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

Ester Faia
Goethe University Frankfurt and CEPR

Monica Paiella
University of Naples Parthenope

This draft: January 2018

Abstract

Despite the lack of delegated monitor and of collateral guarantees P2P lending platforms exhibit relatively low loan and delinquency rates. The adverse selection is indeed mitigated by a new screening technology (information processing through machine learning) that provides costless public signals. Using data from Prosper and Lending Club we show that loans' spreads, proxing asymmetric information, decline with credit scores or *hard information* indicators and with indications from "group ties" *(soft information* from social networks). Also an increase in the risk of bank fragility in the traditional banking sector, controlling for other aggregate factors, increases participation in the P2P markets and reduces their rates (*substitution effect).* We rationalize this evidence with a dynamic general equilibrium model where lenders and borrowers with heterogenous projects choose between traditional bank services, subject to the risk of bank fragility, and P2P markets, which clear at a pooling price due to asymmetric information, but where public signals facilitate screening. An increase in the precision of the signals raises the threshold of unfunded projects, increases entropy and reduces the *Theil-Shannon*-metric for information frictions.

JEL codes: G11, G23.

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

*We thank Inaki Aldasoro, Tobias Berg (dizcussant), Leonardo Gambacorta, Michalis Haliassos, Roman Inderst, Emanuel Moench, Daniel Paravisini (discussant), Silvio Petriconi, Luigi Pistaferri and Victor Rios-Rull for useful discussions and comments. We thank participants at conferences and seminars. We gratefully acknowledge financial support from the Institut Louis Bachelier and from the DFG grant FA 1022/1-2. 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. Lenders can offer a loan with an interest rate based upon borrower information. In less than a decade, this new method of debt finance has grown significantly with a global proliferation of online marketplaces for personal loans, student loans, small business loans, real estate and home improvement loans, auto loans and even loans for cosmetic procedures[1]. Most strikingly, data show that this nascent industry is performing relatively well compared to the traditional banking sector. Despite the absence of delegated monitor and collateral guarantees 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 rather than a new service. Machine learning algorithms process and update information to provide *hard information* signals (credit scores like FICO and similar) at low or zero cost. Moreover, the platform allows investors to gather *soft information* signals such as recommendations or investment choices from "groups of peers".

The focus of our paper is precisely the assessment of the impact of information externalities on loans' spreads. The latter capture both the default risk of the projects as well as information premia, which arise due to asymmetric information. A main result is that signals, of both hard and soft information, mitigate the information premia. Besides this, we also assess the potential substitutability between digital platforms and traditional banking. Much of the increase in participation, hence liquidity, in the platforms appears to be due to the erosion of trust in the traditional banking sector. We examine these issues empirically with data from Prosper and Lending Club[2].

[1] According to research firm Liberum [31], 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.

[2] The first platform has the advantage of providing also soft signals such as recommendations and decisions from groups of friends. The second platform is larger in terms of traded volumes.

These datasets offer a unique opportunity since all information publicly available to investors is also available to the econometrician, thereby eliminating the biases from unobserved heterogeneity. We rationalize the empirical results with a general equilibrium model with information externalities.

In what follows we start by presenting a general equilibrium dynamic model with borrowers and lenders who can turn either to the traditional banking sector or to the P2P platform. Households/lenders in our model solve a portfolio problem and choose between those two forms of investment, demand deposits or equivalently bank short-term liabilities or P2P lending. Borrowers, who wish to fund risky investments, can equally seek funds from banks or on the platform. The general equilibrium perspective allows us to work on the link between information and market equilibrium prices. The dynamic (infinite-horizon) optimization allows us to account also for standard determinants of saving/investment behavior (precautionary saving, inter-temporal substitution, risk insurance). Against this background, the P2P market and the banking sector exhibit peculiar features which determine their relative liquidity (investors' and borrowers' participation) and premia. On the one hand, in P2P lending markets asymmetric information prevents investors from discerning the exact quality of the projects. Loan spreads indeed include information premia. However, costless public signals are available. If signal precision were zero, the market would clear at a pooling price, given by the unconditional mean of projects' success. As signal precision increases the price converges toward the separating full information equilibria. In between a premium for borrowers' adverse selection arises (see Stiglitz and Weiss [46]). On the other hand, the traditional banking sector exhibits fragility. Several events (such as bank failures, runs on deposits, runs on repos, or interbank freezes) might impair banks' liquidity, force early liquidation of projects (in contrast, projects are always brought to maturity on the platform) and haircuts on holders' of short-term liabilities. The relative participation in the two sectors (hence the price/premia at which they clear) depends upon the balance of risks and frictions characterizing the two. We highlight three main results. First, an increase in borrowers' quality (a fall in average default risk) reduces loans' premia on the platform for any degree of information (*selection channel*). Second, by increasing transparency, an increase in public signal precision reduces information premia (and, hence, loan spreads) when the quality of P2P borrowers is higher than average quality of all borrowers (*information channel*). This means that signals are informative when in the platform there is a good selection of borrowers. Third, an increase in the risk of liquidity shocks in the banking sector increases participation in P2P markets and reduces their loan spreads. This is so because both investors, who fear deposit haircuts, and borrowers, who fear early liquidation or a credit crunch, shift to the platform. This in turn increases liquidity and reduces P2P loans' spreads (*substitution channel*). We complement the analytical results with simulations that speak also about the quantitative relevance of the above channels.

Next, we present results from an empirical analysis linking P2P loan spreads to information signals and to the risk of liquidity shortage in the banking sector. We leverage on two novel datasets from Prosper and Lending Club, the two biggest lending platforms worldwide. Prosper has the advantage of providing *soft* information indicators (in addition to *hard* information ones), which allows us to test the role of *social multipliers*. Lending Club is the largest platform in terms of traded volumes and is particularly suitable to perform robustness checks.

In our regressions we include *hard information* signals, such as FICO scores and other creditworthiness measures, *soft information* signals, such as recommendations and investment from "groups of friends", and an indicator of liquidity risk in the traditional banking sector, namely the ratio of currency in the hand of the public to demand deposits which has been shown to rise at banking panic dates (see Gorton [19]). Results provide clear support to the hypothesis that information signals affect loan spreads through both the selection and the information channel. First, a one standard deviation increase in the FICO credit score, which implies an improvement in borrower quality, reduces the lending rate by 4 percentage points. This reduction can occur either because borrowers' quality has improved (higher FICO proxy lower risk) or because more information signals are available, hence investors can better screen borrowers. Second, by exploiting the variability in the signal reporting we are also able to identify the relative size of the selection versus the information channel and to establish that the latter is quantitatively important. Overall, all hard and soft information signals reduce loan spreads also through a reduction of the information premium. Third, we highlight the role of soft indicators. Group membership and recommendations give clear indications of the importance of *social multipliers*. Group membership lowers the loan price by between half and one and a half percentage points. Funding from friends lowers the rate by 2 to 4 points. The fall in loan spreads suggests that *social multipliers* can provide positive externalities here, since the reduction in loan spreads is most likely the result of a reduction in information spreads[3]. This confirms the hypothesis examined in previous studies of the superiority of (transparent) markets over financial intermediaries due to value added by diversity of opinions (see Allen and Gale [2]). Importantly, the overall observed reduction in loan spreads helps to quantify the value that lenders attach to information. Previous literature on the role of information and of social multipliers in determining investment choices usually speaks about the extent to which these factors determine participation (the extensive margin)[4], but are generally silent on the value of information. Last, our empirical analysis shows that an increase in fragility[5] in the banking sector lowers P2P loans' spreads. This is due to an increase in platform participation, hence in its liquidity. This result speaks about a *substitution channel* between the two sectors[6].

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 the model and its results (section 3). In section 4 we review our empirical analysis. Finally, we discuss extensions and policy implications in our conclusions, in 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 need of intermediation by traditional financial institutions. In the aftermath of the crisis, the fragility of the banking system as well as the distrust of investors towards it have been one of the main reasons for the growing popularity of the P2P lending. The decision processes involved in loan origination are given into the hand of private lenders and borrowers, and websites like

[3] Previous studies discussed information externalities linked to the *wisdom of the crowd.* Welch [50], Banerjee [5], Bikhchandani, Hirshleifer and Welch [8] stress the value of information obtained by observing other investors' actions.

[4] Van Nieuwerburgh and Veldkamp [49] show that information acquisition tilts investment toward the asset investors know about. Georgarakos, Haliassos and Pasini [18] demonstrate the role of social circles in debt acquisition. Some theoretical papers show the existence of an information risk premium: see among others Easley and O'Hara[12] or Hughes, Liu and Liu[24].

[5] To measure bank fragility we use alternatively two indicators. The first is the average currency to deposit ratio in the year before, which in the spirit of Gorton[19] is a proxy of bank runs. the second is an indicator of bank failures as reported by the Federal Deposit Insurance Corporation. The second has the advantage of showing much larger state variation. Note that in this regression we control for time and state dummies to control for aggregate factors.

[6] We do not exclude that complementarity between the two sectors might materialize in the future when the technology in the traditional banking sector catches up with the one available in the platforms. However, the data so far seem to support the substitution hypothesis.

Prosper.com offer them a platform to engage with each other. For borrowers, online P2P lending is a way to obtain a loan without a financial institution involved in the decision process. This implies lower costs, no need of collateral guarantees and no risk of early liquidation due to banks' liquidity shortages. However, borrowers have to pay higher returns on the loans when seeking funds on the platform. For lenders, returns are typically attractive compared to current returns on standard investment products offered by commercial banks. Higher platform returns however come together with more information risk as there is no delegated intermediary that screens and monitors the projects. Overall for lenders the fragility of the banking system (i.e. risk of haircuts on bank bonds) together with the possibility offered by the platform of obtaining signals on the quality of the loans are two of the main determinants of the decision to invest in P2P loans as opposed to more traditional liquid assets.

The analysis of P2P markets is relatively recent and is reviewed by Bachmann, Becker, Buerckner, Hilker, Kock, Lehmann, Tiburtius, and Funk [4]. Most of the literature so far has examined either the relationship between borrowers' attributes and listing outcomes or lenders' investment decisions. Pope and Syndor [40] and Ravina [41] look at discrimination in lending on Prosper platform. Duarte, Siegel and Young[11] study the role of appearance for trust in P2P lending relations. Paravisini, Ravina and Rappoport [37] use data from Lending Club to estimate risk aversion from lenders' investment decision on the platform.

A recent growing literature is assessing the extent of asymmetric information and the role of signals. Freedman and Jin [15] show that asymmetric information is widespread in P2P lending markets, but learning by doing for returning lenders partly mitigates the problem. Our paper does not examine the role of learning, but rather focuses on the role of public signals in moderating the information premium within a pooling equilibrium. Iyer, Khwaja, Luttmer and Shue [25] examine the role of interest rates as a signal of creditworthiness and find that the maximum interest rate that borrowers are willing to pay has a larger screening power than the credit score. Kawai, Onishi and Uetake [28] 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. The mechanism proposed in those last two papers are well in line with the type of auction trading and price posting mechanism which characterized the early stages of some platforms.

Currently, on most platforms prices are set through a centralized mechanism. In this setting, private signalling by borrowers is less likely to convey information. Indeed, whereby sending signals is rather costless, it is likely that bad borrowers imitate good borrowers. Our paper focuses on a different set of signals. We consider either public signals conveying hard information, processed through a centralized algorithm, or signals emerging from social interactions (indications from group ties). On the role of soft information there is the paper by Lin, Prabhala and Viswanathan [33] who use data from Prosper and show that "platform friendships" increase the probability of successful funding, lower interest rates on funded loans and are associated with lower ex post default rates.

None of the above papers considers the substitution between traditional banking sector and digital intermediation. The erosion of trust in the traditional banking sector however has played a crucial role in increasing platforms' liquidity. Our paper does examine this aspect.

P2P lending data effectively represent a public credit registry and for this reason they lend themselves to answer lateral research questions. For instance, Liskovich and Shaton[32] examine the role of price mechanisms on borrowers' participation. They focus on the impact on returns of the change in Prosper pricing mechanism from an auction-like system to a centralized price assignment. Bertsch, Hull and Zhang[6] assess the impact on loan demand and supply of the Fed liftoff undertaken in December 16, 2015.

Our paper is related also to the literature examining asymmetric information and the role of social ties in informal lending. Studies in past literature argued that informal lenders can mitigate asymmetric information by exploiting the role of social ties and social learning. See for instance La Ferrara [29], Udry [48], Hoff and Stiglitz [23], Besley and Coate [7]. Our paper contributes evidence in this direction.

Last but importantly, our paper contributes also to the literature comparing markets versus banks (see Allen and Gale[3]). In markets the relevant friction is not moral hazard, but information asymmetry. Investors do not monitor the actions of the borrowers checking the plans' implementation. On the contrary they shall screen projects ex ante based on the available information. A crucial innovation of the P2P markets is the availability of public signals that lessen the information asymmetry. In other words, those markets provide a publicly available credit registry with the full credit history. This is the novel feature that we uncover. Another noteworthy aspect is that P2P platforms are akin to private equity markets, so that ex ante the selection of borrowers includes young and risky entrepreneurs. Despite this, the screening technology and the network externalities facilitate the discovery of information.

3 The Model

Our goal is to characterize price formation and liquidity in P2P markets, per se and vis-a-vis alternative investment/borrowing opportunities. Given this goal and our focus on price formation we shall undertake a general equilibrium perspective to compare returns and liquidity in P2P markets versus banks. Our model is populated by risk-averse households/lenders who wish to invest their savings and risk-neutral borrowers/entrepreneurs who seek funds for risky investment projects. Both can turn either to the traditional intermediation sector or to the P2P lending market. Households/lenders solve an optimal portfolio problem and choose between deposits and P2P lending over an infinite horizon. Risk-aversion and the dynamic portfolio choice setting allow us to retain roles for precautionary saving, inter-temporal substitution and preference for risk-sharing into the investment decisions, hence into the pricing of assets. Indeed, lenders' stochastic discount factor, namely their price of risk, affects the returns of both assets. Participation in both sectors entails risks and benefits for both borrowers and lenders. The relative balance between these risks and benefits determines optimal portfolio shares, which in turn affect sectorial returns.

In our model, banks possess a costly screening technology. Demand depositors or holders of banks' short-term liabilities receive an ex ante flat contractual return and can therefore save the time and costs of monitoring projects. Last, we introduce some fragility in the traditional banking sector, which also serves the purpose of providing a trade-off with respect to the advantage of saving the screening costs. We introduce fragility by assuming that returns from bank short-term liabilities are subject to some liquidation risk, which might come either from bank runs (liquidity shortage) or from direct bank failures (solvency risk). If fragility materialize borrowers face the risk of early project liquidation and returns to depositors or to holders of short-term liabilities are subject to a haircut. As for the digital platform, the absence of a delegated monitor and of a collateral guarantee implies that prices embed information premia. Borrowers indeed have both performing and non-performing loans and they are indistinguishable to lenders. However, the platform provides costless public signals. Loans' spreads declines when the pool of borrowers improves and, under certain conditions, also when signals' precision increases.

3.1 Households/Lenders

There is a continuum of households/lenders who consume and save. They can invest either in bank short-term liabilities or directly on a lending platform. They take consumption/saving decisions and portfolio decisions to maximize 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 given the information set at time zero, and C_t denotes household consumption. Let X_t denote loans on a lending platform, and D_t denote banks' demand deposits or banks' short-term liabilities. We will use the generic term deposits since now on, but we refer to bank short-term liabilities which are callable on demand. For this reason they are subject to liquidity risk or to the risk of haircuts. At time t, in addition to choosing their consumption, households make a portfolio decision and invest a fraction α_t of their wealth W_t in the loan portfolio, X_t, at a price r_t, and a fraction $(1 - \alpha_t)$ into bank deposits. In period t, households receive an exogenous income, Y_t. They also receive proceeds X_{t-1} from previous period investment in the loan portfolio. The loan portfolio can consist either of a single borrower project (or a share), or of a certain proportion of several borrowers' projects. In both cases the return from holding this asset is uncertain and in both cases households cannot fully diversify asset risk. X_t is a non contingent security and its price, r_t, (and hence its expected return) is determined in an environment with asymmetric information and in presence of public signals as we detail below. The expected return of the P2P portfolio, r_t, will be derived later and is linked to the conditional expectation of the pool of projects that the investor is willing to fund. Bank short-term liabilities are subject to the risk of haircust in the event of bank fragility. To model this, we assume that households receive a previously contractible gross return, $\bar{\theta}_{t-1} R^d_{t-1}$, from previous period investment in bank deposits, D_{t-1}. The variable $\bar{\theta}_t$ is meant to capture the haircut tied to losses from defaults or liquidity shortage. We define as ζ_t the probability of bank distress or banks' liquidity shortage[7]. When this

[7]We use in the text those two terms interchangeably as what matter is the general concept of banks' fragility

happens loans are liquidated early at a discount, θ. The average discount, $\bar{\theta}_t = \theta\ \zeta_t + (1 - \zeta_t)$, is transferred onto depositors as an haircut (it is assumed that banks are funded only with deposits, we will return on this point later on). It is important to stress that we are referring to *demand deposits* or banks' short term funding. They may include standard deposits with a cap on insurance or other forms of banks' bonds whose returns are subject to the risk of haircuts in the event of liquidity shortage.

Household inter-temporal budget and wealth constraints read as follows:

$$C_t + r_t X_t + D_t \leq Y_t + X_{t-1} + \bar{\theta}_{t-1} R^d_{t-1} D_{t-1} \tag{2}$$

$$X_t = \alpha_t W_t;\ D_t = (1 - \alpha_t) W_t \quad \forall t \tag{3}$$

Households' optimization delivers the following first order conditions:

$$u'(C_t) r_t = \beta \mathbb{E}_t \left\{ u'(C_{t+1}) \right\} \tag{4}$$

$$u'(C_t) = \beta \bar{\theta}_t R^d_t \mathbb{E}_t \left\{ u'(C_{t+1}) \right\} \tag{5}$$

$$u'(C_t) \left[1 - r_t \right] = \beta \left[\bar{\theta}_t R^d_t - 1 \right] \mathbb{E}_t \left\{ u'(C_{t+1}) \right\} \tag{6}$$

Equation (4) is the first order condition with respect to X_t. Equation (5) is the first order condition with respect to D_t. Equation (6) is the condition for optimal portfolio choice. We derive the arbitrage condition between deposits and P2P investment using equations (4) and (5), which also deliver a link between loans' price r_t and households' pricing of risk, namely the stochastic discount factor:

$$r_t = \beta \mathbb{E}_t \left\{ \frac{u'(C_{t+1})}{u'(C_t)} \right\} = \frac{1}{\bar{\theta}_t R^d_t} \tag{7}$$

Given Harrison and Kreps [21] theorem, equation (7) provides the stochastic discount factor for pricing any asset in the economy. Specifically equation (7) implies that households should receive the same return in *expectations*, whether investing in the traditional banking sector or on the digital platform. The arbitrage condition determines how investment choices change, at the margin, in response to changes in relative risk-return profiles of different investment opportunities.

which might induce losses on depositors' returns. In the empirical analysis we will test with the different metrics for banks' fragility, one tied more to bank runs and one tied more to bank failures.

An increase in liquidation risk in the banking sector shifts investors toward the platform. Notice that the arbitrage condition alone is not sufficient to pin down portfolio shares. The latter are determined by the optimal portfolio allocation condition, (6), according to which the optimal share α_t is obtained at the margin from equating the gross return on demand deposits (weighted by the stochastic discount factor) to that on P2P loans. Like returns in equilibrium, also the optimal portfolio share changes with changes in the risk of liquidity shocks in the traditional banking sector relative to the information risk in the platform.

Before closing this section it is important to notice that the assumption of risk-averse lenders serves the purpose of retaining a role for inter-temporal saving and risk-substitution. The lender stochastic discount factor, which captures precautionary saving motives and inter-temporal substitution, plays a role in the pricing and supply of funds in the platform. Also the desire to limit risk and optimize risk-management justifies the migration across the two sectors.

Last, it shall be noticed that in the above optimization problem we assumed that demand deposit returns are non state contingent. This assumption is sufficient for the purpose of our model and fits well the case of standard deposits. A more general specification would accommodate state contingent returns on demand deposits or bank bonds. In this case also the covariance between the returns on deposits and the price of loans would matter for the comparison of risk-returns profiles between the two sectors. We consider this extension in appendix A.

3.2 Borrowers

In every period there is a continuum of 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. They can seek funds either from banks, for a fraction $(1-\gamma_t)$, or in the P2P lending market, for a fraction γ_t. Borrowers are risk-neutral and finitely lived. Let ξ, denote borrowers' survival probability. By multiplying ξ by the discount factor, β, we obtain borrowers' gross discount factor which is higher than lenders' (notice that ξ can also be interpreted as an exit rate from business). The latter assumption prevents borrowers from saving enough so as to ease up the need of external funding. The assumption of risk-neutrality captures borrowers' higher preference for risk (relatively to lenders). Furthermore we assume limited liability so that all contractible payments vis-a-vis any intermediation sector are bounded below by zero.

Projects deliver uncertain returns at the end of period t. Success probabilities are known to borrowers, but not to banks or to lenders in the P2P market. Banks can screen projects and learn success probabilities by paying a cost. Lenders in the P2P market observe projects' characteristics and receive public signals. To the extent that partial information arise a pooling price will emerge and there will be adverse selection of borrowers. As precision raises, returns approach the full information equilibrium. P2P investors will in general require a premium for lending which depends upon the selection of borrowers and the precision of the signals.

In each period t the project of each borrower can either succeed and deliver a return R_t^I, with probability p^i, or fail and return zero, with probability $1 - p^i$[8]. The assumption of losses bounded by zero, namely the limited liability, ensures that borrowers have risk-shifting incentives.

Like in the standard set-up of hidden information a' la Stiglitz and Weiss [46], we assume that borrowers have different success probability, p^i. The quality of projects is heterogeneous meaning that some projects are more likely to succeed than others. We assume that 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 the individual success probability is known only to the borrower.

In Appendix B we lay down borrowers' dynamic optimization problem and derive the optimality conditions that determine their participation in the platform and the optimal shares, γ_t.

Notice that we exclude the possibility that borrowers may issue corporate bonds in the event of liquidity strains in the banking sector. This restriction is realistic. Indeed, borrowers choosing the lending platform are typically small or risky firms with low collateral and often 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 more akin to private equities.

[8] For ease of notation we are omitting the time index, albeit all relations will hold for every period t. Also, the model can be easily extended to the case of above zero returns in case of default.

3.3 Pricing in P2P Markets

Pricing in the peer-to-peer market reflects the presence of asymmetric information. Since agents are not fully able to discern projects' quality a pooling price emerge for each project. 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 soft information, such as recommendations and investment decisions done by groups of peers. Both type of signals are costless, public and visible to all lenders. Hence, for the purpose of our model, we treat them equally[9]. Notice that we focus on public signals, as opposed to privately sent ones, for two reasons. First, in the current institutional design hard information signals are collected and processed in a centralized way through a machine learning algorithm. They are therefore free from potential conflict of interested parties. Second, private signals (those sent by the borrowers themselves) are unlikely to have an informative content. Since signals are costless bad borrowers can imitate good borrowers; hence projects remain indistinguishable.

Signals are a random variable, σ_i, whose realization is s_i. With probability λ, $s_i = p^i$, i.e. it is identical to the project's true success probability, while with probability $(1-\lambda)$, it yields a totally uninformative value that s_i is randomly drawn from p^i prior distribution, i.e. it is distributed according to the uniform density $\phi_s = \mathbb{U}\left[\bar{p} - \frac{\varepsilon}{2}, \bar{p} + \frac{\varepsilon}{2}\right]$. Signals' precision is captured by the probability λ. Note that our investors, contrary to banks, are unsophisticated. Thereby they do not invest in screening or monitoring to endogenously increase signal precision. The latter is exogenously given and depends on the amount and quality of information linked to the credit history of borrowers as stored through the machine learning algorithm. One can therefore think of λ as a parameter at any given point in time or as a random process that evolves over time (this will be our assumption in the numerical section below).

Signals' distribution can be summarized as follows (we follow Ruckes [43] and Petriconi [39]):

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

[9] 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.

Notice that 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 can be computed from the density of p^i conditional upon the signal. Using the Bayes rule and for any signal realization of the signal, s, we can compute the density of a given project quality, $p^i = p$:

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

Importantly note that there is no sequential sampling among investors and across periods. Indeed in each period t all investors observe the signals at the same time and form beliefs based on the equally and publicly visible and known distribution. Second, there is no sequential sampling across periods, since in every period a new round of projects enter the platform and investors start a fresh a new sampling.

Given the signals' structure in (8), we can 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 $\int_{-\infty}^{\infty} \delta(x) = 1$ 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 $Pr\left[p^i = p \mid \sigma_{i,\lambda} = s\right] = (1-\lambda)\frac{1}{\varepsilon} + \lambda\delta(x)$ and noting that the both the $Pr\left[p^i = p\right]$ and $Pr\left[\sigma_{i,\lambda} = s\right]$ are both uniform, we can compute the conditional expectation of the projects' quality. Hence Bayesian updating of beliefs given the observation of a signal realization of s_i results in the following posterior expectation:

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

Investors will price each project based on the conditional expectation in 10, so that the expected return from each project is given by $\mathbb{E}_t\left[p^i \mid \sigma_{i,\lambda} = s_i\right] R_t^I$. We shall now determine the mass and type of projects that will be funded by the investor. Based on the optimality condition in (5), investors will fund projects that provide conditional expected returns higher or equal to the the return received on deposits. The optimality condition (5) therefore defines the threshold or the marginal project that investors are willing to fund for given signal precision, λ:

$$\mathbb{E}_t \left[p^i = \hat{p} \mid \sigma_{i,\lambda} = s_i \right] R_t^I = \mathbb{E}_t \left\{ \frac{1}{\beta} \frac{u'(C_t)}{u'(C_{t+1})} \right\} = \bar{\theta}_t R_t^d \tag{11}$$

Investors will fund all projects whose conditional expected probability of success is compatible with a return higher or equal to $\bar{\theta}_t R_t^d$. We can identify this project with the the variable, $\hat{p}$. After substituting the average haircut, $\bar{\theta}_t$, we can re-state the above expression for given signal realization s_i as follows:

$$\varpi = \mathbb{E}_t \left[p^i = \hat{p} \mid \sigma_{i,\lambda} = s_i \right] = \frac{\left[\theta \zeta_t + (1 - \zeta_t) \right] R_t^d}{R_t^I} \tag{12}$$

The above condition provides the cut-off for investors' participation in the platform and will be used later on to assess the channels of liquidity in P2P markets. The optimal portfolio of projects funded by the investors contains all projects with expected success probability above the cut-off, ϖ. All projects with expected success probabilities below that threshold will not be funded. In this respect condition (12) also provides a measure of the extent of adverse selection in a market with partial information. Other noteworthy considerations emerge from (12). First, note that, since by (5) $R_t^d \left[\theta \zeta_t + (1 - \zeta_t) \right] = \mathbb{E}_t \left\{ \frac{1}{\beta} \frac{u'(C_t)}{u'(C_{t+1})} \right\}$, condition (12) also captures how the stochastic discount factor affects lenders' supply of funds in the platform. If the stochastic discount raises, namely the price that lenders attach to risk in future contingencies, raises, lenders require higher returns. This implies either that, for given signals' precision, platform returns, R_t^I, shall raise or that the success probability of the marginal borrower shall shift to the right, hence projects' quality shall improve. Second, condition (12) makes explicit that participation in the platform (relatively to the banking sector) depends upon the balance of risks between the two sectors, namely intermediation liquidity risk, ζ_t, and P2P information risk, as captured by λ.

3.4 Loan Premia and The Value of Information

The model features a distribution of lending rates, hence of loan premia. The latter capture two things, the default probabilities of the projects and the information premia. If the investor could perfectly discern the projects for each of them he would accept a loan premia which is given by the exact default rates, namely $1 - p^i$. Under partial information projects are indistinguishable and adverse selection emerges, hence lemons are part of the pool. Under those circumstance the investor requires a premium which is proportional to the value of information. We shall therefore construct a metric for the value of information or the information premium. This will be given by the distance between the probability that a project will not be funded under full information (namely with signal precision $\lambda = 1$) and the one under partial information (with given signal λ). The probability that a project will not be funded corresponds to the probability that its expected probability of success, conditional to the signal, is below the threshold. The latter reads as follows:

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

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 the $\chi_\lambda(\varpi)$ as follows:

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

Given $\chi_\lambda(\varpi)$ the mass of funded projects is $\psi_\lambda(\varpi) = 1 - \chi_\lambda(\varpi)$.

From expression (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, $\chi_\lambda(\varpi)$ if and only if $\varpi \geq \bar{p}$. In other words if the average quality under imperfect information, ϖ, is above the average quality under no information, $\bar{p}$, there is a positive selection of projects. In this case acquiring more information is valuable since it helps to reduce the averse selection and increases investors' willingness to fund projects.

Using $\chi_\lambda(\varpi)$ we can define a metric for the value of information. Since the loan spread includes an information premium, it will be directly related to this metric. We can construct a metric for the value of information as the distance between the probability that a project will not be funded under partial information and the probability that it will not be funded under full information, namely when $\lambda \to 1$. This distance, which reads as $\Theta = \chi_\lambda(\varpi) - \chi_{\lambda=1}(\varpi)$, is akin to the Theil [47] index of information. Indeed $\chi_\lambda(\varpi)$ captures the amount of entropy under partial information and for given signals. As signal precision increases, the dispersion, hence the entropy, of the conditional success probabilities widens, and gets closer to the entropy under full information.

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

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

If the threshold ϖ is to the right of the unconditional average of 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 probability that a project will not be funded, both under full and partial information. In other words there is a positive selection effect. The reduction in $\chi_\lambda(\varpi)$ will be higher under partial information, namely with $0 \leq \lambda \leq 1$. 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 the value of information declines. Furthermore, an increase in precision reduces the distance Θ. This is intuitive. As more precise information is available the difference in the probability of being funded between the case with full and partial information, namely the information premium Θ, declines.

The effect of $\bar{p}$ on Θ captures a *selection* channel, while the effect of λ on Θ captures the *information* channel. Our empirical analysis below will show that public signals reduce the lending rates, hence the loan spreads. In some cases, like for instance for an improvement in the FICO score, the decline in the lending spread might be due to either the selection or the information channel or both. For this reason in section 4.2 we will deepen the analysis and identify the two channels separately by exploiting variability in information reporting across borrowers.

3.5 Banks

Banks hold a costly screening technology. They have to pay a cost μ to assess projects' quality. For simplicity we assume that once the cost is paid banks can learn projects' quality perfectly.

The model can be easily adapted by allowing partial learning on the side of the banks. Generally speaking however, given that their screening technology is less efficient (more costly) than the one available in the platform, it is necessary to give them an advantage in terms of learning abilities to avoid corner solution (zero participation in the banking sector)[10]. Banks are fully competitive, hence they shall just break even. They collect only demand deposits to fund loans, $D_t = L_t$. Hence all projects returns are rebated to investors in demand deposits[11].

We assume that if funded by the bank borrowers cannot extract any rent and must rebate all project returns to the bank. This is so since they are risk-neutral and have no outside option that delivers positive returns. On the platform borrowers rebate all returns to the lenders[12]. Aggregate projects' returns shall solve the banks' break even condition:

$$\bar{p}R_t^I - R_t^d - \mu \leq 0 \tag{15}$$

Note that banks realize returns only if projects are successful, but they have to pay depositors and the screening cost in any case.

As mentioned earlier, banks are subject to the fragility risk due to liquidity shortages or failures in every period t. In all these cases, the bank shall liquidate all funded projects[13] and projects' returns are discounted by a factor $\theta \in (0, 1]$. Given the probability of bank distress, ζ_t, the bank extracts an expected return $\bar{\theta}_t p^i R_t^I$, whereby as before $\bar{\theta}_t$, is equal to $\theta \in (0, 1)$ with probability ζ_t and it is equal to 1 with probability $(1 - \zeta_t)$. Such formulation for the discount factor captures unanticipated changes in the opportunity cost of carrying out outstanding loans. Absent insurance on banks' demand deposits, the expected loss given default is eventually transferred onto depositors, who receive $\bar{\theta}_t R_t^d$ in expectation.

3.6 Substitution, Selection and Information Channel

Below we summarize the determinants of liquidity, as proxied by participation, and loans' spreads on the platform. In particular, we shall assess how the balance of risks and benefits in the two

[10] In practice the participation to the banking sector might be encouraged by other complementary services provided by banks.

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

[12] This assumption can be easily relaxed by assuming that entrepreneur have a positive return from home production.

[13] This is of course conditional to the fact that the bank has exhausted all its loss absorption capacity.

sectors affect funding, investment decisions and returns. In this section we derive three testable implications of the model.

Remark 1 - Substitution channel. *An increase in the risk of a fragility shock in the banking sector raises participation in the platform, hence its liquidity, because it lowers expected defaults and loans' spreads.*

From equation (12) we can see that an increase in banks' liquidity risk, ζ_t, decreases the tolerance cut-off, ϖ, 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 will enlarge. In other words as the risk of a haircut on deposits increases, investors shift to the P2P market. This implies, as per equation (14), a reduction in the probability that a project will not be funded under partial information, namely the $\chi_\lambda(\varpi)$, as long as $\varpi \geq \bar{p}$. If there is a positive selection of projects (the marginal project funded by the investors, ϖ, is to the right of the population mean, $\overline{p}$) an increase in the size of the portfolio induces an increase in the mass of funded projects, $\psi_\lambda(\varpi) = 1 - \chi_\lambda(\varpi)$. Note that, under the assumption $\varpi \geq \bar{p}$, the decline in ϖ also implies a decline in Θ, for any value of λ. The reason for this is as follows. The increase of liquidity or the total supply of funds in the platform induces 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, hence in the lending spreads as proxied by Θ.

Remark 2 - Selection channel. *An increase in the average quality of projects,* $\bar{p}$, *increases platform liquidity, reduces expected information premium,* Θ, *hence loans' spreads to the extent that* $\varpi \geq \bar{p}$.

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

Remark 3 - Information channel. *An increase in signals' precision,* λ, *increases platform liquidity, reduces the information premium , hence the loans' spreads.*

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

3.7 Transmission Mechanism in the Model

So far we have discussed analytically the channels of transmission at work in the model. Now, we quantify the effects through simulations that solve the entire set of model equations simultaneously. We therefore 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.

Before turning to the description of the transmission mechanism a few more remarks are in order. The structure of our model is that of a portfolio allocation model under incomplete markets. It has been noticed in previous studies that an indeterminacy arises in this case in the solution of the non-stochastic steady state and in the first order approximation of the model (see Judd [26]). If there were enough financial assets (which means perfect risk-sharing) it would be possible to identify an equilibrium allocation independent of the financial structure and to derive the implied portfolio allocation. Perfect risk-sharing is of course a rather unrealistic feature, therefore all portfolio allocation models shall deal with such indeterminacy We do so by following the methodology in Roussanov [44], which consists in guessing a model-consistent functional form for the aggregate wealth accumulation, which allows to pin down exactly the portfolio shares devoted to each asset. This allows us to obtain a determinate non-stochastic steady state around which we can simulate the dynamic with a second order approximation.

3.7.1 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, $\frac{C^{1-\sigma}}{1-\sigma}$, with a risk-aversion parameter set to 2 within the range of most of the macro and household finance literature. Next, we need to calibrate the parameters governing the average discount factor for bank bonds. The steady state overall return on banks' bonds, R^d, is equal to $(\beta\bar{\theta})^{-1}$. For the average haircut, Gorton and Metrick [20] report an average value of 17% for the period 2007-2009 for the repo-haircut index:

repos are short-term debt, whose returns and haircuts approximate well bank bonds'. We therefore set $\theta = 0.85$ and $\zeta = 0.2$ so that the average haircut, $\bar{\theta}_t$ is 16% on annual basis. For robustness we run the simulations also for a lower $\theta = 0.65$, which yields a different $\bar{\theta}_t$. This implies that R^d is 4% on quarterly basis, a realistic value based on market fluctuations for subordinated bonds. Next, we set the quarterly average probability of project success, $\bar{p}$, to 97%. This corresponds to a quarterly default rate of 3%[14]. These data are well in line with data for charged-off and delinquency rates on banks' loans from the Fed Board[15] over the last ten years. On an annual basis these numbers yield a default rate of 11%[16] The return R^I is calibrated to deliver realistic returns on the platform and contemporaneously to satisfy the banks' break even condition. 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) upon which $s = \frac{R^d(\theta\zeta + (1-\zeta))}{\lambda R^I} - \frac{1-\lambda}{\lambda}\bar{p}$. The latter depends upon the value of λ. The benchmark value for the latter is set as follows. First, we set the value of R^I using the banks' break even condition, $R^I = \frac{R_t^d + \mu}{\bar{p}}$, and compatibly with a bank margin of 4% on annual basis as in Repullo and Suarez[42]. Note that monitoring costs are calibrated to 15 basis points on quarterly basis. This delivers a value of $R^I = 1.2$. Given this value 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. Specifically for Prosper mean estimated returns range between 10.3% in 2009 to a value of 7.3% in 2014. The equivalent in Lending Club are slightly lower.

3.7.2 Shocks Estimation

We introduce and estimate four benchmark shocks, all in auto-regressive form: an income shock, Y_t, a liquidity shock, ζ_t, a shock to bank returns, R^b, and a shock to the default probability, $(1-p)$. We simulate the model in response to all four shocks and verify that the simulated moments for the volumes and returns of the P2P lending match the equivalent in the data. Furthermore we present impulse responses of the model to the liquidity shock in the banking sector, ζ_t, and to an AR(1) random shock to signal precision, λ. The first impulse response serves the purpose of verifying numerically the existence of the substitution channel discussed above. The second impulse

[14] We checked whether this number is also in line with average default rates on the platform. On Prosper, the default rate of loans issued between 2009 and 2012 is just around 3%.

[15] See Federal Reserve Board data releases on Banks Assets and Liabilities.

[16] The default rate on annual basis is given by $\bar{p} * \bar{p} * \bar{p} * (1-\bar{p}) + \bar{p} * \bar{p} * (1-\bar{p}) + \bar{p} * (1-\bar{p}) + (1-\bar{p}) = 0.1147$.

response provides a numerical assessment of the information channel. Note that the shock to signal precision cannot be directly estimated in the data. Nonetheless its impulse responses provide a qualitative assessment of the information channel.

To calibrate the income shock to households/lenders we follow Curatola and Faia [10], who estimate an auto-regressive income process using PSID data from 1968 to date for families classified as lenders (as defined in Kaplan and Violante [34]). The process has a persistence of 0.6 and a standard deviation of the idiosyncratic component of 0.06. 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 and the standard deviation of the idiosyncratic component is 0.234802. To calibrate bank returns, R^b, we collect quarterly data on the Return on Average Equity for all U.S. Banks again for the period 2006 to 2016, and fit an auto-regressive process. This results in a persistence of 0.878877 and a standard deviation of the idiosyncratic component of 0.085312. To calibrate project success probability shocks, we take the loss rates on the P2P loans from Lending Club for the period 2006 to 2016 and again fit an auto-regressive process. The estimation delivers a persistence of 0.407039 and a standard deviation of the idiosyncratic component of 0.835269.

3.7.3 Quantitative Results

To assess the empirical validity of the model at first we verify whether simulations in response to all four estimated shocks deliver second moments for the prices and quantities of the P2P markets which are in line with the data. To perform the comparison we first compute the second moments of P2P traded volumes and returns in the data. Specifically we take quarterly data (in log terms) on P2P loans from the Prosper and Lending Club platforms for the period 2006Q1 to 2016Q1. We apply the Hodrick-Prescott filter (for the full sample of 10 years of data) and compute the trend of loan volumes and the business cycle component and its moments. The standard deviation of the business cycle component is 0.48 percent, while the persistence is 0.67. We then compute the same statistics using the model and obtain a standard deviation of 0.46 percent and a persistence of 0.51. The model statistics seem to match the data ones well. The difference in the persistence is well explained by the fact that our model includes a simple portfolio decision between bank deposits and

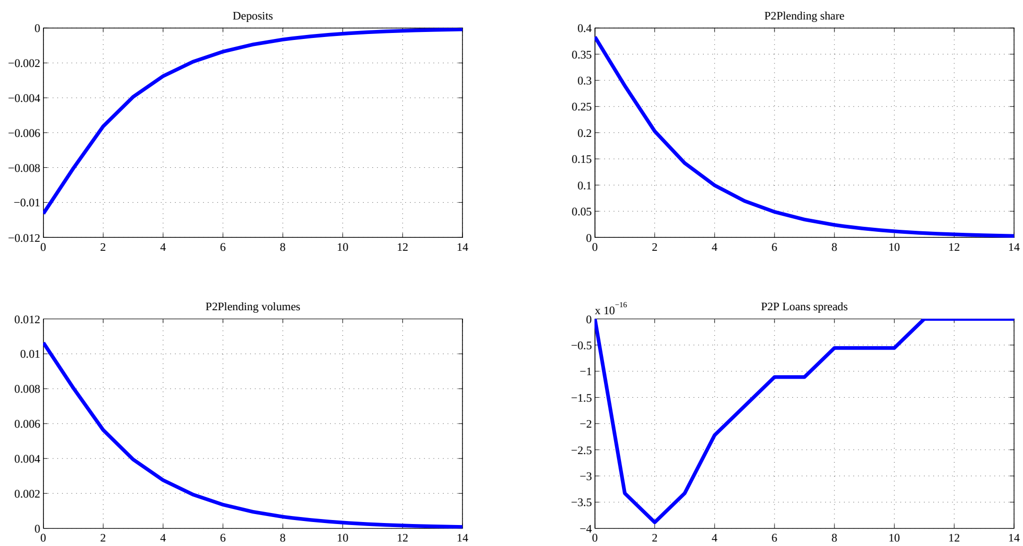

Figure 1: Impulse responses of selected variables to 1% liquidity shock, ζ_t.

P2P lending. Moreover beyond the information friction the market does not include other possible other liquidity frictions. Last, the standard deviation of P2P returns in the model is 7.3%. This is somewhat higher than standard deviation of Prosper mean estimated returns that ranges from a high of 5.2% in 2009.

Next we wish to assess whether some of the channels previously highlighted, more specifically the substitution and the information channels, can be detected even through numerical simulations of the full model equations. To do so we compute impulse responses of selected model variables to the liquidity shock, ζ_t, and to an information shock, modelled as an AR(1) process on λ.

Figure 1 shows the effect of a 1% liquidity shock.

An increase in the probability of a liquidity dry-out in the banking sector shifts households' portfolios out of deposits (first panel on the left) and toward investing in the platform. Both X_t (second panel on the left) and the corresponding wealth share, α_t (first panel on the right), increase. This shock indeed captures the extent of the *loans' substitution channel.* The price that households are willing to pay for platform loans, r_t, increases (not shown for brevity). We define the average loan spread 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[17]. The loan spread decreases (second panel on the right) as now investors require lower returns to participate. Movements in quantities are larger as this shock impacts mainly on the relative liquidity between the two sectors. The movements in the loan spread are a few percentage basis points. This is understandable since in the model ζ does not have a direct impact on P2P loan price, but only an indirect impact through the relative movements of quantities. This result is consistent with the one shown later on in the empirical analysis. The increase in liquidity risk in the banking sector does raise participation and lowers loans' spreads in the platform, but the second effect is rather small.

Figure 2 shows impulse responses of selected variables to a 1% increase in signal precision, λ. This shock captures the *information channel*. For this shock we plot responses for the two values of θ, which allows us to isolate and discuss the impact of the relative risk across the two sectors. Higher liquidity discounts makes the banking sector relatively less attractive. An increase in λimplies that investors are now better at discerning the quality of the loans. Since this reduces adverse selection, investors are willing to pay higher prices. Correspondingly the loan spread declines. Investors require lower premia to compensate for the uncertainty in P2P markets. As a result the lenders' portfolio share invested in the platform raises and deposits fall. Households substitute between different investment opportunities. This last piece of the transmission mechanism highlights another dimension of the substitution channel. Notice that the movements in the shares are large, while the movements in the quantities (deposits) are just few percentage points. This is so since in this model wealth moves relatively little given that there are not other aggregate shocks affecting it.

4 P2P Loans Data - Prosper

We now turn to the data and analyze the determinants of equilibrium P2P loan rates and appraise the predictions of our model regarding the effects of signals and liquidity risk against the data. We start our empirical analysis using data from Prosper (http://prosper.com) which started P2P lending in the US, in February 2006. Currently, Prosper is the second largest platform in the world, after Lending Club[18]. In contrast to other platforms, including Lending Club, in addition to *hard*

[17] 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.

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

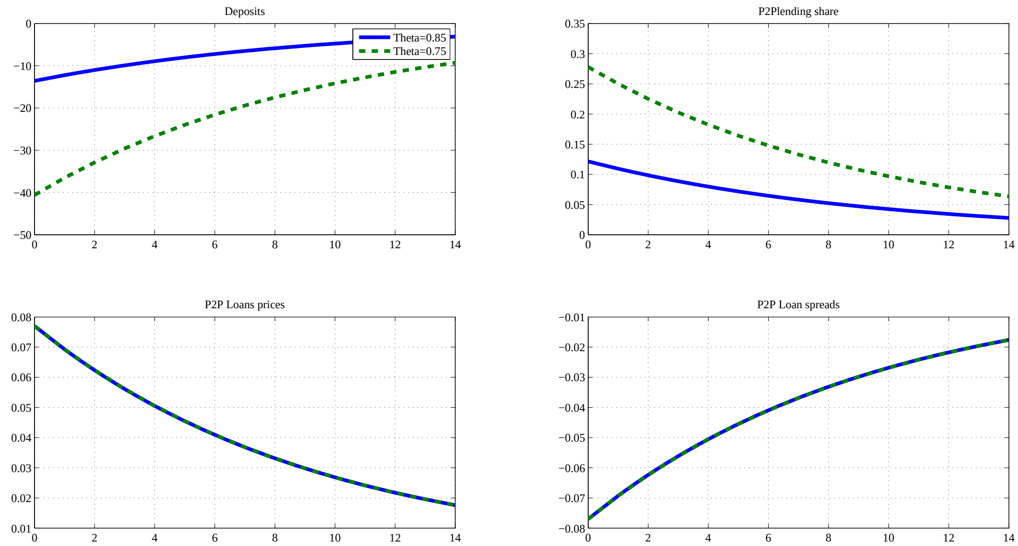

Figure 2: Impulse responses of selected variables to a 1% increase to λ and again for different values of θ.

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 use Prosper data as our benchmark, however in section 4.5 below we perform robustness checks using data from Lending Club as well.

A brief description of the institutional functioning of Prosper and of the type of data is useful. When signing up for a loan on Prosper, borrowers create personal profiles and solicit funding detailing the amount requested, the interest rate, the term and purpose of the loan. The borrower's profile includes independently verified information on his credit history, income and current debts, and a credit grade determined by Prosper based on a proprietary machine learning algorithm (from 2009 onwards). Prosper also creates social networks by linking borrowers in groups (tied by geography, some specific characteristics, common interest, or common loan purpose) and by collecting the endorsements of other Prosper borrowers (friends). In sum information about borrowers takes three forms: signals provided directly from the borrower, hard information signals provided through the machine learning algorithm and soft information signals related to any recommendation or investment by friends. As we shall see below the second and the third appear to be generally more informative that the first. As explained earlier since private signals are costless in this market, they are also not informative since any bad borrower can imitate a good one.

Given the information received, lenders assess and can bid on listings. The loan is funded only if the bids reach the amount sought by the borrower. The maximum length of the bidding period is 2 weeks. Notice that, until mid 2009, Prosper operated a variable rate model and rates were determined through an eBay-like auction. In 2009, Prosper registered with the Security Exchange Commission (SEC) and changed its business model to pre-set rates determined solely by Prosper itself (centralized system). In the regressions below we will always report results for the sample period pre and post 2009. More details on the data and on the Prosper platform are in Appendix C.

4.1 Some Facts

We start by examining simple statistics and data trends that give indications on facts relevant for our analysis.

Tables 1 and 2 report some summary statistics for our data which consist of all loans that have been funded over the period February 2006 to March 11th 2014. Notice that Prosper has grown very quickly over time and tripled its size between its incept and 2013. The first two rows of Table 1 report the average borrower's interest rate and the Annual Percentage Rate of Prosper loans by year. The growth of this sector became faster after the 2007-2009 financial crisis with the ensuing fragility of the banking sector. This is confirmed by inspecting the dynamic of the business cycle component of traded volumes. Figure 3 below shows the Hodrick-Prescott trend and the business cycle component of Prosper volumes. While the trend is steadily increasing, at business cycle frequency volumes increase in periods of fragility of the banking systems (particularly during the period 2007-2009), while they decline in periods of stability of the banking system. This behavior of the business cycle component is indicative of the substitution channel rationalized in our model.

As mentioned earlier in 2009 the pricing system became centralized. In our regressions below we will always repeat estimation before and after this date. One trend which emerges from the estimation is that various signals become more informative in the post 2009 centralized system.

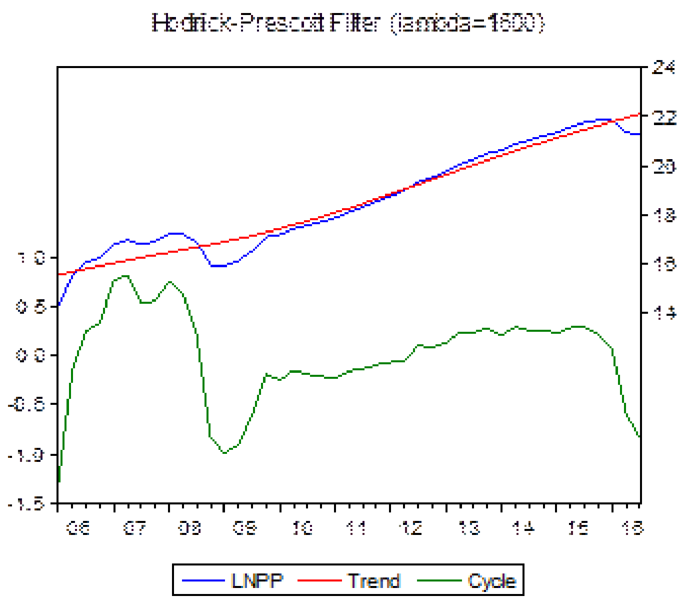

Figure 3: Hodrick_Prescott trend and business cycle component of the volumes traded in Prosper.

Prior to then most of the signals were provided by the borrowers themselves or by groups of their friends. Being costless and in absence of cross-checking with hard information provided independently by the platform, those signals had lower informative content. After 2009 Prosper provided, using the credit history stored in the machine learning algorithm, proprietary credit ratings for each prospective borrowers, ranging from 1 (high risk) to 7 (low risk). Those were based on the company's estimation of the borrower's loss rate (see description in Appendix C). Prosper uses these ratings to assign each listing also an estimated return at issuance corresponding to the difference between the loan yield and the estimated loss rate (reported in Table 2). The possibility of cross-checking the information provided by borrowers and by the network with independent hard information has increased the information value of all signals.

Table 2 provides statistics on returns, liquidity and risk. One noteworthy feature of the data in Table 2 is the sharp decline in lending rates and in borrowers' riskiness (proxied by credit rating measures) over the last years[19]. The decline in rate and risk is already indicative of the

[19] It is worth mentioning that Prosper lending rates tend to be higher than the rates of similar platforms, such as

discipline role of making the credit history public. The decline in the rates has been accompanied by an increase in liquidity. Over time, the average size and the average duration of the loans have increased. Maximum loan amount and length have indeed been raised[20], the time from posting to funding has decreased and the average investment by lenders has gone up. Furthermore, there is a general decline in the risk of loans. Panel (b) of Table 1 reports the loan status as of March 11th, 2014 for all loans that have been funded since Prosper onset, and it gives indications on the risks of the loans in terms of 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 after registration with the SEC[21]. For reference, in 2014, US banks charged off or reported as delinquent 16.6 percent of all consumer loans (18.5 percent in 2013) (Board of Governors of the Federal Reserve [14]). At last, there is a trend of increase in borrowers' quality. Table 2 also summarizes the signals that borrowers can issue to overcome asymmetric information. The information posted on the platform and visible to 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[22] and the number of current delinquencies. Based on these measures, over the sample period considered, the average "quality" of borrowers has improved. Consistently, Prosper in-house credit rating and estimated losses and returns at issuance suggest an improvement in borrowers' reliability[23].

To sum up loan rates and risk have declined and liquidity and borrowers' quality have increased. Overall those trends can be taken as indication that the screening technology employed by these platforms, which consists in the storage and transparent publication of large amount of information on borrowers, improves market efficiency and liquidity.

Table 2 summarizes also soft information-type of signals available to potential lenders. Those

---

Lending Club. However, it charges borrowers with lower fees.

[20] From 25,000 to 35,000, and from 36 to 60 months, respectively.

[21] As an example, 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.

[22] Credit inquiries are requests to check one's credit score. Specifically, when one applies for credit, including for an auto loan, mortgage or credit card potential lenders check their credit history. Large numbers of inquiries typically mean greater risk: 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)

[23] 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 has been stable around 25 percent.

include borrowers' participation in groups, recommendations and investment from Prosper friends and previous loans on Prosper. As explained in the Appendix, any borrower can join one of the several existing groups of Prosper borrowers or propose a new one. 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 in 2014. As to friendships, to create one, borrowers need to send an email to their friends, who must also be Prosper borrowers. Hence, individuals who are friends on Prosper must have at least some offline, non-public information about each other, such as the email address. Like for group participation, also the share of borrowers with recommendations from "Prosper friends" and with "Prosper friends" investing in one's loan fall steadily over the period. The significant drop in network relevance is most likely related to the tremendous growth that Prosper registered over time. Also, after the registration with the SEC, 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 the single 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. Finally, 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 estimations below we will include among the regressors all metrics related to soft information to test the importance of social network in the decision to invest in the platform and in this price formation process. While recommendations from networks have decreased over time, their indications have become more informative in the post-SEC system due to the possibility of cross-checking with hard information. On balance soft information retains a significant role.

At last, in Tables 3, 4 and 5 we divide borrowers in groups based on soft-type of information, such as being part of a group or not, having recommendations from other Prosper borrowers or not, and having borrowed previously on Prosper or not. Two interesting observations emerge. First, borrowers with friends (Table 4) pay lower rates[24]. This gives a clear indication that soft information does play a role for pricing. Second, borrowers belonging to a group (Table 3) appear to be riskier (higher FICO) before 2009. As explained earlier, the centralized pricing system (post 2009) provides hard information signals that anchor expectations of group members. While it is

[24] This evidence is consistent with that of Lin, Prabhala, and Viswanathan [33] who focus on the likelihood of being funded and find that friendships have a much larger impact than group membership.

likely that prior to that date groups' recommendations were based on a subjective and incorrect assessment of risk. Finally, based on Table 5, borrowers with prior Prosper loans appear to pay relatively lower rates in the most recent years covered by the sample despite their FICO and Prosper ratings are not significantly higher than those of first-time borrowers. It is important to note that borrowers with prior loans exhibit lower charged offs and defaults. This fact shows that the availability of a credit history, hence of a record of information publicly available, can play an important role in reducing the information premia, hence the lending rates. The longer the credit history, the stronger is the decline.

4.2 Lending Rates: *Hard* and *Soft* Information Signals

In this section we start by examining the selection and the information channel previously rationalized through our model. To this purpose we report regressions of the lending rate against all type of available signals (private signals, hard information and soft information) with the goal of testing the direction and the significance of their impact. Note that signals indicating a reduction in borrowers risk might induce a decline of the lending rate 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 identify the channels separately in the next section.

Tables 6 and 7 report the results of OLS regressions. In table 6, in the first column, we pool all years and include only loan characteristics (size, term, motive), which are set *ex ante* by borrowers (*private signals*), and dummies for quarter of listing and state of address of the borrower. These variables explain around 23 percent of the variability of the rates.[25] Given the change in the rate setting procedure in 2009, we then split the sample and distinguish between loans posted before and after SEC registration. Interestingly, based on regression adjusted R^2s, the explanatory power of the regressors is much larger after SEC registration, i.e. after Prosper starts setting the rates. In the early years, when rates were set through an auction-like system, most of rates heterogeneity is left

[25] The relationship between the lending rate and the 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 are charged rates that are higher by half and 1 percentage point respectively. Similar results obtain when replacing the left-hand-side variable with effective loan yield, or with Prosper rating, estimated return or estimated loss. Regressions available upon request.

unexplained by this model.[26] This confirms our previous argument. Private signals are generally uninformative since they are cheap and bad borrowers can easily imitate good borrowers. In the post 2009 system however investors can cross-check the information provided by the borrowers themselves against the credit scores provided by Prosper. The possibility of cross-checking with hard signals reduces the borrowers' incentives to mis-report and it increases the information value of private reporting.

In the regressions reported in Table 7, we add variables capturing the 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, Prosper rating, 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 and the analysis suggests that loan rates are decreasing in the FICO score and increasing in number of open credit lines and of credit inquiries. Being delinquent, having a low income, or a high debt to income ratio also increase the rate.[27] This evidence is consistent with the predictions of our model regarding the selection and the information channels. Note that an increase in the FICO score may indicate either that borrowers' quality (the $\bar{p}$ in our model) has improved or that more information is available (higher signal precision, λ, for given borrower's quality). Consistently with the model both decrease the loan rates. In section 4.2 we will separately identify the two channels to asses their relative importance.

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

[26] Results are robust to pooling the data and interacting the regressors with a dummy that takes on value 1 if the loan is funded after SEC registration, 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.

[27] In our sample the FICO score ranges between 500 and 900. A FICO below 600 signals bad credit history. A FICO above 750 signals an excellent history. A one standard deviation (70 point) increase in the FICO score lowers the lending rate by approximately 5 percentage points. As to the other variables, 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.

friends' investment, and for previous borrowing on the platform. In line with the implications of our model for 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 the loan price 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[28]. Once again the institutional change which took place in 2009 seems to increase the informative content of signals. First, the difference in the adjusted R^2 increases significantly in the 'post-SEC registration' years. Second, recommendations from friends not accompanied by actual investment decisions appear to matter only after 2009.[29] Before 2009 investors assigned informative value only to investment decisions that involved friends stakes, but not to cheap talking. After 2009 the possibility of cross-checking the friends recommendations (even in absence of actual decision), with the hard signals provided by the platform, increased the informative value of the network.

Note that the decline in the rates due to group membership persists also after 2009, despite the fact that in the new pricing system group membership becomes visible only after the posting of the loan (hence after the Prosper algorithm must decides on the price). This suggests that the Prosper centralized system is somehow capable of anticipating the information content of group membership. This seemingly puzzling feature can be explained by examining the regression in the last column which includes a dummy for previous Prosper loans. In this regression the dummy for group membership becomes negligible and statistically insignificant and the coefficients for the dummies on recommendations and/or on investment from friends become very small (some become outright insignificant). The dummy for previous borrowing on the platform has instead a large, significant and negative coefficient, suggesting that returning borrowers pay over 4 percentage points less than "new" borrowers, *ceteris paribus*. It is very plausible that a positive correlation exists between the number of Prosper friends and previous funding on the platform and this is part

[28] It is interesting to report that by running the same regressions by omitting credit risk measures, the coefficients of group membership become significant and positive. This suggests an omitted variable bias and is consistent with the hypothesis that group participation may facilitate the funding of risky borrowers, which generates a positive correlation between individual riskiness and group participation as Table 3 suggests.

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

of the information stored in the centralized system.

4.3 Lending Rates: Identifying Selection versus Information Channel

In the previous section we have established that signals of various type have a significant impact on lending rates. Generally speaking the arrival of new information indicating a decline in borrowers risk reduces lending rates. As explained earlier however this decline might be due to either a selection effect (average borrower quality improves and default rates decline) or to an information channel (for given borrowers' quality investors are better able to discern 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 in which we quantify the role of information and signal precision in reducing premia and lending rates. To construct a measure of the signal precision we exploit the variability in the type of *information reporting* among borrowers. Below we perform a series of exercises that follow this direction.

Consider first the case of reporting on borrowers' income. Borrowers with higher income levels pay lower rates (selection effect). 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*. However beyond variability in the income per se, there is also variability in the type of reporting. For instance, some borrowers provide official documentation to support their statement, but over 8 percent of the sample does not do it. Those differences can proxy 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 the interaction terms 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/her income pays over 1 percentage point less than the others. This gives clear indication that higher signal precision also contributes in a statistically significant way to reduce lending rates, hence to increase platform efficiency.

Next, we perform an exercise similar to the one above by using as signal borrowers' 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 is not necessarily associated with more or less risky borrowers[30], however the presence of this information in the reporting increases precision as it can help investors to discern borrowers' quality better (information channel). To verify this in the regression reported in column 2 of table 8 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 that the information channel is quantitatively very important. Noteworthy is that our results give a clear quantification of the value of information in terms of how much returns the investors is willing to give in exchange of an increase in its precision.

Next, in the regression in column 3, we restrict the sample to loans funded before Prosper registered with the SEC, when prices were set through an auction. About 30 percent of borrowers do not report the state of residency. 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 precision, raises the average lending rate by almost 2 percentage points. At last, in the last column of table 8, we consider only the loans funded after registration with the SEC. About 10 percent of borrowers do not provide information regarding the reasons for borrowing. 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 those who omit it.

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.

4.4 Lending Rates and Banking Fragility

Another important aspect rationalized in our model is the substitution channel between different forms of investment. Note that in principle markets and banks could be either be substitutes or complements. If banks eventually adopt a digital screening technology, competition between the two sectors might induce complementarity. Our results below show that the substitution hypothesis

[30] In fact, a borrower could have no credit lines because she never applied for one.

seems to prevail. This is also the reason for which in our model we favoured assumptions inducing substitution. The result is intuitive. Complementarity is indeed more likely to emerge over time when the banking sector engages more heavily into digital investment. The narrative so far instead described the migration of investors to the digital platform as induced by fears of fragility in the banking sector.

To investigate the above hypothesis we introduce among our regressors a proxy for the fragility of the banking sector. In Table 9, regressions include a proxy for bank liquidity risk. We follow Gorton [19] and model the latter as the ratio of currency in the hands of public to demand deposits. This variable reflects the idea that before the occurrence of banking panics depositors demand a large scale transformation of deposits into currency. Gorton [19] shows that historically the ratio of currency to demand deposits has indeed increased at panic dates.

In the first 3 columns of table 9 we add to our regressions in Table 6 the average ratio of currency to deposits in the year before loan listing and the change of the ratio at the listing date with respect to previous year average. In the last 3 columns of the table we include also the variables that capture the signals available to lenders.

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, where increases liquidity induces a decline in the rates. Similarly increases in the currency-deposits ratio, which tend to occur around banking panics, are associated with lower rates. Note that we have included in the regression both a time and state dummy, hence the proxy for bank fragility should not be capturing other aggregate shocks. The coefficients are significant not only statistically, but also economically. We can get a sense of their size by noting that, if the ratio increases by 20 percent, borrowing rates drop by 1 percentage points when we take the sample as a whole, and by 2 percentage points, when we restrict attention to the 2010-2014 period. Interestingly, if we replace these variables with a dummy that takes on value one when, in the quarter before the listing, a bank run occurred (column 5 of the table), we find that the dummy has a negative and statistically significant, albeit small coefficient.[31] The relative small size of the estimates reflects the fact that we are restricting

[31] The panics dates that we consider are: 1) August 2007, when the American firm Countrywide Financial suffered a bank run as a consequence of the subprime mortgage crisis; 2) March 2008, when a bank run began on the securities

the availability of alternative investments to the banking sector alone. In practice investors have richer portfolio choices and asset substitution might take place among different asset classes.

In Table 10 we provide additional evidence of substitution between lending platforms and traditional banks using information on bank failures as reported by the Federal Deposit Insurance Corporation. This additional proxy of banks' fragility is useful since it contains much more inter-state variation than the previous one. 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 (see table E in the appendix for the distribution). We use these data to augment our regressions with dummies that take on value 1 if, in the borrower's State, more than a financial institution failed. The dummies' coefficients are negative and generally significant which suggests that where significant rates of bank failures were recorded investors increasingly turned to the platform, increasing liquidity and hence inducing a decline in P2P loan interests.

Overall, this evidence is consistent with the 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, implying a reduction of the information premia.

4.5 Robustness - Lending Club

To fully assess the robustness of our empirical results we re-estimate the above relations using data from Lending Club, which is the world's largest P2P lending platform in the US. The loan application process and funding are very similar to Prosper. More specific institutional details and data description for this platform can be found in Appendix D.

Table 11 below reports the results for OLS regressions whose specification parallels the ones used for Prosper data. For ease of reading we report also the corresponding Prosper regressions

and banking firm Bear Stearns, which, although it was not an ordinary deposit-taking bank, had financed huge long-term investments by selling short-maturity bonds, making it vulnerable to panic on the part of its bondholders; 3) June 2008, when a warning was issued that U.S. mortgage lender IndyMac Bank might not be viable and a bank run began. In addition to Indy Mac Bank case, in September 2008, Washington Mutual, the largest savings and loan in the United States and the sixth-largest overall financial institution, was shut down due to a massive run. No large banking panic occurred in the period after Prosper registration with the SEC. Hence, we cannot run a similar regression on the 'post-SEC' sample.

from tables 6, 7 and 9. The evidence is very similar. Compared to Prosper, with Lending Club data we obtain more precise estimates due to higher sample size.

In the first column of table 10, we regress the rates only on loan characteristics (size, term, motive) and on dummies for quarter of listing and state of address of the borrower, which explain around 29 percent of the variability of the rates (23 percent in Prosper data regression). In the third column, we add variables capturing the signals that lenders can use to infer borrowers' quality. As explained earlier, Lending Club offers less of those indicators compared to Prosper, mostly since soft information signals (groups recommendations) are 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. These variables improve enormously the ability of the regression model to explain the variability of lending rates. Overall the analysis suggests that loan rates are lower the better the credit grade is. This negative relation is a clear indication of the signalling role of those metrics. Improved transparency increases the success probability that lenders assign to projects with better FICO scores and this results in a decline of the loan spread associated to them. 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 are currently delinquent on other loans, whose income is relatively low and/or 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 in which lenders require higher loan spreads to fund riskier borrowers.

In columns 5 and 7 of the same table, we add to the regression the ratio of currency in the hands of public to demand deposits in the year before loan funding and the change of this ratio at the funding date with respect to previous year average. As explained before those regressors proxy for the fragility risk of the banking sector. With this specification we can investigate again the loans' substitution hypothesis, namely if an increase in risk in the traditional banking sector increases participation in the platform, hence lowers loan rates. Consistent with the predictions of our model and in line with the evidence from Prosper data, we find that the more currency the public held relative to deposits in the year before funding, the lower the equilibrium rates on the platform. Similarly, increases in the currency-to-deposits ratio, which tends to occur around banking panics dates, are associated with 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 asset classes 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 monitor and the potential costs of asymmetric information data suggest that this new intermediation service is performing well relatively to traditional banking[32]. Two are the main externalities that make this market attractive. The first are information externalities. Thanks for the digital technology the lending platform can provide a large variety of costless public information signals that help improve market transparency and facilitate convergence toward a separating equilibrium. Second, in times of bank distress the platforms provide a valuable form of investment substitution that improves risk-sharing possibilities for households-investors. 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.

[32] Recent speeches by policy makers (Constancio [9]) highlight its potential benefits even hinting at the risks that it may pose for the profitability of the banking sector.

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] 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).

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

[6] Bertsch, C., Hull, I. and X. Zhang, (2016). "Fed Liftoff and Subprime Loan Interest Rates: Evidence from the Peer-to-Peer Lending Market. Sveriges Riskbank." Working paper series 319.

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

[8] 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.

[9] 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.

[10] Curatola, G. and E. Faia, (2016). "Divergent Risk-Attitudes and Endogenous Collateral Constraints." CEPR w.p.

[11] 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.

[12] Easley, D. and M. O'Hara, (2004). "Information and the Cost of Capital." *Journal of Finance,* 59(4), 1553-1583.

[13] Faia, E. (2016). "A Note on Credit Risk Transfers and the Macroeconomy." Forthcoming *Macroeconomic Dynamics*. Also ECB w.p. 1256.

[14] The 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)

[15] Freedman, S. and G. Jin, (2016). "Learning by Doing with Asymmetric Information: Evidence from Prosper.com." NBER working paper #16855.

[16] Gale, D., and M. Hellwig. (1985) "Incentive-Compatible Debt Contracts: The One-Period Problem." *Review of Economic Studies,* 52, 647-663.

[17] Galloway, I. (2009). "Peer-to-Peer Lending and Community Development Finance." Community Development Investment Center Working Paper. San Francisco: Federal Reserve Bank of San Francisco. Retrieved from http://ideas.repec.org/p/fip/fedfcw/2009-06.html

[18] Georgarakos, D., M. Haliassos and G. Pasini (2014). "Household Debt and Social Interactions." *The Review of Financial Studies*, Vol. 27(5), pp. 1404-1433.

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

[20] Gorton, G. and A. Metrick, (2009). "Securitized Banking and the Run on Repo." *Journal of Financial Economics,* vol. 104(3), 425-451.

[21] Harrison, J. M. and D. M. Kreps, (1978). "Speculative Investor Behavior in a Stock Market with Heterogenous Expectations." *Quarterly Journal of Economics,* 92, 323-336.

[22] Holmström, B. and J. Tirole, (1997). "Financial Intermediation, Loanable Funds and the Real Sector.". *Quarterly Journal of Economics,* 112, 663-691.

[23] Hoff, K. and Stiglitz, J. E, (1990). "Imperfect Information and Rural Credit Markets–Puzzles and Policy Perspectives." *World Bank Economic Review*, World Bank Group, vol. 4(3), pages 235-250.

[24] Hughes, J.S., J. Liu and J. Liu, (2007). "Information Asymmetry, Diversification and Cost of Capital." *The Accounting Review,* 82(3), 705-729.

[25] 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.

[26] Judd, K. (1998). *Numerical Methods in Economics,* MIT Press.

[27] 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.

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

[29] La Ferrara, Eliana (2003) "Kin Groups and Reciprocity: A Model of Credit Transactions in Ghana." *American Economic Review* 93(5): 1730-1751.

[30] Laudenbach, C., Ungeheuer, M. and M. Weber, (2016). "How to Overcome Correlation Neglect?" Mimeo.

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

[32] Liskovich, I. and Shaton, M.. (2017). "Borrower Selection into Credit Markets: Evidence from Peer-to-Peer Lending." *Behavioral and Experimental Finance eJournal* Vol. 9, No. 3.

[33] 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.

[34] Kaplan, G. and G. Violante, (2014). "A Model of the Consumption Response to Fiscal Stimulus Payments." *Econometrica,* Volume 82, Issue 4, pp. 1199–123.

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

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

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

[38] Parlour, C. and Plantin, G. (2008). "Loan Sales and Relationship Banking." *The Journal of Finance*, Vol. 63 pp1291–1314.

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

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

[41] 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.

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

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

[44] Roussanov, N. (20120). "Diversification and its Discontents: Idiosyncratic and Entrepreneurial Risk in the Quest for Social Status." *Journal of Finance.*

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

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

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

[48] Udry, Cristopher (1994) "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria." *Review of Economic Studies,* 61(3):495-526.

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

[50] Welch, I., (1992). "Sequential Sales, Learning, and Cascades." *Journal of Finance*, 47(2), 695–732.

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

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

5.1 Appendix A. Contingent Deposit Returns

In the main text the households'/lenders' optimization problem was solved by assuming that the general return, R_t^d, has been contracted previously with the bank. This case fits well standard demand deposits. However, it is also possible that households invest in banks' bonds which offer contingent returns. In this case the inter-temporal budget constraint can be written in the more general form as follows:

$$C_t + r_t X_t + D_t \leq Y_t + X_{t-1} + \bar{\theta}_t R_t^d D_{t-1} \tag{16}$$

The first order condition for bank deposits in this case becomes:

$$u'(C_t) = \beta \mathbb{E}_t \left\{ u'(C_{t+1}) \bar{\theta}_{t+1} R_{t+1}^d \right\} \tag{17}$$

The return is now contingent to additional macro risk materializing in period $t+1$. In this case the no-arbitrage between P2P lending and banks' short term funding can be obtained by merging equations (4) and (17) and implies the following:

$$\beta \mathbb{E}_t \left\{ \frac{u'(C_{t+1})}{u'(C_t)} \frac{1}{r_t} \right\} = \beta \mathbb{E}_t \left\{ \frac{u'(C_{t+1})}{u'(C_t)} \bar{\theta}_{t+1} R_{t+1}^d \right\} \tag{18}$$

which can be re-written also as:

$$\mathbb{E}_t \left\{ \frac{u'(C_{t+1})}{u'(C_t)} \frac{1}{r_t} \right\} = \mathbb{E}_t \left\{ \frac{u'(C_{t+1})}{u'(C_t)} \right\} \mathbb{E}_t \left\{ \bar{\theta}_{t+1} R_{t+1}^d \right\} + Cov(\frac{u'(C_{t+1})}{u'(C_t)}, \bar{\theta}_{t+1} R_{t+1}^d) \tag{19}$$

The above expression helps to further elaborate on the channels driving the relative participation in the two sectors. Equation (19) states that, to be attractive, the return of P2P loans shall cover for the additional risk of state contingent bonds. If the covariance between banks' bonds and the market portfolio, whose price is captured by the stochastic discount factor $\frac{u'(C_{t+1})}{u'(C_t)}$, is positive, borrowers on the P2P platform shall offer higher returns to attract lenders. The opposite is true if this covariance is negative.

Under this new scenario the optimal portfolio allocation, namely the first order condition with respect to α_t, reads as follows:

$$\begin{aligned} [1 - r_t] &= \beta \mathbb{E}_t \left\{ \frac{u'(C_{t+1})}{u'(C_t)} \left[\bar{\theta}_{t+1} R_{t+1}^d - 1 \right] \right\} = \\ &= \beta \mathbb{E}_t \left\{ \frac{u'(C_{t+1}) W_{t+1}}{u'(C_t) W_t} \right\} E_t \left\{ \bar{\theta}_{t+1} R_{t+1}^d - 1 \right\} + Cov(\frac{u'(C_{t+1})}{u'(C_t)}, \bar{\theta}_{t+1} R_{t+1}^d - 1) \end{aligned} \tag{20}$$

Equation (20) conveys an intuition similar to that of equation (19). When the covariance between the market return and the return on banks' bonds is positive, households find more difficult to diversify asset risk. Therefore, bonds must pay a premium, $Cov(\frac{u'(C_{t+1})}{u'(C_t)}, \overline{\theta}_{t+1}R^d_{t+1} - 1)$ that brings households close to the certainty equivalent. On the margin, the fraction of wealth invested on the platform, α_t, is determined as the point where net returns from the platform equate net returns from banks' bonds, which include the premium for undiversifiable risk. In equilibrium and since now banks' bonds offer an additional risk premium, for P2P lending to provide a competitive form of investment the returns accruing from the platform investment shall either include an equivalent premium (for given variance of the returns) or offer returns with lower variance. In other words, condition (20) states that the optimal portfolio share shall make households indifferent between the two forms of investment along the risk-return frontier.

6 Appendix B. Borrowers' Optimization Problem

Borrowers fund a fraction, γ_t, of their project with P2P loans, X^b_t, and a fraction $(1-\gamma_t)$ with bank funds, L_t. Borrowers' expected return from the project is $p^i R^I_t$. Aggregation of optimality conditions and wealth accumulation can be done by averaging due to the linearity assumption in the projects' returns and due to the finite lived structure of those agents. For this reason we can skip the index i since now onward. In period t borrowers, whose projects succeed, shall repay a contractual gross return to the bank, which can be curtailed in the event of liquidity shortage. In the same event borrowers shall also repay the gross value of the portfolio to investors in the P2P market At time t borrowers ask for new funds to fund projects, I_t. They obtain, $r_t X^b_t$ from the platform and L_t from the bank. Those funds are part of their revenues. We assume that borrowers can rollover debt to the next period (in any sector) only if the past project succeeded (which happens with probability p), while in case of no repayment they are excluded from future renewal of debt.

Risk neutral borrowers solve their portfolio decision by maximizing the sum of future discounted utilities of consumption, C^b_t:

$$Max_{I_t,\gamma_t}\mathbb{E}_0\left\{\sum_{t=0}^{\infty}(\xi\beta)^t C^b_t\right\} \tag{21}$$

subject to their budget constraint that reads as follows:

$$C_t^b + pX_{t-1}^b + p\left[\zeta_t\theta + (1-\zeta_t)\right]R_{t-1}^I L_{t-1} \leq p(r_t X_t^b + L_t) \tag{22}$$

where:

$$X_t^b = \gamma_t I_t;\ L_t = (1-\gamma_t)I_t \tag{23}$$

Notice that once I_t and γ_t have been optimally determined, X_t^b and L_t follow. Borrowers optimality condition with respect to investment reads as follows:

$$\left[\gamma_t r_t + (1-\gamma_t)\right] = \beta\xi\left[\gamma_t + R_t^I(\zeta_t\theta + (1-\zeta_t))(1-\gamma_t)\right] \tag{24}$$

The above condition states that the expected cost of funding weighted by the share of funds obtained from each sector (right hand side of equation (24)) shall equate the expected benefits from rolling over debt also weighted (left hand side). By maximizing with respect to γ_t (after substituting the optimal wealth shares into the budget constraint) we find that the optimal fraction of P2P lending is obtained by equating the expected net cost of funds in the two markets:

$$(r_t - 1) = \beta\xi\left[1 - \bar{\theta}_{t+1}R_t^I\right] \tag{25}$$

The above condition states that the optimal share γ_t is determined when the net cost from market funding equates the net cost from bank funding.

Since borrowers are risk-neutral, their consumption schedule will result in a corner solution. We can exploit the finite life structure a' la Yaari [51] 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.

7 Appendix C. 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[33], income, and

[33] 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.

current debts. Each listing includes also a credit grade -Prosper Rating- 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[34], and (2) the Prosper Score, computed in-house based on the Prosper population."[35]

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 alumni groups related to a school or an employer, geographically-oriented groups, military groups, medical groups related to medical reasons for borrowing or lending, demographic groups highlighting particular demographics, such as single parents or Hispanics, hobbies for groups targeting people with particular hobbies, religiously oriented groups, and business groups highlighting small businesses, or 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 groups until loan repayment.

In turn, lenders assess and can bid on listings. The minimum bid is set at $25. Until July 2009, Prosper operated a variable rate model. It worked as an eBay-style online 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 interest rates were granted a share of the loan. In 2009, Prosper registered with the SEC and 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 were required to have an FICO credit score of at least 640, while returning borrowers

[34] The borrower actual credit score is not observable by lenders who have access only to her credit score category which is defined over a range of the FICO score.

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

only need a score of 600 to request a loan.[36]

Borrowers are limited to a maximum of two concurrent loans of $25,000 or less each until 2012, $35,000 thereafter[37]. 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. If the monthly payment is late for two or more months, it is sent to a collection agency. 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. More generally, on Prosper, the 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.

8 Appendix D - Lending Club Institutional Data Description

Lending Club enables borrowers to create loan listings on its website by supplying details about themselves and the project that they would like to fund. All loans are unsecured personal loans and can be between $1,000 $35,000. Based on the borrower's credit history and score, on her debt

[36]The original model effectively had the investors setting the rates through an open bidding process. The problem with that original model, that was shut down by regulatory authorities, is that the investors were individuals without any demonstrable skill or 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 unfunded loans and unapproved loans. Additionally, from a regulatory perspective, there were fair lending risks due to no discernible way in which the interest rates were established.

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

to income ratio and desired loan amount, Lending Club assigns loans a letter credit grade that determines payable interest rate and fees. The standard loan period is three years. A five year period is available at a higher interest rate and additional fees. The loans can be repaid at any time without penalty.

Investors can search and browse the loan listings on Lending Club website. They can also decide how much to fund each borrower, with a $25 minimum investment. The interest rate is set by Lending Club. Rates vary between 5% and 26% depending on the credit grade assigned to the loan. Loans credit grades are computed by Lending Club using a proprietary algorithm and correspond to the interest rate that is charged for the loan. They depend on FICO, loan term (shorter term is considered better), and the loan amount requested by the borrower. The range of grades is similar to Prosper's, but Prosper's average interest rate is higher in each group. For example, for credit grade A, it is 4% higher, and that difference increases as the credit grade increases.

Lending Club makes money by charging borrowers an origination fee, which ranges between 1% and 5% depending on the credit grade, and charging investors a 1% service fee on all amounts the borrower pays. To reduce default risk, Lending Club focuses on high creditworthy borrowers, declining approximately 90% of the loan applications it received as of 2012 and assigning higher interest rates to riskier borrowers. Only borrowers with FICO score of 660 or higher can be approved for loans. Most borrowers on Lending Club's website report using their loans to refinance other loans or pay credit card debt (table D1, 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 D2). On average, they have 11 open credit lines; 20 percent have been delinquent on a loan in the two years before borrowing on Lending Club. Borrowers have $73,000 of personal income and takes out an average loan of $14,000 for debt consolidation or for paying off credit card debts. 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%.

9 Appendix E

The information on bank failures comes from the Federal Deposit Insurance Corporation which is often appointed as receiver for failed banks. The original list includes banks which have failed since October 1, 2000. Table E reports the number of failures by state and year for the period covered by Prosper data, i.e. Jan. 2006 to March 2014. Notice that in 2006 no bank failures were recorded.

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 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: 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

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%
Chargedoff	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' 0/1 denotes listings without/with 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 prior 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' 0/1 denotes listings by borrowers who funded their projects 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	Pre-SEC(1)	Post-SEC
Loan size (thousands)	-0.090	-0.078	-0.102
	(0.001)***	(0.003)***	(0.001)***
Loan size (thousands)	0.019	0.025	0.020
	(0.000)***	(0.001)***	(0.000)***
Term (months) (2)	0.011	-	0.012
	(0.000)***		(0.000)***
Debt consolidation(*)	0.004	0.014	0.004
	(0.001)***	(0.002)***	(0.001)***
Home improvement(*)	-0.003	-0.006	-0.003
	(0.001)***	(0.003)*	(0.001)***
Business funding(*)	0.008	0.002	0.010
	(0.001)***	(0.002)	(0.001)***
Adjusted R^2	0.23	0.12	0.28
N	107,549	23,425	84,124

Note: Year dummies and US state dummies are also included. (*) denotes a dummy variable. (1) Pre-SEC refers to loans funded before Prosper registration with the SEC in 2009. (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	Pre-SEC	Post-SEC	Pre-SEC	Post-SEC	Post-SEC
Loan size (thousands)	-0.043	0.016	-0.063	0.018	-0.063	-0.061
	(0.001)***	(0.002)***	(0.001)***	(0.002)***	(0.001)***	(0.001)***
Loan size (thousands)^2	0.010	0.003	0.015	0.002	0.015	0.014
	(0.000)***	(0.001)***	(0.000)***	(0.001)***	(0.000)***	(0.000)***
Term	0.009		0.011		0.011	0.012
	(0.000)***		(0.000)***		(0.000)***	(0.000)***
Debt consolidation(*)	-0.002	-0.001	-0.001	-0.001	-0.001	-0.003
	(0.001)***	(0.001)	(0.001)**	(0.001)	(0.001)**	(0.001)***
Home improvement(*)	0.000	0.001	0.000	0.001	0.000	0.004
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)***
Business funding(*)	0.006	0.003	0.005	0.003	0.006	0.006
	(0.001)***	(0.002)*	(0.001)***	(0.002)*	(0.001)***	(0.001)***
FICO score (hundreds)	-0.070	-0.071	-0.073	-0.071	-0.073	-0.079
	(0.000)***	(0.001)***	(0.000)***	(0.001)***	(0.000)***	(0.000)***
Open credit lines (tens)	0.003	0.005	0.001	0.005	0.001	0.003
	(0.000)***	(0.001)***	(0.001)	(0.001)***	(0.001)	(0.001)***
Credit enquiries (tens)	0.017	0.009	0.024	0.009	0.025	0.030
	(0.000)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Current delinquencies(*)	0.012	0.027	0.008	0.028	0.009	0.009
	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Monthly income	-0.001	0.001	-0.001	-0.000	-0.001	-0.001
(thousands)	(0.000)***	(0.001)	(0.000)***	(0.000)	(0.000)***	(0.000)***
Debt/Income	0.012	0.003	0.027	0.004	0.027	0.029
	(0.001)***	(0.000)***	(0.002)***	(0.000)***	(0.002)***	(0.002)***
Group dummy(*)				-0.005	-0.019	0.000
				(0.001)***	(0.001)***	(0.001)
Recommend + no investm. (*)				0.000	-0.025	-0.004
				(0.001)	(0.002)***	(0.002)*
Recommend + investm. (*)				-0.019	-0.015	-0.008
				(0.002)***	(0.004)***	(0.004)**
Investm.+ no recommend. (*)				-0.045	-0.012	-0.008
				(0.007)***	(0.004)***	(0.004)
Previous Prosper loan(*)						-0.042
						(0.000)***
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	Pre-SEC	Post-SEC
Loan size (thousands)	-0.044	-0.043	0.008	-0.063
	(0.001)***	(0.001)***	(0.002)***	(0.001)***
Loan size (thousands)^2	0.010	0.010	0.003	0.015
	(0.000)***	(0.000)***	(0.001)***	(0.000)***
Term	0.009	0.009		0.011
	(0.000)***	(0.000)***		(0.000)***
Income is verifiable(*)	-0.024			
	(0.009)***			
No open credit lines		0.024		
		(0.003)***		
No US State(*)			0.018	
			(0.009)**	
No reason for borrowing(*)				0.006
				(0.001)***
FICO score (hundreds)	-0.070	-0.070	-0.070	-0.073
	(0.000)***	(0.000)***	(0.001)***	(0.000)***
Open credit lines (tens)	0.002		0.005	0.001
	(0.000)***		(0.001)***	(0.001)
Credit enquiries (tens)	0.017	0.017	0.008	0.024
	(0.000)***	(0.000)***	(0.001)***	(0.001)***
Current delinquencies(*)	0.012	0.011	0.026	0.008
	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Monthly income (thousands)	-0.002	-0.001	-0.001	-0.001
	(0.001)***	(0.000)***	(0.000)	(0.000)***
Income * Inc. is verif. (*)	0.002			
	(0.001)**			
Debt/Income	-0.000	0.012	0.003	0.027
	(0.001)	(0.001)***	(0.000)***	(0.002)***
Debt/Income* Inc. is verif. (*)	0.021			
	(0.002)***			
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 panics and signals

	All	Pre-SEC	Post-SEC	Pre-SEC	Pre-SEC	Post-SEC
Loan size (thousands)	-0.090	-0.078	-0.103	0.018	0.018	-0.061
	(0.001)***	(0.003)***	(0.001)***	(0.002)***	(0.002)***	(0.001)***
Loan size (thousands)^2	0.019	0.025	0.021	0.002	0.002	0.014
	(0.000)***	(0.001)***	(0.000)***	(0.001)***	(0.001)***	(0.000)***
Term (months)	0.011		0.013			0.012
	(0.000)***		(0.000)***			(0.000)***
Debt consolidation(*)	0.005	0.014	0.005	-0.002	-0.001	-0.003
	(0.001)***	(0.002)***	(0.001)***	(0.001)	(0.001)	(0.001)***
Home improvement(*)	-0.003	-0.006	-0.003	0.001	0.001	0.004
	(0.001)***	(0.003)*	(0.001)***	(0.002)	(0.002)	(0.001)***
Business funding(*)	0.008	0.002	0.010	0.003	0.003	0.006
	(0.001)***	(0.002)	(0.001)***	(0.002)*	(0.002)*	(0.001)***
Currency to deposits(1): previous yr average	-0.058	-0.110	-0.055	0.115		-0.044
	(0.010)***	(0.044)**	(0.010)***	(0.042)***		(0.010)**
Currency to deposits: % change	-0.041	-0.016	-0.053	-0.012		-0.093
	(0.014)***	(0.030)	(0.015)***	(0.022)		(0.012)***
Bank run(*) (2)					-0.002	
					(0.001)**	
FICO score (hundreds)				-0.071	-0.071	-0.079
				(0.001)***	(0.001)***	(0.000)***
Open credit lines (tens)				0.005	0.005	0.003
				(0.001)***	(0.001)***	(0.001)***
Credit enquiries (tens)				0.009	0.009	0.030
				(0.001)***	(0.001)***	(0.001)***
Current delinquencies(*)				0.028	0.028	0.009
				(0.001)***	(0.001)***	(0.001)***
Monthly income				-0.000	-0.000	-0.000
(hundreds of thousands)				(0.000)	(0.000)	(0.000)***
Debt/Income				0.004	0.004	0.029
				(0.000)***	(0.000)***	(0.002)***
Group dummy(*)				-0.005	-0.005	0.000
				(0.001)***	(0.001)***	(0.001)
Recommend + no investm. (*)				0.000	0.000	-0.004
				(0.001)	(0.001)	(0.002)*
Recommend + investm. (*)				-0.019	-0.019	-0.008
				(0.002)***	(0.002)***	(0.004)**
Investm.+ no recommend. (*)				-0.045	-0.045	-0.007
				(0.007)***	(0.007)***	(0.004)
Previous Prosper loan(*)				-0.002	-0.002	-0.041
				(0.001)	(0.001)	(0.000)***
Adjustment R^2	0.23	0.12	0.29	0.59	0.59	0.56
N	107,549	23,425	84,124	18,497	18,497	76,899

Note: (1) Data source: Demand Deposits series (Total, Billions of Dollars, Monthly, Seasonally Adjusted) and Currency Component of M1 series (Billions of Dollars, Monthly, Seasonally Adjusted) from Federal Reserve Economic Data (https://fred.stlouisfed.org). (2) ‘Bank run’ is a dummy that takes on value 1 if in the quarter before listing a banking panic occurs in the US. The panics dates we consider are: 1) August 2007, i.e. Q3-2007, when the American firm Countrywide Financial suffered a bank run as a consequence of the subprime mortgage crisis; 2) March 2008, i.e. Q1-2008, when a bank run began on the securities and banking firm Bear Stearns, which, although it was not an ordinary deposit-taking bank, had financed huge long-term investments by selling short-maturity bonds, making it vulnerable to panic on the part of its bondholders; 3) June 26 2008,when a warning was issued that U.S. mortgage lender IndyMac Bank might not be viable and a bank run began. In addition to Indy Mac Bank case, on 25 September 2008, Washington Mutual, the largest savings and loan in the United States and the sixth-largest overall financial institution, was shut down due to a massive run. See note to Table 6. Robust standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01

Table 10 – Lending rates, banking failures and signals

	All	Pre-SEC	Post-SEC	Pre-SEC	Post-SEC
Loan size (thousands)	-0.090	-0.078	-0.103	0.018	-0.061
	(0.001)***	(0.003)***	(0.001)***	(0.002)***	(0.001)***
Loan size (thousands)^2	0.019	0.025	0.021	0.002	0.014
	(0.000)***	(0.001)***	(0.000)***	(0.001)***	(0.000)***
Term (months)	0.011		0.013		0.012
	(0.000)***		(0.000)***		(0.000)***
Debt consolidation$^{(*)}$	0.004	0.014	0.005	-0.001	-0.003
	(0.001)***	(0.002)***	(0.001)***	(0.001)	(0.001)***
Home improvement$^{(*)}$	-0.003	-0.006	-0.003	0.001	0.004
	(0.001)***	(0.003)	(0.001)***	(0.002)	(0.001)***
Business funding$^{(*)}$	0.008	0.001	0.010	0.003	0.006
	(0.001)***	(0.002)	(0.001)***	(0.002)*	(0.001)***
Bank failures$_{mo-1}$$^{(*)\,(1)(2)}$	-0.001	-0.021	0.001	-0.007	0.000
	(0.002)	(0.008)***	(0.002)	(0.005)	(0.001)
Bank failures$_{mo-1}$$^{(*)\,(2)}$	-0.006	-0.027	-0.004	-0.003	-0.003
	(0.002)***	(0.007)***	(0.002)**	(0.005)	(0.001)**
Bank failures$_{mo-1}$$^{(*)\,(2)}$	-0.003	-0.023	-0.001	0.000	-0.002
	(0.002)*	(0.011)**	(0.002)	(0.008)	(0.001)*
FICO score (hundreds)				-0.071	-0.079
				(0.001)***	(0.000)***
Open credit lines (tens)				0.005	0.003
				(0.001)***	(0.001)***
Credit enquiries (tens)				0.009	0.030
				(0.001)***	(0.001)***
Current delinquencies$^{(*)}$				0.028	0.009
				(0.001)***	(0.001)***
Monthly income				-0.000	-0.000
(hundreds of thousands)				(0.000)	(0.000)***
Debt/Income				0.004	0.029
				(0.000)***	(0.002)***
Group dummy$^{(*)}$				-0.005	-0.000
				(0.001)***	(0.001)
Recommend + no investm.$^{(*)}$				0.000	-0.004
				(0.001)	(0.002)*
Recommend + investm.$^{(*)}$				-0.019	-0.008
				(0.002)***	(0.004)**
Investm.+ no recommend.$^{(*)}$				-0.045	-0.008
				(0.007)***	(0.004)*
Previous Prosper loan$^{(*)}$				-0.002	-0.042
				(0.001)	(0.000)***
Adjustment R^2	0.23	0.12	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 k^{th} 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 11 – 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)	-0.027	_-0.090_	-0.002	_-0.043_	-0.026	-0.090	-0.002	-0.043
	(0.000)***	_(0.001)***_	(0.000)***	_(0.001)***_	(0.000)***	(0.001)***	(0.000)***	(0.001)***
Loan size (thousands)^2	0.007	_0.019_	0.001	_0.010_	0.007	0.019	0.001	0.010
	(0.000)***	_(0.000)***_	(0.000)***	_(0.000)***_	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Term	0.020	_0.011_	0.001	_0.009_	0.019	0.011	0.001	0.009
	(0.000)***	_(0.000)***_	(0.000)***	_(0.000)***_	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Debt consolidation(*)	-0.010	_0.004_	-0.001	_-0.002_	-0.010	0.005	-0.001	-0.002
	(0.000)***	_(0.001)***_	(0.000)***	_(0.001)***_	(0.000)***	(0.001)***	(0.000)***	(0.001)***
Credit card payments(*)	-0.023		-0.002		-0.023		-0.002	
	(0.000)***		(0.000)***		(0.000)***		(0.000)***	
Home improvement(*)	-0.015	_-0.003_	-0.002	_0.000_	-0.015	-0.003	-0.002	0.000
	(0.000)***	_(0.001)***_	(0.000)***	_(0.001)_	(0.000)***	(0.001)***	(0.000)***	(0.001)
Business funding(*)	0.014	_0.008_	-0.000	_0.006_	0.014	0.008	-0.001	0.006
	(0.001)***	_(0.001)***_	(0.000)	_(0.001)***_	(0.001)***	(0.001)***	(0.001)*	(0.001)***
FICO score (hundreds)				_-0.070_				-0.070
				_(0.000)***_				(0.000)***
LC grade B(*)			0.038				0.038	
			(0.000)***				(0.000)***	
LC grade C(*)			0.068				0.069	
			(0.000)***				(0.000)***	
LC grade D(*)			0.097				0.098	
			(0.000)***				(0.000)***	
LC grade E(*)			0.126				0.127	
			(0.000)***				(0.000)***	
LC grade F(*)			0.156				0.156	
			(0.000)***				(0.000)***	
LC grade G(*)			0.170				0.170	
			(0.000)***				(0.000)***	
Open credit lines (tens)			-0.001	_0.003_			-0.001	0.003
			(0.000)***	_(0.000)***_			(0.000)***	(0.000)***
Open credit lines (tens)^2			0.001				0.000	
			(0.000)***				(0.000)	
Credit enquiries (tens)			0.006	_0.017_			0.007	0.017
			(0.000)***	_(0.000)***_			(0.000)***	(0.000)***
Current delinquencies(*)			0.001	_0.012_			0.001	0.012
			(0.000)***	_(0.001)***_			(0.000)***	(0.001)***
Annual income			-0.001	_-0.001_			-0.001	-0.001
(thousands)			(0.000)***	_(0.000)***_			(0.000)***	(0.000)***
Debt/Income			0.007	_0.012_			0.006	0.012
			(0.000)***	_(0.001)***_			(0.000)***	(0.001)***
Currency to deposits(1):					-0.008	-0.058	-0.007	-0.029
previous yr average					(0.001)***	(0.010)***	(0.000)***	(0.008)***
Currency to deposits:					-0.043	-0.041	-0.068	-0.031
% change					(0.002)***	(0.014)***	(0.001)***	(0.012)***
Adjusted R2	0.29	_0.23_	0.94	_0.49_	0.27	0.23	0.93	0.49

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 D1 – 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

Table D2 – 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

Table E – US Bank failures by state and year

	Year of closure								
State	2007	2008	2009	2010	2011	2012	2013	2014	Total
AL	0	0	3	1	2	1	0	0	7
AR	0	1	0	1	0	0	0	0	2
AZ	0	0	5	4	3	0	3	0	15
CA	0	5	17	12	4	1	0	0	39
CO	0	0	3	0	6	0	0	0	9
CT	0	0	0	0	0	0	1	0	1
FL	0	2	14	29	13	8	4	0	70
GA	1	5	25	21	23	10	3	0	88
IA	0	0	1	0	1	0	0	0	2
ID	0	0	1	0	0	0	0	1	2
IL	0	1	21	16	9	8	1	1	57
IN	0	0	1	0	1	1	0	0	3
KS	0	1	3	3	1	1	0	0	9
KY	0	0	1	0	0	0	1	0	2
LA	0	0	0	1	1	0	0	0	2
MA	0	0	0	1	0	0	0	0	1
MD	0	0	2	4	0	2	0	0	8
MI	0	1	4	5	2	1	0	0	13
MN	0	1	6	8	2	4	1	0	22
MO	0	2	3	6	1	4	0	0	16
MS	0	0	0	1	1	0	0	0	2
NC	0	0	2	0	2	1	2	0	7
NE	0	0	1	1	1	0	0	0	3
NJ	0	0	2	1	1	1	0	0	5
NM	0	0	0	2	1	0	0	0	3
NV	0	3	3	4	1	0	1	0	12
NY	0	0	1	3	0	0	0	0	4
OH	1	0	2	2	0	0	0	0	5
OK	0	0	1	1	2	1	0	1	6
OR	0	0	3	3	0	0	0	0	6
PA	1	0	1	2	1	2	0	1	8
PR	0	0	0	3	0	0	0	0	3
SC	0	0	0	4	3	2	0	0	9
SD	0	0	1	0	0	0	0	0	1
TN	0	0	0	0	0	3	2	0	5
TX	0	2	5	1	1	0	2	0	11
UT	0	0	2	3	1	0	0	0	6
VA	0	0	1	1	2	0	0	1	5

WA	0	0	3	11	3	0	1	0	18
WI	0	0	1	2	3	0	2	0	8
WV	0	1	0	0	0	0	0	0	1
WY	0	0	1	0	0	0	0	0	1
Total	3	25	140	157	92	51	24	5	497

Source: Federal Deposit Insurance Corporation. Data downloaded on 30th Nov. 2017 from https://www.fdic.gov/bank/individual/failed/banklist.html.
Note: in 2006, no failures were recorded. For 2014, the table includes only the failures occurred in the first quarter (for consistency with Prosper data sample coverage).