JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS

doi:10.1017/S0022109023000789

Reintermediation in FinTech: Evidence from Online Lending

Tetyana Balyuk
Emory University Goizueta Business School
tetyana.balyuk@emory.edu (corresponding author)

Sergei Davydenko
University of Toronto Rotman School of Management
davydenko@rotman.utoronto.ca

Abstract

We document the unique structure of the peer-to-peer lending market. Originally designed as decentralized, the market has become highly, but not fully, reintermediated. The platforms' software now performs essentially all tasks related to loan evaluation, whereas most lenders are passive and automatically fund most applications on offer. Yet unlike banks, and in contrast to theories predicting full reintermediation, the platforms provide detailed loan information, and some active loan pickers coexist with passive investors. We argue that while intermediation attracts unsophisticated passive investors, transparency in the presence of active investors resolves the lending platform's moral hazard problem inherent in intermediated markets.

I. Introduction

Online platforms such as Uber, Airbnb, and eBay bring together buyers and sellers of goods and services over the Internet, reducing search costs in a multitude of very different markets. The rise of financial technology (FinTech) has sometimes been predicted to result in similar developments in the financial sector, allowing providers and users of finance to interact directly without the involvement of banks and other financial intermediaries.[1] Inspired by these ideas, peer-to-peer (P2P)

We thank Christoph Bertsch, Jeffrey Busse, Michele Dathan, Craig Doidge, Alex Dyck, Redouane Elkamhi, Rohan Ganduri, Christoph Herpfer, Julapa Jagtiani, Narasimhan Jegadeesh, Michael King, Florian Koch, Victor Lyonnet, Gonzalo Maturana, Alexandra Niessen-Ruenz, Nagpurnanand Prabhala, Boris Vallee, Christina Wang, and Robert Wardrop; seminar participants at Cass Business School, Scheller College of Business, Goizueta Business School, Rotman School of Management, and FED Board; attendees of the CFIC, Australasian Finance and Banking Conference, Philadelphia FED FinTech Conference, CenFIS (Atlanta FED)/CEAR Conference on Financial Stability Implications of New Technology, NFA, FDIC-JFSR Bank Research Conference, FinTech and Financial Risk Management Conference, FinteQC Conference, Showcasing Women in Finance Conference, and Toronto FinTech Conference; and an anonymous referee for helpful comments and suggestions.

[1]See, for example, "Role of Banks Recedes in Wake of Crisis" (*Financial Times*, June 22, 2014); "Fintech Firms Are Taking on the Big Banks, But Can They Win?" (*The New York Times*, Apr. 6, 2016); and "Digital Disruption: How FinTech Is Forcing Banking to a Tipping Point" (Citibank, Mar. 2016).

lending markets were created to allow consumers to request loans online and creditors to evaluate and directly fund loan applications of their choosing.

Online P2P platforms were originally organized as auctions, like an eBay for consumer loans, but over time they have evolved to take central stage in lending decisions on investors' behalf, replacing financial intermediaries in all but name. Yet unlike bank depositors, P2P lenders retain the ability to evaluate the individual loans they invest in, even though we show that a large majority of them invest passively, outsourcing almost all decision-making to the lending platform's software. In other words, the platform essentially calls all the shots but maintains the infrastructure for active loan picking, which a large majority of investors duly ignores. Why has this market evolved to adopt this seemingly redundant, not-quite-fully reintermediated market model, and what can it teach us about the benefits of and the limits to financial intermediation?

The goal of this article is to document this unique market structure and provide a rationale for it. Theory predicts that the ability to leverage the lending platform's technological expertise in loan evaluation would stimulate reintermediation, which would cause all lenders to adopt passive investment strategies (Vallee and Zeng (2019)). Yet the moral hazard problem inherent to the originate-to-distribute lending model requires maintaining passive investors' trust in the platform's underwriting standards, which can evaporate in adverse circumstances, causing the market to collapse (Thakor and Merton (2019)). In this article, we argue that by accommodating active investment strategies and providing detailed information about its loans, the lending platform allows sophisticated investors to independently ascertain the quality of the loan book, thereby reducing the reliance on trust and mitigating the moral hazard problem. Using an episode of the market's crisis in 2016 as a natural experiment, we show that the presence of active investors helps the platform survive when passive investors' loss of trust in the platform's loans causes them to leave the market.

Using data from Prosper Marketplace (Prosper), one of the largest U.S. P2P platforms, we first show that the U.S. "P2P lending market" is highly centralized, and is neither *P2P* nor a traditional *lending market* in which creditors choose to whom to lend and at what price.[2] First, lending is dominated by institutional investors, whereas retail ("peer") lenders provide less than 10% of the capital. Second, investors are overwhelmingly passive and perform little-to-no loan analysis, agreeing to fund over 98% of loan applications on offer. A large majority of institutional funds is allocated through the "passive" investment pool, whereby lenders prespecify loan portfolio parameters and the platform then automatically buys whole loans on their behalf. Third, investors' passivity contrasts with the active role of the lending platform, which not only sets loan interest rates but also screens out low-quality loans. Thus, the lending platform performs essentially all tasks related to loan origination, except for providing the funds. Nonetheless, investors still have access to detailed information about all individual loans, though

[2]Recently, the term "marketplace lending" has been introduced instead of "P2P." As Morse (2015) points out, "[a]n argument against this rather appealing term is that the amount of additional intermediation that may be optimal in underwriting P2P is now a first-order question." We use the two terms interchangeably.

only a minority exercise the option to actively screen the loans. Moreover, the remaining active investors do not earn higher average returns than those using purely passive strategies.

This market structure is a result of an evolutionary process rather than original design. In Prosper's early years, the platform acted purely as an online meeting place for borrowers and lenders, and loan rates were determined via lenders' auctions, eBay-style. But as the platform became more adept in loan analysis over time, it switched to the reintermediated model by taking over loan screening and pricing, whereas investors became overwhelmingly passive. We find a similar marked shift away from decentralization and toward reintermediation in other P2P lending markets. Nowadays, only 15% of P2P platforms around the world can be characterized as partially decentralized, and none are fully decentralized. Centrally provided expertise tends to replace the wisdom of the crowds.

Yet lending platforms differ from traditional intermediaries such as banks in some important aspects. Notably, they continue to supply investors with detailed information about individual loan applications and provide them with the ability to pick individual loans. Maintaining the infrastructure required for this purpose appears costly and redundant, particularly given that an overwhelming majority of investors are passive, and that even the remaining few active investors do not outperform passive strategies. What purpose is served by this feature of the market?

We hypothesize that the presence of active investors, combined with the platform's transparency about the loans it originates, allows it to mitigate the moral hazard problem that could otherwise threaten the stability of the market. The platform makes money based on the loan volume, yet retains no stake in the loans it originates. As studies of loan securitizations prior to the financial crisis have shown (e.g., Keys, Mukherjee, Seru, and Vig (2010); Purnanandam (2011)), under these conditions, the loan originator may fail to screen out poor-quality loans in order to boost volume. Given the reliance of (predominantly passive) investors on the intermediary's expertise for loan evaluation, the platform may be tempted to relax its lending standards to inflate the volume and thus its fees. We argue that the presence of active investors who can independently ascertain the quality of the loans disciplines the platform and reduces the market's reliance on trust.[3]

We illustrate the stabilizing role of active investors by applying the difference-in-differences (DiD) methodology to the events around the P2P market's near collapse in early 2016, when Moody's announced that it was placing securities backed by Prosper loans on watch for downgrade. The warning concerned loans that were originated more than a year earlier and Moody's eventually withdrew it without taking any action, suggesting that its information content was small or transitory. Nonetheless, in the 4 months following the announcement, loan originations on Prosper fell by 83%, erasing years of exponential growth, and have not fully recovered 5 years later. We show that this drop was overwhelmingly driven by the withdrawal of passive institutional investors, consistent with the loss of trust in

[3]This view is echoed by some public commentators blogging about P2P lending, for example, "a rich available data set allows investors, both new and existing, to explore LC's overall performance from external nonbiased websites, and subsequently to feel good about this investment" (https://www.lendingmemo.com/cutting-open-data-50-lending-club-may-lose).

https://doi.org/10.1017/S0022109023000789 Published online by Cambridge University Press

the quality of Prosper's underwriting standards. By contrast, the announcement had little effect on the lending by active investors, who could independently assess the loan book's quality. We document that in response to the crisis, Prosper increased interest rates and re-evaluated its credit model toward greater conservatism, and loan volume started to recover.[4]

This downgrade warning episode on Prosper suggests that the presence of active investors can help the platform weather a crisis of trust brought about by an extraneous factor that can unnerve passive investors. At the same time, our analysis also shows that active investors make an effort to avoid low-quality loans overpriced by Prosper. To demonstrate this, we apply the DiD methodology to P2P lending in areas affected by natural disasters. While Prosper does not incorporate borrowers' geographical location into its analysis, we find that active, but not passive, investors perform incremental screening to avoid affected loans. Moreover, we also find that active, but not passive, investors are less likely to fund new loans when default rates on the existing loans increase, consistent with learning about longer-term trends in loan performance as they transpire from the data. This monitoring by active investors motivates the platform to screen the loans carefully despite retaining no economic stake in them.

Overall, by providing the raw loan data alongside its credit analysis, the platform can attract not only a large number of passive lenders, but also some active investors, whose independent analysis of the loan book quality both keeps the platform's moral hazard in check over the long term and provides stability in the time of crisis by reducing the market's reliance on trust.

Finally, we show how Prosper resolves the adverse selection problem that may arise when passive investors compete with more sophisticated active investors. In theory, when the two types of investors fund loans from the same pool, active investors' ability to identify the best loans may result in a drop in the average quality of the loans available to the passive investors and drive them away from the market (Vallee and Zeng (2019)). But in practice, Prosper is able to largely mitigate this problem by placing active and passive institutional investors in separate pools and allocating loans randomly to these pools.[5] Given that investors fund almost all loans, this random allocation across pools ensures that average returns for active and passive institutional investors are not significantly different.

Our article contributes to the growing literature on marketplace lending,[6] focusing specifically on the structure of the market and its evolution. Franks, Serrano-Velarde, and Sussman (2021) study price discovery in P2P lending under

[4]Around the same time, LC was accused of misleading investors and also experienced a sharp drop in loan volume. LC's CEO was forced to resign, and our estimates (not reported) show that the platform also re-evaluated its credit model toward greater conservatism in a way similar to Prosper's.

[5]While adverse selection can still affect less skilled investors inside the active institutional pool as well as within the retail pool, this problem is likely of second-order importance, because active institutions have the choice of becoming passive to avoid it, and because the retail pool provides less than 10% of total funds.

[6]A partial list of prominent recent contributions includes Zhang and Liu (2012), Lin, Prabhala, and Viswanathan (2013), Iyer, Khwaja, Luttmer, and Shue (2016), Butler, Cornaggia, and Gurun (2017), Hildebrand, Puri, and Rocholl (2017), Hertzberg, Liberman, and Paravisini (2018), Jagtiani and Lemieux (2019), Tang (2019), Berg, Burg, Gombović, and Puri (2020), Di Maggio and Yao (2021), and Balyuk (2022).

the auction mechanism, and show that as the popularity of passive investment increases and institutional lenders displace small "peers," the auction model may generate less accurate loan prices sensitive to liquidity shocks. Wei and Lin (2016) show that funding rates on Prosper increased when it switched to the posted-price model in 2010. Vallee and Zeng (2019) construct a model of marketplace lending, which predicts that over time the lending platform should switch to posted prices to maximize investor participation, while discouraging loan analysis by active investors by becoming more opaque, in order to protect passive lenders from adverse selection. The resulting equilibrium essentially amounts to full intermediation.

Although the general trend from auctions toward reintermediation that we document is consistent with Vallee and Zeng (2019), the observed market structure remains different in some key aspects. Crucially, the lending platform continues to provide detailed information about its loans, and a fraction of investors remain active. We argue that this feature of the market resolves the moral hazard problem by allowing active investors to monitor the quality of the loan book. Thakor and Merton (2019) predict that maintaining investors' trust is critical to an opaque P2P platform, because in contrast to banks, it cannot fall back on depositors as liquidity providers if adverse developments cause lenders to lose trust in the quality of its loans. Our article shows that as an alternative to maintaining trust, transparency in the presence of active investors can help the platform survive market crises that erode passive investors' trust and cause them to leave the market.

The remainder of the article is organized as follows: Section II outlines the lending process on Prosper and describes our data. Section III documents the current market structure and illustrates the general trend toward reintermediation in P2P lending, both on Prosper and on other platforms around the world. Section IV discusses the moral hazard problem as a counterforce to intermediation, and studies the 2016 market crisis as a shock to trust in lending. Section V looks into the role of active investors. Section VI concludes.

II. Institutional Details and Data

A. P2P Lending in the United States

P2P lending platforms, the earliest type of FinTech lenders, appeared in the United States in 2005. They were created to allow investors to buy small unsecured loans that consumers could request online. The idea proved successful, so much so that FinTech lenders' share in the unsecured personal loan market grew from 5% in 2013 to 36% in 2019 ("Fintech Continues to Disrupt Consumer Lending," Morgan Stanley, referencing TransUnion, Apr. 29, 2020).

P2P lending in the United States has historically been dominated by two major platforms—Prosper and LendingClub (LC)—whose joint market share was 98% in 2014 and 67% in 2018 (https://www.economist.com/finance-and-economics/2014/02/28/banking-without-banks (*The Economist*) and https://brismo.com/market-data (Brismo Analytics) (accessed Jan. 16, 2021)). Our tests are based on data from Prosper, as it is the only platform that provides the details of the investor pools that we need for our analysis.

B. Prosper Marketplace

Prosper was the pioneer of the U.S. P2P loan market, and it is currently one of the three largest P2P lending platforms in the United States. It provides fully amortizing unsecured loans of up to $40,000 to individuals, mostly for debt refinancing. The loans have fixed rates between 5% and 35% and 3- or 5-year maturities. Funding is provided by individual and institutional investors who act as lenders on the platform. Unlike bank depositors, P2P lenders invest in portfolios of specific P2P loans and bear the credit risks of these loans (see Balyuk (2022)).

1. Prosper's Model over Time

We divide Prosper's history into three periods, namely, disintermediation (2005–2010), transition (2011–2012), and reintermediation (2013–2019). We briefly describe these periods here and provide more details in Section B of the Supplementary Material.

For years after its creation in late 2005, Prosper's role in loan origination was mostly limited to providing the infrastructure for borrowers to request loans online and for potential lenders to invest in loans of their choosing. The platform imposed eligibility requirements, collected borrowers' information, and facilitated payments. Loans were allocated to investors via auctions, eBay-style. Lenders evaluated loan applications and manually placed bids on the loans they wanted to invest in. The bids indicated lenders' funding commitments for each loan and the interest rate set by the lenders. Because of the platform's limited role in loan pricing and screening, we call the 2005–2010 period the *disintermediation* period.[7]

In Dec. 2010, Prosper abandoned the auction format and started to assign loan interest rates (often referred to as the "posted" pricing system). The platform also took a more active role in underwriting and collection efforts, verifying borrowers' information and canceling some applications. The change to posted prices marks the platform's gradual transition to centralized decision-making. We call the 2011–2012 period the *transition* period.

In 2013, Prosper introduced a new approach to credit assessment and essentially finalized the transition to the platform-centric model, which has been in use since then. Under this system, Prosper automatically carries out all steps of the lending process, including application handling, data gathering and verification, credit scoring, loan servicing, and collections in case of default.[8] Investors can only choose between funding the loan on the conditions specified by the platform and rejecting it altogether. We refer to the period after 2013 as the *reintermediation* period.

[7]Prosper's data set contains limited data for 2005–2006, and therefore we focus on 2007–2008 in most tests when examining the platform's role during this period. Also, to ensure comparability, we restrict the matched sample to end in Mar. 2017 for tests that require borrower characteristics, because Prosper switched to another credit bureau in Apr. 2017 and changed the set of credit bureau variables it reports.

[8]Although Prosper attempts to collect on charged-off P2P loans and charges debt collection fees (17%–40% of the recovered amount), collection rates have historically been very low. For example, the recovery rate on 88.86% of delinquent Prosper loans that were charged off between July 2009 and Dec. 2014 was only 1.63% (Prosper's 10-K filing from Dec. 31, 2014). Therefore, the default risk is salient for P2P investors.

2. Loan Evaluation

Currently, all steps in the loan origination process on Prosper's side are performed automatically by the platform's software, and manual interventions in the decision-making are expressly prohibited. For a loan application to be listed on the platform, it must satisfy certain eligibility requirements (e.g., the borrower cannot reside in certain states or have a FICO score outside of a certain range). Prosper automatically verifies compliance with these criteria. For loans that satisfy the requirements, the platform uses a proprietary credit-scoring model to evaluate the estimated loss rate (ELR) from default as a function of the borrower's credit report and self-reported data.[9] The ELR is a sufficient statistic that summarizes Prosper's assessment of the loan's default risk and, on any given date, fully determines the interest rate on the loan as well as the rating assigned to the loan by Prosper.

Priced loans are posted on Prosper's web platform, and at that point become eligible for investment. Investors can decide whether to fund a loan or not, but they do not set loan interest rates. We refer to this funding decision as *loan screening by investors.* As investors do not know the identity of the borrowers, they have no inside information about the loan quality, nor can they perform additional due diligence.

At the same time, before funds are transmitted to borrowers, for a subset of listed loan applications Prosper conducts a "post-funding review," whereby the platform attempts to automatically verify certain self-reported variables, such as income and employment, and performs some additional proprietary loan analysis. If as a result of these checks it concludes that the loan's risk is too high for the interest rate assigned by the pricing algorithm (perhaps due to suspected fraud or misreporting), then the loan application is "canceled" by Prosper.[10] We refer to this function of canceling suspicious loans as *loan screening by the platform.*

A loan application fails to result in loan origination if it is withdrawn by the borrower, expires unfunded, or is canceled by the platform as part of its screening process. In practice, the most important of these factors by far is loan cancellations by Prosper, which account for 87.4% of failed applications.

To summarize, loan pricing is conducted exclusively by the platform. By contrast, both the platform and the (active) investors can perform loan screening (i.e., decide which loans should remain unfunded). Investors screen loans by providing or denying funding, whereas Prosper does so by conducting a review and canceling loan applications it deems untrustworthy. Of note, Prosper performs

---

[9]While in early years loan listings included "soft" or nonstandard information, such data were later removed from the platform, so that currently loan analysis is based exclusively on "hard" variables, reducing subjectivity and facilitating automated quantitative modeling (see, e.g., Duarte, Siegel, and Young (2012), Lin et al. (2013), Iyer et al. (2016), and Balyuk (2022)).

[10]According to Prosper, the algorithm is supposed to flag loans for which the borrower's self-reported income is "highly determinative" of the Prosper rating. Prosper reportedly verified income and/or employment information self-reported by borrowers for 58% of the originated loans on a unit basis and approximately 72% of originated loans on a dollar basis between July 13, 2009 and Mar. 31, 2016. Subsequent to such verification, Prosper canceled 11% of loan applications solely on the grounds of inaccurate or insufficient information (Prosper Prospectus dated May 24, 2016). Therefore, less than half of all cancellations can be attributed to possible fraud or misstatements by borrowers.

https://doi.org/10.1017/S0022109023000789 Published online by Cambridge University Press

this review *after* it lists applications on its platform, presumably to speed up the loan origination process.[11] As we report below, in practice, investors are overwhelmingly passive, and therefore whether Prosper conducts its screening before or after investors commit the funding for the loans appears inconsequential. Indeed, agreeing to fund almost all loans on offer signals investors' confidence that Prosper will screen out untrustworthy ones in due time.

3. Investment Pools

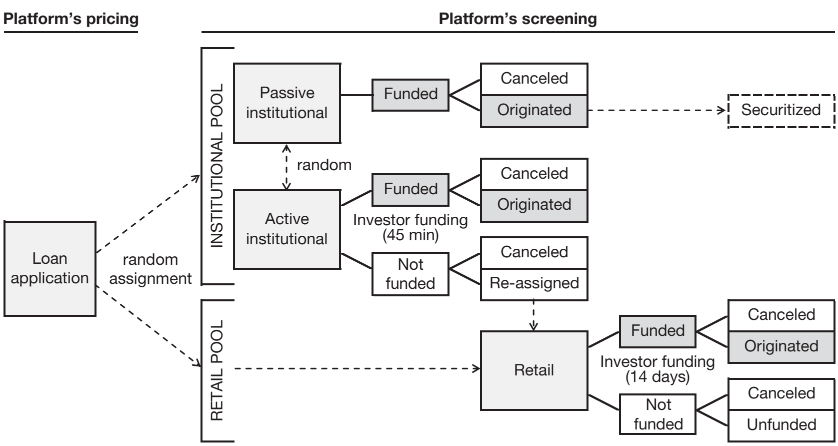
We summarize Prosper's lending process in Figure 1. A loan can be funded through one of three investor pools. Retail lenders invest in whole loans or fractions thereof ("crowdfund" the loans) through the so-called "Fractional" (*retail*) investment pool. They can decide which individual loans they are willing to fund, in the spirit of "P2P" lending, or they can instruct Prosper to invest on their behalf using tools that Prosper provides for this purpose (e.g., Quick Invest and AutoInvest).

In addition to retail investors, lenders on Prosper include institutions, which typically fund loans in full through the "Whole Loan" pool, first introduced in Apr. 2013. Institutional investors pursue active investment strategies by reviewing individual loans and choosing which ones to fund through the "Whole Active" (also known as *active institutional*) pool. In Nov. 2013, Prosper also introduced a passive investment possibility for institutions by creating a separate "Whole Passive" (*passive institutional*) pool, where institutions may invest by instructing Prosper to automatically fund loans with certain characteristics on their behalf. Passive institutional lenders typically specify very broad criteria for their desired investments

FIGURE 1

Lending Process on Prosper

Figure 1 shows the main steps in the loan origination process on Prosper Marketplace.



[11] Anecdotal evidence suggests that most cancellations occur after investors have reviewed loan applications and decided on funding, although some may occur while applications are still open for funding. See, for example, https://prosper.zendesk.com/hc/en-us/articles/208500716-What-happens-after-I-make-my-investments.

and Prosper then automatically allocates the investors' money to loans that meet these criteria. Reportedly, some passive institutional investors, such as BlackRock and Citigroup, further securitize Prosper loans.

Prosper purports to randomly allocate new loan applications across the three investor pools, based on the platform's estimate of the relative investor demand in each pool. During our sample period, Prosper had a policy of "recycling" loan applications rejected by institutional investors by offering them again for funding to retail investors in the fractional pool. These "recycled" loans were clearly marked as such, so that investors could easily avoid them if they wanted.[12]

C. Data Description

1. Prosper's Data Sets

To facilitate investors' loan analysis and decision-making, Prosper makes two data sets available on its website, one detailing loan applications ("listings") and one describing the subsequent performance of originated loans. The listings data include application dates, details of requested loans (e.g., amount and maturity), borrower-reported data (e.g., income and employment), and detailed data from consumers' credit bureau reports (Experian before Mar. 2017 and TransUnion thereafter),[13] which Prosper obtains at the time of application. Also included is information about the investment pool, percent funded, and application status (i.e., originated, canceled, expired, or withdrawn) for each application, as is the loan origination date. After applying several data filters,[14] our data set consists of 1,478,138 loan applications submitted between Feb. 2007 and Mar. 2019, which resulted in 1,076,881 originated loans. The loan data set contains information pertaining to originated loans, such as repayment status and amount repaid. We observe the subsequent performance of the loans through Dec. 2018.

To relate subsequent loan returns to borrower characteristics for originated loans, it is necessary to merge loan performance and listing details files, which are not linked in Prosper. To this end, we match loans in the two data sets on the variables common to both (i.e., loan origination date, amount, rate, Prosper rating, and maturity). If there is a unique match, the loan is classified as matched. If there are several listings that can be matched to a particular loan or vice versa, we classify all these observations as unmatched. This approach allows us to match 95% of loans before 2011, but this proportion decreases in later years as the number of listings

[12]We discuss random allocation of loan applications and "recycled" loans in greater detail in Section V.D.

[13]Unfortunately, it is not possible to precisely map variables in credit reports that Prosper obtains from the two credit bureaus to one another because of differences in the sets of variables and their definitions. As a result, we restrict the sample period for tests involving borrower characteristics to end in Mar. 2017.

[14]To select our sample, we first remove any duplicates from each data set, retaining only the last entry for each loan. We also remove all applications that do not meet Prosper's eligibility criteria, loans without a Prosper rating or with a Prosper score over 11 (the highest possible), as well as 1-year maturity loans, which Prosper stopped originating in Apr. 2013. We also exclude loan applications from Iowa, Maine, North Dakota, and Puerto Rico because borrowers residing in these states were prohibited from using Prosper for a substantial portion of our sample period and their applications could not result in loan origination.

https://doi.org/10.1017/S0022109023000789 Published online by Cambridge University Press

rises, so that only 55% of loans originated after 2016 are matched uniquely. Our matched loan sample consists of 550,577 originated loans. We restrict our analysis to this matched sample in those tests that combine variables from the listings and performance data sets. However, few of our tests rely on data from the two data sets simultaneously. Therefore, to maximize the sample size, we use the full performance and listings data sets for most of the analysis.[15] Section C of the Supplementary Material discusses the possibility of a bias due to imperfect matching. As best as we can ascertain, the matching bias, if any, is negligible and does not affect any of our results.

2. Descriptive Statistics

Table 1 provides descriptive statistics for all applications and originated loans on Prosper. The median loan size is \$12,000. Loans in our sample are amortized over either 36 or 60 months, with the former comprising 71% of all loans. The median interest rate on originated loans is 14.1%, and the median ELR is 5.99%. Prosper loans are primarily requested for debt repayment and debt consolidation purposes (72.6%). Other loan purposes include home improvement

TABLE 1

Descriptive Statistics

Table 1 reports loan and borrower characteristics for loan applications and for originated P2P loans. The sample is 2007–2019. See Section I of the Supplementary Material for the definitions of variables.

	Applications				Matched Loans			
	Mean	Median	Std. Dev.	N	Mean	Median	Std. Dev.	N
	1	2	3	4	5	6	7	8
Panel A. Loan Characteristics								
LOAN_AMOUNT	13,264	12,000	8,099	14,78,138	12,693	10,000	8,421	5,50,577
LOAN_MATURITY	42.9	36.0	10.9	14,78,138	44.0	36.0	11.3	5,50,577
INTEREST_RATE	15.7%	14.1%	6.8%	14,78,138	15.3%	14.1%	6.4%	5,50,577
ESTIMATED_LOSS_RATE	6.89%	5.99%	4.23%	14,78,138	6.64%	5.99%	3.84%	5,50,577
RATING (1-7)	4.47	5.00	1.55	14,78,138	4.54	5.00	1.46	5,50,577
CONSOLIDATION_PURPOSE	72.6%	—	—	14,78,138	72.3%	—	—	5,50,577
Panel B. Borrower Characteristics								
FICO_SCORE	703.5	689.5	40.0	13,83,377	706.3	709.5	40.4	5,16,268
SCOREX_SCORE	707.7	712.5	60.5	9,88,856	710.5	712.5	59.2	3,43,908
EMPLOYED	81.4%	–	–	14,78,138	84.1%	–	–	5,50,577
SELF_EMPLOYED	7.9%	–	–	14,78,138	6.2%	–	–	5,50,577
YEARS_EMPLOYED	9.09	6.17	10.96	14,68,987	9.17	6.33	10.64	5,49,333
MONTHLY_INCOME	6,367	5,253	4,192	14,78,131	6,204	5,200	3,906	5,50,575
DEBT_TO_INCOME	0.26	0.24	0.70	14,41,782	0.26	0.24	1.14	5,40,197
MORTGAGE	32.2%	–	–	14,78,138	31.4%	–	–	5,50,577
TRADES_LAST_6M	0.98	1.00	1.18	14,78,081	0.96	1.00	1.16	5,50,552
ACTIVE_CREDIT_CARDS	4.77	4.00	2.88	13,83,319	4.68	4.00	2.85	5,16,242
CREDIT_CARD_BALANCE	5,490	4,172	4,351	13,83,319	5,554	4,272	4,312	5,16,242
TRADES_EVER_DELIQUENT	8.01%	–	–	14,78,081	7.83%	–	–	5,50,552
PAST_BANKRUPTCIES	0.17	0.00	0.38	13,83,319	0.17	0.00	0.38	5,16,242
PRIOR_LOAN	18.1%	—	—	14,78,138	24.4%	—	—	5,50,577
Panel C. Loan Performance								
REALIZED_RETURN	NA	—	—	—	6.51%	10.26%	15.04%	3,44,288
REALIZED_DEFAULT_RATE	NA	—	—	—	6.08%	—	—	5,50,577

[15]Specifically, we use the listing-loan matched sample in Panel C of Table 1, columns 5–8 of Table 4, Table 7, and Table D.1 in the Supplementary Material. Everywhere else we use the full data sets after applying the filters described in 14.

(9.3%), medical bills payment (2.9%), small business purposes (2.1%), and large purchases (1.7%). The median borrower's FICO (SCOREX) score is 689.5 (712.5), and the median self-reported annual income is $63,036. For comparison, the median FICO score in the general population during this period was 700, and the per-capita income in the United States in 2016 was $43,183, according to the Bureau of Labor Statistics.

There are few differences between the characteristics of loan applications and those of originated loans evident from Table 1. One notable exception is that applications from borrowers that have at least one prior Prosper loan are more likely to be successful. Thus, the platform's algorithms are less likely to cancel a loan application if the borrower already has an outstanding Prosper loan, presumably because the borrower was screened previously.

III. Market Structure and Its Evolution

A. Current Market Structure

Table 2 summarizes some key variables that describe the structure of the market in the current (reintermediation) period and the role played by the platform and investors in loan evaluation. In addition to the aggregate numbers, we report statistics by Prosper rating, which ranges from the safest AA category (ELR $<$ 2%) to the riskiest HR category (ELR $>$ 15%).

Panel A of Table 2 reports the proportion of the funds provided through the three investment pools. The first thing to note is that institutional investors dominate the market. Institutions supply 91.8% of funds, whereas retail "peers" contribute a mere 8.2% of the funding. A large majority of institutional investors are explicitly passive: 75.9% of funding is extended through the passive institutional pool, corresponding to 82.6% of all institutional funds. The fractions of institutional

TABLE 2

Platform's Screening, Investor Pools, and Funding Rates

Table 2 reports the fraction of each investment pool, application funding rates, and cancellation rates, for all loan applications and by Prosper rating. The sample period is 2013–2019. See Section I of the Supplementary Material for the definitions of variables. All fractions are value-weighted (by funded loan amount).

	All Ratings	AA	A	B	C	D	E	HR
	1	2	3	4	5	6	7	8
Panel A. Platform's Pricing and Screening								
CANCELLATION_RATE	24.42%	20.56%	24.22%	24.79%	25.09%	26.01%	23.87%	23.83%
INTEREST_RATE	15.24%	6.81%	9.50%	12.38%	16.59%	22.38%	27.62%	31.13%
ELR	6.59%	1.42%	3.18%	5.12%	7.49%	10.44%	13.43%	16.85%
Panel B. Allocation of Funds Across Investment Pools								
PASSIVE_INST	75.9%	76.7%	76.9%	75.8%	76.1%	74.2%	74.6%	71.2%
ACTIVE_INST	15.9%	12.9%	15.8%	17.5%	16.7%	15.6%	11.9%	7.1%
RETAIL	8.2%	10.4%	7.3%	6.7%	7.2%	10.2%	13.5%	21.7%
Panel C. Funding Rates								
PASSIVE_INST	99.6%	99.6%	99.6%	99.6%	99.6%	99.5%	99.6%	99.6%
ACTIVE_INST	90.6%	78.4%	84.6%	95.3%	94.5%	92.1%	93.7%	79.4%
RETAIL	91.8%	89.7%	86.6%	91.8%	92.6%	93.1%	95.3%	96.9%

funding and passive investing are universally high across the rating categories, except for a somewhat higher fraction of retail funds invested in the riskiest HR-rated loans.

Panel B of Table 2 reports the proportion of loan applications that investors agree to fund. Overall, 98.6% of loan applications receive funding, meaning that only 1.4% of them are rejected by the investors. The funding rate is close to 100% in the passive institutional pool, which is 3 times as large as the other two pools combined. But even active institutional lenders, which can select the individual loans to invest in, agree to fund 90.6% of the loans. This fraction is even higher for retail investors, at 91.8%. Looking at the funding rates in detail, we find the fraction of funded loans to vary somewhat across credit quality and investor pools, possibly reflecting investors' preferences and loan-picking skills, but the funding rates are very high across the board.

Investors' passivity stands in sharp contrast to the active role played by the platform. As described above, Prosper's software not only estimates the expected loss rate due to default (the ELR) and sets the loan interest rate, but also screens out loans on the extensive margin by canceling applications that in its judgment may be riskier than they appear. Panel C of Table 2 reports that Prosper cancels on average 24.4% of the loans. Across the different ratings, Prosper screens out between one-fifth and one-quarter of the loans. Comparing these cancellation rates with the investors' loan rejection rate of 1.4%, it can be seen that it is Prosper that does almost all of the loan screening, in addition to pricing the loans.

Thus, the investors not only have no ability to set loan interest rates, but also largely forgo the right to screen the loans on the extensive margin. As a result, their decision is not so much about which loans to buy, but rather whether to participate in the market or not. Those who do decide to participate by and large simply buy almost all loans offered to them at prices set by Prosper. Most investors fully outsource all loan evaluation functions to the platform's algorithms. Dominated by passive institutional investors, this market structure is a far cry from the original vision of "P2P" lending.

B. Transition Toward Reintermediation

The current centralized market structure is a result of an evolutionary process rather than an original design. In this section, we document the evolution of some key market characteristics through its transition toward reintermediation.

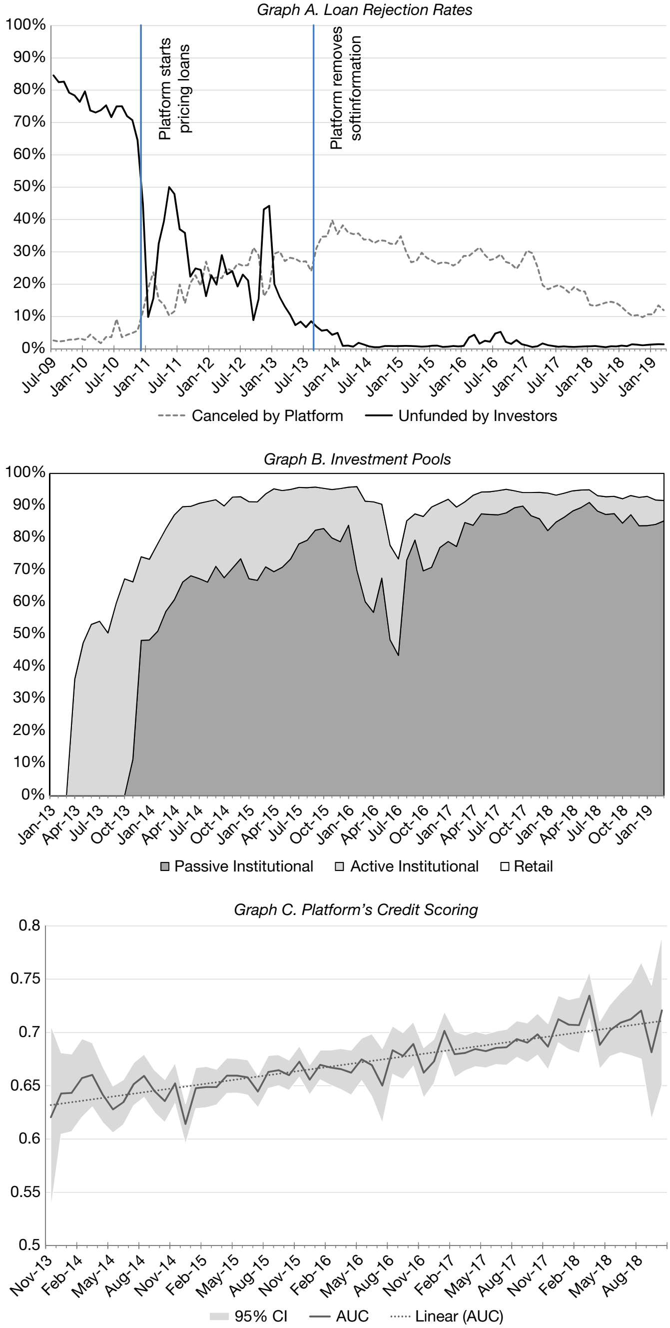
Graph A of Figure 2 compares the trends in the proportion of loans screened out by investors (by refusing to provide funding) and by the platform (through cancellations). During the disintermediation period, average funding rates were below 30%. It was not until 2013 that investors' funding rates increased to over 90%, and eventually reached their current levels in excess of 98%.[16] At the same time, Prosper stepped up cancellations dramatically when it abandoned the auction

[16]We note that unfunded loans dropped rapidly after Prosper switched from the auction model to posted prices in Dec. 2010. Wei and Lin (2016) also find this pattern. However, they explain the increase in funding by higher borrowing rates, whereas we argue that investors' loan screening was replaced by Prosper's. The platform's decision to step up cancellations likely increased investor confidence in funding the loans.

FIGURE 2

Trends in Rejection Rates, Investment Pools, and Platform's Credit Scoring

Figure 2 shows loan rejection rates, investment pools, and Prosper's credit scoring over time. Graph A contrasts the rates of loan rejection (i.e., extensive margin of screening) by the P2P lending platform and by investors. Graph B shows the evolution of the three investor pools after these pools were created in 2013. Graph C shows the area under curve (AUC), which measures the ability of Prosper's estimated loss rate to discriminate between defaulting and nondefaulting loans. For each month, AUC is calculated using all loans originated in the passive investment pool in that month. Also shown are the linear time trend for the series and the 95% confidence intervals.



model and switched to platform-based pricing in Dec. 2010. In effect, at that point, loan screening by Prosper almost entirely replaced that by lenders.

Interestingly, loan cancellation rates peaked at 40% in late 2013, but have been trending down since then. They averaged 13.5% in 2018–2019, compared with 31.3% in 2014–2015. By itself, this trend may suggest that Prosper has been increasingly unwilling to screen loans prudently, but it is also consistent with gradual improvements in loan evaluation technology that allow for more accurate pricing of risk, obviating the need for loan denials. Nonetheless, denials by the platform are still an order of magnitude more frequent than those by investors.

Graph B of Figure 2 plots the composition of the three investor pools since the institutional pool was created in Apr. 2013 and subsequently split into the passive and active pools in Nov. 2013. With the exception of the period of the market crisis in 2016, which we discuss in detail in Section IV, the share of passive institutional funds has been trending up, and reached 85.8% of total funds by 2018–2019.[17] By contrast, active institutional investment has declined from 21.8% in 2014–2015 to 7.2% currently. The proportion of retail funding went up briefly to 30.3% in July 2016, driven by the drop in institutional credit. By the end of 2016, the trend was reversed, with only 7.7% of loans funded through the retail investor pool throughout 2017–2019. The growth in passive funding volume outpaced the decline in active funding, which suggests that new passive investors have been attracted to the market. Rather than screening loans themselves, these investors are apparently comfortable with Prosper's software performing these functions on their behalf.

Graph C of Figure 2 illustrates how the quality of Prosper's credit risk model changed over time.[18] To assess the cross-sectional accuracy of the model, we compute the area under curve (AUC) using the ELR as the sufficient statistic that summarizes Prosper's assessment of the loan's default risk.[19] The AUC for loans originated in the passive pool in 2013–2018 is 0.662, which indicates relatively good accuracy. For comparison, this statistic is only 0.612 when FICO bins are used as the default predictor. Moreover, the AUC has been increasing steadily, suggesting improvements in Prosper's default model. In fact, the lower bound of the confidence interval for AUC was higher in 2017–2018 than its upper bound in 2013.[20]

[17]Although we do not have data on the use of passive strategies by retail investors, indirect evidence suggests that a sizable and increasing fraction of them have been using Auto Invest and other tools that allow them to automatically purchase fractions of loans that satisfy certain criteria. As an indicator of the use of Auto Invest, we exploit the fact that it evaluates loan applications for inclusion in retail investors' portfolios at a few specific points in time every day when loan applications are listed. We estimate the proportion of loan applications that are funded within the first 3 minutes of loan listing times, and find that the share of "passively funded" loans for the smallest loan size in the retail pool has grown from less than 2% in 2011 to 26.5% in 2014 and 48.3% in 2019, with a steady upward trend. The details of these tests are available from the authors.

[18]We examine the effectiveness of Proper's cancelations in Section D of the Supplementary Material.

[19]AUC, or the area under the receiver operating characteristic curve, is a standard measure of classification accuracy for binary classifiers such as predictors of default (see Stein (2007)). It varies from 0.5 for a random classifier to 1.0 for a perfect predictor that makes no classification errors. Iyer et al. (2016) note that, "AUC of 0.6 or greater is generally considered desirable in information-scarce environments, and AUCs of 0.7 or greater are the goal in more information-rich environments…. a 0.01 improvement in AUC is considered a noteworthy gain in the credit scoring industry."

[20]Applying a similar methodology to LC's data, Vallee and Zeng (2019) find that the accuracy of its credit model also increased over time.

To summarize, Prosper's technical expertise in loan evaluation has been growing over time. The trend toward greater use of passive strategies is consistent with the growing number of investors relying on Prosper's loan analysis. Given that passive investors rarely reject loan applications, this trend has resulted in funding rates in excess of 98.5%. But as investors all but ceased to perform any active role in loan evaluation, Prosper stepped up loan cancellations to screen out low-quality loans. In effect, as the market transitioned toward reintermediation, Prosper's expertise displaced that of investors, a large majority of whom have become passive.

C. Intermediation in Other P2P Platforms

The tendency toward the platform-centric market model is not unique to Prosper or the U.S. P2P lending market. To establish general trends regarding centralization in P2P lending, we manually collect historical information on the organization of other P2P platforms in the United States and Europe. We start with the list of platforms from the International P2P Lending Volumes Report by P2P-Banking.com and supplement it with information on other major platforms mentioned in the press.[21] Our final sample consists of 23 platforms in the United States and Europe, excluding Prosper. We obtain historical data on the business models of these platforms from their websites, financial statements, and the Internet, including historical snapshots through Wayback Machine.

Figure 3 shows the number of major P2P platforms in the United States and Europe over time. As can be seen, the proportion of platforms that use centralized loan pricing has been growing consistently. None of the platforms in our sample can currently be characterized as fully decentralized, and only 15% employ a mixed, partially decentralized model. The transition of several platforms from decentralized to centralized or mixed systems (e.g., Funding Circle and ThinCats in the United Kingdom, Bondora in Estonia, Linked Finance in Ireland, Upstart in the United States) and the emergence of new P2P lending platforms with the centralized system (e.g., Twino in Latvia and Bolden in France) contribute to these trends.[22] Specifically, 9 out of 23 platforms in our international sample transitioned from decentralized to centralized or mixed systems (in different years), 13 were created with the centralized system in place (9 created in 2013–2015), and 1 had a mixed system.

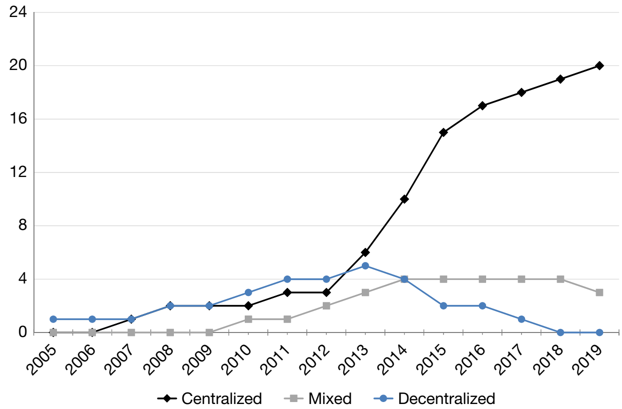
[21] We apply two filters to the list of P2P lending platforms. First, we keep only platforms with at least \$2 million in monthly originations (as of Feb. 2018). This restriction ensures that we examine platforms of meaningful size and with functional business. It also allows us to focus on platforms that are more likely to disclose information about their operations. Second, we exclude platforms that sell loans from loan originators. Although these platforms are frequently classified as P2P platforms, they are in fact online trading venues where originators resell loans, which they have screened. Thus, these platforms can be thought of as secondary markets for already originated loans.

[22] These centralized and mixed pricing models employ different rate-setting approaches, such as application-specific rates based on credit risk algorithms (e.g., Nucleus in the United Kingdom, Twino in Latvia, October in France, and Upstart in the United States) or the same rates per bucket of applications based on loan grade and term (e.g., Linked Finance in Ireland). Although the latter mechanisms differ from Prosper's rate setting, the common feature is that P2P loan rates are increasingly set in a fashion similar to how traditional intermediaries price loans.

FIGURE 3

Number of P2P Platforms by Type

Figure 3 shows the number of P2P lending platforms operating in the United States and Europe over time, by type. We classify P2P lending as *centralized* if the platform performs loan evaluation and pricing (possibly in negotiations with borrowers), *decentralized* if investors determine loan rates (e.g., via auctions), and *mixed* if loan rates are partly determined by the platform (e.g., when auction-based and centralized pricing coexist or when the platform adjusts the loan rate for supply and demand within rate bounds).



Although public data on the proportions of institutional investors and passive strategies on P2P lending platforms are very fragmented, we briefly summarize the anecdotal evidence we have been able to collect. We first note that institutional investors dominate marketplace lending not only on Prosper, but also elsewhere, including LC in the United States (75%), Funding Circle in the United Kingdom (68%), and October in France (68%). We then examine the presence of passive investors. Similar to the dominance of institutional investors, we observe the dominance of passive investment strategies in other P2P lending platforms inside and outside the United States. For example, Funding Circle in the United Kingdom eliminated the possibility of manual loan picking from its platform in Sept. 2017; only investment in select loan portfolios is allowed. On October's platform in France, institutional investors can only invest through a fund. Yet it is not uncommon for P2P platforms to maintain active loan selection options for investors. For example, the robo-investment plan on Renrendai in China (Uplan) constitutes only 67% of all investments. While 96% of investments on OnDeck in the United States are passive (e.g., investments in funds and securitizations), the platform still maintains a marketplace where active institutional investors can form their own loan portfolios.

Overall, there has been a clear trend toward abandoning any possibility of price discovery by investors in favor of prices set by the platform. As a result, marketplace lending has been increasingly intermediated by technology-based agents inside and outside the United States.

D. Deviations from Full Intermediation

The original decentralization of online lending and its subsequent reintermediation are generally consistent with the theory. In the model by Vallee and

Zeng (2019), in its early years, the P2P lending market is decentralized, and the lending platform relies on the wisdom of the crowds for loan analysis. But as the platform's expertise grows, at some point, it becomes possible even for unsophisticated investors to earn adequate returns by fully relying on the platform's loan analysis. At this point, the market equilibrium switches to a fully centralized model. Specifically, the platform is predicted to decrease the provision of raw loan data, so that even sophisticated investors cannot profit from in-house loan analysis. As a result, all investors become fully passive and henceforth resemble traditional bank depositors, with no information about and no discretion over the individual loans.

While the prediction that over time the platform should take a more active role in loan evaluation while investors should become passive is consistent with our findings, the observed market structure differs from theoretical predictions and from that of most other intermediated markets in several key aspects. We point out two deviations from the fully reintermediated equilibrium, which differentiate P2P lending platforms from traditional intermediaries.

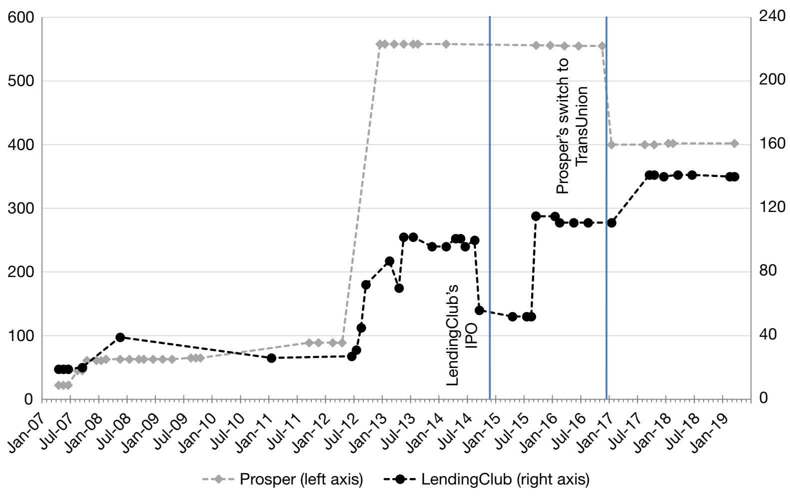
First, a fraction of investors still maintain the ability to actively select the loans that they want to fund. We discuss the behavior of active investors in detail in Section V. Second, in contrast to banks, Prosper continues to provide investors with detailed information about the individual loans.

To document the trends in the provision of loan data, we manually collect data on the number of variables describing loans originated previously and those on offer at any given time, which Prosper and LC provided to investors over time. The results are shown in Figure 4, and further details can be found in the Appendix.

FIGURE 4

Trends in Information Provision

Figure 4 shows the evolution of the number of variables per originated loan that Prosper (left axis) and LendingClub (right axis) provide to investors in downloadable historical data files. We manually collect data from historical data files, loan dictionaries, and blog posts and exclude variables with all missing values for a particular month. The loan application data provided to investors for real-time investing display similar trends (not shown).



https://doi.org/10.1017/S0022109023000789 Published online by Cambridge University Press

Contrary to the predictions of Vallee and Zeng (2019), over time, both platforms have increased, rather than reduced, their data provision to investors. Specifically, during the disintermediation period, Prosper disclosed 22 variables to investors, but this number grew to over 400 variables during the reintermediation period. The number of variables did go down from the previous high of 555 in Jan. 2017, but this drop was caused by the fact that at that time Prosper switched the credit bureau it partnered with from Experian to TransUnion, and the variables collated by the latter are somewhat different (but still exhaustive). We observe a similar general trend toward greater transparency for LC, which between 2014 and 2019 increased the number of variables disclosed by 40%.[23]

The hybrid, not-quite-fully reintermediated market structure that we describe looks puzzling and uneconomical. Passive investors provide more than three quarters of the total funding, fully rely on Prosper's analysis, and earn returns not statistically different from those of active investors. Why then does the lending platform maintain the (presumably costly) infrastructure that enables investors to evaluate and select loans for investment, even though in practice only a very small fraction of loan applications are rejected? One could argue that if Prosper switched to the fully intermediated model, the few remaining active investors would adopt passive strategies rather than abandon the platform, and their returns would hardly change as a result. Yet there seem to be factors that discourage full reintermediation in P2P lending. In Section IV, we study the benefits of this structure.

IV. Moral Hazard and a Study of the 2016 Crisis

Given that lending platforms make money by charging loan origination fees, but typically maintain little or no skin in the game, they may be tempted to relax lending standards to boost volume, which can in turn destabilize the market. This moral hazard problem is fundamental in financial intermediation (Diamond (1984)), and its effects in originate-to-distribute debt markets, such as those for securitized products, have been identified as one of the causes of the Great Financial Crisis (e.g., Keys et al. (2010), Purnanandam (2011)). Similar to originate-to-distribute lenders, P2P lending platforms sell most loans they originate to investors instead of keeping them on the balance sheet.[24] This raises the question: When investors rely on the lending platform's assessment of the loans, who "watches the watcher"?

We hypothesize that it is the active investors who perform this monitoring role. Thakor and Merton (2019) construct a model of P2P lending, and argue that

[23]Vallee and Zeng (2019) cite an episode in Nov. 2014, when LC decreased the number of variables it provided to investors, as evidence in support of their prediction that lending platforms will actively try to discourage active strategies by becoming less transparent. In reality, the Nov. 2014 decrease in data provision proved to be a temporary aberration in anticipation of LC's IPO, and the decline in the number of variables provided was fully reversed within 1 year.

[24]Prosper does not typically invest in loans it originates, and LC invested in only 0.16% of the total value of its loans between 2007 and 2019.

investors who cannot independently verify the loan quality must *trust* the platform's analysis before they agree to lend. If the loans underperform, investors' trust can be eroded, and they may abandon the platform and precipitate its collapse, even if the underperformance is due to bad luck rather than the platform's negligence. We argue that if enough information is available about the loans, active investors can perform their own loan credit analysis instead of relying on trust. If they can confirm that the sufficient quality of the loan book is being maintained, their presence can serve as a buffer during episodes when passive investors' trust is eroded, making them abandon the platform.

We provide evidence in support of this conjecture by documenting the events around the market's near collapse in the first half of 2016, focusing in particular on the differential reaction of active versus passive investors. Later, we also demonstrate that active investors react to trends in the performance of outstanding loans detectable from the loan data. Together, these tests show that the presence of active investors can play a stabilizing role, while their monitoring can incentivize the platform to provide high-quality loan analysis.

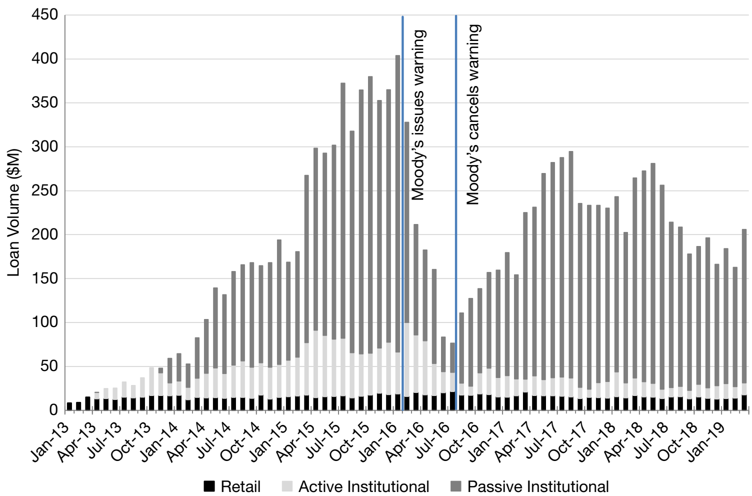
A. The 2016 Crisis as a Natural Experiment

On Feb. 11, 2016, the rating agency Moody's placed Citibank's securitizations of Prosper loans under review for downgrade, and revised upward its loss expectations for Blackrock's securitizations. The downgrade warning from Moody's precipitated a collapse in new loan originations. Over the first half of 2016, new loan volume on Prosper fell by 83%, erasing several years' worth of cumulative growth. As shown in Figure 5, and as we discuss in detail below, the drop in funding

FIGURE 5

Loan Originations on Prosper

Figure 5 shows the volume of newly originated loans by month. The vertical lines represent the 2016 crisis, from the initiation of Moody's downgrade warning on Feb. 11, 2016 to the warning's cancellation on July 15, 2016.



https://doi.org/10.1017/S0022109023000789 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109023000789 Published online by Cambridge University Press

was disproportionally large in the passive institutional pool. At the same time, the fraction of loans that failed to secure funding increased substantially. Moody's withdrew its warning in July 2016, and no downgrade took place. Following the withdrawal, loan volume increased, but it has not fully recovered as of the time of this study.

We use Moody's warning as a quasi-exogenous shock to trust in Prosper's loan quality in a DiD framework. Several factors make this episode an interesting laboratory for a DiD analysis. First, the shock was unanticipated. The possibility of this move had not been mentioned in media reports before the actual announcement. We also find no evidence of any pre-trends related to the shock in the data. Second, this announcement referred to loans that were originated and securitized almost a *year before*. Therefore, the announcement pertained to loans that were already held by investors, rather than to new originations, and it should have had no direct effect on newly originated loans other than through the channels we focus on. Third, the rich data on loan applications, funded loans, and platform's actions (e.g., pricing and cancellations) allow us to distinguish between the responses of different sets of actors (i.e., investors vs. the platform) to the shock, accounting for possible changes in the applicant pool.[25]

Notes on the Empirical Methodology

We conduct several sets of DiD analyses in two windows around the shock: ±2 months and ±6 months. Passive investors are the treatment group, and active investors are broadly the control group. One set of tests examines the probability of a loan remaining unfunded by investors. These regressions are estimated at the loan level. They contain ELR fixed effects to account for differences in the credit quality in different months. They also control for the loan rate because Prosper adjusted the ELR–interest rate mapping during this period (see Section IV.C).[26] We cluster standard errors at the ELR level because of the possible correlation of loan risk within ELR bins, but the results are robust to clustering at the month-of-origination level (see Petersen (2009) and Abadie, Athey, Imbens, and Wooldridge (2023) for guidance on clustering). Loan-level regressions are also used to study changes in the quality of loans that investors fund after the shock. For this purpose, we examine measures of loan quality (i.e., ELR and FICO scores) and loan performance (i.e., realized returns and realized default rates) as the dependent variables in loan-level DiD analyses. Additionally, we conduct a series of loan-level heterogeneity tests where we interact the treatment indicator with measures of loan types and funding volume. One empirical challenge with loan-level regressions around Moody's warning is that Prosper purports to balance loan supply and demand based on its estimates of the relative investor demand in each pool. Thus, one should expect the platform to react to unexpected withdrawal of treated

[25]We report the results of the parallel trend analyses and falsification tests in Section IV.B to support the validity of this shock.

[26]As discussed earlier, ELR is a function of credit bureau data and borrower self-reported data. We do not include these other variables together with the ELR in regressions to avoid multicollinearity issues.

investors by allocating more applications to untreated investment pools. In untabulated tests, we find evidence of such adjustments during the crisis. Such a reallocation should mechanically lower funding rates in these untreated pools, unless active investors can promptly adjust their loan demand upward in response to the increase in loan supply.

While this adjustment mechanism likely increases the total volume of loans originated within the platform, it renders analyses of funding rates within separate investor pools uninformative. Therefore, we do not test for changes in loan funding rates within each investment pool. Instead, we conduct a set of DiD analyses of funding volumes across the three investor pools. To measure funding volume and to account for possible composition effects, we aggregate the data for these tests at the ELR–month (or ELR–month–investor pool) level. We include ELR fixed effects and interest rate controls in these regressions and cluster standard errors at the ELR level. We provide the specific regression specifications we are estimating in Section IV.B.

B. Investors' Reaction to the Announcement

Figure 2, discussed earlier, suggests that the fraction of loans that failed to secure funding increased substantially during the 2016 crisis (Graph A) and that the share of loans originated through the passive institutional pool dropped sharply during this period (Graph B). We now study these developments in more detail. To this end, we regress the probability that a loan application fails to secure funding from investors on month fixed effects, and graphically plot the coefficients. We estimate the following regression specification:

$$(1)\qquad \text{UNFUNDED}_{it} = \alpha_0 + \alpha_1 \text{INTEREST\_RATE}_{it} + \gamma_{elr} + \sum_{t=\text{Aug-15}}^{\text{Jul-16}} \beta_t \delta_t + \varepsilon_{it},$$

where UNFUNDED is an indicator of loan i not receiving enough commitments from investors for the loan to originate and Prob(UNFUNDED) is the estimated probability of a loan remaining unfunded. INTEREST_RATE is the interest rate on the loan set by the platform, γ represents ELR fixed effects, δ represents month-of-origination t dummies, and ε is the error term. We cluster standard errors at the ELR level. The reference month is July 2015 (omitted).

In Graph A of Figure 6, the solid line plots the coefficients for the month fixed effects, which represent the difference between the probabilities of a loan remaining unfunded in each month relative to its level in July 2015 (the omitted dummy). As can be seen, loan rejection rates were stable before Moody's announcement, but jumped in Feb. 2016 by 3 percentage points (pp), controlling for the ELR bin and borrower rate. To put this number in perspective, the average loan rejection rate in the 6 months prior to the announcement was 0.8%, so that this drop in funding represents a 275% increase in loan rejections.

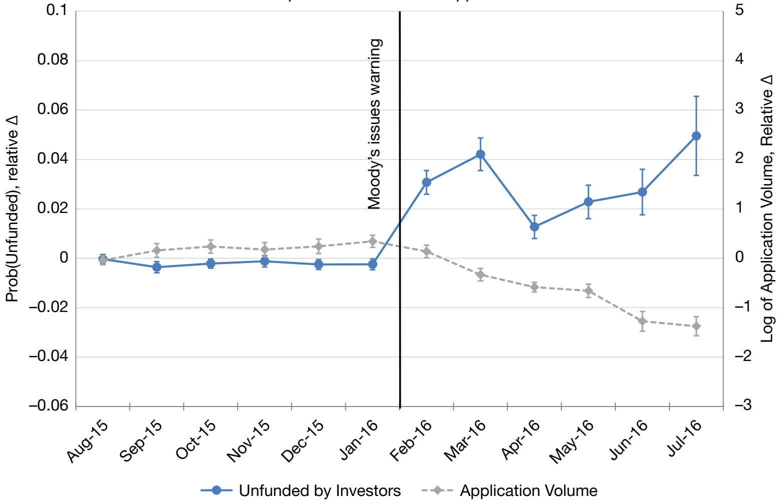
It is possible that loan originations on Prosper dropped not only because of higher rejection rates by investors, but also because of the lower loan supply by borrowers, particularly because Prosper raised loan interest rates and decreased its

FIGURE 6

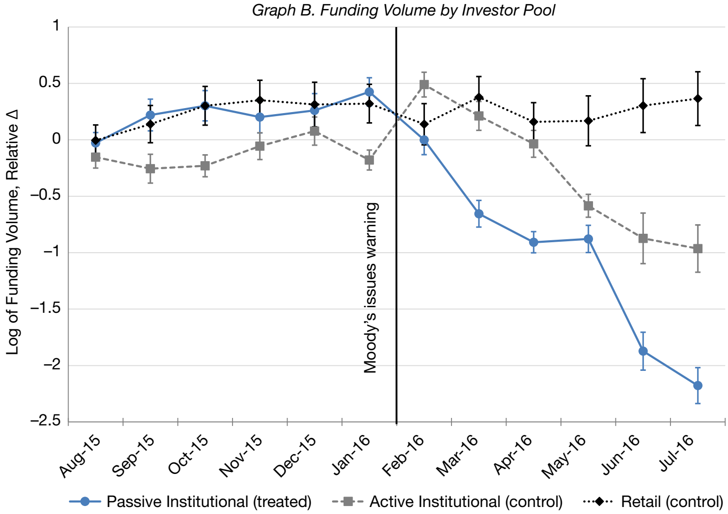
Event Study Analyses Around Moody's Downgrade Warning

Figure 6 plots the regression coefficients of the month fixed effects from the event study analyses around Moody's downgrade warning. The dependent variables of these regressions are the loan-level probability of not receiving investor funding (left axis) and the natural logarithm of monthly application volume in each estimated loss rate (ELR) bin (right axis) in Graph A and the natural logarithm of monthly loan funding volume by different types of investors in each ELR bin in Graph B. The reference month is July 2015. The vertical line corresponds to Moody's downgrade warning on Feb. 11, 2016. Each regression includes ELR fixed effects and controls for the interest rate. The error bars represent the 90% confidence intervals of the respective regression coefficients, where standard errors are clustered at the ELR level.

Graph A. Unfunded Loan Applications



Graph B. Funding Volume by Investor Pool



marketing efforts after the shock. To isolate this effect, we estimate the changes in loan supply by regressing the natural logarithm of the monthly application volume in each ELR bin on month fixed effects, controlling for the borrowing rate. As before, the coefficients of month dummies β_t are of interest. The regression equation is as follows:

https://doi.org/10.1017/S0022109023000789 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109023000789 Published online by Cambridge University Press

$$(2)\quad \text{LOG\_OF\_APPLICATION\_VOLUME}_{elr,t} = \alpha_0 + \alpha_1 \text{INTEREST\_RATE}_{elr,t} + \gamma_{elr} + \sum_{t=\text{Aug}-15}^{\text{Jul}-16} \beta_t \delta_t + \varepsilon_{elr,t},$$

where LOG_OF_APPLICATION_VOLUME is the natural logarithm of monthly application volume in each ELR bin. INTEREST_RATE is the mean interest rate on loans within the same ELR bin as set by the platform. Other variables and the clustering level are defined as in equation (1).

The monthly coefficients are shown by the dashed line in Graph A of Figure 6. While loan supply dropped during the crisis, the decline only started in Mar. 2016; the coefficient for February is small and does not significantly differ from the coefficient for January, as opposed to the large and significant coefficient for the probability of the loan remaining unfunded. Thus, the drop in funding rates on Prosper preceded the drop in loan applications after the platform raised its rates, suggesting that it was the investors' reaction rather than that of borrowers that precipitated the drop in loan volume.[27]

1. Passive Versus Active Investors' Response

We next compare the reaction to the crisis of lenders in different investment pools. We hypothesize that active investors, who can evaluate loans as they appear on the platform, should be less sensitive to Moody's warning regarding the quality of loans originated sometime prior, particularly given that the warning was later withdrawn and as such proved unwarranted. By contrast, passive lenders are more likely to base their decisions to participate in the market on public reports, and the warning by Moody's could undermine their trust in the quality of Prosper's loan book.

Graph B of Figure 6 plots the evolution of loan funding by the passive institutional, active institutional, and retail investors around Moody's downgrade warning. Similar to changes in loan supply, we estimate changes in investor loan demand by regressing the natural logarithm of the monthly funding volume in each ELR bin on month fixed effects, controlling for the borrowing rate. The regression specification is as follows:

$$(3)\quad \text{LOG\_OF\_FUNDING\_VOLUME}_{elr,t} = \alpha_0 + \alpha_1 \text{INTEREST\_RATE}_{elr,t} + \gamma_{elr} + \sum_{t=\text{Aug}-15}^{\text{Jul}-16} \beta_t \delta_t + \varepsilon_{elr,t},$$

[27]We also conduct these tests using daily rather than monthly data, and the conclusions are similar. These results can be found in Section E of the Supplementary Material. Furthermore, we verify that neither the increase in the federal funds rate (Dec. 16, 2015) nor the subsequent raise of interest rates by LC (Dec. 22, 2015) affected funding rates on Prosper. The effect of the rate increase by Prosper (Feb. 16, 2016), if any, should have a positive effect on funding rates, which means that the rate of loan rejections by investors might have been even higher in the absence of the rate increase.

https://doi.org/10.1017/S0022109023000789 Published online by Cambridge University Press

where LOG_OF_FUNDING_VOLUME is the natural logarithm of monthly loan application volume within each ELR bin funded by the respective pool of investors. Other variables and the clustering level are identical to those in equation (2). We estimate equation (3) separately for each investor pool (i.e., passive institutional, active institutional, and retail).

Graph B of Figure 6 shows that retail lending remained stable throughout, active institutional lending declined somewhat starting in May 2016, but passive institutional lending dropped precipitously in February and continued to drift downward for several months thereafter. Thus, the collapse in loan originations precipitated by Moody's announcement was driven by passive institutional investors, whereas active investors continued to provide funding at roughly the same level as before. This evidence is consistent with the hypothesis that an adverse development can make passive investors lose trust in the platform's loan analysis that underpins their decision to lend (Thakor and Merton (2019)). This loss of trust can cause them to withdraw from the market and threaten the platform's survival. In contrast, because active investors can independently ascertain the quality of the loans and do not need to rely on trust, their presence can mitigate the effect of the shock and help the platform survive.

To confirm these results in a more formal setting, we apply the DiD regression analysis around the downgrade warning. We estimate the following regression specifications:

$$\text{UNFUNDED}_{it} = \alpha_0 + \alpha_1 \text{POST}_t + \alpha_2 \text{INTEREST\_RATE}_{it} + \gamma_{elr} + \varepsilon_{it}, \tag{4a}$$

$$\begin{aligned} \text{LOG\_OF\_FUNDING\_VOLUME}_{elr,t} = {} & \alpha_0 + \alpha_1 \text{POST}_t \\ & + \alpha_2 \text{INTEREST\_RATE}_{elr,t} \\ & + \gamma_{elr} + \varepsilon_{elr,t}, \end{aligned} \tag{4b}$$

$$\begin{aligned} \text{LOG\_OF\_FUNDING\_VOLUME}_{elr,t} = {} & \alpha_0 + \alpha_1 \text{POST}_t \times \text{PASSIVE\_INST} \\ & + \alpha_2 \text{POST}_t \times \text{ACTIVE\_INST} \\ & + \alpha_3 \text{PASSIVE\_INST} \\ & + \alpha_4 \text{ACTIVE\_INST} \\ & + \alpha_5 \text{INTEREST\_RATE}_{elr,t} \\ & + \gamma_{elr} + \delta_t + \varepsilon_{elr,t}, \end{aligned} \tag{4c}$$

where PASSIVE_INST is an indicator for the loan funded in the passive institutional pool and ACTIVE_INST is an indicator for the loan funded in the active institutional pool. Loans funded in the retail pool serve as the reference group. Other variables and the clustering level are the same as in equations (1) and (3). Equations (4a) and (4b) are single-difference tests, and equation (4c) is a DiD regression.

TABLE 3

Investor Response to Moody's Warning: DiD Analyses

Table 3 reports the results of the DiD analyses around Moody's downgrade warning on Feb. 11, 2016. The dependent variable in columns 1 and 2 is the loan-level probability of not receiving investor funding, and the analyses are at the loan level. The dependent variable in columns 3–6 is the natural logarithm of monthly loan funding volume by different types of investors in each estimated loss rate (ELR) bin, and the analyses are at the ELR–month level. See Section I of the Supplementary Material for the definitions of variables. The reference group is loans funded in the retail investor pool. The estimates of the intercept and fixed effects are omitted for brevity. Standard errors are clustered at the ELR level. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. *t*-statistics are presented in parentheses.

	Prob(UNFUNDED)		LOG_OF_FUNDING_VOLUME			
	(−2 m; +2 m)	(−6 m; +6 m)	(−2 m; +2 m)		(−6 m; +6 m)	
Event Window	1	2	3	4	5	6
POST	0.042***	0.033***	−0.358***		−0.683***	
	(13.59)	(13.30)	(−9.86)		(−19.37)	
POST × PASSIVE_INST				−0.32***		−0.99***
				(−10.91)		(−28.84)
POST × ACTIVE_INST				0.773***		0.038
				(21.56)		(0.75)
PASSIVE_INST				2.66***		2.61***
				(45.92)		(50.42)
ACTIVE_INST				0.75***		0.68***
				(7.07)		(6.25)
INTEREST_RATE	−1.02***	−1.22***	−0.29	7.72	−18.90***	4.17
	(−2.73)	(−7.25)	(−0.03)	(0.94)	(−4.89)	(0.91)
ELR FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month FEs	No	No	No	Yes	No	Yes
N	134,339	328,926	260	777	676	2,019
Adj. R^2	0.016	0.011	0.855	0.920	0.874	0.862

Table 3 reports the results for two event windows: a 4-month window (i.e., +2 m/−2 m) in columns 1, 3, and 4 and a 12-month window (i.e., +6 m/−6 m) in columns 2, 5, and 6. The single-difference regressions in columns 1 and 2 show that loan funding rates went down after Moody's downgrade warning. The probability of a loan remaining unfunded increased by 4.2 pp over the 4-month window and 3.3 pp over the 12-month window, a fourfold increase compared with less than 1% of loan remaining unfunded before the shock. The results on funding volume in columns 3 and 5 have a similar interpretation. Column 3 shows that the monthly volume of funded loans in each ELR bin decreased by an average of 30.2% between Feb. and Mar. 2016, compared with funding volume in Dec. 2015 and Jan. 2016. The decrease in funding from Feb. to June 2016 is more substantial, at 49.2% (column 3).[28]

Columns 4 and 6 of Table 3 report the results of the DiD analyses, where we use the retail pool as the reference group. As expected, the decrease in the volume of loans funded in the passive investor pool is large and statistically significant. The coefficient of the interaction term between the post-Feb. 2016 indicator and the passive pool indicator implies a decrease of 27.4% (62.8%) in passive investor funding over 2 (6) months following the shock.[29] By contrast, funding by active institutional investors does not seem to decline between Feb. and June 2016, on average. If anything, active institutional investors funded more loans in Feb. to Mar. 2016 compared with 2 months before, possibly snapping some loans forgone

[28]We calculate these changes as $e^{0.36} - 1 = 0.3023$ and $e^{0.68} - 1 = 0.4924$.

[29]As above, the calculations are $e^{0.32} - 1 = 0.2739$ and $e^{0.99} - 1 = 0.6284$.

TABLE 4

Investor Response to Moody's Warning: Additional DiD Analyses

Table 4 reports the results of the additional DiD analyses around Moody's downgrade warning on Feb. 11, 2016. The dependent variables are the ELR, FICO_SCORE, REALIZED_RETURN, and DEFAULT_PROBABILITY in columns 1 and 2; 3 and 4; 5 and 6; and 7 and 8, respectively. The analyses are at the loan level and only include funded loans in columns 1–4 and originated loans in columns 5-8. See Section I of the Supplementary Material for the definitions of variables. The reference group is loans funded in the retail investor pool. The estimates of the intercept and fixed effects are omitted for brevity. Standard errors are heteroskedasticity-consistent (robust) standard errors in columns 1 and 2 and standard errors clustered at the ELR level in columns 3–8. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. *t*-statistics are presented in parentheses.

	ELR		FICO_SCORE		REALIZED_RETURN		Prob(DEFAULT)	
	(−2 m; +2 m)	(−6 m; +6 m)	(−2 m; +2 m)	(−6 m; +6 m)	(−2 m; +2 m)	(−6 m; +6 m)	(−2 m; +2 m)	(−6 m; +6 m)
Event Window	1	2	3	4	5	6	7	8
POST × PASSIVE_INST	0.0901**	0.0263	0.0764***	0.0375**	0.0025	0.0036	−0.0043	−0.0083
	(2.07)	(0.70)	(3.95)	(2.36)	(0.76)	(1.22)	(−0.57)	(−1.24)
POST × ACTIVE_INST	0.0865*	−0.0885**	0.0303	−0.0201	−0.0044	0.0010	0.0098	−0.0038
	(1.65)	(−2.11)	(1.06)	(−0.82)	(−0.97)	(0.31)	(0.93)	(−0.52)
PASSIVE_INST	−1.1916***	−0.8368***	−0.0571***	−0.0553***	0.0099**	0.0030	−0.0156*	−0.0041
	(−22.11)	(−24.36)	(−3.17)	(−4.22)	(2.23)	(1.16)	(−1.69)	(−0.72)
ACTIVE_INST	−1.2246***	−0.8786***	−0.0613***	−0.0477***	0.0136***	0.0054*	−0.0262**	−0.0083
	(−20.62)	(−23.26)	(−2.95)	(−3.59)	(2.99)	(1.88)	(−2.50)	(−1.18)
INTEREST_RATE			−2.326	−2.570	0.540	0.269**	−0.400	0.052
			(−0.56)	(−1.15)	(1.32)	(2.22)	(−0.45)	(0.19)
ELR FEs	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	131,509	323,699	131,509	323,699	39,258	101,701	39,258	101,702
Adj. R^2	0.007	0.006	0.397	0.374	0.003	0.003	0.051	0.050

by passive institutional investors.[30] Analyzing pre-shock descriptive statistics by investor pool, we find that the loan and borrower characteristics across the three investor pools were similar before Moody's downgrade warning. (These results are reported in Section E of the Supplementary Material.) Thus, we should not expect the differential response of active versus passive investors to the shock that we find to be driven by pre-shock composition differences across these pools. These results are consistent with uninformed passive investors being more affected by the news about possible lax screening by the P2P platform than informed investors were.

In Table 4, we examine additional outcomes related to how the pool of funded loans changed after the shock, within investor pools. Did passive investors forgo riskier loans and were these loans picked by active investors? Did the performance of loans funded within different pools change as a result? To answer these questions, we estimate additional DiD regressions at the loan level for funded loans. We run the following regression specifications:

$$\begin{aligned}(5)\quad \text{LOAN\_RISK\_OR\_PERFORMANCE}_{it} = {} & \alpha_0 + \alpha_1 \text{POST}_t \times \text{PASSIVE\_INST} \\ & + \alpha_2 \text{POST}_t \times \text{ACTIVE\_INST} \\ & + \alpha_3 \text{PASSIVE\_INST} \\ & + \alpha_4 \text{ACTIVE\_INST} \\ & + \alpha_5 \text{INTEREST\_RATE}_{it} \\ & + \gamma_{elr} + \delta_t + \varepsilon_{it},\end{aligned}$$

[30] In unreported analyses, we also find immediate (partial) reversal of investments by passive, but not active, institutional investors after the withdrawal of Moody's downgrade warning in July 2016.

where LOAN_RISK_OR_PERFORMANCE represents measures of loan risk (e.g., ELR and FICO score) or loan performance (e.g., realized return and realized default). Other variables and the clustering level are as in equations (1), (3), and (4). We do not include ELR fixed effects or the interest rate control in the specifications with ELR as the dependent variable.

Columns 1 and 2 of Table 4 report the results for the ELR as the sufficient risk statistic estimated by Prosper. The results are mixed and depend on the horizon. We find that loans funded by both passive and active institutional investors in the first 2 months appear somewhat riskier (i.e., higher ELRs), possibly because Prosper responded to the shock by raising interest rates on its loans (see Section IV.C) and this increase was larger for loans with higher credit risk, making riskier loans more attractive to investors. However, this result reverses when we consider the first 6 months after the shock. Looking at which loans investors fund within each ELR bin in columns 3 and 4, we find that institutional investors fund loans with higher traditional (FICO) scores. It is noteworthy that this result is conditional on the ELR and the interest rate. It appears that passive investors shy away somewhat more from borrowers with lower FICO scores compared with active investors, possibly due to flight to quality. We find no effects on the realized loan returns (columns 5 and 6) or the realized defaults (columns 7 and 8), which suggests that investors could not beat Prosper's credit risk algorithms by adjusting their loan funding criteria after the Moody's warning.

2. Heterogeneity in Investors' Response

We supplement our main DiD tests with heterogeneity analyses to further establish that the withdrawal of funding by (mostly passive) investors was driven by news about loan quality. Recall that passive institutional investors typically specify broad criteria for their desired investments (see Section II.B.3). When trust in the platform's credit assessment is eroded, we expect that passive investors should shy away from the riskiest loans. These are the loans with higher ELRs and thus lower Prosper ratings. We define a Prosper loan to be a high-risk loan if it was rated as D, E, or HR by Prosper. In a similar vein, longer-term loans are riskier. Prosper's algorithms assign higher interest rates to these loans, all else equal, and it also takes longer for defaults on long-term loans to materialize, which we confirm by examining default hazard by loan maturity (not reported). Therefore, we expect investors to reduce funding of long-term loans to a larger extent after a shock to investor trust such as Moody's downgrade warning.

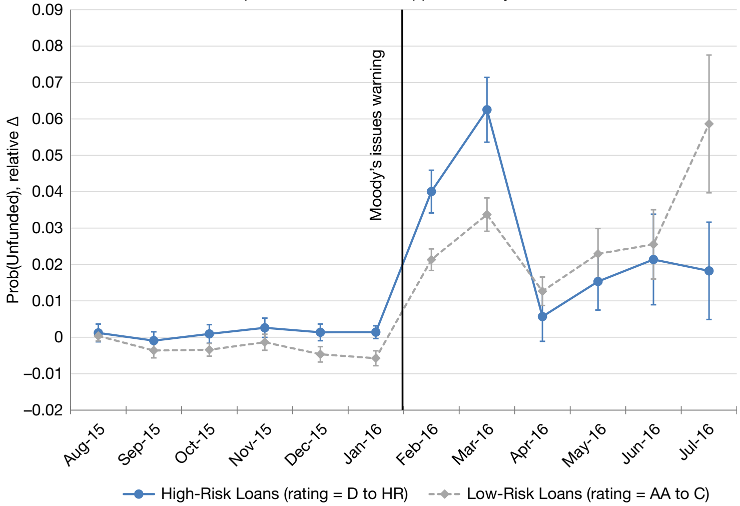
Figure 7 illustrates the patterns in the data pertaining to this heterogeneity analysis. The figure plots the coefficients of month fixed effects from equation (1), where each regression is run on a subsample of loans based on loan risk and maturity, as described above. As shown in Graph A, funding rates for loans of different risks were similar before Moody's downgrade warning, but they diverged sharply in the month of the warning, with riskier loans becoming significantly more likely to remain unfunded by investors over the first 2 months. Funding rates became statistically indistinguishable across the two groups of loans in Apr. 2016, with some reversal in July 2016. Similarly, 60-month loans became less likely to receive investor

FIGURE 7

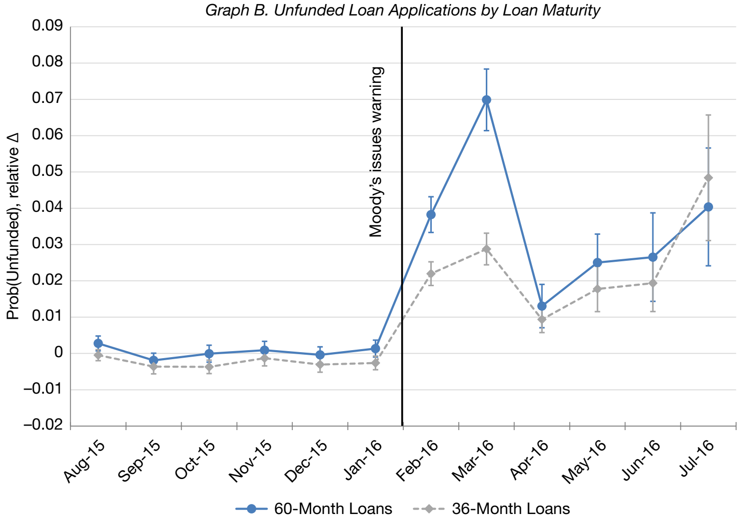
Heterogeneity Analyses Around Moody's Downgrade Warning

Figure 7 plots the regression coefficients of the month fixed effects from heterogeneity analyses around Moody's downgrade warning. The dependent variables of these regressions are the loan-level probability of not receiving investor funding. The reference month is July 2015. The vertical line corresponds to Moody's downgrade warning on Feb. 11, 2016. Each regression includes estimated loss rate (ELR) fixed effects and controls for the interest rate. Standard errors are clustered at the ELR level.

Graph A. Unfunded Loan Applications by Loan Risk



Graph B. Unfunded Loan Applications by Loan Maturity



funding after the shock, especially in the earlier part of the analysis window, compared with 36-month loans and compared with funding rates for both types of loans before the shock.

Regressions in Table 5 corroborate these results. The regression specification includes ELR and month fixed effects, and it controls for changes in the ELR–interest rate mapping. The regression equations are as follows:

TABLE 5

Investor Response to Moody's Warning: Heterogeneity Analyses

Table 5 reports the results of the heterogeneity analyses around Moody's downgrade warning on Feb. 11, 2016. The dependent variable is the loan-level probability of not receiving investor funding. See Section I of the Supplementary Material for the definitions of variables. The reference group is loans with Prosper rating = AA in columns 1 and 3, and it is 36-month loans in columns 2 and 4. The estimates of the intercept and fixed effects are omitted for brevity. Standard errors are clustered at the ELR level. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. *t*-statistics are presented in parentheses.

	UNFUNDED			
	(−2 m; +2 m)		(−6 m; +6 m)	
Event Window	1	2	3	4
POST × HIGH_RISK_LOAN	0.051***		0.036***	
	(11.07)		(5.67)	
POST × LONG_TERM_LOAN		0.039***		0.021***
		(11.99)		(10.60)
LONG_TERM_LOAN		−0.000090		0.003162***
		(−0.09)		(4.73)
INTEREST_RATE	−2.08***	−0.44	−2.28***	−1.12***
	(−5.42)	(−1.60)	(−5.74)	(−4.47)
ELR FEs	Yes	Yes	Yes	Yes
Month FEs	Yes	Yes	Yes	Yes
N	134,339	134,339	328,926	328,926
Adj. R^2	0.016	0.017	0.013	0.014

$$\text{(6)} \qquad \text{UNFUNDED}_{it} = \begin{array}{l} \alpha_0 + \alpha_1 \text{POST}_t \times \text{LOAN\_TYPE} + \alpha_2 \text{LOAN\_TYPE} \\ + \ \alpha_3 \text{INTEREST\_RATE}_{it} + \gamma_{elr} + \delta_t + \varepsilon_{it}, \end{array}$$

where LOAN_TYPE corresponds to a particular loan type (e.g., high-risk loans). Other variables and clustering are as in equations (1) and (4a). We drop the level of LOAN_TYPE in specifications where it is subsumed by ELR fixed effects.

We find that riskier loans were 5.1 (3.6) pp less likely to receive investor funding relative to safer loans in the first 2 (6) months after Moody's downgrade warning. Likewise, long-term loans are 3.9 (2.1) pp less likely to be funded compared with short-term loans over the first 2 (6) months. Consistent with investors' concerns about loan quality being the channel behind investor withdrawal from Prosper's P2P lending market,[31] the results reported in Table 5 suggest that investors became more wary of funding loans for which mistakes in quality assessment can lead to higher defaults or materialize with a greater lag.

3. Falsification Tests

We conduct a series of placebo tests in the pre-2016 period, both for the main DiD analyses and for the tests of heterogeneous treatment effects. The regression specifications are the same as in our earlier DiD analyses, but all dates are shifted back by 1 year. Specifically, we use Feb. 2015 as the month of the placebo shock. We present a brief summary of these tests here, and the details can be found in Section F of the Supplementary Material.

[31]Note that the results reported in the previous section suggest that it is passive investors who withdrew the funding during the crisis, while the lending by the better-informed active investors was little affected.

https://doi.org/10.1017/S0022109023000789 Published online by Cambridge University Press

We find no discernible effects of the placebo shock on funding rates, suggesting that the decrease in investor funding following Moody's downgrade warning in Feb. 2016 was not seasonal in nature. We also do not observe significant differences in funding rates of different loan types around the placebo shock. In our placebo regression analysis, the coefficients are mostly small and insignificant, and do not follow the patterns documented in Tables 3 and 5. Overall, the placebo tests show the absence of the effects that we document for the 2016 crisis.

One might also be concerned that the drop in investors' funding on Prosper in Feb. 2016 may have been caused by market-wide trends rather than the warning from Moody's. Of note, in Apr. 2016, a scandal broke out at LC, and one could argue that it might have been partially anticipated and had spillover effects on Prosper.[32] To address these concerns, we compare monthly loan originations on Prosper with those of LC during this period (Section F of the Supplementary Material). We observe parallel trends in loan volume for the two platforms up to Jan. 2016, but a large divergence in February and March, driven by a sharp decrease in Prosper's originations. These tests suggest that the observed investors' reaction at Prosper was triggered by Moody's warning in February, and that investors on LC only reacted to its own crisis in April.

C. Prosper's Reaction to the Crisis

Although the warning by Moody's concerned loans that were originated almost a year prior to the announcement, the withdrawal of institutional investors prompted Prosper to undertake a series of corrective actions, including several adjustments to loan rates and an increase in the rate of loan denials to riskier borrowers. The platform also announced major changes to its credit model, which were enacted in June 2016.

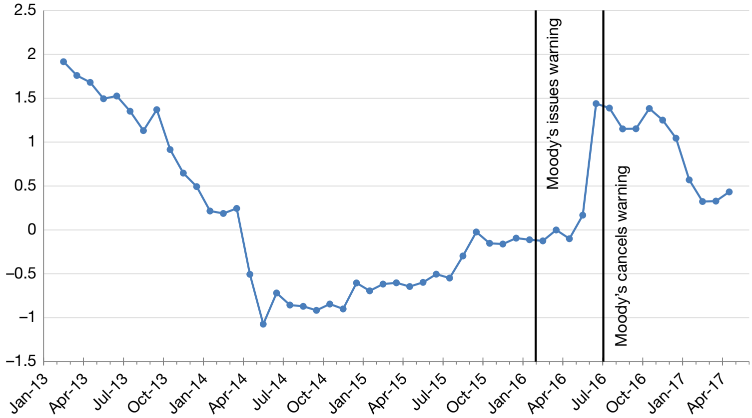
To better understand the evolution of the lending standards in the run-up to the crisis and Prosper's reaction to it, we estimate the changes in the "conservatism" of its credit model using the following procedure. First, for each loan within each monthly cohort, we regress Prosper's ELR on an exhaustive set of loan and borrower characteristics (excluding those that are functions of the ELR, such as rating and interest rate) using the ridge regression approach, which searches through combinations of possibly correlated explanatory variables to identify the model that generates the most accurate out-of-sample prediction (see, e.g., Zou and Hastie (2005)). Second, using the model estimated for month $t-1$, we predict the ELR for loans originated in month t. Third, we compute the difference between the actual ELR and our prediction from the "stale" model. This difference is positive when the actual ELR is higher than what it would have been if it were calculated using the previous month's credit model, that is, when the credit risk model is adjusted toward higher conservatism, generating higher ELRs for the same loan characteristics. Conversely, this difference is negative when the lending standards are relaxed. Next, we average these differences across all loans originated in month t, which

[32]LC was accused of violating investor instructions by issuing P2P loans that did not meet buyers' criteria and of poor governance due to the undisclosed interest of its CEO in a fund that was buying LC loans, allegedly to boost demand. At that time, LC's loan volume halved, as many institutional investors, including banks, put their loan orders on hold or exited the platform. LC's CEO resigned.

FIGURE 8

Platform's Conservatism over Time

For each month t, Figure 8 shows the difference between each loan's estimated loss rate (ELR) assigned by Prosper and the fitted ELR predicted for that loan by the model that was in place in month $t-1$, averaged over all loans originated in month t, cumulated over time, and normalized to be 0 in Mar. 2016. For any two dates, the difference in this measure represents the difference in the ELR that would have been generated for an average loan by the credit models that were in place on those two dates. Positive values correspond to greater conservatism (higher ELR holding loan characteristics constant). The details of the procedure used to construct this graph are given in Section IV.C. The vertical lines represent the 2016 crisis, from the initiation of Moody's downgrade warning on Feb. 11, 2016 to the warning's cancellation on July 15, 2016.



yields our estimate of the change in conservatism from month $t-1$. Finally, we cumulate this measure over a period of time to compute the total change in strictness over that period. We normalize this measure to be zero in Mar. 2016, so that the model's strictness in that month serves as the reference point.

Figure 8 shows the results. Between Mar. 2013 and June 2014, the cumulated month-on-month prediction error went down by almost 3 points. This implies that, roughly speaking, for an average loan originated in Mar. 2013, the ELR would have been almost 3 pp lower if it were estimated using the credit model that was in place in June 2014, indicating a substantial decline in conservatism. Of course, these adjustments to the model may have been made in response to the excessive conservatism that Prosper's model exhibited earlier. In untabulated analysis, we find that realized loan returns at that time consistently exceeded estimates, and the difference was at its maximum in mid-2013. Hence, the ELR estimates for subsequent cohorts may have been scaled down to compensate. Figure 8 further suggests that the model was gradually adjusted toward greater conservatism after July 2014.

Of particular interest is the sharp increase in the model's conservatism in June 2016. As can be seen, the new credit model that Prosper enacted in response to the crisis bumped up the ELR by about 1.5 pp, on average. We have also compared the ELR-to-interest rate mapping, and found that as part of the model recalibration, Prosper increased interest rates for most ELR bins by up to 7% of their pre-crisis level. Coupled with the tightening of the ELR model documented above, the net effect was to increase interest rates for some borrowers by up to 3 pp. In untabulated tests, we find that subsequently realized returns on newly originated loans became better aligned with Prosper's predictions.

https://doi.org/10.1017/S0022109023000789 Published online by Cambridge University Press

It is worth noting that despite these dramatic measures by Prosper designed to bring investors back on board, some of which were initiated as early as February, loan volume continued to drop until after Moody's withdrew its warning in July 2016. This illustrates the persistent effect that salient signals can have on investors' trust, particularly given the delay with which improvements in the quality of the loans that the platform originates can manifest themselves through lower defaults later. For this reason, it may be difficult for the platform to regain investors' trust quickly, which reinforces the importance of maintaining stable demand from active investors that can help the platform to survive the crisis.

V. Active Investment Strategies

As discussed previously, the presence of active investors is a distinctive feature of P2P lending, which stands in contrast with other intermediated markets. This section reports several additional results that are consistent with active investors keeping the platform's moral hazard in check. We also show that after controlling for the loan risk, active investors on average do not outperform passive strategies. Finally, we discuss the adverse selection problem that may arise when active and passive investors coexist within the platform.

A. Are "Active" Investors Active?

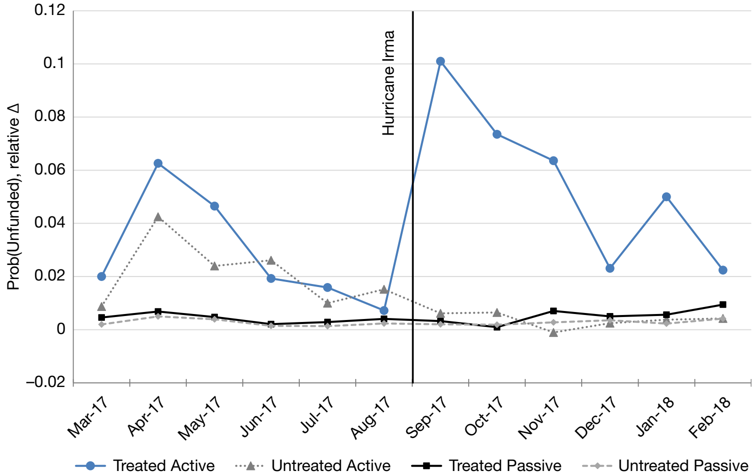
We first test whether active investors perform incremental credit analysis, using natural disasters as an exogenous shock to loan quality. Prosper does not price in disaster risk because the platform does not include borrowers' geographical location in its credit model to avoid the risk of violating fair-lending laws. If active (i.e., active institutional and retail) investors indeed make informed decisions, they should screen out risky "disaster loans," which are likely to have high expected losses.

Using event study analysis around Hurricane Irma in Sept. 2017, we first examine the dynamics of rejection rates by active and passive investors (Figure 9). We find that active investors reject loans requested by borrowers in areas hit by Hurricane Irma (i.e., treated active) at a much higher rate than in any of the six months preceding the hurricane. By contrast, funding rates of disaster loans by passive investors (i.e., treated passive) remain virtually unchanged. In addition to showing rejection rates for applications in disaster areas (i.e., treated), we show rejection rates by active and passive investors for applications in areas *not* hit by the hurricane (i.e., untreated). We find that the funding rates in these other areas remain practically unchanged both in active pools and the passive investment pool after the hurricane. Of note, rejection rates are similar for treated and untreated loan applications within active and passive pools before the event, which supports the parallel trend assumption. The DiD regression analysis confirms that the probability of not receiving investor funding is higher for disaster applications after the hurricane, but only for active institutional pools, in which loans are open to investor screening (the details of these tests can be found in Section G of the Supplementary Material). We conclude that at least some "active" institutional investors do perform incremental loan analysis instead of relying passively on Prosper's assessment.

FIGURE 9

Event Study Analysis Around Hurricane Irma

Figure 9 plots the regression coefficients of the month fixed effects from the event study analyses around Hurricane Irma. The dependent variable is the loan-level probability of not receiving investor funding, in the active (institutional and retail) and passive (institutional) investor pools. The coefficients are from distinct regressions estimated within the *treated* applications (i.e., disaster applications) and the *untreated* ones (i.e., regular applications). The reference period is Feb. 2017. The vertical line corresponds to Hurricane Irma in Sept. 2017. Each regression includes estimated loss rate fixed effects and controls for the interest rate.



B. Monitoring by Active Investors

While active investors are less reliant on trust in lending, they are likely to be more sensitive to the deterioration in the loan performance detectable from the loan data supplied by the platform. To test this hypothesis, we regress the probability of a loan remaining unfunded on the proportion of all outstanding loans that defaulted in the previous month. We estimate these regressions for all investors, separately for active and passive investor pools, and in pooled regressions with interaction terms. The specifications are as follows:

$$\text{(7a)} \qquad \text{UNFUNDED}_{it} = \alpha_0 + \alpha_1 \text{MONTHLY\_DEFAULT\_RATE}_{t-1} + \alpha_2 \text{INTEREST\_RATE}_{it} + \gamma_{elr} + \varepsilon_{it},$$

$$\text{(7b)} \qquad \text{UNFUNDED}_{it} = \alpha_0 + \alpha_1 \text{MONTHLY\_DEFAULT\_RATE}_{t-1} \times \text{ACTIVE} + \alpha_2 \text{ACTIVE} + \alpha_3 \text{INTEREST\_RATE}_{elr,t} + \gamma_{elr} + \delta_t + \varepsilon_{it},$$

where MONTHLY_DEFAULT_RATE is the annualized default rate on all outstanding loans on Prosper in a given month and ACTIVE is an indicator for the loan funded in an active (i.e., active institutional or retail) pool. Other variables are defined as in equations (1) and (4a). We follow Petersen (2009) in clustering standard errors at the month-of-origination level given that our main independent variable of interest is MONTHLY_DEFAULT_RATE.

TABLE 6

Investor Response to Realized Defaults on Prosper

Table 6 reports the results of loan-level regressions of investor funding on the default rate on all outstanding matched loans on Prosper in the preceding month. See Section I of the Supplementary Material for the definitions of variables. The estimates of the intercept and fixed effects are omitted for brevity. Standard errors are clustered at the listing month level (the results are similar if we double-cluster at the listing month and ELR level). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. *t*-statistics are presented in parentheses.

	Prob(UNFUNDED)			
	All	Active	Passive	Pooled
	1	2	3	4
Panel A. Apr. 2013 to Jan. 2019				
$MONTHLY\_DEFAULT\_RATE_{t-1}$	0.139	0.981***	−0.030*	
	(1.35)	(3.59)	(−1.71)	
$MONTHLY\_DEFAULT\_RATE_{t-1} \times ACTIVE$				0.82***
				(3.19)
ACTIVE				−0.014
				(−1.05)
INTEREST_RATE	0.414***	0.838***	−0.024*	−0.041
	(3.71)	(3.86)	(−1.81)	(−0.96)
ELR FEs	Yes	Yes	Yes	Yes
Month FEs	No	No	No	Yes
N	1,345,246	347,417	997,829	1,345,246
Adj. R^2	0.002	0.009	0.000	0.028
Panel B. Excluding Feb. to Jul. 2016				
$MONTHLY\_DEFAULT\_RATE_{t-1}$	0.106	0.908***	−0.030*	
	(1.08)	(3.15)	(−1.71)	
$MONTHLY\_DEFAULT\_RATE_{t-1} \times ACTIVE$				0.73***
				(2.71)
ACTIVE				−0.014
				(−0.99)
INTEREST_RATE	0.4946***	1.1195***	-0.225	0.0059
	(5.03)	(10.60)	(-1.65)	(0.17)
ELR FEs	Yes	Yes	Yes	Yes
Month FEs	No	No	No	Yes
N	1,235,857	303,846	932,011	1,235,857
Adj. R^2	0.003	0.011	0.000	0.025

The results are reported in Table 6. Column 1 of Panel A suggests that overall loan rejection rates are not related to recent defaults. However, the regression in column 2, which only includes active (institutional and retail) investors' loans, reports that loan rejections by these investors are higher following an increase in default rates on outstanding loans. A 1-standard-deviation increase in the annualized monthly default rate on Prosper loans is associated with a 0.98-pp higher probability of loans remaining unfunded by this category of investors. This increase constitutes a 27.01% increase relative to the mean probability of failure to fund of 4.06% between Apr. 2013 and Jan. 2019. Column 3 reports that there is no such effect for passive institutional investors. The coefficient of the monthly default rate is negative, economically small, and only marginally significant. Finally, column 4 reports the results for a regression with interaction terms. This specification is more stringent because we are able to include month fixed effects, which absorb any differences in funding rates across time. Our findings remain unchanged. Active investors are significantly more responsive to recent defaults on the platform than passive ones are. The results are similar when we exclude the 2016 crisis (Panel B).

To summarize, active investors incorporate the data about recent loan performance in their decisions, and withdraw funding when rising default rates in the data signal a potential deterioration in the loan book quality. At the same time, as shown in Section IV.B.1, lending by active investors is not much affected by extraneous signals that can scare away uninformed lenders by eroding their trust in the platform's loan analysis. Thus, monitoring by active investors should moderate the platform's incentives to inflate loan volume by relaxing its lending standards, while their reliance on hard data can help the platform survive a crisis of trust.

C. Investment Performance Across Pools

In this section, we test whether active strategies are profitable on average, by regressing realized returns and default rates on indicator variables designating the investor pool in which the loan was originated. We follow Abadie et al. (2023) in clustering standard errors at the month–rating level because Prosper randomizes applications across investor pools at any given time within Prosper rating (see Section V.D). The results of these regressions are reported in Panel A of Table 7. The univariate regression of column 1 suggests that average returns are higher in the active institutional pool than in the passive institutional pool. However, this result is driven by the changing composition of investors and differences in borrower quality over time, and disappears once we control for the time variation in loan quality by including origination cohort (i.e., month–rating) fixed effects in column 2. Controlling for the ELR and the interest rate (to account for changes in the ELR–interest rate

TABLE 7

Loan Returns and Risk by Investment Strategy

Table 7 tests whether realized returns and the probability of default are different between the passive and active institutional pools (Panel A) as well as original and recycled retail loans (Panel B). The sample consists of loans originated in 2013–2017. We use the matched listing-loan sample in this table (all columns, for consistency). See Section I of the Supplementary Material for the definitions of variables. Standard errors are clustered at the month–rating level. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. *t*-statistics are presented in parentheses.

	REALIZED_RETURN			Prob(DEFAULT)		
	1	2	3	4	5	6
Panel A. Active Institutional Versus Passive Institutional						
ACTIVE_INST	0.00164**	0.00055	0.00050	0.00135	−0.00023	−0.00044
	(2.03)	(0.75)	(0.68)	(0.44)	(−0.13)	(−0.25)
INTEREST_RATE			0.30			0.48
			(1.47)			(1.17)
Month-rating FEs	No	Yes	Yes	No	Yes	Yes
ELR FEs	No	No	Yes	No	No	Yes
N	255,474	255,474	255,474	255,928	255,928	255,928
Adj. R^2	0.000	0.009	0.010	0.000	0.046	0.048
Panel B. Original Retail Versus Recycled Retail						
RECYCLED_RETAIL	−0.014***	−0.014***	−0.014***	−0.013	0.034***	0.037***
	(−4.84)	(−4.53)	(−4.46)	(−1.49)	(5.06)	(5.57)
INTEREST_RATE			−0.083			1.060
			(−0.16)			(1.06)
Month–rating FEs	No	Yes	Yes	No	Yes	Yes
ELR FEs	No	No	Yes	No	No	Yes
N	28,470	28,470	28,470	28,477	28,477	28,477
Adj.R^2	0.001	0.022	0.024	0.000	0.065	0.069

mapping) does not change this result (column 3). Likewise, we do not find statistically significant differences in default rates across the two pools (columns 4–6).[33]

Interestingly, Prosper's practice of recycling loans rejected by institutional investors by relisting them in the retail pool allows us to gage the quality of those rejected loans that end up being funded by the retail investors, a counterfactual not usually observable for rejected loans. We find that 80% of the recycled loans do get financed in the retail pool, and that these loans subsequently underperform other loans by 1.4% (Panel B of Table 7). Although by rejecting these loans institutional investors improve their returns, the rejection rate is too low to make the average returns significantly higher than those in the passive pool.

On reflection, even though the absence of significant differences in loan performance between active and passive investors may appear surprising, it is a natural outcome of the very low loan rejection rates documented earlier, coupled with the random allocation of loans to the different pools. Put differently, with average funding rates close to 91%, active investors' behavior is not very different from funding passively all the loans offered to them by Prosper, and hence their returns are similar to passive investors' returns.

Thus, the quality of Prosper's risk assessment and loan pricing appears to be high enough to compel even active institutional investors to fund over 90% of loans offered to them. This ensures that the average active investor does not outperform passive lenders, and explains the observed popularity of passive strategies. Furthermore, as the accuracy of Prosper's model increases over time, any remaining opportunities for outperformance should shrink further. It is not surprising that, according to Figure 2, the proportion of active institutional investors decreased by two-thirds between 2013 and 2019, and currently stands at only 7.2%. Yet, notwithstanding the lack of outperformance, the active institutional pool survives, implying that some investors find it worthwhile to retain the ability to actively pick loans and the platform sees benefits in maintaining its active investment pools.

D. Resolving the Adverse Selection Problem

A central friction in the Vallee and Zeng (2019) model is the adverse selection that sophisticated active investors can impose on passive investors financing loans from the same pool. The model predicts that it is in order to alleviate this adverse selection problem that the platform should withdraw information about its loans, making active loan analysis infeasible, and precipitating a switch to the fully reintermediated market model. As we document above, contrary to these predictions, in reality, information is not withdrawn, and active loan pickers coexist with passive investors. But how?

As described previously, Prosper physically separates active and passive investors into different pools. The number of loans with a certain rating allocated to each pool reflects the anticipated investors' demand within that pool, but conditional on rating, exactly *which* loans are allocated to each pool is determined randomly, with probabilities proportional to the desired size of that pool. In other

[33]The results are similar when we use month-of-origination instead of month–rating categories as well as in pooled regressions with dummies for active institutional, passive institutional, and recycled retail loans (not reported).

words, Prosper manages the pool sizes and their rating composition, but loan characteristics conditional on rating are supposed to be similar across the pools. By allocating the loans randomly, Prosper can avoid allegations of “unfair treatment” of investors in different pools (especially the retail pool). Using the approach suggested by Sacerdote (2001) and Guryan, Kroft, and Notowidigdo (2009), we test for the random assignment to the three pools conditional on rating, and are unable to reject the hypothesis that it is indeed random (Section H of the Supplementary Material).

The combination of the physical separation of active and passive investors into different pools and the random assignment of loans to the pools can be a simple way to prevent passive investors from being taken advantage of by sophisticated loan pickers, which in turn allows the platform to avoid full reintermediation and preserve the monitoring role played by the active investors.

We note several potentially attenuating factors. First, the coexistence of active and passive retail investors within the same pool can still impose adverse selection on passive retail investors. This is unlikely to have a large effect on the overall loan volume because the retail pool is small (8% of funds). Second, there may be heterogeneity in skill among active institutional investors, which could negatively affect the less skilled lenders and potentially drive them away from the market. However, these investors always have the option of switching to passive strategies, which, as we show below, generate returns similar to the average returns in the active pool. While it is possible that some investors may be hurt by adverse selection, by revealed preference they must perceive some offsetting advantages of staying in the active pool. Third, the practice of recycling loans from institutional pools to the retail pool does arguably impose some adverse selection on retail investors by including in the mix loans deemed unworthy of funding by institutions. As documented in Section V.C, recycled loans have significantly lower investment returns. However, recycled loans amount to only 1.3% of the total funds and 13.1% of the retail pool after 2013 (when the institutional pools were created). In recent years, this fraction has been even lower, at 0.07% of the total funds and 0.9% of the retail pool since 2017.

Overall, placing active and passive investors into separate pools provides a simple way for them to coexist on the same platform without the former imposing adverse selection on the latter and driving them away from the market.

VI. Summary and Conclusions

P2P lending markets were originally designed as disintermediated, but became much more centralized over time. Today, the lending platforms’ software performs all essential functions of traditional lending officers, including loan evaluation, pricing, screening, and servicing. The ability to outsource loan analysis to the highly technologically skilled platform has attracted a large number of passive investors. The segregation of loan pickers into separate investment pools, coupled with the high and growing accuracy of the credit assessment by the platform, allows them to earn returns similar to those of active investors.

Yet a number of investors still retain the ability to actively select individual loans, even though in practice they end up funding almost all of them and do not earn superior returns. We argue that the presence of active investors who are provided with detailed loan information and can independently ascertain the quality of the loan book allows the lending platform to survive a market crisis that erodes passive investors' trust and causes them to withdraw funding (Thakor and Merton (2019)). By relying on data rather than trust, active investors stabilize the market, and their ongoing monitoring incentivizes the lending platform to continue to provide high-quality loan analysis, keeping in check the moral hazard problem that plagues intermediaries under the originate-to-distribute market model.

It is noteworthy that some active investors explicitly recognize that their focus is on monitoring rather than superior returns. In the words of Bryce Mason, the founder of P2P-Picks, "I have come to view the main purpose of independent modeling [as] to be protective, guarding against a situation where originators grade their own loans and where these definitions may vary over time or even suddenly change or degrade. There may still be some opportunity to outperform an index, but it is modest and secondary to the protective element."[34]

Our findings are instructive in the context of the vast theoretical literature that rationalizes the existence of financial intermediaries (see Gorton and Winton (2003) for a comprehensive review), and in particular the problem of monitoring the intermediary (Diamond (1984)). In contrast to banks, P2P platforms do not take deposits, perform maturity transformation, improve risk-sharing, or monitor loans after origination. Instead, our evidence is consistent with scale economies and technological improvements in loan evaluation as a key benefit of intermediation (Millon and Thakor (1985), Boyd and Prescott (1986)). Our findings demonstrate that transparency can be an alternative to trust in lending (Thakor and Merton (2019)).

Appendix. The Supply of Loan Data over Time

Both Prosper and LC provide a number of variables describing the loans originated previously ("historical data") and those on offer at any given time ("real-time data"). The majority of the variables are those reported for the borrower by a credit bureau; in addition, the data include borrower's self-reported information (income, employment status, etc.), loan characteristics, and, for previously originated loans, the loan's performance to date.

We manually collect historical data on the number of variables that Prosper and LC made available to investors between 2007 and 2019, including borrower characteristics, loan terms, and platform-specific metrics. To this end, we search for any publicly available vintage versions of Prosper and LC data sets, dictionaries with definitions of variables, official press releases, and investor blog posts. We perform an extensive search on the Internet and its archived copies in Wayback Machine. This process allows us to identify 69 distinct vintage data sets and data dictionaries for Prosper and 57 such district documents for LC, as well as dozens of relevant press releases and blog posts. For each document, we collect the maximum number of

[34]https://www.lendingmemo.com/cutting-open-data-50-lending-club-may-lose.

variables with nonmissing values that the platforms provided to investors in each month of each year.[35]

We carefully examine the dates of publication of these materials, effective dates of data dictionaries, dates of the last loan application or loan in each vintage data set, and other relevant information, to reconstruct the timeline of data provision by the platforms. We distinguish between variables that investors likely had access to for real-time decision-making (e.g., credit attributes) and those that appear in historical data files, which can be used for backtesting and monitoring (e.g., loan repayments and defaults). As an additional step, we check the quality of hand-collected data by comparing it to the number of variables with nonmissing values inferred from the data that we use for the main tests. Our hand-collected data align closely with the estimates based on the data set.[36]

Figure 4 of the main text plots the number of variables that Prosper and LC have provided to investors in each month-year for loans originated previously (i.e., the historical data records).[37] The graph shows a clear general trend toward more information provision over time. From 2007 to 2019, the number of variables reported by LC increased from 19 to 140, and for Prosper, it went up from 22 to as many as 402 variables. Of particular interest is the dramatic jump in information provision by Prosper in Dec. 2012 (from 89 to 558), which occurred just before Prosper launched a "beta version" of its institutional pool, as well as the slightly more gradual increase by LC (from 27 to 87) over several months following the launch of its own institutional pool in Sept. 2012. This evidence not only shows that both platforms provide a wealth of data about their loans and become more transparent over time, but also suggests that in doing so they may be responding to the demands of institutional investors.

Two episodes of an apparent reduction in the data provision evident from Figure 4 are worth discussing. First, between Dec. 2016 and Mar. 2017, as part of the ongoing re-evaluation of its credit model, Prosper switched the credit bureau it partnered with from Experian to TransUnion. Because the two bureaus collect and report a somewhat different set of variables, for loans originated after this period, Prosper started reporting some new data items and ceased to report some old ones, with the net effect that the overall number of variables decreased by 28%. Nonetheless, even after this drop, Prosper has been providing over 400 variables, likely spanning the information obtainable from TransUnion's reports. The sheer quantity of data disclosed makes it implausible that the decrease in information provision was due to the platform's attempt to limit transparency.

[35]If we find several documents pertaining to the same month-year with a different number of variables, such as in cases when platforms provide full data sets to registered users but redacted data sets for public posting, we use the full data sets, consistent with our focus on information that investors most likely had access to when constructing their credit models at the time.

[36]We do not make this procedure the primary basis of our estimates because of concerns that platforms may have back-filled some of the variables in more recent versions of historical data. For example, ELR appears back-filled in the data set we have at hand for the period prior to its implementation in July 2012.

[37]Although the information available for real-time analysis of loans on offer is somewhat different from that on previously originated loans (in particular, because the latter includes information on payments after the loan was originated), the time trends in the number of variables are similar in both series, and hence for brevity, we only describe the historical data sets. An important potential exception is the decrease in the number of variables provided by LC's historical file that took place in 2014, which Vallee and Zeng (2019) focus on. As best as we could ascertain, this drop in the historical records was not mirrored in the real-time data feed (see 38).

Second, Figure 4 also shows that in Nov. 2014 LC decreased the number of variables provided from 100 to 56. This is the episode that the empirical analyses in Vallee and Zeng (2019) focus on. However, as illustrated in Figure 4, this decline was fully reversed within 1 year, and since then, LC has further increased the number of variables by 40%.[38]

Supplementary Material

To view supplementary material for this article, please visit http://doi.org/10.1017/S0022109023000789.

References

Abadie, A.; S. Athey; G. W. Imbens; and J. Wooldridge. "When Should You Adjust Standard Errors for Clustering?" *Quarterly Journal of Economics*, 138 (2023), 1–35.

Balyuk, T. "FinTech Lending and Bank Credit Access for Consumers." *Management Science*, 69 (2022), 555–575.

Berg, T.; V. Burg; A. Gombović; and M. Puri. "On the Rise of FinTechs – Credit Scoring Using Digital Footprints." *Review of Financial Studies*, 33 (2020), 2845–2897.

Boyd, J. H., and E. C. Prescott. "Financial Intermediary-Coalitions." *Journal of Economic Theory*, 38 (1986), 211–232.

Butler, A. W.; J. Cornaggia; and U. G. Gurun. "Do Local Capital Market Conditions Affect Consumers' Borrowing Decisions?" *Management Science*, 63 (2017), 4175–4187.

Di Maggio, M., and V. Yao. "FinTech Borrowers: Lax Screening or Cream-Skimming?" *Review of Financial Studies*, 34 (2021), 4565–4618.

Diamond, D. W. "Financial Intermediation and Delegated Monitoring." *Review of Economic Studies*, 51 (1984), 393–414.

Duarte, J.; S. Siegel; and L. Young. "Trust and Credit: The Role of Appearance in Peer-to-Peer Lending." *Review of Financial Studies*, 25 (2012), 2455–2484.

Franks, J.; N. Serrano-Velarde; and O. Sussman. "Marketplace Lending, Information Aggregation, and Liquidity." *Review of Financial Studies*, 34 (2021), 2318–2361.

Gorton, G., and A. Winton. "Financial Intermediation." In *Handbook of the Economics of Finance*, Vol. 1. G. Constantinides, M. Harris, and R. Stulz, eds. Amsterdam, Netherlands: Elsevier (2003), 431–552.

Guryan, J.; K. Kroft; and M. J. Notowidigdo. "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments." *American Economic Journal: Applied Economics*, 1 (2009), 34–68.

Hertzberg, A.; A. Liberman; and D. Paravisini. "Screening on Loan Terms: Evidence from Maturity Choice in Consumer Credit." *Review of Financial Studies*, 31 (2018), 3532–3567.

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

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

Jagtiani, J., and C. Lemieux. "The Roles of Alternative Data and Machine Learning in Fintech Lending: Evidence from the LendingClub Consumer Platform." *Financial Management*, 48 (2019), 1009–1029.

Keys, B. J.; T. Mukherjee; A. Seru; and V. Vig. "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans." *Quarterly Journal of Economics*, 125 (2010), 307–362.

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

[38]Furthermore, although public records surrounding this drop in the data provision by LC are surprisingly patchy, online forum discussions at the time suggest that it may have been limited to the historical data files and the web interface, and did not affect the API data feed used by institutional and sophisticated retail investors, nor the real-time data on listed loans.

Millon, M. H., and A. V. Thakor. "Moral Hazard and Information Sharing: A Model of Financial Information Gathering Agencies." *Journal of Finance*, 40 (1985), 1403–1422.

Morse, A. "Peer-to-Peer Crowdfunding: Information and the Potential for Disruption in Consumer Lending." *Annual Review of Financial Economics*, 7 (2015), 463–482.

Petersen, M. A. "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches." *Review of Financial Studies*, 22 (2009), 435–480.

Purnanandam, A. "Originate-to-Distribute Model and the Subprime Mortgage Crisis." *Review of Financial Studies*, 24 (2011), 1881–1915.

Sacerdote, B. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics*, 116 (2001), 681–704.

Stein, R. M. "Benchmarking Default Prediction Models: Pitfalls and Remedies in Model Validation." *Journal of Risk Model Validation*, 1 (2007), 77–113.

Tang, H. "Peer-to-Peer Lenders Versus Banks: Substitutes or Complements?" *Review of Financial Studies*, 32 (2019), 1900–1938.

Thakor, R. T., and R. C. Merton. "Trust in Lending." Working Paper, Massachusetts Institute of Technology (2019).

Vallee, B., and Y. Zeng. "Marketplace Lending: A New Banking Paradigm?" *Review of Financial Studies*, 32 (2019), 1939–1982.

Wei, Z., and M. Lin. "Market Mechanisms in Online Peer-to-Peer Lending." *Management Science*, 63 (2016), 4236–4257.

Zhang, J., and P. Liu. "Rational Herding in Microloan Markets." *Management Science*, 58 (2012), 892–912.

Zou, H., and T. Hastie. "Regularization and Variable Selection via the Elastic Net." *Journal of the Royal Statistical Society*, 67 (2005), 301–320.