P2P Lending: Information Externalities, Social Networks and Loans' Substitution*

Ester Faia
Goethe University Frankfurt and CEPR

Monica Paiella
University of Naples Parthenope

This draft: October 2018

Abstract

Despite the lack of delegated monitor P2P lending exhibits relatively low loan and delinquency rates. Adverse selection is mitigated by a new screening technology that provides costless public signals. Using data from Prosper and Lending Club we show that loans' spreads, proxing asymmetric information, decline with hard and soft information indicators, such as credit scores and measures of social networks. Also, an increase in bank fragility risk increases participation in P2P markets and reduces rates (substitution effect). We rationalize this evidence with a dynamic general equilibrium model, featuring asymmetric information in P2P lending and whereby public signals improve the information efficiency content of loan rates. In the model investors choose between traditional bank investment and P2P loans by comparing the relative risk of bank defaults with the risk of mis-pricing of P2P loans.

JEL codes: G11, G23.

Keywords: peer-to-peer lending, heterogenous projects, pooling equilibria, signals, Bayesian updating, value of information, bank fragility.

*We thank participants and discussants at various conferences and seminars. We gratefully acknowledge financial support from the DFG grant FA 1022/1-2 and Institut Louis Bachelier. Correspondence to: Ester Faia, Department of Money and Macro, Theodor Adorno Platz 3, 60323, Frankfurt am Main, Germany. Phone: +49-69-79833836. E-mail: faia@wiwi.uni-frankfurt.de.

1 Introduction

The online lending model emerged in 2005 as ordinary people began fulfilling the capital needs of anonymous borrowers over the Internet. Peer-to-peer (P2P) lending platforms are markets for consumers' debt where lenders and borrowers match and trade directly, hence in absence of intermediation. Borrowers describe the purpose of their loan request and provide information about their current financial situation. Furthermore, the platform provides hard information, such as credit scores, to lenders. Some platforms provide also soft information, such as recommendations from peers. Lenders can fund loans whose rates are based upon borrower and loan characteristics and on the signals available. In less than a decade, this new method of debt finance has grown significantly with a global proliferation of online marketplaces for several types of loans.[1] Most strikingly, data show that this nascent industry is performing well compared to the traditional banking sector. In fact, despite the absence of delegated monitors platforms exhibit low loan and delinquency rates. As we argue below, crucial for platforms is the availability of costless public signals that facilitate screening and help to mitigate the lemons' market adverse selection (see Akerlof [1]). P2P lending is indeed an innovation in screening technology rather than a new service.

The focus of our paper is on the information efficiency of platforms as captured by loan spreads, and on the possible substitutability between traditional banking and P2P markets. We use data from Prosper and Lending Club, the two biggest P2P platforms worldwide. These datasets offer a unique opportunity of analysis since all information publicly available to investors is also available to the econometrician, thereby eliminating the biases from unobserved heterogeneity. Using these datasets, we find evidence that signals mitigate adverse selection and reduce information premia. We also find evidence of a shift from traditional banking to the platforms, the more so when bank fragility raises. We rationalize the empirical results through a dynamic general equilibrium model with borrowers and lenders interacting in two intermediation sectors, given by traditional banks and P2P markets wherein asymmetric information is mitigated by signals whose precision is stochastic.

[1] According to research firm Liberum [21], in 2015, the online lending industry surpassed $28 billion in the US and Europe, and reached $157 billion in China. Currently, P2P lending makes up just less than 2.5% of the US unsecured personal lending market and a mere one tenth of one percent of the overall lending market. Venture Capital firm, Foundation Capital, predicts that by 2025, $1 trillion in loans will be originated in this manner globally.

Briefly, our model features optimizing risk-averse lenders that invest in bank deposits, which are subject to an haircut in the event of bank distress, and P2P loans, which are subject to the risk of borrowers' default. Borrowers are risk-neutral heterogenous agents who seek funds, either from banks or on the platform, for investing in risky projects whose default rates are stochastically distributed. Projects' quality is unknown to lenders. Banks can learn it by paying a cost. Lenders in P2P markets are unable to discern projects' default rates[2], but can observe public signals[3] that allow them to form conditional expectations of the success probability of each project[4]. The latter expectations are then used to price loans. In equilibrium a distribution of loan returns, as conditional expectations on signals, emerge. Lenders only fund projects whose expected return (conditional on the signal on the probability of default) is at least as high as the expected return on deposits. The arbitrage condition between expected returns on P2P loans and on deposits determines the allocation of funds between the two sectors. In that respect information affects the portfolio choice of investors[5]. Furthermore, signals precision improves lenders' conditional forecast of default rates and brings the separating equilibrium closer to full information efficiency.

Given this set-up, by solving analytically the model, we obtain three main predictions regarding the determinants of P2P loan prices, which we then test empirically. First, we establish a *selection channel* such that an increase in the average quality of projects reduces the adverse selection premia and lowers prices in P2P markets, by inducing a downward shift in the distribution of default rates. Second, an increase in transparency, as captured by signal precision, reduces information premia, hence loan rates (*information channel*)[6]. However, signals are truly informative, and the latter effect materializes, only when the platform features a good selection of borrowers, i.e. when the marginally funded borrower is on the right of the average borrower. Third, an increase in the risk of liquidity shocks in the banking sector increases the risk of haircuts on deposits, thereby shifting both lenders and borrowers to the platform. As participation in P2P markets increases, loan rates fall (*substitution effect*).

[2] Along the tradition of Stiglitz and Weiss [32].

[3] The role of public signals for market pricing, particularly so for innovative firms, has been considered also in past theoretical works by Allen, Morris and Shin[4], Bacchetta and van Wincoop[6] and recently by Tinn[34].

[4] In this we follow Petriconi [27] and Ruckes [30].

[5] The role of information acquisition for potfolio choice is studied in a general set-up by Peress[26].

[6] Our work focuses on the role of signals to ease asymmetric information. There is also work by Chemla and Tinn[12] that examines the role of learning on firms' incentives toward moral hazard.

Our empirical analysis captures the three channels through an econometric specification that links P2P loan rates to proxies for borrowers' quality (selection channel), dummies capturing borrowers' heterogeneity in information reporting (information channel), and proxies for the risk of liquidity shortages or fragility in the banking sector (substitution channel).

Specifically, in our regressions, we include hard information signals, such as FICO scores and other creditworthiness measures, and our results provide clear support to the hypothesis that signals affect loan rates. As an example, a one standard deviation increase in the FICO score, which implies an improvement in borrower quality, reduces the lending rate by 4 percentage points (over 20 percent of the mean rate). This reduction can result both from the selection channel (borrowers' quality has improved) and from the information channel (more signals are available, hence investors can better screen borrowers). To disentangle the two, we exploit the variability in signal reporting and find that, on the margin, borrowers who report more information pay lower premia.

A captivating aspect of the importance of information for pricing can be further explored with Prosper data which provide lenders also with soft information signals, such as recommendations and investment from friends and membership in groups of borrowers. Including those signals in the regressions allows us to test for and quantify any positive externality linked to *social multipliers*.[7] We find that group membership lowers loan rates by between half and two percentage points. Funding from friends lowers the rate by two to four points. This evidence supports the hypothesis examined in previous studies of the superiority of transparent markets over financial intermediaries due to the value added by diversity of opinions.[8] However, the effect is quantitatively smaller than that of hard signals. This implicitly suggests that social multipliers exist, but might be of quantitatively limited importance.

Finally, using data on 497 bank failures from the US Federal Deposit Insurance Corporation as proxies for bank fragility, we find that borrowers in US States where significant rates of bank closures were recorded pay significantly lower rates, which is consistent with our hypothesis that bank fragility increases participation in the platform, i.e. with the hypothesis of bank-platform

[7] Previous studies discussed information externalities linked to the *wisdom of the crowd*. Banerjee [7] is one. Bikhchandani, Hirshleifer and Welch [9] stress the value of information obtained by observing other investors' actions.

[8] See Allen and Gale [2].

substitutability.

Nevertheless, overall, this evidence is consistent with our model prediction that when there are signs of fragility in the banking system more borrowers and lenders turn to P2P platforms. This leads to lower equilibrium rates which reflect not only higher demand and higher supply, but also the presence of a larger share of good projects on the online market, with a reduction of information premia.

The rest of the paper is organized as follows. In the next section we give an account of P2P markets institutional design and provide a comparison to the literature. We then describe our model and its results (section 3). In section 4 we review our empirical analysis. Finally, we discuss extensions and policy implications in our conclusions (section 5).

2 Institutional Background and Related Literature

Online peer-to-peer loans owe their origin to the growing popularity of online communities. They essentially transfer the idea of personal credit to the Web. In this kind of lending model there is no intermediation by traditional financial institutions. In the aftermath of the recent financial crisis, the fragility of the banking system as well as the distrust of investors towards it have been one reason for the growing popularity of P2P lending. Here, the decision process involved in loan origination is given into the hands of private lenders and borrowers, and websites like Prosper or Lending Club offer them a platform to engage with each other. For borrowers, online P2P lending is a way to obtain a loan without turning to a financial institution, that extracts informational monopoly rents, and to eliminate the risk of early liquidation due to bank crises. In the platform loan rates are proportional to the (perceived) project quality, hence in presence of signalling good borrowers might find profitable to turn to the platform. For lenders, returns on the platforms are typically attractive compared to returns on deposits. Lenders however are unable to discern projects' quality exactly, but can only estimate it. In the traditional banking system investors face the risk of bank defaults or bank runs, while in the platform they have to bear losses of non-performing loans. Their relative participation in the two sectors is decided by balancing those risks. Overall, for both borrowers and lenders, the balance of costs and benefits between the two forms of investment determines their participation in the two sectors and the relative returns and premia.

As we document below, over time we have observed declining P2P loan and delinquency rates, a sign that platforms feature some information efficiency and can provide a valid alternative to the traditional intermediation sector. Our paper focuses on the role of information in driving prices and on the substitution between traditional banking and platform lending.

The analysis of P2P markets is relatively recent[9], but the literature is growing and several different aspects of FinTech have been examined. Here, we reference only the studies that are more related to our work. Specifically, few, mainly empirical papers have been assessing asymmetric information and the role of signals. These include Kawai, Onishi and Uetake [20] who estimate a model where borrowers can signal privately low default risk by posting low reserve interest rates. They show that adverse selection destroys as much as 16% of total surplus, but up to 95% can be restored with signaling.[10] The mechanism just described is well in line with the type of auction trading and price posting mechanism characterizing the early stages of some platforms. However, over time pricing on all platforms has moved to a centralized system. In this new institutional arrangement, prices/returns are conditional upon both private information provided by borrowers and hard information signals processed by machine learning algorithms, which increase platforms' information efficiency. This is consistent with the evidence from our empirical analysis which uncovers an increase in the impact of all signals on loan pricing after such change. For this reason, our model and empirical analysis will envisage a role for public signals beyond private ones.

On soft information in the FinTech industry there is the work by Chemla and Tinn[12] who develop a model that rationalizes lending through crowd-funding where crowdfunding allows firms to learn about total demand from a limited sample of target consumers. Learning creates a valuable real option as firms invest only if updated expectations of total demand are sufficiently high. This is particularly valuable for firms facing uncertainty about consumer preferences. Most importantly learning also enables firms to overcome moral hazard, even if diverting funds is costless. In our analysis we focus on the role of information for mitigating adverse selection using both a model-based analysis and an empirical assessment. On signals on FinTech platforms, there is also the

[9] A review of institutional aspects can be found in Bachmann, Becker, Buerckner, Hilker, Kock, Lehmann, Tiburtius, and Funk [5].

[10] For P2P lending Iyer, Khwaja, Luttmer and Shue [18] also examine empirically the role of interest rates as a signal of creditworthiness and find that the maximum rate that borrowers are willing to pay has a larger screening power than the credit score.

empirical paper by Lin, Prabhala and Viswanathan [22]. Using data from Prosper, they show that "platform friendships" lower interest rates on funded loans and are associated with lower ex post default rates. These authors do not consider other signals (private or hard), nor they examine substitution between P2P loans and other forms of credit. Indeed, none of the above papers considers the substitution between digital intermediation and traditional banking despite the fact that the post-crisis fragility of traditional financial institutions has played a crucial role in increasing platforms' popularity. Our paper does examine also this aspect. Specifically, in our paper we study both the importance of signals and the issue of stubsitution from an empirial perspective as well as with a model employing a signalling structure that has been previously used in a different context, specifically bank funding, by Petriconi[27]. P2P lending data provide an ideal environment to study these issues, as they provide lots of information about loan and lender characteristics. In our empirical analysis we also exploit the quasi-experiment nature of the shift from an auction-like to a centralized type of pricing mechanism and find that this contributed to increase transparency, hence market efficiency.

There are a number of related topics from past literature upon which our paper touches upon, such as the impact of signals and information acquisition on investors' portfolio decisions. The impact of information acquisition on investors' portfolio choice has been examined previously by Peress[26]. Besides this, our paper is related to the literature examining asymmetric information and the role of social ties in informal lending[11]. It contributes also to the literature comparing markets and banks (see Allen and Gale[3]). In markets, the relevant friction is information asymmetry as investors, mostly unsophisticated lenders, have to screen projects by themselves. A crucial innovation of P2P markets is the availability of public signals that lessen this information asymmetry. Notice that P2P platforms are akin to private equity markets, where the ex ante selection of borrowers includes young and risky entrepreneurs. Despite this, we show that public signals facilitate the revelation of information and determine the equilibrium heterogenous distribution of loans' returns[12]. Past works have also examined the role of public signals in equity markets, particularly for innovative firms (see works by Allen, Morris and Shin[4], Bacchetta and van Wincoop[6] and

[11] See for instance Besley and Coate [8].

[12] The relevance of news for the heterogenous distribution of stock returns has been highlighted also in the past by Fang and Peress[15].

recently Tinn[34]).

Last, a number of empirical papers on FinTech have examined the link between borrowers' attributes and listing outcomes. Ravina [28] looks at discrimination in lending on Prosper platform. Duarte, Siegel and Young[14] study the role of appearance for trust in P2P lending relations. Others use P2P data to infer investors' risk attitudes from investment decisions (see Paravisini, Ravina and Rappoport [25] with data from Lending Club). Our paper is less linked to this household finance perspective and more related to pricing mechanisms under asymmetric information and signalling and their consequences in general equilibrium.

3 The Model

The model is a dynamic general equilibrium model with borrowers and lenders interacting in two intermediation sectors, given by traditional banks and P2P markets. The dynamic general equilibrium perspective allows us to determine the flows of funds to each sector resulting from the forces of demand and supply and of arbitrage, and to assess the conditions and the extent upon which households substitute between the two.

The model is populated by borrowers and households/lenders and features one-side heterogeneity on the borrowers' side. Lenders are homogenous, risk-averse[13] patient households who save and can invest in bank deposits and in a portfolio of P2P loans. Households' saving is determined by the balance between their stochastic discount factor and asset returns. The fraction of savings invested into each sector is determined by the arbitrage condition between the return on deposits and the return on the portfolio of P2P loans. Both assets are risky. Hence, households shall balance those risks on the margin. Specifically, bank deposits are subject to a risk of run or liquidity dry-out in the banking sector which would result in an haircut on deposit rates. On the other side, each P2P loan is subject to the risk of default which households cannot quantify exactly, but can only form conditional expectations about, given an imprecise public signal. Signals allow us to assess the degree of information efficiency of the platform. Hence, signal precision also determines the premium that P2P borrowers shall pay to compensate lenders for the uncertainty regarding the risk of default.

[13]Risk-aversion allows us to retain a role for precautionary saving and intertemporal substitution. Both are important drivers of the stochastic discount factor and, hence, of households' willingness to engage in lending.

Borrowers are risk-neutral and seek funds for risky projects, whose idiosyncratic success probabilities are heterogenous and stochastically distributed. Notice that from now we will always use the index i to label the P2P projects. Borrowers can fund their projects through banks or on the platform. Borrowers' decision to participate in the platform versus resorting to bank credit is done by comparing the loan services to be paid in each sector.

Banks have access to a costly screening technology to determine exactly borrowers' risk of default. Hence, banks charge borrowers a rate that breaks-even with the cost of screening and the return banks have to pay on deposits. On the platform, borrowers pay project-specific returns, which depend on lenders' expected probability of project success conditional onto the signal.

Notice that the model features credit rationing because not all projects will be funded. On the platform, given the presence of signals, separating equilibria arises. The latter converge to the full-information equilibrium as signal precision is maximum. P2P lenders will only fund projects whose expected return, conditional on the signal about the probability of success, is higher than the expected return on deposits.[14] Equally, banks will only fund projects whose success probability, correctly discerned through screening, allows them to break-even. Those conditions allow us to determine the fraction of borrowers funded in each sector and the ones not funded at all.

3.1 Households/Lenders

In our model, risk-averse households/lenders pool their resources and make their consumption/saving decisions and investment decisions by maximizing the following lifetime expected utility:

$$\mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t u(C_t) \right\} \tag{1}$$

where $\mathbb{E}_0$ is the expectation operator conditional on the information as of time zero, and C_t denotes household consumption. In period t, households receive exogenous income Y_t. Households can invest their saving in demand deposits, D_t, and in P2P loans, X_t

Bank deposits, D_t, are one period bonds that promise to pay a non-contingent gross real rate of return $\bar{R}_{t+1}$ in the absence of bank default. However, banks can default either because of

[14] The presence of signals allows us to generate separating equilibria on the platform. Still credit rationing (not all P2P loans applicants are funded) can arise in separating equilibria. This is a well known result shown also in Wette [38].

insolvency, namely bank assets are too low to pay depositors, or because depositors run the bank and force early project liquidation. In the event of default, at $t+1$, depositors receive a fraction $\varkappa_{t+1} \in [0,1)$ of the promised return. We can interpret the fraction $\varkappa_{t+1}$ as an haircut applied to deposits. We can keep a time index to the haircut as this fits the case of uncertain returns, but the fraction can be easily assumed to be constant.

Let ς denote the probability of bank defaults. Given ς and $\varkappa_{t+1}$ we can write the gross return, R^d_{t+1}, on the deposit contract as:

$$R^d_{t+1} = \begin{cases} \bar{R}_{t+1} & \text{with probability } 1-\varsigma, \\ \varkappa_{t+1}\bar{R}_{t+1} & \text{with probability } \varsigma. \end{cases} \tag{2}$$

Households can also invest in a portfolio of heterogenous P2P loans, each succeeding with probability p^i. Let $R^{\mathcal{P}_t}_{t+1}$ denote the uncertain return on such portfolio. Households just choose how much of their savings are to be allocated to deposits and to P2P loans. Due to asymmetric information not all P2P loans' applications will be accepted. In the next section, we will determine the size of the loan portfolio by designing a pooling scheme through which the loan funded at the margin yields a return that, in expectations, is just equal to the expected return on one extra unit of deposits.

Households choose processes $\{C_t, D_t, X_t\}^{\infty}_{t=0}$ by maximizing the sum of future discounted utilities, (1), subject to the following budget constraint:

$$C_t + X_t + D_t \leq Y_t + R^{\mathcal{P}_t}_t X_{t-1} + R^d_t D_{t-1} \tag{3}$$

Let us define the stochastic discount factor between any period t and $t+1$ as $\Lambda_{t,t+1} = \frac{\beta u'(C_{t+1})}{u'(C_t)}$. Then, the first order condition for bank deposits reads as follows:

$$1 = \left[(1-\varsigma)\mathbb{E}_t(\Lambda_{t,t+1} \mid no\ def) + \varsigma\mathbb{E}_t(\Lambda_{t,t+1}\varkappa_{t+1} \mid def)\right]\bar{R}_{t+1} \tag{4}$$

where $\mathbb{E}_t(\Lambda_{t,t+1} \mid no\ def)$ and $\mathbb{E}_t(\Lambda_{t,t+1} \mid def)$ are the expected values conditional on no bank default or default at time $t+1$.

Similarly, we can derive the first order condition with respect to X_t which is as follows:

$$1 = \left[\mathbb{E}_t(\Lambda_{t,t+1}R^{\mathcal{P}_t}_{t+1})\right] \tag{5}$$

First order conditions (4) and (5) reveal the role of households' preferences and risk-attitudes for investment in both deposits and P2P lending. Indeed, in both cases, a change in the stochastic discount factor, which captures the extent of inter-temporal substitution and of precautionary savings, requires to be balanced by changes in returns or in their risk for households to maintain the same level of investment.

3.2 Determining the P2P Loan Portfolio through a Pooling Scheme

Projects funded on the platform are heterogeneous and, like in the standard set-up of hidden information a' la Stiglitz and Weiss [32], have different success probabilities such that, in each period, they can either succeed and pay a non-contingent, gross, real rate of return R^I_{t+1} with idiosyncratic probability p^i, or fail and pay zero with probability $1-p^i$.[15] Then, the random return on the i^{th} P2P project is:

$$R^i_{t+1} = \begin{cases} R^I_{t+1} & \text{with probability } p^i, \\ 0 & \text{with probability } 1-p^i. \end{cases} \tag{6}$$

Households cannot discern the exact probability of success, but form an expectation of such probability based on a signal, $\sigma_{i,\lambda}$ they may receive. We denote this estimated probability by $\pi^i = \mathcal{E}_t \left[p^i \mid \sigma_{i,\lambda} \right]$, where $\mathcal{E}_t$ denotes a Bayesian expectations. The exact expression for this will be derived later on in section 3.4.

Households fund a portfolio of loans some of which have expected returns that are higher than those on deposits, while others are lower. Specifically, households will fund all loans whose returns satisfy the following "break even" condition with respect to the returns on deposits:

$$\left[\pi^i \mathbb{E}_t(\Lambda_{t,t+1} \mid suc)\right] R^I_{t+1} \geq \left[(1-\varsigma)\mathbb{E}_t(\Lambda_{t,t+1} \mid no\ def) + \varsigma \mathbb{E}_t(\Lambda_{t,t+1} \varkappa_{t+1} \mid def)\right] \bar{R}_{t+1} \tag{7}$$

where $\mathbb{E}_t(\Lambda_{t,t+1} \mid suc)$ is the expectation conditional on the project succeeding at time $t+1$.

Condition (7) acts like an arbitrage condition and determines the allocation of household saving between the two assets and it will be used later on, in section 3.5, to determine the size of the P2P loan portfolio which will be financed in equilibrium. Besides this, condition (7) highlights some of the forces related to the main channels we are interested in. Specifically, it shows that if the risk

[15] For ease of notation we are omitting the time index, albeit probabilities may be time varying. Also, the model can be easily extended to the case of above zero returns in case of default.

of bank default, ς, raises, the expected return from deposits falls. Hence, households shift their savings to the platform until arbitrage between deposits and loans is satisfied. Similar substitution occurs if the haircut, $\varkappa_{t+1}$, or the return on deposits, $\bar{R}_{t+1}$, fall. Also, an increase in P2P loans' success probabilities raise expected returns from the platform, and moves savings towards that.

3.3 Borrowers

In every period there is a continuum of heterogenous borrowers indexed by $i \in [0,1]$. Each of them wishes to fund a project of scale I_t. Borrowers do not have internal funds. Some borrowers obtain funds from the platform, and some from banks. Some projects with very low success probability won't be funded at all. The fraction of unfunded projects is determined in section 3.4, while the equilibrium for P2P loans and bank credit is determined in section 3.8.

Borrowers are risk-neutral and finitely lived. Let ξ, denote borrowers' survival probability or their exit rate from business. By multiplying ξ by the discount factor, β, we obtain borrowers' gross discount factor which is higher than lenders' one. The last assumption prevents borrowers from saving enough so as to ease up the need of external funding.[16] The assumption of risk-neutrality captures borrowers' higher preference for risk relative to lenders. Furthermore, we assume limited liability so that all contractible payments are bounded below by zero. The assumption of limited liability ensures that borrowers have risk-shifting incentives.

As explained earlier, projects are risky and have different success probabilities such that they can succeed with probability p^i and deliver a return R_t^I, or fail and return zero. Individual probabilities are distributed according to a uniform density, ϕ_p, such that $p^i \in \mathbb{U}\left[\bar{p} - \frac{\varepsilon}{2}, \bar{p} + \frac{\varepsilon}{2}\right]$, where $\bar{p}$ is the unconditional mean. The corresponding cumulative density is denoted by Φ_p. This distribution is the same for all borrowers and it is publicly known, whereas individual success probabilities are known only to borrowers. As already mentioned, banks can screen projects and learn success probabilities at a cost. Instead, lenders in the P2P market observe projects' characteristics and receive public signals on projects' probability of success. With asymmetric information, if no signals were available, a pooling equilibrium would emerge with all projects yielding the same return. With signals, separating equilibria emerge. The sequence of separating equilibria, each identified by their

[16] Since borrowers are risk-neutral, their consumption schedule results in a corner solution. We can exploit the finite life structure a' la Yaari [36] and establish that consumption takes place when borrowers exit business. Hence $C_t^b = (1-\xi)W_t^b$, where W_t^b represents borrowers' net wealth at the time when they exit business.

signal precision, converges, in the limit, to the full information equilibrium if signals are perfectly informative. In equilibria with partial information, investors form Bayesian expectations about the distribution of success probabilities and the projects' distribution of returns reflects that.

Notice that we restrict attention to borrowers that fund themselves either through the bank or through the platform. Issuance of corporate bonds might represent an additional source of funding. We simplify and exclude this possibility, although the model could be extended to include this. The exclusion is nevertheless a realistic assumption in the context of our paper. Indeed, borrowers choosing P2P platforms are typically small and risky firms with a short history in business, hence with little reputation. This implies that they would have little chances of obtaining funds on traditional equity or bond markets. All in all, P2P markets are akin to private equities.

3.4 Pricing in P2P Markets

Pricing in the peer-to-peer market reflects the presence of asymmetric information. Although full information is never possible, digital markets offer the possibility of gathering costless signals. These signals convey both hard information (e.g. FICO scores, data on current delinquencies and the debt-to-income ratio) that exploit machine learning algorithms for processing and updating, and in some cases also soft information, such as recommendations and investments by peers. Both type of signals are public and visible to all lenders at no cost. Hence, for the purpose of our model, we treat them equally.[17] Notice that we focus on signals which are equally and publicly available to all, rather than on signals privately provided by borrowers. The reason is that, in the current institutional arrangement, the first count much more for pricing. Indeed, prices/returns are currently determined centrally and conditional upon all public information which is processed by a machine learning algorithm.[18]

We model signals as random variables, σ_i, whose realization is s_i. Signals' distribution can be

[17] In principle, soft signals might produce negative externalities to the extent that they convey biased information inducing herd behavior. However recommendations by peers are also coupled with actual peer decisions which convey useful information on the actual quality of a borrower project.

[18] Up to 2009 loan returns were determined through an auction-like mechanism and borrowers' credit history was limited. Auction-based pricing reflected to some extent borrowers' private signals. However, information consisting in costless web-reporting is generally less informative than public signals which are based on events that are publicly observable. In fact, bad borrowers can easily imitate good ones in their reporting. In our empirical analysis indeed we find that, after 2009, the centralized pricing mechanism, which more efficiently conditions on all available signals, coupled with the longer history of credit scores, generally increased information efficiency of P2P markets by making the impact of all signals more significant.

summarized as follows (we follow Ruckes [30] and Petriconi [27]):

$$\sigma_{i,\lambda} = \begin{cases} s_i = p^i & \text{with probability } \lambda, \\ s_i \sim \mathbb{U}\left[\bar{p} - \frac{\varepsilon}{2}, \bar{p} + \frac{\varepsilon}{2}\right], & \text{with probability } (1-\lambda) \end{cases} \tag{8}$$

In words, with probability λ, $s_i = p^i$, i.e. the signal conveys the project's true success probability, while with probability $(1-\lambda)$, it is a random draw from p^i prior distribution and it is totally uninformative. Probability λ captures signals' precision. Notice that our investors, contrary to banks, are unsophisticated. Thereby, they do not invest in screening or monitoring to increase signal precision. The latter is exogenously given and depends on the amount and quality of information as a whole on borrowers' credit history. One can therefore think of λ as a parameter (possibly, at any given point in time) or as a random process that evolves over time (this will be our modeling approach in the numerical simulation that we present in Appendix 6). Given the above, signals are distributed as a uniform, such that $\sigma_{i,\lambda} \sim \mathbb{U}\left[\bar{p} - \frac{\lambda\varepsilon}{2}, \bar{p} + \frac{\lambda\varepsilon}{2}\right]$.

Once they receive a signal s_i, lenders can update their estimate of projects' success probabilities by computing the conditional mean $\mathbb{E}_t\left[p^i \mid \sigma_i = s_i\right]$. The latter is based on the density of p^i conditional upon the signal. Using Bayes rule and for any realization of the signal, s, we can determine the density of a given project's quality, $p^i = p$:

$$\phi_{p|s} = Pr\left[p^i = p \mid \sigma_{i,\lambda} = s\right] = \frac{Pr\left[\sigma_i = s \mid p^i = p\right] Pr\left[p^i = p\right]}{Pr\left[\sigma_i = s\right]} \tag{9}$$

We assume that, in every period t, all borrowers issue a signal. Prospective lenders observe these signals and form their beliefs on the probability of success of borrowers' projects. Given the signals' structure in (8), we can then compute the conditional density function:

$Pr\left[\sigma_{i,\lambda} = s \mid p^i = p\right] = (1-\lambda)\frac{1}{\varepsilon} + \lambda\delta(s-p)$. Let's define $x = s-p$. The function $\delta(x)$ is the Dirac function which goes to infinite if $x = 0$ with $\int\limits_{-\infty}^{\infty} \delta(x) = 1$ and it is equal to zero when $x \neq 0$. The corresponding cumulative function is $\Phi_{s,t} = Pr\left[\sigma_{i,\lambda} \leq s \mid p^i = p\right] = (1-\lambda)(\frac{1}{2} + \frac{s-\bar{p}}{\varepsilon}) + \lambda\mathcal{H}(s-p) = \lambda\mathcal{H}(p-s) + (1-\lambda)\Phi(p)$, where $\mathcal{H}(s-p)$ is the Heavisde step function. The Heavside step function is equal to zero if $x < 0$ and it is equal to one if $x = 0$, where again $x = s-p$. Given (9) the conditional density is $Pr\left[p^i = p \mid \sigma_{i,\lambda} = s\right] = (1-\lambda)\frac{1}{\varepsilon} + \lambda\delta(x)$. Noticing that the both the $Pr\left[p^i = p\right]$ and $Pr\left[\sigma_{i,\lambda} = s\right]$ are uniform, we can compute the conditional expectation of projects' quality. Hence, given a signal, the Bayesian updating of beliefs results in the following posterior expectation:

$$\mathcal{E}_t \left\{ p^i \mid \sigma_{i,\lambda} \right\} = \pi^i = \lambda s_i + (1-\lambda)\overline{p} \tag{10}$$

Investors will price each project based on the conditional expectation in (10), so that the expected return from each project is given by $\mathcal{E}_t \left\{ p^i \mid \sigma_{i,\lambda} \right\} R_t^I = \pi^i R_t^I$.

3.5 Threshold for Projects Funded in the Platform

We can now determine the mass and type of projects that households will fund on the platform. This turns into determining the marginal loan that is funded in equilibrium or the threshold for the success probability that truncates the distribution of funded projects from the rejected ones. In a general equilibrium context households' arbitrage can be used to determine this threshold. Specifically, the threshold corresponds to the loan that on the margin and in expectation delivers the same return as one extra unit of deposits. Substituting the estimated success probability derived in section 3.4 into the arbitrage condition in (7), we obtain:

$$\begin{aligned} = \quad & \left[\mathcal{E}_t \left\{ p^i = \hat{p} \mid \sigma_{i,\lambda} = s_i \right\} \mathbb{E}_t(\Lambda_{t,t+1} \mid suc) \right] R_{t+1}^I = \\ & [(1-\varsigma_t)\mathbb{E}_t(\Lambda_{t,t+1} \mid no\ def) + \varsigma_t \mathbb{E}_t(\Lambda_{t,t+1} \varkappa_{t+1} \mid def)] \overline{R}_{t+1} \end{aligned} \tag{11}$$

where $\hat{p}$ denotes the marginal borrower. Upon defining $\varpi_t = \mathcal{E}_t \left[p^i = \hat{p} \mid \sigma_{i,\lambda} = s_i \right]$, we can write equation (11) in a more compact form:

$$\varpi_t = \frac{[(1-\varsigma_t)\mathbb{E}_t(\Lambda_{t,t+1} \mid no\ def) + \varsigma_t \mathbb{E}_t(\Lambda_{t,t+1} \varkappa_{t+1} \mid def)] \overline{R}_{t+1}}{\mathbb{E}_t(\Lambda_{t,t+1} \mid suc) R_{t+1}^I} \tag{12}$$

Expression (12) is the main equilibrium condition of the model and summarizes how signals' precision, the moments of projects' distribution and the risk of bank fragility determine the marginal borrower on the platform. For this reason, equation (12) will be used later on to discuss the selection, the information and the substitution channels driving lenders' and borrowers' decisions.

A few remarks are in order. First, the projects funded on the platform have expected success probabilities above the cut-off, ϖ_t. All projects with expected success probability below that threshold will not be funded. If the equilibrium ϖ_t increases, less projects are funded. Second, condition (12) makes explicit that participation in the platform, relatively to the banking sector, depends upon the balance of returns-risks between the two sectors. To fix ideas, if for given information efficiency, λ, banking fragility ς raises the numerator in (12) falls and so does ϖ_t, hence lenders fund more projects in the platform. On reverse, for given risk of bank failure, technological improvements that increase platforms' information precision, as captured by λ, raise expected rates (denominator of ϖ_t) and, hence, participation in the P2P business.

3.6 Loan Premia and The Value of Information

In the presence of positive signal precision, the separating equilibrium features a distribution of loan returns/premia which depend on projects' default probabilities (adverse selection premium) and on the precision of signals (information premium). It is useful to disentangle the two.

If the investor could perfectly discern projects' quality, she would be require for each project a loan premium just equal to its default probability, namely $1-p^i$. Under partial information, projects are indistinguishable and adverse selection emerges, hence lemons are part of the pool. Under such circumstance, the investor requires an additional premium which is proportional to the value of information. We can construct such an information premium using the Theil [33] index, which is given by the distance between the probability that a project will not be funded under partial information (with given signal λ) and the equivalent probability under full information (namely when $\lambda \to 1$).

The probability that a project will not be funded under partial information corresponds to the probability that its expected probability of success, conditional on the signal, is below the threshold for funding. The latter reads as follows:

$$\begin{aligned} \chi_\lambda(\varpi_t) &= \Pr(\mathcal{E}_t \left[p^i \mid \sigma_{i,\lambda}\right] \leq \varpi_t) = \\ &= \Pr(\lambda\sigma_{i,\lambda} + (1-\lambda)\bar{p} \ \leq \varpi_t) = \\ &= \Pr(\sigma_{i,\lambda} \leq \frac{\varpi_t - (1-\lambda)\bar{p}}{\lambda}) \end{aligned} \tag{13}$$

The above also identifies the mass of projects that won't be funded at all, which is $\psi_\lambda(\varpi_t) = 1 - \chi_\lambda(\varpi_t)$.

Given the distribution function for $\sigma_{i,\lambda} \sim \mathbb{U}\left[\bar{p} - \frac{\lambda\varepsilon}{2}, \bar{p} + \frac{\lambda\varepsilon}{2}\right]$, we can re-write $\chi_\lambda(\varpi_t)$ as follows:

$$\chi_\lambda(\varpi_t) = \begin{cases} 0 & \text{if } , \varpi_t \leq \bar{p} - \frac{\lambda\varepsilon}{2} \\ \frac{\varpi_t - \bar{p}}{\lambda\varepsilon} + \frac{1}{2} & \bar{p} - \frac{\lambda\varepsilon}{2} \leq \varpi_t \leq \bar{p} + \frac{\lambda\varepsilon}{2} \\ 1, & \text{if } \varpi_t \geq \bar{p} + \frac{\lambda\varepsilon}{2} \end{cases} \tag{14}$$

From condition (14), one can see that an increase in the average quality of projects, $\bar{p}$, or in the precision of the signal, λ, reduces the probability that a project will not be funded $(1 - \chi_\lambda(\varpi_t))$ if and only if $\varpi_t \geq \bar{p}$. In other words, if the average quality under imperfect information, ϖ_t, is above the average quality under no information, $\bar{p}$, there is a positive selection of projects. In this case, more precise signals are valuable because they help to reduce averse selection and increase the equilibrium mass of funded projects.

Given $\chi_\lambda(\varpi_t)$, the Theil index for the value of information then reads as follows:

$$\Theta = \chi_\lambda(\varpi_t) - \chi_{\lambda=1}(\varpi_t) \tag{15}$$

Notice that $\chi_\lambda(\varpi_t)$ captures the amount of entropy among the funded projects under partial information and for given signals. As signal precision increases, the dispersion or entropy widens, getting closer to the entropy under full information.

Lemma 1. *The information premium, $\Theta = \chi_\lambda(\varpi_t) - \chi_{\lambda=1}(\varpi_t)$, decreases in p* and λ *to the extent that $\varpi_t \geq \bar{p}$.*

Proof. Let's focus on the interval $\bar{p} - \frac{\lambda\varepsilon}{2} \leq \varpi_t \leq \bar{p} + \frac{\lambda\varepsilon}{2}$. In this interval $\Theta = \chi_\lambda(\varpi_t) - \chi_{\lambda=1}(\varpi_t) = \left[\frac{\varpi_t - \bar{p}}{\lambda\varepsilon} + \frac{1}{2}\right] - \left[\frac{\varpi_t - \bar{p}}{\varepsilon} + \frac{1}{2}\right] = \left[\frac{1}{\lambda} - 1\right]\left[\frac{\varpi_t - \bar{p}}{\varepsilon}\right]$. As $0 \leq \lambda \leq 1$ an increase in the average quality $\bar{p}$ decreases the information premium to the extent that $\varpi_t \geq \bar{p}$. Similarly, when $\varpi_t \geq \bar{p}$, an increase in signal precision, λ, reduces the absolute value of the distance Θ, and, hence, the information premium.

If the threshold ϖ_t is to the right of the unconditional average success probability in the population, it means that investors expect a positive selection of projects. Under those circumstances an increase in the average quality of borrowers, as captured by $\bar{p}$, reduces the mass of unfunded projects both under under full information, $\chi_{\lambda=1}(\varpi_t)$, and under partial information, $\chi_\lambda(\varpi_t)$. Hence, when

$\varpi_t \geq \bar{p}$ there is generally a positive selection effect. However, for given increase in $\bar{p}$ the reduction in the dispersion under partial information $\chi_\lambda(\varpi_t)$ is higher. Hence, an increase in $\bar{p}$ also reduces the information premium Θ and consequently the lending spread. Intuitively, if the average quality of projects improves, it is worth less to gather information. Besides this, an increase in precision, λ, also reduces the distance Θ. This is intuitive. As more precise information is available, projects' dispersion under partial information approaches the dispersion under full information. As a result the distance, Θ, declines.

The effect of $\bar{p}$ on Θ captures a *selection* channel, while the effect of λ on Θ captures the *information* channel. In our empirical analysis below, we will show that signals reduce lending rates, hence loan spreads, and, by exploiting the variability in information reporting across borrowers, we are able to identify the relative relevance of the selection and information channels.

3.7 Banks

In this section we describe the pricing of bank loans, whereas the fraction of borrowers that apply for bank loans and are eventually funded is determined in section 3.8.

To simplify the model environment, we devise a parsimonious set-up for the banking sector. There is one (or one type of) bank that has access to a costly screening technology. The bank shall pay a cost, μ, per project to assess their quality. Once this cost is paid, the bank learns quality perfectly.[19] The bank acts in a competitive environment, hence it shall just break even. It collects only demand deposits to fund loans, $D_t = L_t^b$ and all projects returns, net of screening costs, are then rebated to investors in demand deposits[20].

Given the above, in every period t, borrower i shall pay the bank loan services, $R_t^{i,L}$, that allow the bank to break even with the costs of deposits and the cost of monitoring:

$$R_t^{i,L} = p^i R_t^I \geq \bar{R}_t + \mu \tag{16}$$

Condition (16) implies the assumptions that borrowers make zero profit, as all expected returns are rebated to the bank.[21]

[19] The model can be easily adapted to the case of partial learning by banks.

[20] Banks are akin to mere monitoring technology, hence there is no conflict of interest between bank managers and outside financiers.

[21] This is true also of the projects' returns funded through the platform. Even here the model can be easily extended to assume that entrepreneurs can extract some rents.

From equation (16), we obtain that the bank will funds all projects whose success probability satisfies the following:

$$\tilde{p}^i \geq \frac{\bar{R}_t + \mu}{R_t^I} \tag{17}$$

The above margin condition is used in the next section to assess the extent of banks' supply of funds to entrepreneurs.

3.8 A Note on the Equilibrium of Flows

At this stage we are able to determine the fraction of projects that is funded in equilibrium in the two sectors, which results from the demand and supply of funds.

Let us start to examine the equilibrium quantities traded and cleared in P2P markets. From section 3.5, households fund all projects whose success probability is above:

$$\varpi_t = \frac{\Sigma_t \bar{R}_{t+1}}{\mathbb{E}_t(\Lambda_{t,t+1} \mid suc) R_{t+1}^I} \tag{18}$$

where for notational convenience we have set $\Sigma_t = [(1 - \varsigma_t)\mathbb{E}_t(\Lambda_{t,t+1} \mid no\ def) + \varsigma_t \mathbb{E}_t(\Lambda_{t,t+1} \varkappa_{t+1} \mid def)]$, which is the expected loss given default.

Borrowers will demand funds until P2P rates equalize bank loan services:

$$\varpi_t \mathbb{E}_t(\Lambda_{t,t+1} \mid suc) R_{t+1}^I \leq R_t^{i,L} = R_t^d + \mu \tag{19}$$

The intersection of demand and supply implies that the distribution of projects funded on the platform is truncated from above and from below as follows:

$$\Sigma_t \bar{R}_{t+1} \leq \varpi_t \mathbb{E}_t(\Lambda_{t,t+1} \mid suc) R_{t+1}^I \leq \bar{R}_{t+1} + \mu \tag{20}$$

Projects whose success probability is such that $\varpi_t \mathbb{E}_t(\Lambda_{t,t+1} \mid suc) R_{t+1}^I \leq \Sigma_t \bar{R}_{t+1}$ will not be funded at all. For projects whose success probability is such that $\varpi_t \mathbb{E}_t(\Lambda_{t,t+1} \mid suc) R_{t+1}^I \geq \bar{R}_{t+1} + \mu$ borrowers will file an application to the bank, as here they pay lower loan service.

Notice that the presence of a P2P market with signals allows funding of an additional fraction of borrowers with respect to an economy where only banks operate. The need to pay a screening cost reduces the quantity of loans that banks can serve in equilibrium. To see this note that an increase in banks' costs, μ, raises the upper limit of the interval (20) and implies that more projects are funded on the platform.

Notice also that the substitution and information channels discussed above affect also the bounds of the set of projects that are funded on the platform. To see this it is useful to substitute equation (10) into condition (20) and focus on the left hand side of the interval:

$$\frac{\Sigma_t \bar{R}_{t+1}}{(\lambda s_i + (1-\lambda)\overline{p})} \leq \mathbb{E}_t(\Lambda_{t,t+1} \mid suc) R^I_{t+1} \tag{21}$$

From condition (21), and when $s_i > \overline{p}$, i.e. when the signal points towards a success probability greater than average, an increase in signal precision, λ, lowers the threshold, ϖ_t, hence, more projects are funded on the platform. Again, an increase in market transparency makes it easier for lenders to discern projects' quality and reduces the extent of credit rationing channelled through the platform. Besides this, an increase in bank fragility, ς, by reducing Σ_t, also reduces the lower bound of the interval and raises the volumes traded in the platform. The reason is that an increase in the risk of default within the banking sector shifts households towards platform investment.

We shall now derive the equilibrium volumes serviced by the traditional banking sector. First, we have seen that the bank funds only those projects whose success probability satisfies $p^i \geq \frac{\bar{R}_t + \mu}{R^I_t}$. Second, borrowers apply for bank loans only if the loan services therein are lower than the rates to be paid on the platform, i.e. $R^{i,L}_{t+1} \leq \pi^i R^I_{t+1}$. The intersection of demand and supply implies that also the distribution of projects funded by the bank is truncated above and below as follows:

$$\bar{R}_{t+1} + \mu \leq R^{i,L}_{t+1} \leq \pi^i R^I_{t+1} \tag{22}$$

The above implies that borrowers who pay lower rates at the bank will turn to it for funding. However, banks will only fund projects' application that allows them to break even.

3.9 Substitution, Selection and Information Channel

Given all the previous discussion we now derive three testable implications of the model.

Remark 1 - Selection channel. *An increase in the average quality of projects, $\bar{p}$, increases platform liquidity, reduces the expected information premium, Θ, and, hence loans' spreads to the extent that $\varpi_t \geq \bar{p}$.*

Remark 2 follows from Lemma 1 and the considerations related to it.

Remark 2 - Information channel. *An increase in signals' precision, λ, increases platform liquidity, reduces the information premium, and, hence, loans' spreads.*

Remark 3 follows from Lemma 1 and the considerations related to it.

Remark 3 - Substitution channel. *An increase in the risk of bank fragility shifts investors' participation toward the platform. This, in turn, lowers the premia that lenders require to hold P2P loans.*

From equation (12) we can see that an increase in banks' liquidity risk, ς_t, decreases the tolerance cut-off, ϖ_t, at which investors provide funds in the platform. Investors fund all projects that have an expected success probability above the cut-off. Since the cut-off now declines, the mass of P2P loans expands. In other words, as the risk of an haircut on deposits increases, investors shift to the P2P market. This implies, as per equation (14) and as long as as long as $\varpi_t \geq \bar{p}$, a reduction in the probability that a project will not be funded at all, $\psi_\lambda(\varpi_t) = 1 - \chi_\lambda(\varpi_t)$. Interestingly, notice that, when $\varpi_t \geq \bar{p}$, the decline in ϖ_t also implies a decline in the information premium Θ, for any value of λ. The reason for this is as follows. The increase in liquidity or in the total supply of funds on the platform induce an increase in the number of funded projects, all of which have a higher probability of success than the population average. This implies a fall in the conditional default probability and, hence, in the spread, Θ.

Despite the general equilibrium dimension, the model is actually tractable enough that we could obtain these three testable implications using analytics. For completeness, in Appendix 6, we quantify these effects through simulations that solve the entire set of model equations simultaneously. We calibrate the model and simulate it subject to a set of shocks estimated on data from US P2P platforms Prosper and Lending Club, and from the banking system. Using impulse response functions we then discuss the three channels highlighted above, namely the selection, the substitution and the information channel. Besides this, we verify whether the simulated model matches second moments and autocorrelation of P2P lending volumes, using again data from Prosper and

Lending Club. While our main results are valid independently of a calibrated solution, it is of interest to confirm our results quantitatively.

4 Empirical Analysis

We now turn to the data to analyze the determinants of equilibrium P2P loan rates and appraise empirically the predictions of our model regarding the effects of signals and liquidity risk. We start using data from Prosper (http://prosper.com) which initiated P2P lending in the US, in February 2006. Our data include all loans that have been funded over the period February 2006 to March 11th, 2014. Currently, Prosper is the second largest platform in the world,[22] after Lending Club. In contrast to other platforms, including Lending Club, in addition to *hard* credit information, Prosper provides prospective lenders also with *soft* information on its listings, which makes this data set particularly suitable for our purposes. For this reason we run most of our analysis using Prosper data. However, in section 4.5 below, we perform robustness checks using data from Lending Club as well.

4.1 Prosper

A brief description of the institutional functioning of Prosper is in order. When signing up for a loan on Prosper, borrowers create personal profiles and solicit funding detailing the interest rate, the amount requested and the term of the loan. Besides this, borrowers' profiles include three types of additional information. First, there are signals provided directly from the borrower, including the purpose of the loan, home ownership, employment status and occupation. Second, there are hard information signals resulting from processing by the machine learning algorithm, including independently verified information on the borrower credit history, income and current debts, and a credit grade determined by Prosper and based on a proprietary algorithm (from 2009 onwards). Third, there are soft information signals coming from the social networks that Prosper creates by linking borrowers in groups (tied by geography, common interest, common loan purpose or some specific characteristic) and by collecting and making public the endorsements of other Prosper borrowers ('friends'). As we shall see below, the second and the third type of information appear to

[22] With issues of almost $4bn new loans in 2015.

be generally more informative than the first. As explained earlier, since private signals are costless in this market, they are also not informative because any bad borrower can imitate a good one.

Given the signals, lenders assess and can bid on the listings. Loans are funded only if the bids reach the amount sought by borrowers. The maximum length of the bidding period is 2 weeks. Until 2010, Prosper loan rates were determined through an eBay-like auction. In 2009, Prosper registered with the Security Exchange Commission (SEC) and afterwards changed its business model to pre-set rates determined solely by Prosper itself (centralized system). More details on the platform are in Appendix 7.

4.2 Rates, Volumes and Risk of Prosper Loans

We start by examining simple statistics and data trends pointing toward facts that are relevant for our analysis.

Tables 1 and 2 report some summary statistics for our data. Table 1 focuses in particular on the trends in loan volumes and rates, which give first indications of the channels operating in those markets. Notice that Prosper has grown very quickly over time and tripled its size between its incept and 2013 when a whole loan program was lunched. The jump in volumes occurred at a time when US banks experienced great fragility as the large number of bank failures reported by the Federal Deposit Insurance Corporation shows.[23] This suggests a possible shift of flows which is consistent with the substitution channel of our model. For what concerns loan rates, they exhibit a tendency to increase in the first half of the sample period possibly due to an excess demand by borrowers or shortage of funds Since 2011 rates have been falling steadily which may reflect, among other things, a fall in default and information premia.[24] Over time, the maximum loan size and duration have increased and this has resulted in larger average loans and longer duration.[25] Also, the time from posting to funding has decreased and the average investment by lenders has gone up. All this suggests increased market liquidity. As to the reason for borrowing, currently most loans are for consolidating debt from other sources. At the onset, a non negligible share was

[23] We use data on bank failures in section 4.6 when we consider the substitutability between P2P platforms and traditional financial institutions.

[24] It is worth mentioning that Prosper lending rates tend to be higher than the rates of similar platforms, such as Lending Club. However, it charges borrowers with lower fees.

[25] Indeed, the maximum loan amount and length have been raised from $25,000 to $35,000, and from 36 to 60 months, respectively.

to fund business activities. The recent decline is most likely due to the fact that over the recent past, a large number of 'special purpose' platforms have been set up, including ones specialized in business lending, and borrowers have increasingly turned just to those platforms specialized in catering to their specific needs.

Table 2 summarizes the signals that borrowers issue and help to mitigate asymmetric information. The information posted on the platform and visible to all prospective investors includes borrowers' credit scores as provided by official credit rating agencies (FICO scores), the number of open credit lines, the number of credit enquiries[26] and the number of current delinquencies. Over the sample period considered, the trend of all those measures indicates an improvement in the average "quality" of borrowers, beyond the tighter eligibility criteria imposed on perspective borrowers by the SEC. Consistently, Prosper in-house credit rating[27] and estimated losses and returns at issuance suggest an improvement in borrowers' reliability[28]. The decline over time in borrowers' riskiness (as proxied by credit rating measures) and lending rates is indicative of the disciplining role of making the credit records and history public.

Last, we observe a general decline also in ex post loan riskiness. Panel (b) of Table 1 reports the status as of March 11th, 2014 of all loans that have been funded since Prosper onset, and it gives the frequency of defaults. The share of loans classified as 'Charged off', i.e. 120 days or more past due, or in 'Default' was relatively high at the onset of the platform, but has fallen significantly, even allowing for the fact that many loans had not reached maturity. The decline has been sharper after the registration with the SEC. Indeed, of all loans issued in 2010, which had all reached maturity by 2014, the share that was not fully paid back amounts to 17 percent. If we take as reference consumer loans in 2014, the share that US banks charged off or reported as delinquent

[26] Credit inquiries are requests to check one's credit score. Specifically, when one applies for credit, including auto loans, mortgages or credit cards, potential lenders check their credit history. A large number of inquiries typically means greater risk. Indeed, people with six inquiries or more on their credit reports are eight times more likely to declare bankruptcy than people with no inquiries on their reports. (http://www.myfico.com/crediteducation/questions/inquiry-credit-score.aspx)

[27] Using prospective borrowers' credit history and information stored in its system, Prosper provides proprietary credit ratings for each listing, ranging from 1 (high risk) to 7 (low risk). Such ratings are based on the estimation of borrowers' loss rate (see description in Appendix 7). Prosper then uses these ratings to assign each listing also an estimated effective yield corresponding to the difference between the loan yield and the estimated loss rate (reported in Table 2).

[28] Table 2 reports also borrowers' average monthly income, which has been increasing over time, and their debt-to-income ratio at the time the listing, which instead has been stable at around 25 percent.

amounted to 16.6 percent (18.5 percent of all consumer loans in 2013) (Board of Governors of the Federal Reserve [10]).

To sum up, Tables 1 and 2 indicate a contemporaneous increase in volumes, liquidity and borrowers' quality as well as a decline in rates and risk. All together those trends are suggestive of the importance of information for those markets and of how the screening mechanism of these platforms, which consists in the storage and transparent publication of large amounts of information on borrowers, contributes to market efficiency and liquidity.

4.3 The Role of Soft Information Signals

Table 2 summarizes the soft information-type of signals available to potential lenders on Prosper. They include borrowers' participation in groups, recommendations and investment from "Prosper friends" and previous Prosper loans. The share of borrowers who are part of a group was very high, almost 70 percent, at the onset of the platform, but drops to less than 1 percent at the end of our sample period. As to friendships, to create one, borrowers need to send an email to their friends, who must also be active on Prosper. Hence, individuals who are friends on Prosper must have at least some offline, non-public information about each other, such as an email address. Like for group participation, also the share of borrowers with recommendations or investment from "Prosper friends" falls steadily over the period. The significant drop in network relevance is most likely related to the tremendous growth that Prosper registered over time. Besides this, after SEC registration and with the introduction of the whole loan program, investment from institutional investors has grown quickly and this has resulted in a drop in the number of investors funding each loan (and an increase in the size of individual investments). In the last two years covered by our sample more than half of the loans were funded by a single lender, most likely an institution (Table 1). For what concerns previous experience on the platform, a non-negligible share of listings is from returning borrowers (almost 20 percent in 2013). In our regressions below we will include all metrics related to soft information with the goal of testing the importance of social networks in the decision to invest in the platform and in this price formation process.

In Tables 3, 4 and 5 we split borrowers based on soft-type of information, such as being part of a group or not (Table 3), having recommendations from other Prosper borrowers or not (Table 4), and having borrowed previously on Prosper or not (Table 5). Few interesting observations emerge. First, before SEC registration, borrowers belonging to a group appear to be riskier (higher FICO). Second, borrowers with friends pay lower rates.[29] Last, borrowers with previous Prosper loans (Table 5) appear to pay relatively lower rates in the most recent years covered by the sample despite their FICO and Prosper ratings being not significantly higher than those of first-time borrowers. Also, they exhibit fewer charged offs and defaults. This suggests that the availability of a public credit history is important, even beyond the availability of soft signals, in reducing information premia and lending rates. The longer the credit history is, the stronger the decline is.

4.4 Lending Rates, *Hard* and *Soft* Information Signals

In this section we start examining the selection and information channels previously rationalized through our model. To this purpose we report regressions of lending rates on all types of available signals (private signals, hard information and soft information) with the goal of testing the direction and the significance of their impact. As explained earlier, signals indicating a reduction in borrowers' risk might induce a decline of lending rates due to a selection effect (better borrowers' quality and lower default rates) or due to an increase in information precision (investors can better discern borrowers and require lower premia). In this section we will not identify separately those two effects, but assess them jointly. We will distinguish the channels in the next section.

Before going forward it is important to mention that, given the change in the rate setting procedure from an auction-type to a centralized system, in our analysis we always repeat estimation splitting the sample based on the type of rate setting.

Tables 6 reports the results of OLS regressions of lending rates on loan characteristics (size, term, motive), as reported by borrowers (private signals), and dummies for year-quarter of listing and for borrower's state of address. In the first column, where we pool all years, these variables explain around 23 percent of lending rates' variability.[30] Noteworthy, when we split the sample, based on regressions' adjusted R^2s, the explanatory power of the regressors is higher when Prosper

[29] This evidence is consistent with that of Lin, Prabhala, and Viswanathan [22] who focus on the likelihood of being funded and find that friendships have a much larger impact than group membership.

[30] The relationship between lending rates and size of the loan is U-shaped with a minimum at $4,700. Increasing the term by 1 month raises the rate by 1 percentage point. Loans for debt consolidation and for business funding are charged rates that are higher by half and one percentage point respectively. Similar results obtain when replacing the left-hand-side variable with effective loan yields, or with Prosper ratings, estimated returns or estimated losses. Regressions available upon request.

starts setting rates centrally.[31] However, most heterogeneity remains unexplained. This confirms our previous argument that a centralized pricing system, which more efficiently conditions on all available signals, and a longer credit history increase information efficiency.

In the regressions in Table 7, we add variables capturing the *public* signals that lenders can use to infer borrowers' quality. In columns (1) to (3) we include only variables conveying *hard* information, which consist of quantitative and verifiable signals, such as borrowers' FICO score, number of open credit lines, number of credit inquires, a dummy for any current delinquencies, monthly income, and debt-to-income ratio. These variables improve substantially the ability of the regression model to explain the variability of lending rates. The analysis suggests that loan rates are decreasing in the FICO score: a one standard deviation (70 point) increase in the FICO score lowers the lending rate by approximately 4 percentage points (20 percent of the mean rate).[32] Being delinquent, having a large number of open credit lines and of credit inquiries, having a low income, or a high debt to income ratio also increase rates.[33]

In the last three columns of table 7, we add in variables conveying additional information of a *softer* type, i.e. whose value and bearing have a major subjective component. These consist of dummies for participation in a group, for endorsement from Prosper friends with or without Prosper friends' investment, and for previous borrowing on the platform. In line with the implications of our model regarding the impact of information and with the studies on the role of networks mentioned earlier, once credit risk is controlled for, being part of a group lowers loan rates, by between half and two percentage points. When it comes to friendships, we find that rates are lower for borrowers with funding from friends (with or without recommendation) by up to 4.5 percentage points before 2009, and by up to 1.5 percentage points after 2009.[34] The decline in the significance of those soft signals

[31] Results are robust to pooling the data and interacting the regressors with a dummy that takes on value 1 if the loan is funded when rates are set centrally, instead of splitting the sample. Also, the evidence is similar if we take as left-hand-side variable loans' APR, Proper rating, Prosper estimated return or Prosper estimated loss. Regressions available upon request.

[32] In our sample the FICO score ranges between 500 and 900. A FICO below 600 signals poor credit history. A FICO above 750 signals an excellent history.

[33] Being delinquent on some other account raises the rate by 1 to 3 percentage points. A monthly income higher by one standard deviation ($8000) lowers the rate by 1 percentage point. Increasing the debt-to-income ratio by one standard deviation raises the rate by up to 1.5 points.

[34] It is interesting to notice that if we omit credit risk measures from the regression, the coefficient of the group membership dummy becomes significant and positive. This suggests an omitted variable bias and is consistent with the hypothesis that group participation may facilitate funding of risky borrowers, which generates a positive correlation between individual riskiness and group participation as Table 3 suggests.

might also be due to the decline in the incidence of group participation and of recommendations as evident from table 2.[35]

Last, it is interesting to notice that social network variables are significant also in the regressions on the sample of loans whose rates are set centrally, despite the fact that these variables become visible only after the posting of the loan, hence after Prosper algorithm sets the interest. This suggests that Prosper pricing algorithm is somehow capable of anticipating the information content of these variables. This seemingly puzzling feature can be explained by examining the regression in the last column of the table which includes a dummy for previous Prosper loans. In this regression the dummy for group membership becomes negligible and statistically insignificant and the coefficients of the dummies for recommendations and/or on investment from friends become very small (some become outright insignificant). Instead, the dummy for previous borrowing on the platform has a large, significant and negative coefficient, suggesting that returning borrowers pay over four percentage points less than "new" borrowers, *ceteris paribus.* It is very plausible that a positive correlation exists between joining a group and the number of Prosper friends, on the one hand, and previous funding on the platform, on the other, and this is part of the information stored in the centralized system.

4.5 Lending Rates and Signals: Identifying Selection versus Information Channel

In the previous section we have established that signals of various types have a significant impact on lending rates. Generally speaking the availability of information suggesting relatively low borrowers' risk reduces lending rates. As explained earlier, however, this decline might be due to either a selection effect (higher borrower quality reduces default rates) or to an information channel (for given borrowers' quality, the more the information is, the better the investors are at discerning borrowers). In this section we wish to quantify how much of the decline is due to the information channel relatively to the selection channel.

Table 8 reports the results of OLS regressions where we quantify the role of information and signal precision in reducing premia and lending rates. To construct a measure of signal precision

[35] All results are robust to running regressions on all listings and interacting the right-hand-side variables with dummies for loans priced by Prosper and to introducing Prosper in-house credit rating measures in the regressions.

we exploit the variability in the type of *information reporting* among borrowers.

We consider first the case of income. From the regression in column (1) of table 7, borrowers whose income is at the 75^{th} percentile of the distribution pay 50 basis point less than borrowers at the 25^{th} percentile, *ceteris paribus*. Hence, borrowers with higher income levels pay lower rates (selection effect). However beyond variability in income per se, there is also variability in the type of reporting. In fact, most borrowers provide official documentation to support their statement, but over 8 percent of the sample does not. This difference can proxy for signals' precision. In column 1 of table 8, we augment the benchmark regression of table 7 with a dummy that takes on value 1 if the borrower income can be verified and with interactions of this dummy with income and with the debt-to-income ratio. The dummy has a large, negative and statistically significant coefficient. To get a sense of the magnitude of this effect, we can compare two borrowers with identical median income and debt-to-income ratio but different reporting. The borrower that provides documentation about his income pays over 1 percentage point less than the other, consistent with the hypothesis that higher signal precision contributes in a statistically significant way to reduce lending rates.[36]

Next, we consider credit lines. In the regression in column (1) of table 7, this variable has a positive and statistically significant coefficient. More credit lines are associated with more risk, hence with higher lending rates (selection effect). However, about 1 percent of borrowers has no credit lines open. Having no credit lines open cannot be associated with being more or less risky because one could have no credit lines because she has never applied or because her applications have been turned down. Instead, having some credit line open can help investors to discern borrowers' quality (information channel). To verify this, in the regression reported in column 2, we replace the number of credit lines with a dummy that takes on value 1 if the borrower has no credit lines open (value 0 if she has any). The dummy has a positive, extremely large and significant coefficient and implies that borrowers without any open credit line pay on average 2.4 percentage points more than those who do have them. This result is important since it shows again that the information channel is quantitatively very important. Interestingly, our results give a clear quantification of the value of information in terms of how much returns investors are willing to give up in exchange of an increase in its precision.

[36]Notice that strategic non-reporting of this information is not an issue in that what matters is the fact that for some borrowers the available information is perceived as less reliable (whatever the reason).

Next, in the regression in column 3, we consider the state of residency which over 30 percent of borrowers do not report before Prosper registered with the SEC.[37] In this case we add a dummy that takes on value 1 if the state of residency is missing. The absence of this piece of information, by lowering the precision of the signals available, raises the average lending rate by almost 2 percentage points. Last, in the fourth column of table 8, we consider the reasons for borrowing, which is missing for about 10 percent of the loans posted after 2009. Similarly to the case of residency reporting, we add a dummy to single them out. Lending rates appear to be higher (by over half percentage points) for borrowers who omit this information.[38]

Overall, this evidence confirms that the information channel per se is quantitatively very important in driving the decline in lending rates, hence in increasing market efficiency.

4.6 Lending Rates and Banking Fragility

Another important prediction of our model concerns the substitutability between different forms of investment. In principle, markets and banks could be either substitutes or complements. If banks eventually adopt a digital screening technology, competition between the two sectors might induce complementarity. However, our empirical analysis, discussed below, suggests that substitution is prevailing. This is also the reason why in our model we favoured assumptions inducing substitution. Indeed, complementarity is more likely to emerge over time when the banking sector engages more heavily into digital investment. Instead, so far, the narrative has been of a migration of investors to the digital platform due to fears of fragility in the banking sector.

To investigate this issue we add to our regressions a proxy for bank fragility, based on information on bank failures as reported by the Federal Deposit Insurance Corporation. Over the period considered, 497 bank failures were recorded in the US, with most occurring between 2009 and 2011. 42 states experienced at least one bank failure. We use these data to augment our regressions with three dummies that take on value 1 if, in the borrower's state, more than one financial institution failed the month before, or two months before, or three months before the loan

[37] After SEC registration, information on the state is available for all borrowers.

[38] The regressions testing the significance of non reporting the state of residency and the reason for borrowing use observations that are dropped from the sample used for all other regressions. Including these observations in the previous regressions does not affect the size and significance of the coefficients of the other variables of interest.

[38] The Federal Deposit Insurance Corporation is the agency that is typically appointed as receiver for failed banks. The original list includes all banks which have failed since October 1, 2000.

was funded. Notice that we have included in the regressions both year-quarter and state dummies to capture other aggregate shocks. Results are in Table 9. The dummies' coefficients are negative and generally significant which suggests that, where significant rates of bank failures were recorded, agents increasingly turned to the platform, increasing liquidity and hence inducing a decline in P2P loan interests. The reduction in rates is between 20 and 30 basis points which is an admittedly small effect, most likely due to the fact that we are restricting the availability of instruments that are alternative to the banking sector. In practice, agents have more choices and asset substitution might take place among different asset classes.

Nevertheless, overall, this evidence is consistent with our model prediction that when there are signs of fragility in the banking system more borrowers and lenders turn to P2P platforms. This leads to lower equilibrium rates which reflect not only higher demand and higher supply, but also the presence of a larger share of good projects on the online market, with a reduction of information premia.

4.7 Robustness - Lending Club

To fully assess the robustness of our results we have repeated the analysis using data from US Lending Club, which is the world's largest P2P lending platform. Loan application processes and funding are very similar to Prosper. Specific institutional details and data description for this platform are in Appendix 8.

Table 10 below reports the results of OLS regressions whose specifications parallel the ones used with Prosper data. For ease of reading we report also the corresponding Prosper regressions from tables 6, 7 and 9. Compared to Prosper, Lending Club data yield more precise estimates due to higher sample size, but the evidence is very similar.

In the first column of table 10, we regress rates only on loan characteristics (size, term, motive) and on dummies for quarter-year of listing and state of address of the borrower, which explain

[38] We have measured bank fragility also using the (state-invariant) average ratio of currency in the hands of public to demand deposits in the year before the loan was posted. This variable reflects the idea that before the occurrence of banking panics depositors demand a large scale transformation of deposits into currency. Gorton [16] shows that historically the ratio of currency to demand deposits has indeed increased at panic dates. Consistent with the predictions of our model, we find that the more currency the public held relative to deposits in the year before listing, the lower the equilibrium rates. In other words, as the fragility of the banking sector increases, investors migrate to the platform, which increases liquidity and lowers the rates. Regressions are available upon request.

around 29 percent of the variability of lending rates (23 percent in Prosper data regression, second column). In the third column of the table, we add in variables capturing the signals that lenders can use to infer borrowers' quality. As explained earlier, Lending Club offers fewer indicators compared to Prosper, because soft information is missing. As hard signals, we include dummies for Lending Club credit grade (in the absence of borrowers' FICO score), the number of open credit lines and the number of credit inquires, a dummy for any delinquencies in the previous two years, annual income, and the debt-to-income ratio. Like with Prosper data, these variables improve enormously the ability of the regression model to explain the variability of lending rates. Also, the analysis suggests that rates are lower the better the credit grade is, which is a clear indication of the signalling role of those metrics. We also find that lending rates are higher the larger the number of credit enquiries and of open credit lines. They are higher for borrowers who report delinquencies on other loans, whose income is relatively low and their debt-to-income ratio is relatively high. This is again all consistent with the fact that public signals facilitate convergence to a separating equilibrium where lenders require higher loan spreads to fund riskier borrowers.

In columns 5 and 7 of the same table, we add to the regression the dummies for financial institution failures in the state of residency of the borrower, one, two and three months before loan funding which proxy for liquidity shortage and fragility risk in the banking sector. Consistent with the predictions of our model regarding bank-platform substitutability, and in line with the evidence from Prosper data, we find that borrowers living in states that experience financial institutions' failures pay lower rates. Again, like in Prosper data regressions, the estimated effects are relatively small possibly because we consider only bank deposits and P2P loans and do not allow for other instruments which could be involved in the substitution.

5 Conclusions

P2P lending has experienced an impressive growth over the recent years and has penetrated most markets including high growth economies like China. Despite the lack of delegated monitors, and the potential costs of asymmetric information, data suggest that this new intermediation service is performing well relatively to traditional banking[39]. Two are the main externalities that

[39] Recent speeches by policy makers (Constancio [13]) highlight its potential benefits even hinting at the risks that it may pose for the profitability of the banking sector.

make this market attractive. The first are information externalities. Thanks to digital technology, lending platforms can provide a large variety of costless public information signals that help improve market transparency and facilitate price discovery in separating equilibria. Second, in times of bank fragility, platforms provide a valuable substitute that improves risk-sharing for investors and borrowers. We explore those two hypotheses through a theoretical and an empirical analysis, the latter based on data from the two biggest lending platform in the US, namely Prosper and Lending Club. Importantly, our empirical analysis allows us to quantify the value of information since we are able to quantify the reduction in rates that investors are willing to accept in exchange of additional signals.

Our results have several important implications. First, they show that transparency in debt markets helps to improve its liquidity and efficiency. This has important policy implications. Transparency is a prerequisite of equity markets, while debt markets are typically opaque. After the recent financial crisis, which originated in debt markets, a debate surged on whether enforcing higher transparency in debt market might help to improve stability and efficiency. Lending platforms provide a good experiment in this sense. Second our results also speak on the importance of fostering the emergence and growth of markets offering alternative funding and investment opportunities with respect to the traditional banking sector. This argument is at the core of important policy debates such as the one on the creation of a capital market union in Europe.

References

[1] Akerlof G.(1970). "The Market for lemons: Quality uncertainty and the market mechanism." *Quarterly Journal of Economics,* 84(3), pp. 488-500.

[2] Allen, F. and D. Gale, (1999). "Diversity of Opinion and Financing New Technologies." *Journal of Financial Intermediation,* 8, 68–89.

[3] Allen, F. and D. Gale, (2000). *Comparing Financial Systems,* MIT Press.

[4] Allen, F., S. Morris, and H. S. Shin, (2006). "Beauty Contests and Iterated Expectations in Asset Markets." *Review of Financial Studies*, 19(3): 719–52.

[5] Bachmann, A., Becker, A., Buerckner, D., Hilker, M., Kock, F., Lehmann, M., Tiburtius, P. and B. Funk (2011). "Online Peer-to-Peer Lending – A Literature Review." *Journal of Internet Banking and Commerce*; Vol. 16(2).

[6] Bacchetta, P. and E. van Wincoop, (2008). "Higher Order Expectations in Asset Pricing." *Journal of Money, Credit, and Banking*, 40(5): 837–66.

[7] Banerjee, A. V., (1992). "A Simple Model of Herd Behavior." *Quarterly Journal of Economics,*107(3), 797–817.

[8] Besley, Timothy and Stephen Coate (1995) "Group Lending, Repayment Incentives and Social Collateral." *Journal of Development Economics,* Vol. 46, 1-18.

[9] Bikhchandani, S., D. Hirshleifer, and I. Welch, (1992). "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades." *Journal of Political Economy,* 100(5), 992–1026.

[10] Board of Governors of the Federal Reserve System (2016). Charge-Off and Delinquency Rates on Loans and Leases at Commercial Banks. http://www.federalreserve.gov/releases/chargeoff/ (downloaded on Oct. 20, 2016).

[11] Bonfim, D. and J. Santos, (2017). "The importance of deposit insurance credibility." AEA Meetings.

[12] Chemla, G. and K. Tinn, (2018). “Learning through Crowd-funding.” Mimeo, Imperial College.

[13] Constancio, V. (2016). “Challenges for the European Banking Industry.” Speech delivered the Conference on “European Banking Industry: what’s next?”, organized by the University of Navarra.

[14] Duarte, J., S. Siegel and L. Young, (2012). “Trust and Credit: The Role of Appearance in Peer-to-Peer Lending.” *Review of Financial Studies,* 25(8), 2455-2483.

[15] Fang, L. and J. Peress, (2009). “Media Coverage and the Cross-Section of Stock Returns.” *Journal of Finance*, 64(5), 2023-2052.

[16] Gorton, G. (1988). “Banking Panics and Business Cycles.” *Oxford Economic Papers*, New Series, vol. 40(4), 751-781.

[17] Gorton, G. and A. Metrick, (2009). “Securitized Banking and the Run on Repo.” NBER w.p. 15223.

[18] Iyer, R., Khwaja, A., Luttmer, E. and K. Shue (2015). “Screening Peers Softly: Inferring the Quality of Small Borrowers.” *Management Science*, Vol. 62(6) pp. 1554 - 1577.

[19] Kaufmann, C., M. Weber, and E. Haisley, (2015). “The Role of Experience Sampling and Graphical Displays on One’s Investment Risk Appetite.” *Management Science,* 59, 323–340.

[20] Kawai, K., Onishi, K. and Uetake, K. (2016). “Signaling in Online Credit Markets.” Mimeo. Yale University.

[21] Leech, C. (2015). “Direct Lending: Finding value/minimizing risk.” Downloadable from: http://www.liberum.com/alternative-finance/opinion-insight (on October 17, 2016).

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

[23] OECD (2015). Financing SMEs and entrepreneurs 2015. An OECD Scoreboard. OECD Publishing.

[24] Prime Meridian Capital Management (Accessed October 17, 2016 - http://www.pmifunds.com/what-is-p2p-lending/).

[25] Paravisini, D., V. Rappoport and E. Ravina (2016). "Risk Aversion and Wealth: Evidence from Person-to-Person Lending Portfolios." Forthcoming *Management Science*.

[26] Peress, J. (2004). "Wealth, Information Acquisition and Portfolio Choice." *Review of Financial Studies*, 17(3), 879-914.

[27] Petriconi, S. (2016). "Bank Competition, Information Choice and Inefficient Lending Booms." Mimeo, Bocconi university.

[28] Ravina, E. (2012). "Love and Loans: The Effect of Beauty and Personal Characteristics in Credit Markets." Unpublished. Available at SSRN: http://ssrn.com/abstract=1101647.

[29] Repullo, R. and J. Suarez, (2013). "The Procyclical Effects of Bank Capital Regulation." *Review of Financial Studies*, vol. 26(2), pages 452-490.

[30] Ruckes, M. (2004). "Bank competition and credit standards." *The Review of Financial Studies*, 17(4):1073-1102.

[31] Spence, M. (1973). "Job Market Signaling." *Quarterly Journal of Economics*, Vol. 87(3), pp. 355-374.

[32] Stiglitz, J. and A. Weiss, (1981). "Credit Rationing in Markets with Incomplete Information." *American Economic Review*, 71(3), pp. 393-410.

[33] Theil, H. (1967). *Economics and Information Theory*, Rand McNally and Company - Chicago.

[34] Tinn, K. (2010). "Technology Adoption with Exit in Imperfectly Informed Equity Markets." *American Economic Review*, vol. 100(3), pages 925-57.

[35] Van Nieuwerburgh, S. and L. Veldkamp, (2010). "Information Acquisition and Under-diversification." *Review of Economic Studies,* 77(2), 779-805.

[36] Yaari, M. (1965). "Uncertain Lifetime, Life Insurance and the Theory of Consumer." *Review of Economics Studies,* 32, 137-50.

[37] Zopa.com (Accessed October 17, 2016 - https://www.zopa.com/about).

[38] Wette, H. (1983). “Collateral in Credit Rationing in Markets with Imperfect Information: Note.” *American Economic Review,* vol. 73, issue 3, 442-45.

6 Appendix A. Simulated Model

In section 3, we have derived analytically and discussed the channels of transmission at work in the model. Now, we quantify the effects through simulations that solve the entire set of equations of the model simultaneously. We proceed by calibrating the model and simulating it subject to a set of shocks. Specifically, we simulate the effects of shocks to the signal precision, λ, to the average success probability, $\bar{p}$, whereby shocks to the latter capture the selection channel through first order stochastic changes in the distribution of projects, and to the aggregate project return, R^I. Last, to capture the substitution channel, we shock the probability of bank default, ς.

The impulse responses to those shocks allow us to rationalize further our data evidence through graphics and through the lenses of a full model solution. The parameters of the model are chosen as close as possible to their data equivalent. To add further realism to the quantitative simulations we also estimated the shocks on data from US P2P platforms Prosper and Lending Club, and from the banking system.

We start by detailing our parameter choice and shock estimation. We then present the impulse response functions of selected variables and discuss the transmission mechanisms.

6.0.1 Calibration

Time is taken to be quarters. The discount factor, β, is set to 0.99, so as to induce a risk-free rate of 4% on annual basis. We assume a CRRA utility for households/lenders, $\frac{C^{1-\sigma}}{1-\sigma}$, with a risk-aversion parameter set to 2 which is within the range of most of the macro and household finance literature. We experiment also with higher values and notice that the transmission of shocks is robust. For the average haircut, Gorton and Metrick [17] report an average value of 17% for the period 2007-2009 for the repo-haircut index, where repos are short-term debt, whose returns and haircuts approximate well bank bonds'. We therefore set $\theta = 0.9$[40] and $\zeta = 0.2$ which yield an average haircut, $\bar{\theta}_t$ of 18% on annual basis.

We then set the average probability of success so as to satisfy our condition $\varpi_t \geq \bar{p}$. The return R^I is calibrated as follows. First, we set the value of R^I using the banks' break even condition,

[40] Bonfim and Santos[11] who examined severe bank run episodes in Europe find haircust at 9.9%. We use this value as reference to calibrate θ.

$R^I = \frac{R_t^d + \mu}{\widetilde{p}}$, and compatibly with an annual bank margin of 4% as in Repullo and Suarez[29]. Notice that monitoring costs are calibrated to 15 basis points on quarterly basis. This delivers a value of $R^I = 1.2$ (per quarter). The expected return of the P2P loans is $\left[\lambda s + (1-\lambda)\bar{p}\right] R^I$. We obtain a value for s using condition (11) in the steady state. The latter depends upon the value of λ. We shall set the benchmark value for λ so as to obtain an overall expected return on P2P loans of 1.073 percent. This value is in line with those on Prosper and Lending Club loans. Specifically, for Prosper, mean estimated returns range from 10.3% in 2009 to a value of 7.3% in 2014. The equivalent in Lending Club are slightly lower.

As to the shocks, we model them as autoregressive processes. For the shock to the project success probability, we take the loss rates of Lending Club loans for the period 2006 to 2016 and fit an auto-regressive process. The estimation delivers a persistence of 0.407039. Next, for the shock to the return on banks' liabilities,we collect quarterly data on the return on the equity of all U.S. banks again for the period 2006 to 2016, and fit an auto-regressive process. This yields a persistence of 0.878877. We cannot calibrate the changes in the precision of signals directly, because in the data it corresponds to a qualitative increase in transparency or in the number of the signals available. Hence, for this shock, impulse responses are indicative of the qualitative trends in the transmission mechanism. Last, to calibrate the liquidity shock we take the quarterly LIBOR-OIS spread from Bloomberg (difference between the London interbank offered rate and the overnight swap index) for the period 2006 to 2016 and again fit an AR(1) process. The resulting persistence is 0.762562.

To conclude, notice that our model implies that households' portfolio choice is between deposits and P2P loans. To make sure that indeterminacy does not arise in steady state[41] we guess a model-consistent functional form for the aggregate wealth accumulation. This allows to pin down exactly the investments in each asset.

6.0.2 Quantitative Results

Figure 1 shows the impulse responses of selected variables to a 1% increase in signal precision, λ. This shock captures the *information channel*. An increase in λ (first panel) implies that investors are now better at discerning loans' quality. The threshold for the marginal funded project (second

[41] This possibility exists as different allocation between the two assets may be compatible with the same set of first order conditions.

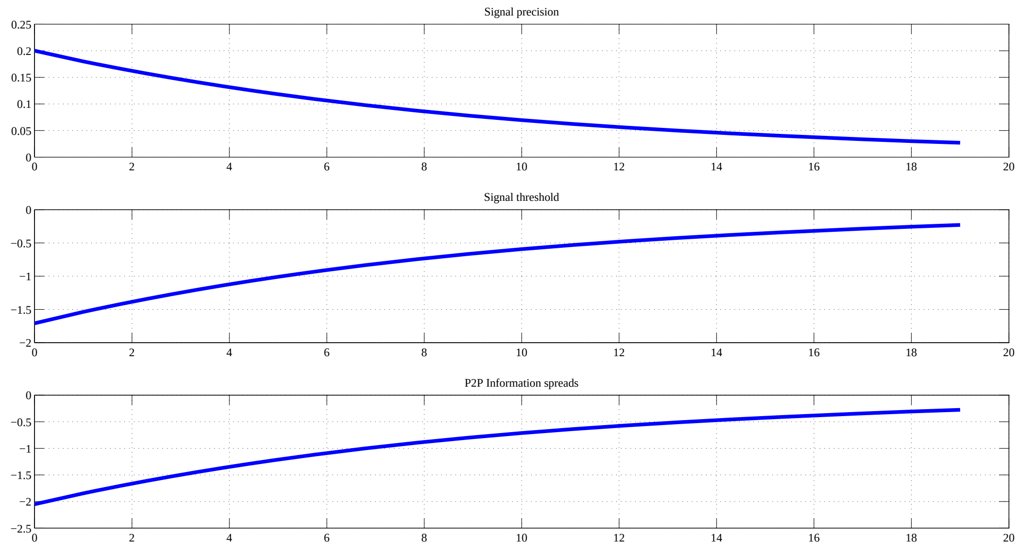

Figure 1: Impulse responses of selected variables to a 1% increase to signal precision λ.

panel) drops and households are willing to fund more projects. We also compute, and show in the third panel, the average information spread, defined as the difference between the conditional expected returns $\left[\lambda s + (1-\lambda)\bar{p}\right] R^I$ under full information, namely with $\lambda = 1$, and under partial information[42]. As expected, it also declines because the severity of the information asymmetry is reduced.

Figure 2 shows the impulse responses to an increase in the average loan quality, $\bar{p}$. This represents a shift upward toward a probability distribution which first-order stochastically dominates the initial one. Intuitively, the quality of all projects increases, for given signal precision. Hence, this shocks captures purely the selection channel. An increase in $\bar{p}$ (first panel) results in a decline of the threshold for the marginally funded project (because the average projects' default rate has declined) and also of the loan information spread, for any degree of signal precision. Like before, households are now willing to fund more projects on the platform.

Figure 3 shows the impulse responses to an increase in projects' return, R^I. Responses are rather intuitive. Now. the expected return on the platform raises, and, hence, the loan spreads

[42]Note that this is different, albeit related to the individual loan spreads, $\chi_{\lambda=1}(\varpi) - \chi_{\lambda}(\varpi)$, described in the analytical section 3.3.

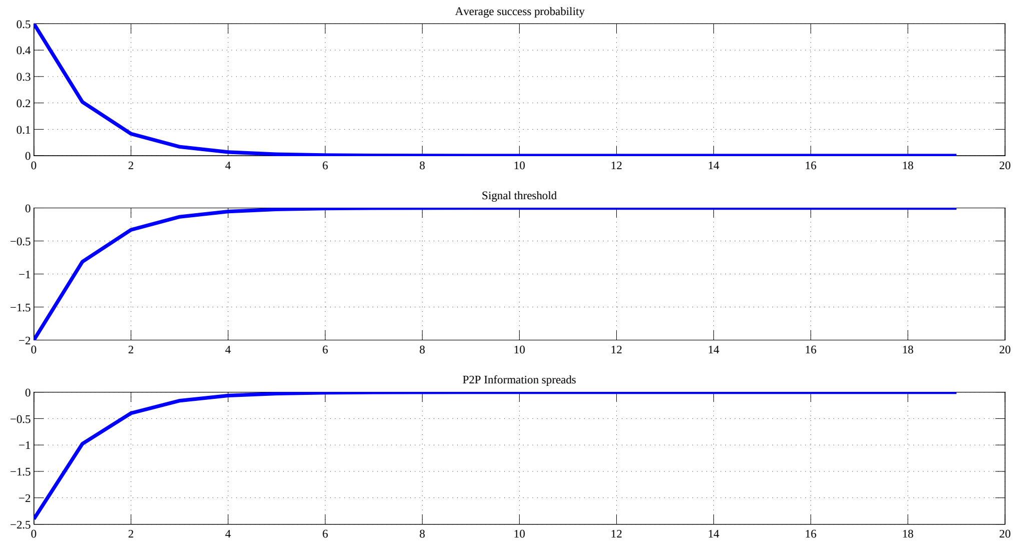

Figure 2: Impulse responses of selected variables to a 1% increase to average sucess probability, $\bar{p}$.

requested by households fall. The threshold for the marginal projects falls too and households fund more projects.

Finally, figure 4 shows the impulse responses to changes in the probability of bank defaults. This simulation is meant to capture the substitution channel. For this reason, in this case, we plot the returns on deposits and the return on the P2P loan funded at the margin. As risk in the banking sector raises, banks have to offer higher returns to depositors. Still savings shift to the relatively less risky P2P sector and this reduces the marginal return there.

7 Appendix B. Prosper Data Description

Prosper (http://prosper.com) is an online platform for peer-to-peer lending, which opened up on February 5th, 2006, in the US. Borrowers create personal profiles and solicit loans via online listings detailing the amount requested, the interest rate, the term and purpose for the loan. The borrower's profile includes independently verified information on his credit history[43], income, and

[43] Each listing includes hard credit data such as lower and upper values delimiting the range of the borrower's credit score as provided by a consumer credit rating agency, number of credit inquiries in the last six months, the number of open credit lines, revolving accounts and credit account records, and number of accounts delinquent.

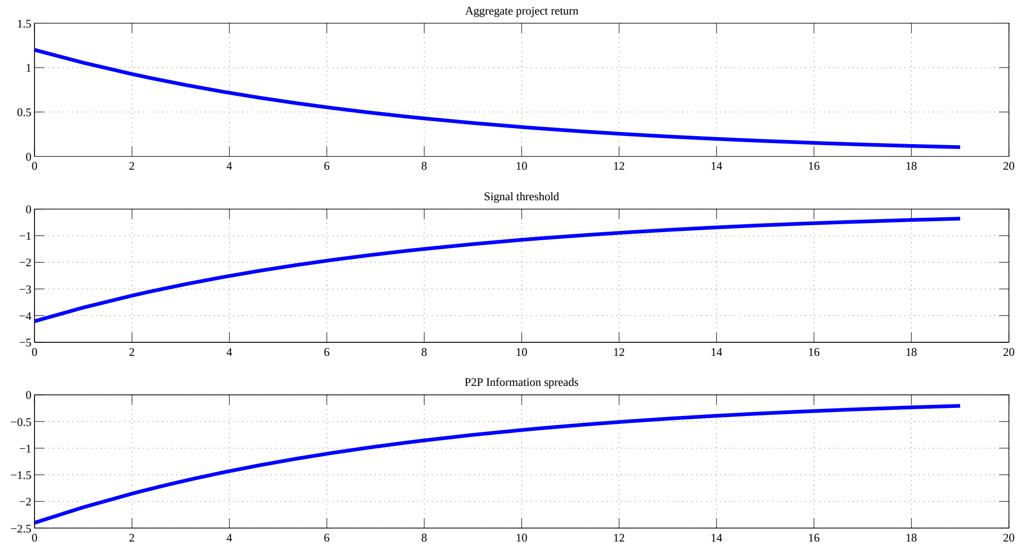

Figure 3: Impulse responses of selected variables to a 1% increase to aggregate project return, R^I.

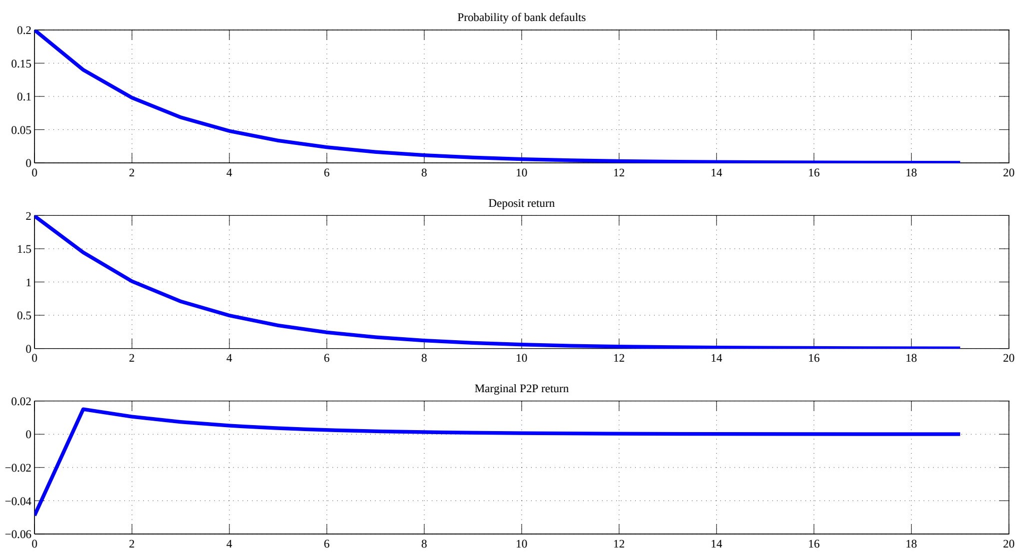

Figure 4: Impulse responses of selected variables to a 1% increase to probability of bank default, ς.

current debts. Each listing includes also a Prosper Rating, i.e a credit grade that goes from 7 (label AA) to 1 (label HR - "High Risk") determined by Prosper based on a proprietary algorithm. This rating depends on two scores: (1) the credit score, obtained from an official credit reporting agency, and (2) Prosper Score, computed in-house based on the Prosper population.[44]

Prosper also creates social networks by linking borrowers in groups (tied by geography, common interest, or common loan purpose) and collecting the endorsements of other Prosper members (friends). For what concerns groups, any borrower can join one of the several groups in the Prosper marketplace or propose one. There are groups related to a school (alumni) or an employer, geographically-oriented groups, military groups, medical groups related to medical reasons for borrowing, demographic groups highlighting some particular demographics, such as single parents or Hispanics, groups targeting people with particular hobbies, religiously-oriented groups, and business groups for small businesses, or for new business development. Admission to a group is based on eligibility criteria. Some groups, such as employee or university alumni groups, require verification of the qualifying criteria, whereas others have looser joining criteria. An individual can be a member of only one group at a time and borrowers cannot leave or change group until loan repayment.

In turn, lenders assess and can bid on listings. The minimum bid is set at $25. At its incept Prosper operated a variable rate model. It worked as an eBay-style auction marketplace. In their bid, lenders would specify the investment and the minimum interest they were willing to accept. If the total bid amount exceeded the amount requested, those lenders asking the lowest interests were granted a share of the loan. In 2009, Prosper registered with the SEC and afterwards it changed its business model to use pre-set rates determined solely by an own algorithm evaluating each prospective borrower's profile and credit risk. Under the new approach, lenders no longer determine the loan rate in an auction. Instead, they simply choose whether or not to invest at the rate which Prosper's loan proprietary pricing algorithm assigns to the loan after analyzing the borrower's credit report and financial information. Following the SEC registration, new prospective borrowers are required to have a FICO credit score of at least 640, while returning borrowers only need a score of 600 to request a loan.[45]

[44] Prosper Ratings are available for loans originated after July 2009.

[45] The original model effectively had the investors setting the rates through an open bidding process. The problem

Borrowers are limited to a maximum of two concurrent loans of $25,000 or less each until 2012, $35,000 thereafter.[46] Fees of 1 to 5 percent of the loan amount are deducted from the loan, depending on the borrower's risk profile and loan duration. Loans amortize over a 36-month period until 2009, a maximum 60-month period thereafter. Repayments are in monthly installments that are automatically deducted from a borrower's bank account. Prosper loan status is defined as 'Completed' if the loan has been fully paid off, and as 'Current' if payments are made on time and as agreed. When the borrower misses payments, loans are classified as 'Past due'. When loans are 120 days past due (i.e., 4 payments are missing), Prosper moves them to 'Charged off' status. A loan designated as 'Charged off' is due in full immediately. It can also be sold to outside debt collectors. For investors, the entire balance moves into a charged off balance and is assumed to be lost. There are some cases where, after a loan is charged off, the borrower may still make some payments. About 16% of Prosper's charged off loans have had some level of recovery. A loan is tagged as in 'Default' in case of delinquency, bankruptcy or death. Bankruptcy is the most prominent reason for a loan to be tagged as defaulted. Some of the loans in 'Default' ultimately get settled or paid in part or in full. Delinquencies are reported to credit report agencies and affect the borrower's credit score. Borrowers who default on their loans are not allowed to borrow using Prosper.com again.

8 Appendix C. Lending Club Institutional Data Description

Lending Club is the world largest platform for peer-to-peer lending. All loans are unsecured personal loans and are between $1,000 and $35,000. Based on the borrower's credit history and score, on her debt-to-income ratio and desired loan amount, Lending Club assigns loans a letter credit grade to which a specific interest rate and fees are associated. Rates vary between 5% and 26% depending on the credit grade. The range of grades is similar to Prosper's, but Prosper's average interests

with that original model, that was shut down by regulatory authorities, is that the investors were individuals without any demonstrable skill and there were no controls in place to ensure that they had sufficient resources to take on the risk. Additionally, that model had severe inefficiencies because of a lack of large institutional investors that could immediately fund a whole loan. This left severe cash drag issues with many loans going unfunded. Additionally, from a regulatory perspective, there were severe lending risks due to no discernible way in which the interest rates were set.

[46] Currently, loans are funded if they obtain 70 percent or more of the amount requested. Originally, no partial funding was allowed.

are higher. For example, for credit grade A, the rate is 4% higher, and that difference increases as credit grade worsens. The standard loan period is three years. A five year period is available at a higher interest rate and additional fees. Loans can be repaid in full at any time without penalty.

Investors can search and browse the loan listings on Lending Club website. The minimum investment is $25.

Lending Club makes money by charging borrowers an origination fee, which ranges between 1% and 5% depending on the credit grade, and charging lenders a 1% service fee on all amounts the borrower pays. To reduce default risk, Lending Club focuses on high creditworthy borrowers, with a FICO score of 660 or more, declining approximately 90% of the loan applications as of 2012.

Summary statistics for this data set are in tables C1 and C2. Most borrowers on Lending Club's website report using their loans to refinance other loans or pay credit card debt (table C1, panel (a)). Like for Prosper, the share of this type of borrowers has increased over time, whereas the share who intend to fund their business has sharply declined. Borrowers debt-to-income ratio (excluding mortgage) has increased over time and gone from 10 percent in 2007 to 18 percent in 2014 (table C2). On average, borrowers have 11 open credit lines; 20 percent have been delinquent on a loan in the two years before borrowing on Lending Club. The mean personal income is $73,000 and the average loan is $14,000. Investors funded $5 mln in loans in 2007 up to $3.5 bln in 2014. The nominal average interest rate was 12% in 2007 and 14% in 2014. The default rate is around 11%.

Table 1 – Summary Statistics

Panel (a)

Year of the loan	2006	2007	2008	2009	2010	2011	2012	2013	2014
Borrower lending rate	0.191	0.177	0.186	0.193	0.213	0.230	0.220	0.184	0.153
	(0.069)	*(0.064)*	*(0.085)*	*(0.091)*	*(0.098)*	*(0.079)*	*(0.077)*	*(0.061)*	*(0.054)*
Borrower annual percentage rate (APR)	0.201	0.186	0.204	0.216	0.239	0.262	0.253	0.214	0.182
	(0.070)	*(0.066)*	*(0.089)*	*(0.095)*	*(0.106)*	*(0.086)*	*(0.082)*	*(0.065)*	*(0.059)*
Estimated effective yield				0.103	0.106	0.213	0.201	0.162	0.134
				(0.052)	*(0.055)*	*(0.074)*	*(0.071)*	*(0.054)*	*(0.048)*
Size of loans	4763	7050	6022	4355	4767	6692	7834	10545	11912
	(4404)	*(6126)*	*(5400)*	*(4070)*	*(3714)*	*(4273)*	*(5527)*	*(6575)*	*(6684)*
Term (months)	36	36	36	36	36	37	43	45	44
Time for funding (median)	9	11	10	14	12	10	8	6	5
Lenders' investment: mean/median	231/96	126/58	122/45	98/29	55/35	286/78	381/89	5,762/3,000	9,131/9,000
No. of investors: mean/median	57/36	127/92	136/95	146/93	134/103	80/55	82/53	56/1	29/1
Loans funded by single investor (%)	2	1	1	1	<1	1	2	51	75
Loans for debt consolidation (%)		42	46	47	48	48	74	79	42
home improvement (%)		5	9	10	11	11	6	4	5
business (%)		16	11	10	11	9	4	3	16
other (%)		37	34	33	30	32	16	14	37
# observations	5,906	11,460	11,552	2,047(1)	5,652	11,228	19,553	33,910	11,734(2)

Notes: Mean values, unless specified otherwise. Standard deviations in parentheses. (1) In 2009, loan issuance was suspended during Prosper registration with the SEC. (2) Loans issued from January to March 11th.

Panel (b) – Loan status (as of March 31st 2014)

Year of the loan	2006	2007	2008	2009	2010	2011	2012	2013	2014	Total
Completed	61%	61%	67%	85%	83%	49%	28%	7%	1%	34%
Current	-	-	-	-	-	29%	54%	89%	99%	49%
Past Due (1-120 days)	-	-	-	-	-	3%	4%	3%	-	2%
Charged-off	16%	26%	24%	11%	14%	16%	12%	1%	-	11%
Defaulted	23%	14%	9%	4%	3%	3%	2%	0%	-	4%
	100%	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*

Table 2 – Hard and soft information about borrowers

Year of the loan	2006	2007	2008	2009	2010	2011	2012	2013	2014
Mean FICO score[1]	609	654	674	715	714	709	711	708	703
Number of open credit lines		8	8	9	8	8	8	10	11
		(5)	*(5)*	*(5)*	*(5)*	*(5)*	*(5)*	*(5)*	*(5)*
Number of credit inquiries	11	10	8	6	4	4	4	4	4
	(12)	*(11)*	*(8)*	*(5)*	*(4)*	*(4)*	*(4)*	*(4)*	*(4)*
Borrowers w/ current delinquencies (%)	52	39	23	11	14	21	20	15	10
Prosper credit rating				4.286	3.837	3.552	3.688	4.258	4.718
				(1.937)	*(1.985)*	*(1.710)*	*(1.829)*	*(1.468)*	*(1.387)*
Estimated loss				0.075	0.093	0.097	0.091	0.073	0.062
Estimated return at issuance				0.103	0.103	0.115	0.110	0.088	0.073
Debt-income ratio	0.249	0.431	0.254	0.228	0.230	0.251	0.264	0.264	0.259
	(0.737)	*(1.318)*	*(0.342)*	*(0.152)*	*(0.299)*	*(0.402)*	*(0.464)*	*(0.243)*	*(0.113)*
Monthly income	4,744	4,654	4,619	5,092	5,291	5,660	5,710	6,161	6,336
	(5,207)	*(4,711)*	*(3,705)*	*(3,225)*	*(4,099)*	*(8,544)*	*(13,350)*	*(5,664)*	*(4,382)*
Borrowers who are in a group (%)	70	51	14	11	9	5	3	1	1
Borrowers with recomm. from Prosper friends (%)		17	18	8	6	3	2	1	<1
Borrowers with invest. from Prosper friends (%)		6	7	5	4	1	1	<1	<1
$ investment from friends (cond. on having friends)		939	1017	713	773	572	429	233	298
Borrowers with previous Prosper loans (%)	-	4	15	43	34	34	28	19	10
# observations	5,906	11,460	11,552	2,047	5,652	11,228	19,553	33,910	11,734

Notes: (1) Prosper data do not include borrowers' exact FICO score, but only a lower and an upper value representing the range of the borrower's credit score as provided by a consumer credit rating agency. The table reports the mean of these values. See also notes to Table 1.

Table 3 – Borrowers who are part of a group

	2006-2008		2009-2010		2011		2012		2013		2014	
	In group		In group		In group		In group		In group		In group	
	0	1	0	1	0	1	0	1	0	1	0	1
Loan status												
Completed	66%	58%	83%	87%	49%	56%	28%	33%	7%	11%	1%	1%
Current	0%	0%	0%	0%	29%	26%	54%	53%	90%	84%	99%	99%
Past due (1-120 days)	0%	0%	0%	0%	3%	2%	4%	4%	3%	3%	0%	0%
Charged off or defaulted	34%	42%	17%	13%	19%	16%	14%	10%	1%	1%	0%	0%
	100%	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*
Borrower APR	0.193	0.201	0.232	0.244	0.263	0.252	0.254	0.226	0.214	0.191	0.183	0.161
	(0.081)	*(0.071)*	*(0.104)*	*(0.102)*	*(0.086)*	*(0.089)*	*(0.818)*	*(0.080)*	*(0.065)*	*(0.075)*	*(0.059)*	*(0.064)*
Estimated return			0.105	0.092	0.116	0.111	0.110	0.105	0.089	0.083	0.073	0.068
			(0.050)	*(0.060)*	*(0.032)*	*(0.034)*	*(0.029)*	*(0.031)*	*(0.019)*	*(0.023)*	*(0.014)*	*(0.016)*
Prosper rating			4.010	3.429	3.550	3.598	3.676	4.179	4.251	4.799	4.722	5.209
Mean FICO score[(1)]	669	627	716	699	710	697	712	695	709	702	703	702
Time for funding (median)	10	11	12	13	10	9	8	7	6	7	5	6
# observations	17,252	11,666	6,953	746	10,676	552	19,066	487	33,499	412	11,666	67

Notes: (1) Prosper data do not include borrowers' exact FICO score, but only a lower and an upper value representing the range of the borrower's credit score as provided by a consumer credit rating agency. The table reports the mean of these values. See also notes to Table 1.

Table 4 – Borrowers with recommendations from Prosper friends

	2007-2008 w/friends(*)		2009-2010 w/friends		2011 w/friends		2012 w/friends		2013 w/friends		2014 w/friends	
	0	1	0	1	0	1	0	1	0	1	0	1
Loan status												
Completed	63%	66%	83%	88%	49%	58%	28%	37%	7%	14%	0%	0%
Current	0%	0%	0%	0%	29%	29%	54%	52%	89%	83%	99%	100%
Past due (1-120 days)	0%	0%	0%	0%	3%	1%	5%	4%	3%	3%	1%	0%
Chargedoff or defaulted	37%	34%	17%	12%	19%	12%	13%	7%	1%	<1%	0%	0%
	100%	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*
Borrower APR	0.197	0.192	0.233	0.231	0.263	0.248	0.254	0.220	0.214	0.184	0.183	0.150
	(0.077)	*(0.078)*	*(0.104)*	*(0.104)*	*(0.086)*	*(0.088)*	*(0.082)*	*(0.076)*	*(0.065)*	*(0.075)*	*(0.059)*	*(0.068)*
Estimated return			0.105	0.081	0.116	0.111	0.110	0.105	0.089	0.080	0.073	0.065
			(0.050)	*(0.059)*	*(0.032)*	*(0.034)*	*(0.029)*	*(0.033)*	*(0.019)*	*(0.024)*	*(0.014)*	*(0.017)*
Prosper rating			3.974	3.614	3.549	3.690	3.680	4.295	4.253	4.948	4.723	5.522
Mean FICO score(1)	652	663	715	706	709	696	711	696	708	705	703	708
Friends who invest (%)	1	45	2	42	1	20	0	14	0	11	0	13
# observations	25,887	3,031	7,272	427	10,938	290	19,299	254	33,679	232	11,710	23

Notes: (*) The dummy 'w/friends' takes on value 1 if the borrower has any recommendations from friends. It is available from 2007. (1) Prosper data do not include borrowers' exact FICO score, but only a lower and an upper value representing the range of the borrower's credit score as provided by a consumer credit rating agency. The table reports the mean of these values. See also notes to Table 1.

Table 5 – Borrowers with previous loans through Prosper

	2007-2008		2009-2010		2011		2012		2013		2014	
	Prior loans(*)		Prior loans		Prior loans		Prior loans		Prior loans		Prior loans	
	0	1	0	1	0	1	0	1	0	1	0	1
Loan status												
Completed	63%	69%	82%	86%	49%	50%	28%	30%	6%	10%	1%	1%
Current	0%	0%	0%	0%	27%	32%	54%	53%	91%	85%	99%	99%
Past due (1-120 days)	0%	0%	0%	0%	3%	3%	4%	5%	3%	3%	0%	0%
Chargedoff or defaulted	37%	31%	18%	14%	21%	15%	14%	12%	1%	2%	0%	0%
	100%	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*
Borrower APR	0.194	0.203	0.224	0.249	0.271	0.245	0.263	0.227	0.218	0.199	0.185	0.171
	(0.078)	*(0.086)*	*(0.103)*	*(0.103)*	*(0.086)*	*(0.085)*	*(0.082)*	*(0.075)*	*(0.063)*	*(0.071)*	*(0.059)*	*(0.058)*
Estimated return			0.107	0.097	0.117	0.111	0.111	0.107	0.089	0.086	0.073	0.070
			(0.047)	*(0.056)*	*(0.032)*	*(0.032)*	*(0.029)*	*(0.028)*	*(0.018)*	*(0.022)*	*(0.014)*	*(0.014)*
Prosper rating			4.255	3.424	3.458	3.731	3.502	4.178	4.163	4.658	4.697	4.973
Mean FICO score(1)	664	665	725	697	718	693	719	689	710	703	703	702
# observations	26,700	2,218	4,985	2,804	7,361	3,867	14,165	5,388	27,444	6,466	10,556	1,177

Notes: (*) The dummy 'Prior loans' takes on value 1 if the borrower funded some project on Prosper before. (1) Prosper data do not include borrowers' exact FICO score, but only a lower and an upper value representing the range of the borrower's credit score as provided by a consumer credit rating agency. The table reports the mean of these values. See also notes to Table 1.

Table 6 – OLS regressions of lending rates on loan characteristics

	All	Auction pricing(1)	Prosper pricing
Loan size (thousands)	-8.989	-7.757	-10.180
	(0.102)***	(0.261)***	(0.109)***
Loan size (thousands)	1.862	2.517	2.038
	(0.033)***	(0.102)***	(0.034)***
Term (months) (2)	1.132	-	1.248
	(0.019)***		(0.019)***
Debt consolidation(*)	0.448	1.349	0.454
	(0.060)***	(0.167)***	(0.063)***
Home improvement(*)	-0.318	-0.588	-0.291
	(0.099)***	(0.338)*	(0.103)***
Business funding(*)	0.812	0.149	1.006
	(0.104)***	(0.220)	(0.115)***
Adjusted R^2	0.23	0.12	0.28
N	107,549	23,425	84,124

Note: The left-hand-side variable is the lending rate in percentage points. Quarter-year dummies and US state dummies are also included. (*) denotes a dummy variable. (1) 'Auction pricing' refers to loans whose rates are set by lenders, which was Prosper pricing mechanism until 2010. 'Prosper pricing' refers to loans whose rates are set centrally by Prosper before posting. (2) Until 2010, all loans had a 36 month term. Robust standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01

Table 7 – OLS regressions of lending rates on loan characteristics and signals

	All	Auction pricing	Prosper pricing	Auction pricing	Prosper pricing	Prosper pricing
Loan size (thousands)	-4.336	1.640	-6.331	1.766	-6.292	-6.082
	(0.091)***	(0.208)***	(0.099)***	(0.207)***	(0.099)***	(0.093)***
Loan size (thousands)^2	1.037	0.260	1.485	0.234	1.478	1.390
	(0.028)***	(0.081)***	(0.029)***	(0.080)***	(0.029)***	(0.028)***
Term	0.941		1.108		1.105	1.162
	(0.017)***		(0.017)***		(0.017)***	(0.016)***
Debt consolidation$^{(*)}$	-0.231	-0.098	-0.130	-0.120	-0.125	-0.287
	(0.051)***	(0.120)	(0.055)**	(0.119)	(0.055)**	(0.051)***
Home improvement$^{(*)}$	0.040	0.090	0.021	0.057	0.043	0.392
	(0.082)	(0.227)	(0.086)	(0.227)	(0.085)	(0.079)***
Business funding$^{(*)}$	0.605	0.316	0.508	0.297	0.560	0.635
	(0.094)***	(0.158)**	(0.108)***	(0.157)*	(0.108)***	(0.101)***
FICO score (hundreds)	-7.011	-7.057	-7.295	-7.123	-7.342	-7.890
	(0.039)***	(0.063)***	(0.046)***	(0.063)***	(0.046)***	(0.043)***
Open credit lines (tens)	0.329	0.524	0.060	0.539	0.074	0.314
	(0.047)***	(0.080)***	(0.055)	(0.079)***	(0.055)	(0.052)***
Credit enquiries (tens)	1.658	0.894	2.413	0.887	2.454	2.977
	(0.049)***	(0.055)***	(0.061)***	(0.054)***	(0.061)***	(0.060)***
Current delinquencies$^{(*)}$	1.152	2.733	0.833	2.797	0.867	0.856
	(0.051)***	(0.101)***	(0.057)***	(0.100)***	(0.057)***	(0.054)***
Monthly income	-0.059	-0.001	-0.052	-0.005	-0.053	-0.048
(thousands)	(0.015)***	(0.009)	(0.015)***	(0.009)	(0.015)***	(0.013)***
Debt/Income	1.197	0.326	2.742	0.364	2.738	2.907
	(0.075)***	(0.048)***	(0.228)***	(0.047)***	(0.229)***	(0.243)***
Group dummy$^{(*)}$				-0.512	-1.938	-0.026
				(0.092)***	(0.148)***	(0.146)
Recommend + no investm. $^{(*)}$				0.039	-2.493	-0.420
				(0.128)	(0.222)***	(0.217)*
Recommend + investm. $^{(*)}$				-1.884	-1.466	-0.838
				(0.160)***	(0.411)***	(0.395)**
Investm.+ no recommend. $^{(*)}$				-4.478	-1.198	-0.809
				(0.732)***	(0.423)***	(0.419)*
Previous Prosper loan$^{(*)}$						-4.158
						(0.048)***
Adjustment R^2	0.49	0.59	0.51	0.59	0.51	0.56
N	95,396	18,497	76,899	18,497	76,899	76,899

Note: See note to Table 6. Robust standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01

Table 8 – Lending rates and signal precision

	Income has been verified	No open credit lines	No state of residency	No reason for borrowing
	All	All	Auction pricing	Prosper pricing
Loan size (thousands)	-4.376	-4.263	1.806	-6.341
	(0.090)***	(0.091)***	(0.196)***	(0.099)***
Loan size (thousands)^2	1.049	1.015	0.183	1.486
	(0.028)***	(0.028)***	(0.076)**	(0.029)***
Term	0.947	0.942		1.109
	(0.017)***	(0.017)***		(0.017)***
Income is verifiable(*)	-2.406			
	(0.925)***			
No open credit lines		2.436		
		(0.274)***		
No US State(*)			1.360	
			(0.703)*	
No reason for borrowing(*)				0.630
				(0.091)***
FICO score (hundreds)	-6.989	-6.962	-6.997	-7.296
	(0.039)***	(0.039)***	(0.059)***	(0.046)***
Open credit lines (tens)	0.225		0.480	0.060
	(0.048)***		(0.076)***	(0.055)
Credit enquiries (tens)	1.675	1.731	0.838	2.416
	(0.048)***	(0.049)***	(0.051)***	(0.061)***
Current delinquencies(*)	1.154	1.094	2.594	0.832
	(0.051)***	(0.051)***	(0.094)***	(0.057)***
Monthly income (thousands)	-0.205	-0.053	-0.002	-0.052
	(0.069)***	(0.013)***	(0.008)	(0.015)***
Income * Inc. is verif. (*)	0.155			
	(0.070)**			
Debt/Income	-0.017	1.246	0.332	2.736
	(0.116)	(0.075)***	(0.039)***	(0.228)***
Debt/Income* Inc. is verif. (*)	2.121			
	(0.186)***			
Motives for borrowing	Y	Y	Y	Y
Adjustment R^2	0.93	0.49	0.94	0.51
N	95,396	95,396	20,213	76,899

Note: Motives for borrowing include the three dummies for borrowing for debt consolidation, home improvement or business funding. See note to Table 6. Robust standard errors in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 9 – Lending rates, banking failures and signals

	All	Auction pricing	Prosper pricing	Auction pricing	Prosper pricing
Loan size (thousands)	-8.990	-7.779	-10.269	1.769	-6.082
	(0.102)***	(0.262)***	(0.109)***	(0.207)***	(0.093)***
Loan size (thousands)^2	1.862	2.527	2.065	0.231	1.390
	(0.033)***	(0.102)***	(0.033)***	(0.080)***	(0.028)***
Term (months)	1.133		1.254		1.163
	(0.019)***		(0.019)***		(0.016)***
Debt consolidation$^{(*)}$	0.450	1.386	0.467	-0.126	-0.287
	(0.060)***	(0.167)***	(0.063)***	(0.119)	(0.051)***
Home improvement$^{(*)}$	-0.317	-0.552	-0.297	0.070	0.392
	(0.099)***	(0.341)	(0.103)***	(0.227)	(0.079)***
Business funding$^{(*)}$	0.815	0.142	1.003	0.294	0.637
	(0.104)***	(0.221)	(0.115)***	(0.157)*	(0.101)***
Bank failures$_{mo-1}^{(*)\,(1)(2)}$	-0.093	-2.083	0.072	-0.696	0.013
	(0.167)	(0.796)***	(0.171)	(0.539)	(0.140)
Bank failures$_{mo-2}^{(*)\,(2)}$	-0.554	-2.708	-0.373	-0.337	-0.316
	(0.157)***	(0.681)***	(0.162)**	(0.484)	(0.134)**
Bank failures$_{mo-3}^{(*)\,(2)}$	-0.293	-2.299	-0.106	0.021	-0.220
	(0.154)*	(1.142)**	(0.156)	(0.790)	(0.128)*
FICO score (hundreds)				-7.121	-7.890
				(0.063)***	(0.043)***
Open credit lines (tens)				0.544	0.316
				(0.079)***	(0.052)***
Credit enquiries (tens)				0.889	2.976
				(0.054)***	(0.060)***
Current delinquencies$^{(*)}$				2.796	0.854
				(0.100)***	(0.054)***
Monthly income				-0.004	-0.048
(hundreds of thousands)				(0.009)	(0.013)***
Debt/Income				0.364	2.907
				(0.047)***	(0.243)***
Group dummy$^{(*)}$				-0.480	-0.022
				(0.094)***	(0.146)
Recommend + no investm. $^{(*)}$				0.050	-0.421
				(0.128)	(0.217)*
Recommend + investm. $^{(*)}$				-1.884	-0.817
				(0.160)***	(0.394)**
Investm.+ no recommend. $^{(*)}$				-4.472	-0.779
				(0.732)***	(0.418)*
Previous Prosper loan$^{(*)}$				-0.178	-4.158
				(0.131)	(0.048)***
Adjustment R^2	0.23	0.13	0.29	0.59	0.56
N	107,549	23,425	84,124	18,497	76,899

Note: (1) Data source: Federal deposit insurance corporations (https://www.fdic.gov/bank/individual/failed/banklist.html). (2) 'Bank failures$_{mo-k}$' is a dummy that takes on value 1 if in the kth month before listing more than one bank failed in the State of the borrower. See note to Table 6. Robust standard errors in parentheses. * *p*<0.1; ** *p*<0.05; *** *p*<0.01

Table 10 – Lending Club vs. Prosper: OLS regressions

	Tb. 6, col. (1)		*Tb 7, col. (1)*		*Tb. 9, col. (1)*			
	Lending Club	*Prosper*	Lending Club	*Prosper*	Lending Club	*Prosper*	Lending Club	*Prosper*
Loan size (thousands)	-2.671	*-8.989*	-0.241	*-4.336*	-2.670	*-8.990*	-0.203	*-4.463*
	(0.026)***	*(0.102)****	(0.008)***	*(0.091)****	(0.026)***	*(0.102)****	(0.008)***	*(0.092)****
Loan size (thousands)^2	0.723	*1.862*	0.073	*1.037*	0.723	*1.862*	0.065	*1.044*
	(0.007)***	*(0.033)****	(0.002)***	*(0.028)****	(0.007)***	*(0.033)****	(0.002)***	*(0.029)****
Term	1.952	*1.132*	0.129	*0.941*	1.952	*1.133*	0.113	*0.983*
	(0.006)***	*(0.019)****	(0.002)***	*(0.017)****	(0.006)***	*(0.019)****	(0.002)***	*(0.017)****
Debt consolidation$^{(*)}$	-0.976	*0.448*	-0.122	*-0.231*	-0.976	*0.450*	-0.108	*-0.267*
	(0.020)***	*(0.060)****	(0.006)***	*(0.051)****	(0.020)***	*(0.060)****	(0.006)***	*(0.051)****
Credit card payments$^{(*)}$	-2.284		-0.223		-2.284		-0.187	
	(0.022)***		(0.006)***		(0.022)***		(0.007)***	
Home improvement$^{(*)}$	-1.530	*-0.318*	-0.175	*0.040*	-1.530	*-0.317*	-0.162	*0.146*
	(0.030)***	*(0.099)****	(0.009)***	*(0.082)*	(0.030)***	*(0.099)****	(0.009)***	*(0.083)**
Business funding$^{(*)}$	1.403	*0.812*	-0.008	*0.605*	1.403	*0.815*	-0.002	*0.636*
	(0.054)***	*(0.104)****	(0.016)	*(0.094)****	(0.054)***	*(0.104)****	(0.016)	*(0.095)****
FICO score (hundreds)				*-7.011*				*-6.869*
				*(0.039)****				*(0.039)****
LC grade B$^{(*)}$			3.756				3.794	
			(0.005)***				(0.005)***	
LC grade C$^{(*)}$			6.826				6.867	
			(0.005)***				(0.005)***	
LC grade D$^{(*)}$			9.713				9.756	
			(0.006)***				(0.006)***	
LC grade E$^{(*)}$			12.629				12.658	
			(0.008)***				(0.009)***	
LC grade F$^{(*)}$			15.558				15.587	
			(0.014)***				(0.014)***	
LC grade G$^{(*)}$			16.996				16.998	
			(0.029)***				(0.029)***	
Open credit lines (tens)			-0.112	*0.329*			-0.068	*0.320*
			(0.011)***	*(0.047)****			(0.012)***	*(0.046)****
Open credit lines (tens)^2			0.017				0.001	
			(0.004)***				(0.004)	
Credit enquiries (tens)			0.622	*1.658*			0.759	*1.635*
			(0.018)***	*(0.049)****			(0.019)***	*(0.049)****
Current delinquencies$^{(*)}$			0.072	*1.152*			0.072	*1.253*
			(0.004)***	*(0.051)****			(0.004)***	*(0.051)****
Annual income (thousands)			-0.066	*-0.059*			-0.068	*-0.055*
			(0.008)***	*(0.015)****			(0.008)***	*(0.014)****
Debt/Income			0.701	*1.197*			0.583	*1.193*
			(0.025)***	*(0.075)****			(0.026)***	*(0.075)****
Bank failures$_{mo-1}$$^{(*)}$					-0.043	*-0.093*	-0.021	*-0.111*
					(0.050)	*(0.167)*	(0.018)	*(0.141)*
Bank failures$_{mo-2}$$^{(*)}$					-0.119	*-0.554*	-0.023	*-0.241*
					(0.049)**	*(0.157)****	(0.017)	*(0.135)**

Bank failures$_{mo-3}$ $^{(*)}$					-0.033	*-0.293*	-0.044	*-0.196*
					(0.047)	*(0.154)**	(0.017)***	*(0.132)*
Adjusted R2	0.29	*0.23*	0.94	*0.49*	0.29	*0.23*	0.93	*0.48*
N	466,287	*107,549*	466,258	*95,396*	466,284	*107,549*	466,255	*95,396*

Note: Year dummies and US state dummies are also included. Robust standard errors in parenthesis * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table C1 – Summary Statistics for Lending Club

Panel (a)

Year of the loan	2007	2008	2009	2010	2011	2012	2013	2014
Borrower lending rate	0.118	0.121	0.124	0.120	0.122	0.136	0.145	0.138
	(0.027)	(0.025)	(0.027)	(0.035)	(0.041)	(0.044)	(0.044)	(0.043)
Size of loans	8224	8829	9834	10528	12048	13462	14708	14870
	(6115)	(5753)	(5994)	(6599)	(8169)	(8087)	(8099)	(8438)
Term (months)	36	36	36	42	44	40	42	43
Loans for debt consolidation (%)	35	41	42	46	49	58	60	61
credit card payment (%)	14	17	12	12	13	19	24	24
home improvement (%)	6	6	7	8	8	5	5	6
business (%)	10	5	7	4	4	3	1	1
other (%)	36	31	32	31	26	15	9	9
# observations	600	2,393	5,281	12,537	21,721	53,367	134,756	235,629

Notes: Standard deviations in parentheses.

Panel (b) – Loan status

Year of the loan	2007	2008	2009	2010	2011	2012	2013	2014	Average
Completed	34.33%	54.95%	78.05%	80.15%	80.31%	78.80%	57.93%	31.36%	61.99
Current	0	0	0	0	4.42%	5.19%	27.59%	56.49%	11.71
Past Due (1-120 days)	58.17%	34.73%	10.70%	7.99%	0.23%	0.40%	1.66%	3.09%	26.29
Chargedoff or defaulted	7.50%	10.32%	11.25%	11.86%	15.03%	15.62%	12.81%	9.05%	0.01
	100%	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*

Panel (c) - LC loan grade

Year of the loan	2007	2008	2009	2010	2011	2012	2013	2014
A (%)	13	13	23	23	26	20	13	15
B (%)	16	25	27	29	30	35	33	26
C (%)	24	24	26	22	18	22	28	28
D (%)	16	18	15	15	13	14	15	18
E (%)	17	12	6	8	8	6	7	9
>E (%)	14	8	3	4	4	3	4	3
	100	100	100	100	100	100	100	100

Notes: A-grade borrowers are the safest.

Table C2 – Information about borrowers on Lending Club

Year of the loan	2007	2008	2009	2010	2011	2012	2013	2014
Median LC credit grade	C	C	B	B	B	B	B	B
Number of open credit lines	9	10	9	9	9	11	11	12
	(5)	(5)	(4)	(4)	(4)	(4)	(5)	(5)
Number of credit inquiries	3	2	1	1	1	1	1	1
	(4)	(3)	(2)	(2)	(1)	(1)	(1)	(1)
Borrowers w/ current delinquencies (%)	17	15	10	11	11	14	17	21
Debt-income ratio	10.7	13.2	12.5	13.1	13.8	16.7	17.2	18.0
	(7.3)	(7.4)	(6.6)	(6.6)	(6.7)	(7.6)	(7.6)	(8.0)
Annual income	64390	65196	69241	69511	69456	69720	73236	74854
	(63812)	(61410)	(62092)	(86439)	(47597)	(58655)	(48828)	(55548)
# observations	600	2,393	5,281	12,537	21,721	53,367	134,756	235,629