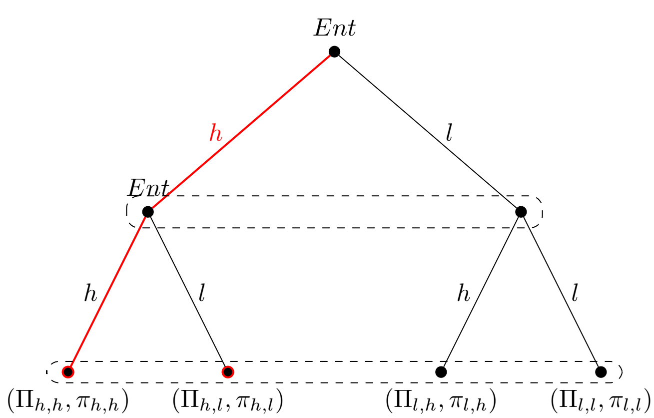
$$\frac{g(x_2|h,h)(c_h - c_l)}{g(x_2|h,h) - g(x_2|h,l)} + \mathbb{E}(\Pi(h,h,K_1^M,K_2^M,S_M(z))) + p \cdot C = \frac{f(K_1^M,K_2^M)}{f_1(K_1^M,K_2^M)}$$

It is clear that $\mathbb{E}(\Pi(K_1^M,K_2^M,S_M(z))|h,h) > \mathbb{E}(\Pi(K_1^N,K_2^N,S_N(z))|h,h)$ because VC must be better off if she monitors. Also, since $2g(x_1|h,l) < g(x_1|h,h) + g(x_1|l,l)$ is a necessary condition for VC to monitor(will be shown later), by Monotone Likelihood Ratio Property we previously assumed along with our previous result $x_2 > x_1$, we have $\frac{2g(x_2|h,l)}{g(x_2|h,h)} - \frac{g(x_2|l,l)}{g(x_2|h,h)} < 1$. Therefore, immediately there is, $\frac{g(x_2|h,h)}{g(x_2|h,h) - g(x_2|h,l)} > \frac{2g(x_1|h,h)}{g(x_1|h,h) - g(x_1|l,l)}$.

Since $C \geq 0$ and $f(\cdot,\cdot)$ is concave, it is obvious that $K_1^M > K_1^N$ and $K_2^M > K_2^N$. $\square$

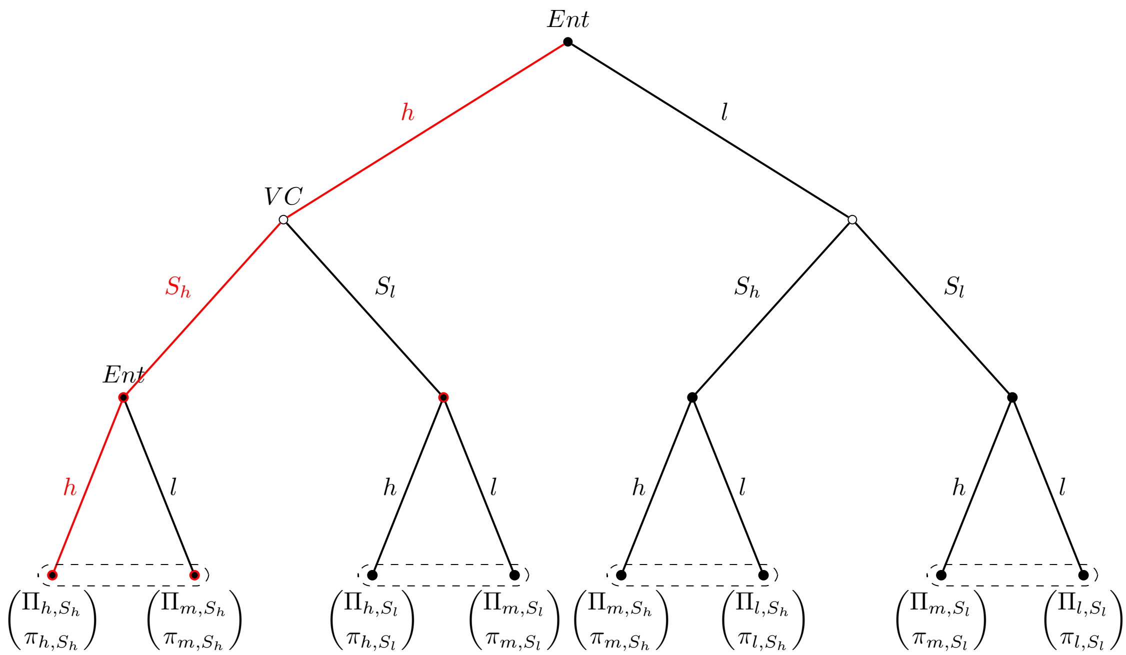
Appendix D Game Extensive Structure

D.1 No monitoring



Note: VC's information set is represented by the dashed boxes.

D.2 Deterministic Monitoring



Appendix E General Case for T>2

In general, the total number of financing stages be $T \leq \infty$. Monitoring provides incentives for the first $T-1$ periods. The initial menu contract specifies seeding stage investment K_1 and encapsulates a set of the bundles $\{S^i(z), K^i\}_0$, and grants the VC to choose a proper subset after each periods of monitoring recursively. That is, $\{S^i(z), K^i\}_t \subset \{S^i(z), K^i\}_{t-1}$. $\{S^i(z), K^i\}_T$ must be a singleton set.

Lemma 1. *If number of monitoring periods is T, then the cardinality of $\{S^i(R)\}_t$ $\forall t = 0, \ldots, T-1$ must be strictly greater than $T-t$ and it can be infinite.*

Claim: if a bounded mapping $S(z)$ satisfies the constraint, $a < |\int f(S^i(z), z)d\nu(z)| < b$, and if $f(\cdot,\cdot)$ is continuous in its first argument and $\nu(\cdot)$ is a well-defined probability measure on the space of Z, then there are infinitely many $S(z)$ also satisfy the constraint.

Proof of lemma 1. f is continuous in the first argument, such that for every $\varepsilon > 0$, $\exists\ \delta$ such that $\forall S(R) \in B(S^i(R), \delta)_{||L^1||}$ there is $B(\int f(S(R), R)d\nu(R), \varepsilon) \subset (a, b)$. $\square$

Claim: The cardinality of $\{S^i(z)\}_t\ \forall t = 0, \ldots, T-1$ must be strictly greater $T-t$. Or in other words, the cardinality of $\{S^i(z)\}_t$ must be strictly larger than remaining periods of monitoring.

Proof. If $T - t = 1$ and assume $|\{S^i(z)\}_t| \leq 1$, the contract is not contingent and the monitoring at $t = T$ cannot provide incentives.

Assume $T - t = n$, and $|\{S^i(z)\}_t| \geq n + 1$, then when $T - t = n + 1$, for its next monitoring to provide incentives, $|\{S^i(z)\}_t| > n + 1$, subsequently, $|\{S^i(z)\}_t| \geq n + 2$. The subset chosen must be a proper subset so that monitoring provides incentives, i.e. $\{S^i(z)\}_t \subsetneq \{S^i(z)\}_{t-1}$ $\square$

Appendix F Risk Preference

In this section, I study how adding risk aversion would affect monitoring intensity and the contract.

F.1 the VC risk neutral, and the entrepreneur risk averse

Let the VC be risk neutral and the entrepreneur be risk averse. Denote the entrepreneur's utility $u(\cdot)$, concave and increasing. In the following, we solve the contracts under the special case where $u(x) = \log(x + 1)$.[24]

[24] Note that inequality 4 has a slight change due to the utility function. See the inequality below

$$\begin{aligned} &\mathbb{E}(u(\pi(h, h, K_1, K_2, S(z)))) \\ \geq &\max\{\mathbb{E}(u(\pi(h, l, K_1, K_2, S(z)))), \mathbb{E}(u(\pi(l, h, K_1, K_2, S(z)))), \mathbb{E}(u(\pi(l, l, K_1, K_2, S(z)))), 0\} \end{aligned} \tag{4*}$$

F.1.1 No Monitoring

$$\max_{(S(z),K_1,K_2)} \mathbb{E}(\Pi(K_1, K_2, S(z)|h, h)) \tag{38}$$

$$\text{s.t. } (4^*)\ ,\ 0 \le S(z) \le 1.$$

The sharing rule of the contract is:

$$S(z) = \mathbb{1}\{z \ge \min(\widetilde{x_1}, \widetilde{x_2})\}, \tag{39}$$

where $\int_{\widetilde{x_1}}^{1} \log(zf(K_1, K_2))(g(z|h, h) - g(z|l, l))dz = 2(c_h - c_l)$,
and $\int_{\widetilde{x_2}}^{1} \log(zf(K_1, K_2))(g(z|h, h) - g(z|h, l))dz = c_h - c_l$

Capital Investments:

$$\max_{(K_1,K_2)} \int_0^{\widetilde{x_1}(K_1,K_2)} zf(K_1, K_2)d\ G(z|h, h) - K_1 - K_2 \tag{40}$$

FOC on K_1:

$$\widetilde{x_1} g(\widetilde{x_1}|h, h)\frac{\partial \widetilde{x_1}}{\partial K_1} \cdot f(K_1, K_2) + \int_0^{\widetilde{x_1}} zg(z|h, h)dz \cdot f_1(K_1, K_2) - 1 = 0 \tag{41}$$

Apply Implicit Function Theorem, and we get:

$$\frac{\partial \widetilde{x_1}}{\partial K_1} = \frac{f_1(K_1, K_2) \int_{\widetilde{x_1}}^{1} (g(z|h, h) - g(z|l, l))dz}{\widetilde{x_1} f(K_1, K_2) \log(\widetilde{x_1} f(K_1, K_2)) \cdot (g(\widetilde{x_1}|h, h) - g(\widetilde{x_1}|l, l))} \tag{42}$$

Combining two equations above, we can solve for K_1. Solving for K_2 is an analogue.

Similarly other ICs and IRs are adjusted, i.e. inequality (6), (8), (10), (11), (13) and (15)

F.1.2 Random Monitoring

$$\max_{(p,S(z),S'(z),K_1,K_2,K_2')} \mathbb{E}(U(\Pi(K_1,K_2,S(z)|h,h))) \tag{43}$$

$$\text{s.t. (6)—(12), (15), } 0 \leq p \leq 1 \text{ }, 0 \leq S(z) \leq 1 \text{ }, 0 \leq S'(z) \leq 1$$

The sharing rule of the contract is:

$$S(z) = \mathbb{1}\{z \geq \widetilde{x_2}\}, \tag{44}$$

where $\int_{\widetilde{x_2}}^{1} \log(zf(K_1,K_2))(g(z|h,h) - g(z|h,l))dz = (c_h - c_l)$

The alternative Sharing rule when $K_2 = K_2'$:

$$S(z) = \mathbb{1}\{z \geq \widetilde{x_2'}\}, \tag{45}$$

where $\int_0^{\widetilde{x_2'}} zg(z|h,l)dz = \int_0^{\widetilde{x_2}} zg(z|h,h)dz$.

The optimal monitoring intensity is:

$$p^* = \frac{\int_{\widetilde{x_2}}^{\widetilde{x_1}} \log(f(K_1,K_2)z)(g(z|h,h) - g(z|l,l))dz}{\int_{\widetilde{x_2}}^{\widetilde{x_2'}} \log(f(K_1,K_2)z)g(z|l,l)dz} \tag{46}$$

Capital Investments:

$$\max_{(K_1,K_2)} \int_0^{\widetilde{x_2}(K_1,K_2)} zf(K_1,K_2)d\ G(z|h,h) - K_1 - K_2 \tag{47}$$

FOC on K_1:

$$\widetilde{x_2} g(\widetilde{x_2}|h,h)\frac{\partial \widetilde{x_2}}{\partial K_1} \cdot f(K_1,K_2) + \int_0^{\widetilde{x_2}} zg(z|h,h)dz \cdot f_1(K_1,K_2) - 1 = 0 \tag{48}$$

Apply Implicit Function Theorem, and we get:

$$\frac{\partial \widetilde{x_2}}{\partial K_1} = \frac{f_1(K_1, K_2)\int_{\widetilde{x_2}}^{1}(g(z|h,h) - g(z|h,l))dz}{\widetilde{x_2} f(K_1, K_2)\log(\widetilde{x_2} f(K_1, K_2)) \cdot (g(\widetilde{x_2}|h,h) - g(\widetilde{x_2}|h,l))} \tag{49}$$

F.2 the VC risk averse, and the entrepreneur risk neutral

Now we let the VC be risk averse, who has utility $U(\cdot)$, increasing and concave. That is, $U'(\cdot) > 0$ and $U''(\cdot) < 0$. Specifically, we delineate contracts where $U(x) = \log(x+1)$. Since we have derived that random monitoring weakly dominates deterministic monitoring, we no longer consider the latter case.

F.2.1 No Monitoring

$$\max_{(S(z), K_1, K_2)} \mathbb{E}(U(\Pi(K_1, K_2, S(z)|h,h))) \tag{50}$$

$$\text{s.t. } (4)\ ,\ 0 \leq S(z) \leq 1$$

Under a special case where $U(x) = \log(x+1)$, the share is

$$S(z) = 1 - \frac{\dfrac{g(z|h,h)}{\lambda(g(z|h,h) - g(z|l,l))} - 1}{f(K_1, K_2)z} \tag{51}$$

where the Lagrangian multiplier on the entrepreneur's IC is

$$\lambda = \frac{1}{\int_0^1 f(K_1, K_2)z(g(z|h,h) - g(z|l,l))dz - 2(c_h - c_l)} \tag{52}$$

F.2.2 Random Monitoring

$$\max_{(p,S(z),S'(z),K_1,K_2,K_2')} \mathbb{E}(U(\Pi(K_1,K_2,S(z)|h,h))) \tag{53}$$

$$\text{s.t. (6)—(12), (15), } 0 \le p \le 1 \text{ , } 0 \le S(z) \le 1 \text{ , } 0 \le S'(z) \le 1$$

Under the special case where $U(x) = \log(x+1)$, the share is

$$S(z) = 1 - \frac{\frac{g(z|h,h)}{\lambda(g(z|h,h) - g(z|h,l))} - 1}{f(K_1,K_2)z} \tag{54}$$

where the Lagrangian multiplier on the entrepreneur's IC is

$$\lambda = \frac{1}{\int_0^1 f(K_1,K_2)z(g(z|h,h) - g(z|h,l))dz - (c_h - c_l)} \tag{55}$$

The alternative sharing rule and optimal monitoring intensity:

$$S'(z) = 1 - \frac{\beta g(z|h,l)}{f(K_1,K_2)g(z|l,l)} \tag{56}$$

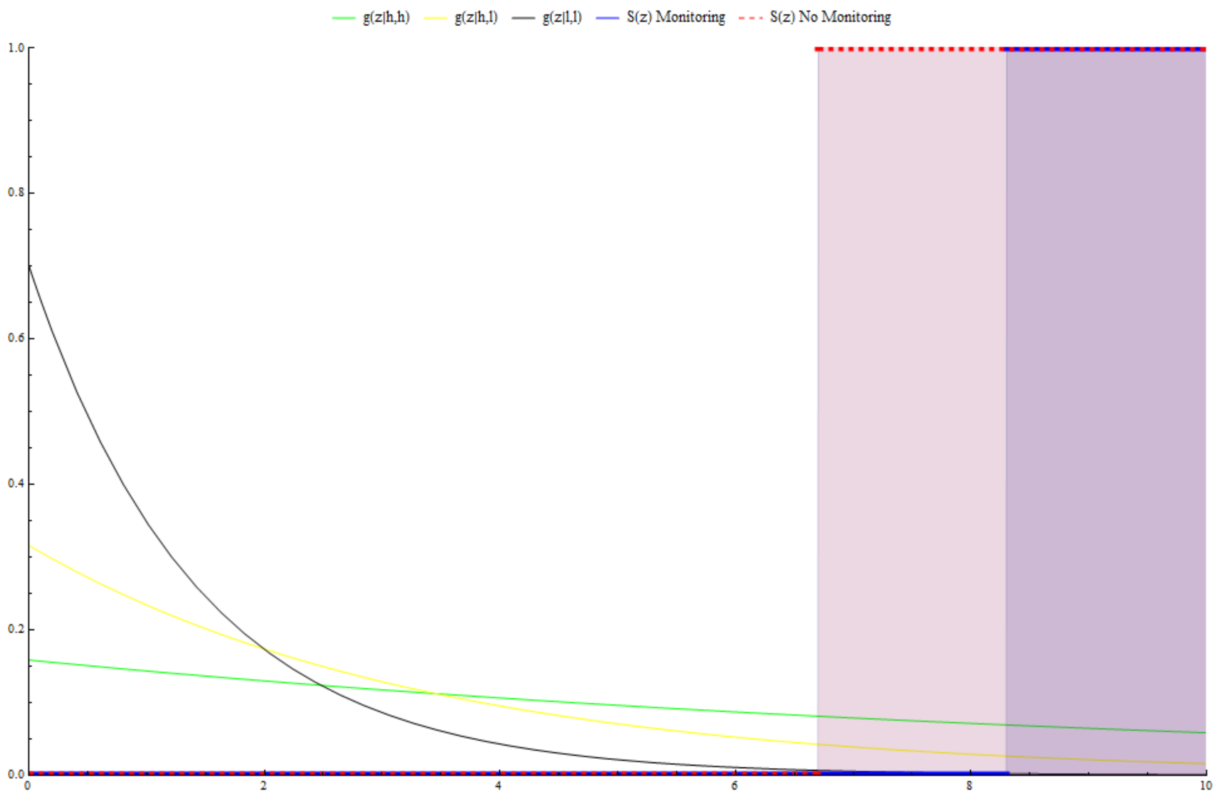
where the Lagrangian multiplier on the VC's IC is:

$$\beta = \frac{\int_0^1 \log(\frac{g^2(z|h,h)g(z|l,l)}{g^2(z|h,l)(g(z|h,h) - g(z|h,l))})dz}{\alpha} \tag{57}$$

The optimal monitoring intensity is:

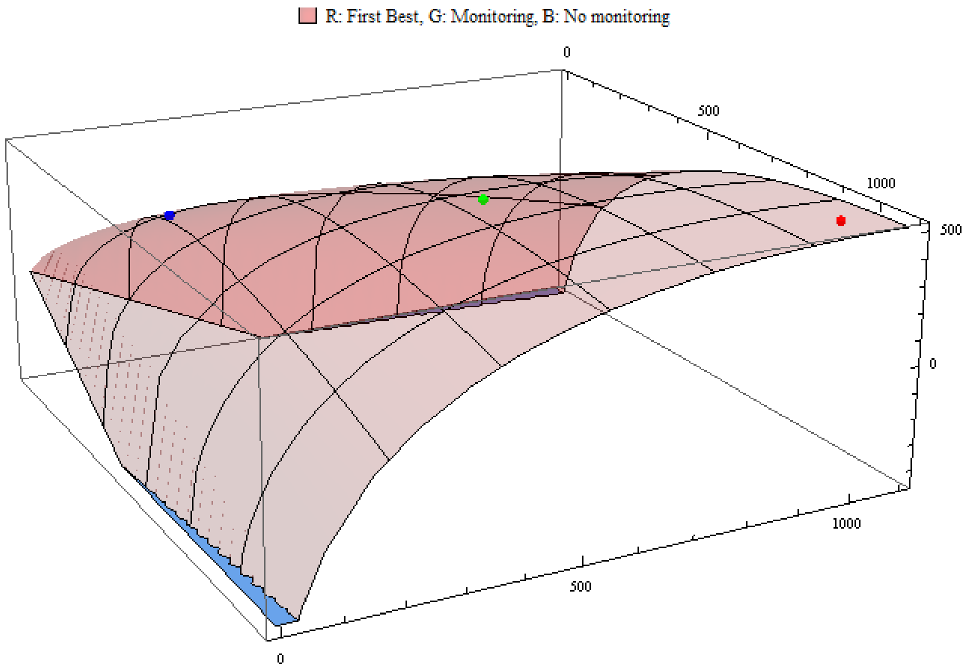
$$p^* = \frac{\frac{2(c_h - c_l)}{f(K_1,K_2)} - \int_0^1 S(z)(g(z|h,h) - g(z|l,l))dz}{\int_0^1 (S(z) - S'(z))g(z|l,l)dz} \tag{58}$$

Figure 3: Sharing Rules



The 3 curves in Figure 1 shows an example of the density functions of the productivity shock. Monotone Likelihood Ratio property is satisfied. The sharing rules are two indicator functions, demonstrated by the shadings. The VC receives higher share if she monitors, and thus the entrepreneur's (or the Residual Claimant's) share is lowered.

Figure 4: Capital Investments and Project Value



The curve above illustrates the net value of the project with capital investments on the X-Y plane

$$\int_0^1 zg(z|h,h)f(K_1,K_2)dz - K_1 - K_2 \tag{59}$$

The red dot indicates the socially optimal outcome (a.k.a the First Best). The blue dot refers to the outcome where the VC does not monitor. If necessary conditions for monitoring are satisfied, the outcome will be improved to the green dot. Agents' payoffs are determined upon the sharing rules as well as the monitoring expense.