Internet Appendix for Reintermediation in FinTech: Evidence from Online Lending

Tetyana Balyuk
Emory University Goizueta Business School
tetyana.balyuk@emory.edu

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

B. Further details on Prosper's history

The timeline of various developments affecting Prospers' operations is presented in Figure B.1. Prosper began operations on November 1, 2005 as an online platform where consumers could borrow from their friends and family. On February 21, 2006, the company announced its public launch as a P2P lending marketplace. The funding model that Prosper originally used is often referred to as Prosper 1.0 (pre-"quiet period" model).

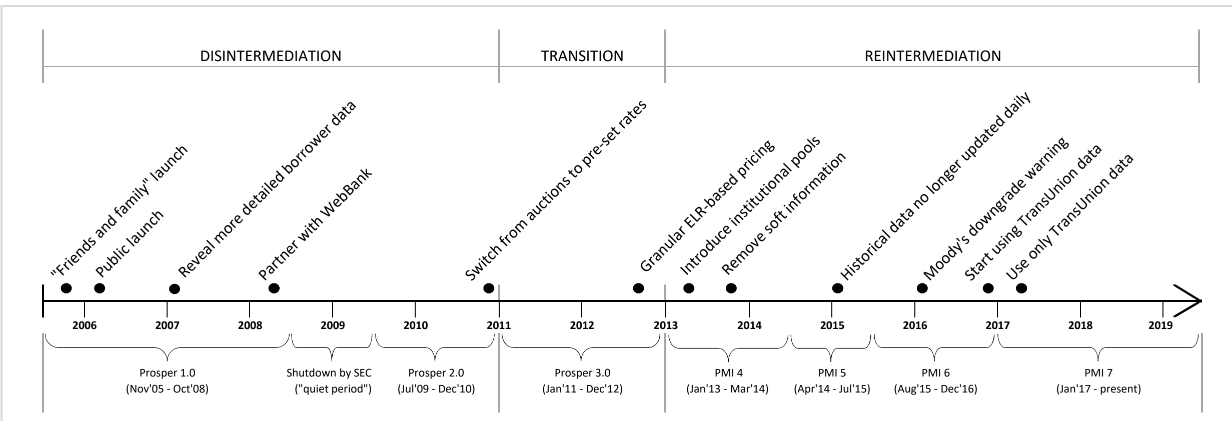
Under Prosper's initial model, the platform played a passive role as a loan facilitator, providing the infrastructure (similar to trading venues, such as Electronic Communication Networks) and loan ratings (similar to rating agencies), which were solely based on a credit bureau's score. The platform did little borrower screening other than imposing eligibility requirements, collecting borrower information, and facilitating loan repayment (2005–2008), and subsequently also detecting fraud and verifying income (2009–2010). Interest rates were determined via auctions, eBay-style. The interest rate was the maximum rate that a borrower was willing to pay up to a state-mandated ceiling (in case of automated funding) or the maximum rate that lenders were willing to accept on "winning bids" (in case of auction bidding). Lenders could either bid on each loan or invest via a "standing order" (a passive investment strategy). Lenders willing to accept lower rates often outbid other lenders.

Regulatory scrutiny of the P2P loan market caused Prosper to cease operations from October 15, 2008 to July 13, 2009 ("quiet period"). Upon reopening, Prosper transformed its loans into securities ('borrower-dependent notes') and made changes to its lending process. The platform's new funding model is often referred to as Prosper 2.0 (post-"quiet period" model). Prosper's platform continued to operate as an essentially disintermediated loan market with the platform doing

FIGURE B.1

Timeline

This graph shows the evolution of Prosper's P2P lending platform from its launch on November 1, 2005 to present. The timeline highlights three periods in the platform's evolution (top), significant changes in the operation of the market (middle), and major credit model changes (bottom).



little borrower screening, other than fraud detection and income verification. The interest rate was determined in an auction process. Prosper set the minimum rate on a loan based on its rating, whereas borrowers determined the maximum rate. Investors could place manual bids on P2P loans or use Prosper's "portfolio plan system" (a passive strategy). Because of the platform's passive role in loan evaluation and screening, we call the 2005–2010 period the *disintermediation* period.

In December 2010, Prosper made a major change to its platform by switching from an auction model to a pre-set rate model for determining the loan rates. The platform not only started pricing loans but also undertook a more active role in loan screening by stepping up its verification, loan cancellation, and collection efforts. In July 2012, the platform also introduced more granular pricing based on the Estimated Loss Rate (ELR), which is the annualized expected loss of loan principal due to borrower's default. The change to posted prices marks the platform's gradual transition to a reintermediated P2P loan market. We call the 2011–2012 period the *transition* period.

The ELR summarizes Prosper's assessment of the loan's default risk and fully determines the interest rate on the loan. Prosper estimates the ELR are based on a consumer credit bureau's score (such as FICO or SCOREX) and the platform's own analysis of historical losses on P2P loans. Prosper employs the following three-step procedure to evaluate the ELR. First, using its historical data on P2P loan defaults in conjunction with the borrower's self-reported information and credit bureau data, the platform estimates the probability of the loan becoming 60+ days past due within 12 months of the application date. This probability determines the loan's Prosper Score, ranging from 1 to 11. Second, the platform computes historical loss rates for each combination of the Prosper score and the FICO score (more precisely, one of 12 discrete FICO score bins, FICO 599-619, 619-639,. . . 829-850). Third, Prosper adjusts these 'base' loss rates based on a few variables it deems highly predictive of borrower risk, which at different times included the maturity of the loan, the debt-to-income ratio, and whether the borrower has previously had a Prosper loan. The ELR equals the base loss rate plus any adjustments. The stated reason for combining FICO scores with its own analysis of P2P loan defaults is that while cancellation scores are based on credit analyses of a huge number of consumers, borrowers in P2P markets may not be representative of the general population, and thus a credit score reflecting past P2P default experiences can be incrementally informative. The ELR is mapped into one of seven Prosper ratings, which range from AA (the safest) to HR (high risk). This mapping, loan interest rates, and the model Prosper uses to estimate the ELR are adjusted periodically.

Prosper made a number of substantial changes to its business model in 2013. The platform introduced a new credit model, launched separate investment pools for institutions (April 2013), and started using the FICO score instead of SCOREX (January–September 2013). From 2013 to present (2021), Prosper has been making periodic adjustments to its ELR algorithms and the ELR–interest rate mapping. In September 2013, the platform removed borrower pictures and narratives from applications and eliminated the ability of investors to ask borrowers questions. Thus, soft information that used to be available in the P2P loan market was phased out and is no longer used. In January 2015, Prosper reduced the frequency of updating its historical loan origination data from daily updating to updating with a several months' lag, ostensibly to protect its proprietary credit model. At the same time, data on performance of originated loans provided to investors remained available at daily updating frequency.[1] We refer to the period after 2013 as the *reintermediation* period because of the platform's active role as an intermediary in the P2P loan market.

C. Matched versus unmatched loans

Loan performance and listing files are not linked in Prosper data. Yet, our tests of realized returns and defaults rely on merging the two data sets, which we accomplish by matching loans on the common variables (i.e., loan origination date, loan amount, interest rate, Prosper rating, and loan maturity). In what follows, we compare the characteristics of matched and unmatched loans and discuss the direction and magnitude of possible bias that imperfect matching may introduce to loan performance tests in the time series and across investor pools.

Table C.1 compares characteristics of matched and unmatched loans. While most differences are statistically significant as measured by Wilcoxon rank-sum tests (not reported), they are

[1]Prosper's latest credit model is based on TransUnion data. Prosper switched from Experian to TransUnion in April 2017. Of note, LC has also been making some major changes to its pricing and screening.

small in economic terms, with few exceptions (e.g., loan amount, purpose). Together, these differences point to matched loans being somewhat safer than unmatched ones. Matched loans have lower rates and ELRs and hence higher ratings. These loans are provided to borrowers with somewhat higher credit scores (i.e., FICO, SCOREX) who are less likely to be self-employed or to have prior delinquencies. A higher fraction of these borrowers have a prior loan from the platform. These differences suggest that we may be underestimating average realized default rates in our analyses, although the bias is unlikely substantial.

To better understand the possible bias from imperfect matching, we use Prosper's loan performance files to examine the evolution of the AUC based on the interest rate as the classifier, as well as the time series of defaults and interest rates on matched and unmatched loans.[2] Prosper's loan files do not contain borrower characteristics. Neither do they report ELRs. However, given that the ELR fully determines the interest rate on the loan on any given date (Section **??**), the cross-section of loan rates should closely mimic the cross-section of the ELR at each point in time.

Figure C.1 plots the AUC for matched loans in Panel A and the AUC for unmatched loans in Panel B, by month-of-origination cohort. We first note that the AUC in Panel A closely resembles the AUC based on the ELR as the classifier reported in Panel C of Figure **??**. This is to be expected given that the AUC in Figure **??** is constructed based on matched loans only. Comparing the AUC for matched loans with the AUC for unmatched loans, we observe an almost identical trend and similar month-on-month variation in AUCs. In fact, the correlation between the two series is 81.6%. These results suggest that the quality of our matching procedure does not signifi-

[2] We remove one-year maturity loans, for consistency with our main analyses.

TABLE C.1

Characteristics of matched versus unmatched loans

This table reports loan and borrower characteristics for matched and unmatched loans in 2007–2019. See Section I for the definitions of variables.

	Matched Loans				Unmatched Loans			
	Mean	Median	SD	N	Mean	Median	SD	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Loan characteristics								
LOAN_AMOUNT	12,693	10,000	8,421	550,577	13,625	12,000	7,522	526,304
LOAN_MATURITY	44.0	36.0	11.3	550,577	41.9	36.0	10.3	526,304
INTEREST_RATE	15.3%	14.1%	6.4%	550,577	15.5%	13.5%	7.0%	526,304
ESTIMATED_LOSS_RATE	6.64%	5.99%	3.84%	550,577	6.76%	5.99%	4.07%	526,304
RATING (1-7)	4.54	5.00	1.46	550,577	4.50	5.00	1.59	526,304
CONSOLIDATION_PURPOSE	72.3%	—	—	550,577	75.6%	—	—	526,304
Panel B: Borrower characteristics								
FICO_SCORE	706.3	709.5	40.4	516,268	701.1	689.5	39.0	516,742
SCOREX_SCORE	710.5	712.5	59.2	343,908	704.4	712.5	61.0	320,025
EMPLOYED	84.1%	–	–	550,577	83.1%	–	–	526,304
SELF_EMPLOYED	6.2%	–	–	550,577	10.1%	–	–	526,304
YEARS_EMPLOYED	9.17	6.33	10.6	549,333	9.36	6.50	10.8	525,612
MONTHLY_INCOME	6,204	5,200	3,906	550,575	6,491	5,417	4,055	526,304
DEBT_TO_INCOME	0.26	0.24	1.14	540,197	0.26	0.25	0.12	514,055
MORTGAGE	31.4%	–	–	550,577	29.9%	–	–	526,304
TRADES_LAST_6M	0.96	1.0	1.16	550,552	1.01	1.0	1.20	526,281
ACTIVE_CREDIT_CARDS	4.68	4.0	2.85	516,242	4.81	4.0	2.87	516,719
CREDIT_CARD_BALANCE	5,554	4,272	4,312	516,242	5,461	4,171	4,293	516,719
TRADES_EVER_DELIQUENT	7.83%	3.00%	10.93%	550,552	8.22%	4.00%	11.07%	526,281
PAST_BANKRUPTCIES	0.17	0	0.38	516,242	0.17	0	0.39	516,719
PRIOR_LOAN	24.4%	—	—	550,577	18.7%	—	—	526,304

cantly bias the results on the quality of the platform's pricing model, either in the cross-section or in the time-series.

We next plot realized default rates on matched and unmatched loans by month-of-origination cohort in Panel A of Figure C.2. Similar to the results in Table C.1, we find that matched loans are somewhat safer than unmatched ones because matched loans have on average lower realized default rates. Most differences, however, are in the 2011–2013 period, and the two series converge starting from 2014. The average difference in defaults of 7.5% throughout 2011–2013 suggests that we may be overestimating realized returns during this period unless higher defaults on unmatched loans are priced in the interest rates.

We turn to interest rates in Panel B of Figure C.2. Indeed, similar to realized defaults, we observe divergence in interest rates on matched and unmatched loans in 2011–2013. The divergence shrinks toward the later part of 2013 and disappears in 2014 onward. The average difference in defaults of 5.8% throughout 2011–2013 suggests that the difference in defaults is largely priced in by the platform and if we do overestimate realized loan returns in 2011–2013, this overestimation is likely modest. Our back-of-the-envelope calculations suggest an overestimation of returns of around $(7.5 - 5.8) \times 0.352 = 0.6\%$, where 0.352 is the fraction of loans that remain unmatched in 2011–2013. Correcting realized returns for this potential bias better aligns realized returns with Prosper's estimates in 2011–2013 and does not invalidate any of the conclusions from our analyses based on this figure.

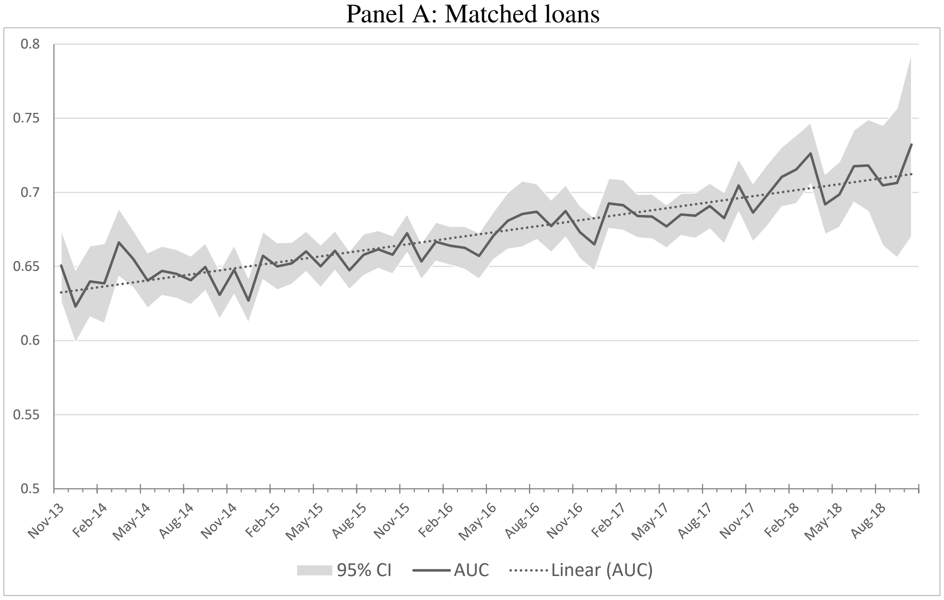
Last, we attempt to assess whether our matching procedure introduces any bias into a comparison of realized returns and defaults by investment strategy. An important complication in this assessment is that we cannot attribute unmatched loans to specific investment pools and thus cannot test for differences in realized returns and defaults on unmatched loans across investment strategies. Instead, we rely on suggestive evidence based on *expected* returns and defaults. We relate expected returns and defaults on loans originated in different investment pools to the indicator for whether a loan is matched or unmatched. The results of these analyses are reported in Table C.2.

FIGURE C.1

AUC for matched and unmatched loans

This graph shows AUC, which measures the ability of Prosper's interest rate to discriminate between defaulting and nondefaulting loans, for matched (Panel A) and unmatched (Panel B) loans. For each month, AUC is calculated using all loans originated in all investment pools in that month. Also shown are the linear time trends for each series and the 95% confidence intervals.

Panel A: Matched loans



Panel B: Unmatched loans

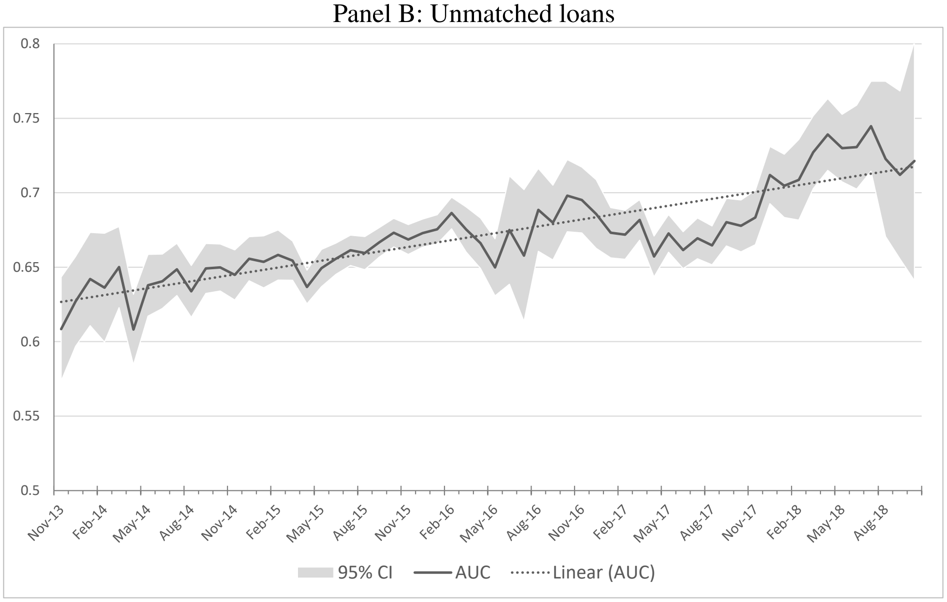
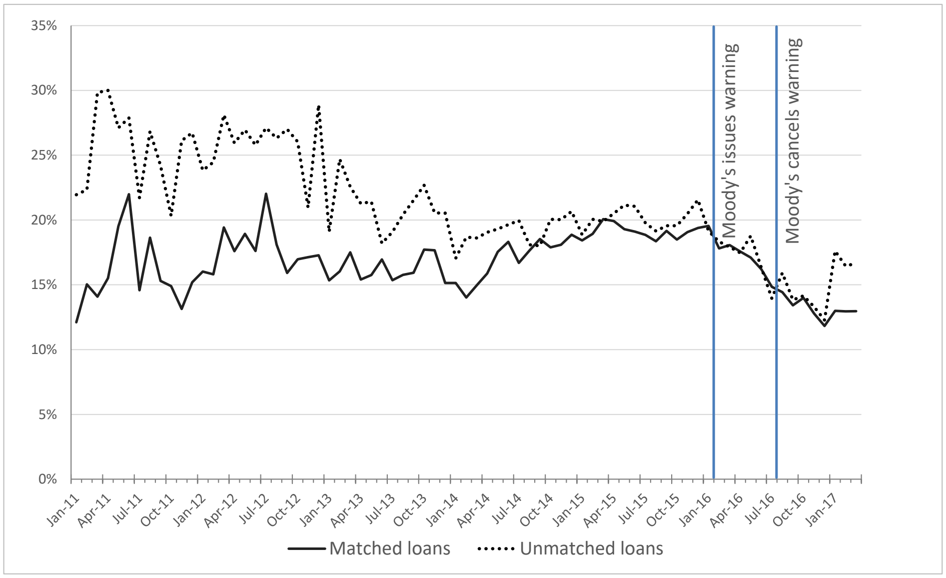


FIGURE C.2

Default and interest rates on matched and unmatched loans by month-of-origination cohort

This graph compares realized default rates (Panel A) and interest rates (Panel B) on matched and unmatched loans originated in the same calendar month. The vertical lines represent the 2016 crisis, from the initiation of Moody's downgrade warning on February 11, 2016 to the warning's cancellation on July 15, 2016.

Panel A: Probability of Default



Panel B: Interest Rate

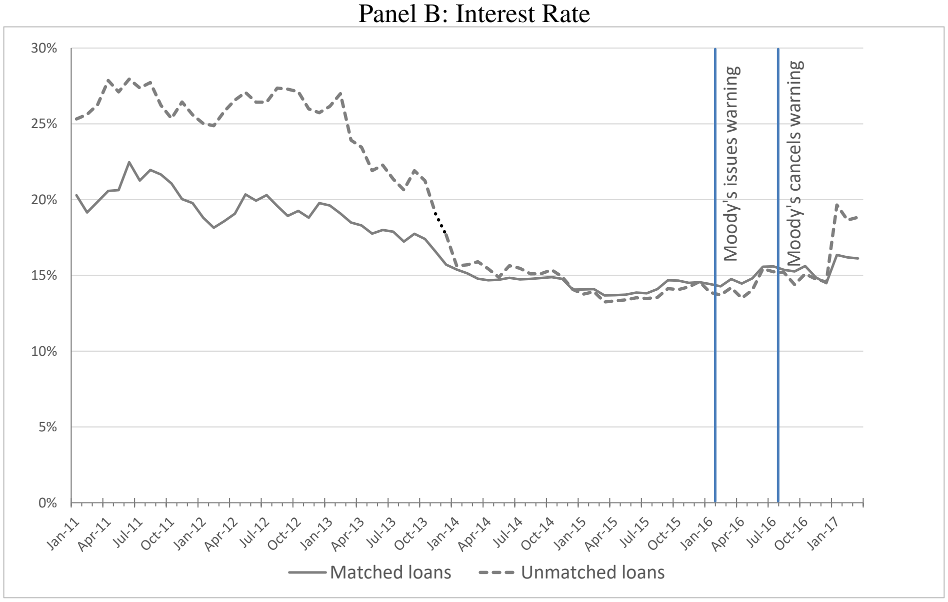


TABLE C.2

Assessment of matching bias by investment strategy

This table tests whether expected returns and defaults are *differentially* related to an indicator for a matched loan in the passive and active institutional pools (Panel A) as well for 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). 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.

	Prob(Match)					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Active Institutional versus Passive Institutional						
ACTIVE_INST×EXPECTED_RETURN	0.21	-0.037	-0.052			
	(1.28)	(-0.45)	(-0.64)			
EXPECTED_RETURN	1.51***	0.92***	0.83***			
	(6.15)	(6.96)	(5.93)			
ACTIVE_INST×EXPECTED_DEFAULT				-0.0020	0.041	0.030
				(-0.02)	(0.70)	(0.51)
EXPECTED_DEFAULT				-0.17	-0.21***	-0.20***
				(-1.13)	(-3.66)	(-3.52)
ACTIVE_INST	0.0085	0.0069	0.0084	0.026***	0.0016	0.0025
	(0.61)	(0.96)	(1.18)	(2.76)	(0.37)	(0.57)
INTEREST_RATE			3.36***			3.78***
			(3.53)			(3.97)
Month-Rating FEs	NO	YES	YES	NO	YES	YES
ELR FEs	NO	NO	YES	NO	NO	YES
N	532,219	532,219	532,219	532,219	532,219	532,219
Adj. R^2	0.007	0.054	0.057	0.001	0.053	0.057
Panel B: Original Retail versus Recycled Retail						
RECYCLED_RETAIL×EXPECTED_RETURN	0.86*	0.40	0.56**			
	(1.70)	(1.65)	(2.37)			
EXPECTED_RETURN	0.62*	0.58***	0.42**			
	(1.76)	(3.21)	(2.25)			
RECYCLED_RETAIL×EXPECTED_DEFAULT				-0.12	-0.25	-0.14
				(-0.50)	(-1.43)	(-0.83)
EXPECTED_DEFAULT				-0.65***	-0.066	-0.036
				(-3.40)	(-0.63)	(-0.34)
RECYCLED_RETAIL	-0.080**	-0.041**	-0.057***	-0.026	0.0032	-0.0061
	(-2.11)	(-2.04)	(-2.92)	(-1.51)	(0.26)	(-0.50)
INTEREST_RATE			3.42***			3.57***
			(2.84)			(2.96)
Month-Rating FEs	NO	YES	YES	NO	YES	YES
ELR FEs	NO	NO	YES	NO	NO	YES
N	51,586	51,586	51,586	51,586	51,586	51,586
Adj. R^2	0.003	0.079	0.085	0.002	0.078	0.085

We find that matched loans are associated with higher expected returns and lower expected defaults. However, the coefficients of interaction terms between expected returns and an indicator for active institutional pool and between expected defaults and an indicator for active institutional pool in Panel A are not statistically significant. The absence of statistical significance implies that the matching bias, if any, does not differ between active and passive institutional pools. It should not affect the comparison of loan performance in these pools. Several of the coefficients of interaction terms between expected returns and an indicator for recycled retail loans in Panel B are positive and significant. If anything, this significance suggests that the underperformance of recycled loans could be even higher than we estimate. Thus, the effect of matching quality on our results is likely negligible.

D. The effectiveness of Prosper's screening

In this section, we assess the effectiveness of Prosper's screening along the extensive margin (i.e., loan denial). Prosper's algorithms flag a fraction of loan applications as possibly fraudulent or excessively risky, and attempt to automatically verify certain critical information, such as borrower's income and employment status. Specifically, Prosper's screening algorithm may cancel a loan if it determines that the likelihood of default by the borrower may be materially greater than the one implied by the initially assigned Prosper rating. To formally assess the quality of Prosper's loan screening, one would need to investigate if loans that were canceled would have resulted in higher losses had they been extended. Unfortunately, this counterfactual is generally unobservable, so the evidence is only indirect.[3]

Our tests of the effectiveness of Prosper's cancellations are based on the applicants who, having had their loan application canceled, reapply for another loan within a month.[4] For such borrowers, Prosper re-uses their previous credit report, but their self-reported information may be different. We first compare loan and borrower characteristics for such re-applying borrowers appearing on their old (canceled) and new (resubmitted) loans applications. We find that even though the applicants are the same and their applications are close in time, the proportion of applicants who admit to being self-employed increases from 6% to 9%, and their median monthly income is decreased by $988. Thus, these applicants are likely untruthful on the first application, which is detected by Prosper and leads to the loan cancellation. As a result of the revision in the application data, the loans' assessed ELR increases by 82.5 basis points on average, which results in a 0.32 points lower Prosper rating, 1.3 pp higher interest rate, and 11.6% smaller loan size.

Further regression analyses show that, controlling for the ELR, re-submitted loans that receive funding are significantly more likely to default (Table D.1). We find default rates on these loans to be 3.2 pp higher than those for other borrowers with the same ELR in the same month. In other words, the loans from the previously screened out applicants are too risky even for the newly

---

[3]For example, Carmichael (2017) identifies borrowers who applied for a P2P loan both through Prosper and through LC. He finds that applicants whose Prosper loans were canceled but who nevertheless received a loan from LC are more likely to default given their rating, implying that canceled loans are riskier than they appear.

[4]We focus on borrowers who reapply within 30 days to mitigate concerns that the borrower's true creditworthiness or their assessment by Prosper are affected by other unobservable factors. In the sample, 77.1% of borrowers who reapply after the application is canceled do so within this period, perhaps to ensure that Prosper uses their prior credit report instead of requesting a new one.

TABLE D.1

Resubmitted applications and default risk

This table examines default risk of borrowers whose applications were canceled, resubmitted, re-evaluated by the platform, and subsequently funded by investors. The sample period is 2011–2017 and consists of matched observations, i.e., observations with non-missing loan performance data. See Section I for the definitions of variables. *Old ELR* refers to the ELR on the canceled loan application. *New ELR* refers to the ELR on the resubmitted loan application. The estimates of the intercept and fixed effects are omitted for brevity. Standard errors are clustered at the listing month level. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. t-statistics are presented in parentheses.

	Prob(Default)			
	1	2	3	4
RESUBMITTED	0.057***	0.047***	0.039***	0.032***
	(21.68)	(18.19)	(16.43)	(13.34)
LOG_OF_INCOME		-0.033***		-0.033***
		(-18.97)		(-19.07)
DEBT_TO_INCOME		0.10***		0.095***
		(11.22)		(10.64)
EMPLOYED		-0.041***		-0.041***
		(-12.87)		(-12.82)
SELF_EMPLOYED		0.0033		0.0012
		(0.79)		(0.29)
LONG_TERM_LOAN		0.038***		0.037***
		(8.79)		(8.65)
Old ELR FEs	YES	YES	NO	NO
New ELR FEs	NO	NO	YES	YES
Month FEs	YES	YES	YES	YES
N	298,372	297,926	298,372	297,926
Adj. R^2	0.047	0.053	0.047	0.053

assigned, higher ELR. Had these loans been issued at the ELR assessed originally on the first application before it was canceled, the resulting default rate would have been 4.7 points higher than on other loans with the same ELR. These findings confirm that loans that are canceled by Prosper are indeed riskier than others with the same assessed risk. However, Prosper appears too lenient when borrowers who were previously screened out reapply for another loan, and does not fully compensate for their higher risk when assigning the new ELR.[5]

E. Screening around Moody's downgrade warning

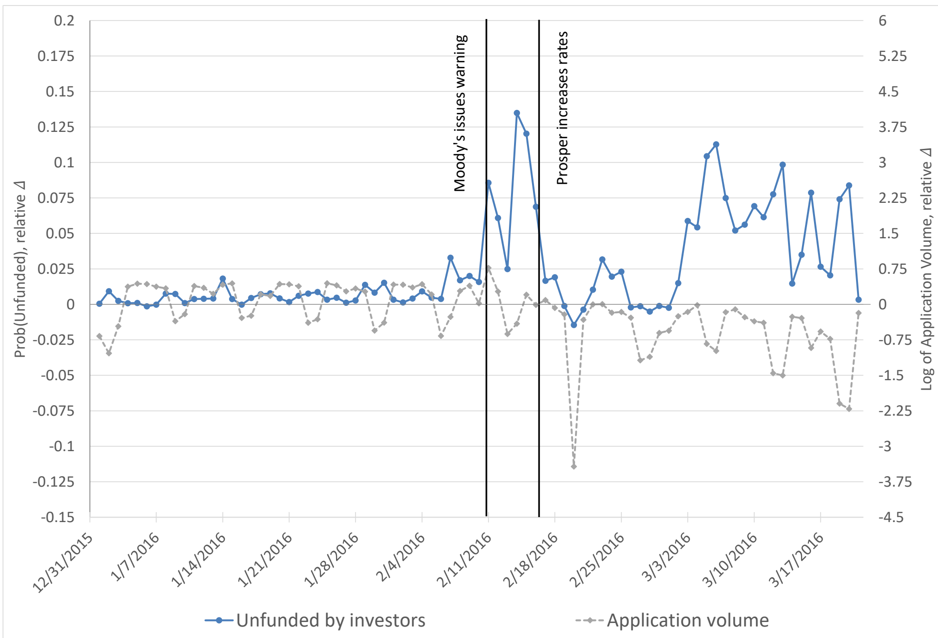
Below we provide additional results pertaining to investor screening and loan origination volume around Moody's downgrade warning on securitizations backed by Prosper loans in February 2016. We first examine daily rejection rates by investors and loan application volume around February 11, 2016 when Moody's issued the warning. Figure E.1 shows a sharp increase in rejection rates from a mean of 0.71% in January 1, 2016 – February 10, 2016 to a mean of 9.90% in February 11, 2016 – February 15, 2016. In an attempt to bring investors back on board, Prosper raised interest rates across 49 out of 52 ELR categories in a manner than increases the risk premium and expected investor returns on February 16, 2016. However, Prosper's reevaluation of its interest rate policies and subsequent tightening of the lending standards only had a temporary effect on investor funding. Loan rejection rates by investors climbed back continued in early March and remained significantly higher than before the shock throughout March. Figure E.1 also shows

[5] In interpreting realized performance statistics documented in this paper, one should keep in mind that we obtain them using more data and longer histories than were available at the time when Prosper's algorithms were calibrated. Thus, Prosper's model parameters chosen at the time may have been adequate conditional on the available data.

FIGURE E.1

***Daily-level* event study analyses around Moody's downgrade warning**

This graph plots regression coefficients of day fixed effects from two distinct regressions. The dependent variables of these regressions are the loan-level probability of not receiving investor funding (left axis) and the natural logarithm of daily within-ELR application volume (right axis). The reference date is December 31, 2015. Each regression includes ELR fixed effects (which absorb the intercept) and controls for the interest rate and loan maturity.



that the application volume on Prosper started to decline only after Prosper raised interest rates, which suggests that investor reaction to Moody's downgrade warning precipitated the drop in loan volume.

TABLE E.1

Descriptive statistics by investor pool in the pre-2016 period

This table reports loan and borrower characteristics for loan applications and for originated P2P loans matched to listings, by starting investor pool. The sample is from 2013–2015. See Section I for the definitions of variables.

	Passive Institutional				Active Institutional				Retail			
	Applications (N=407,719)		Matched Loans (N=127,731)		Applications (N=143,406)		Matched Loans (N=50,590)		Applications (N=54,099)		Matched Loans (N=19,995)	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
	1	2	3	4	5	6	7	8	9	10	11	12
Panel A: Loan characteristics												
LOAN_AMOUNT	13,985	12,000	12,644	10,000	13,796	12,000	12,541	10,000	11,844	10,000	11,072	10,000
LOAN_MATURITY	43.6	36.0	44.6	36.0	43.8	36.0	44.6	36.0	44.7	36.0	45.5	36.0
INTEREST_RATE	14.3%	13.5%	14.4%	13.7%	14.4%	13.5%	14.6%	14.0%	16.6%	16.0%	16.5%	16.0%
ESTIMATED_LOSS_RATE	6.22%	5.49%	6.27%	5.74%	6.00%	5.49%	5.99%	5.49%	7.09%	6.49%	6.91%	6.49%
RATING (1-7)	4.69	5.00	4.66	5.00	4.78	5.00	4.78	5.00	4.34	4.00	4.41	4.00
CONSOLIDATION_PURPOSE	76.8%	–	76.6%	–	77.0%	–	76.8%	–	74.5%	–	74.5%	–
Panel B: Borrower characteristics												
FICO_SCORE	700.1	689.5	700.6	689.5	702.2	689.5	702.6	689.5	698.2	689.5	698.7	689.5
SCOREX_SCORE	705.2	712.5	705.8	712.5	709.3	712.5	710.7	712.5	703.7	712.5	705.9	712.5
EMPLOYED	83.4%	–	86.4%	–	85.0%	–	88.0%	–	85.1%	–	87.4%	–
SELF_EMPLOYED	7.6%	–	6.2%	–	6.7%	–	5.4%	–	6.6%	–	5.6%	–
YEARS_EMPLOYED	9.19	6.58	9.21	6.75	9.17	6.67	9.26	6.92	8.98	6.50	9.15	6.75
MONTHLY_INCOME	6,247	5,250	5,883	5,000	6,384	5,417	6,073	5,250	6,044	5,000	5,876	5,000
DEBT_TO_INCOME	0.26	0.25	0.26	0.25	0.26	0.25	0.26	0.24	0.26	0.25	0.26	0.25
MORTGAGE	46.9%	–	48.1%	–	49.0%	–	50.8%	–	48.2%	–	50.5%	–
TRADES_LAST_6M	0.95	1.0	0.95	1.0	0.89	1.0	0.86	1.0	0.88	1.0	0.85	1.0
ACTIVE_CREDIT_CARDS	4.74	4.0	4.60	4.0	4.79	4.0	4.65	4.0	4.64	4.0	4.54	4.0
CREDIT_CARD_BALANCE	5,256	3,913	5,016	3,750	5,301	3,935	5,084	3,775	4,921	3,587	4,817	3,533
TRADES_EVER_DELIQUENT	8.58%	4.00%	8.81%	4.00%	8.18%	4.00%	8.33%	4.00%	8.66%	4.00%	8.78%	4.00%
PAST_BANKRUPTCIES	0.19	0	0.20	0	0.19	0	0.20	0	0.21	0	0.22	0
PRIOR_LOAN	7.9%	–	11.6%	–	8.4%	–	12.6%	–	9.5%	–	14.5%	–

We next compare characteristics of applications and originated loans in the pre-2016 period across passive institutional, active institutional, and retail investors. Table E.1 shows that the loan and borrower characteristics across the three investor pools are similar. Such similarity is to be expected given Prosper's random assignment of applications across investor pools at any given point in time, which we test in Section H. The remaining differences are likely due to a combination of changes in the borrower pool and differences in the growth of these investor pools over time. Likewise, loan and borrower characteristics are similar across the three pools, with the exception that retail pool loans appear somewhat riskier, although a larger fraction of borrowers in this pool obtained a Prosper loan before.

F. The 2016 crisis: Falsification tests

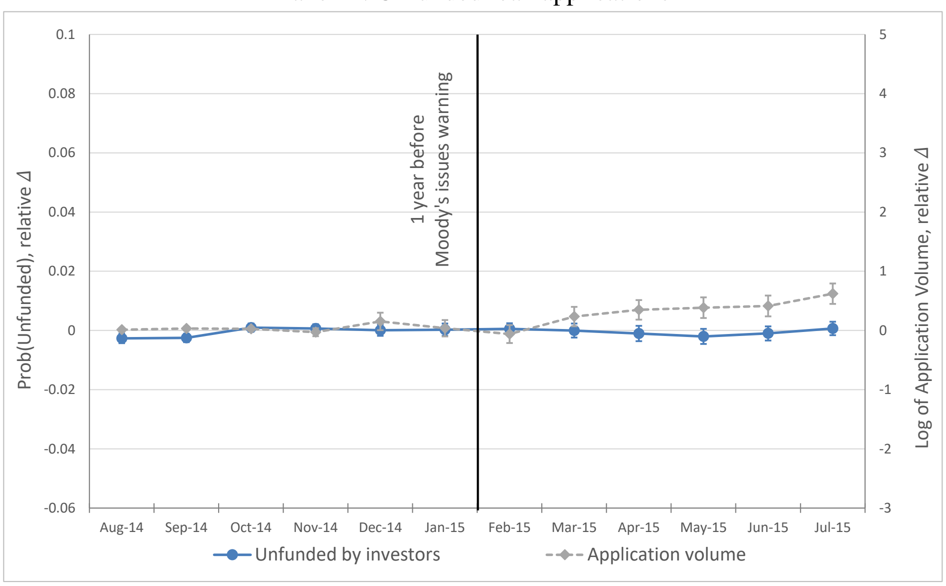
In this section, we report the results on the falsification tests described in Section **??** of the main text. Figure F.1 and F.2 show that there are no discernible effects of the placebo shock on funding rates. We see some increase in funding volumes, specifically funding by passive and active institutional investors, because the market was growing at a rapid pace around that time and Prosper's marketing was targeted at institutions. Importantly, these patters show that the decrease in investor funding following Moody's downgrade warning in February 2016 is not seasonal in nature because we do not observe a similar drop in funding volume during the same months in the preceding year. We do not observe significant differences in funding rates of different loan types around the placebo shock in Figure F.2, which adds validity to our empirical design.

FIGURE F.1

Placebo tests: Event study analyses before Moody's downgrade warning

This graph plots regression coefficients of the month fixed effects from the event study analyses one year before 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 ELR bin (right axis) in Panel A and the natural logarithm of monthly loan funding volume by different types of investors in each ELR bin in Panel B. The reference month is July 2014. The vertical line corresponds to one year before Moody's downgrade warning on February 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.

Panel A: Unfunded loan applications



Panel B: Funding volume by investor pool

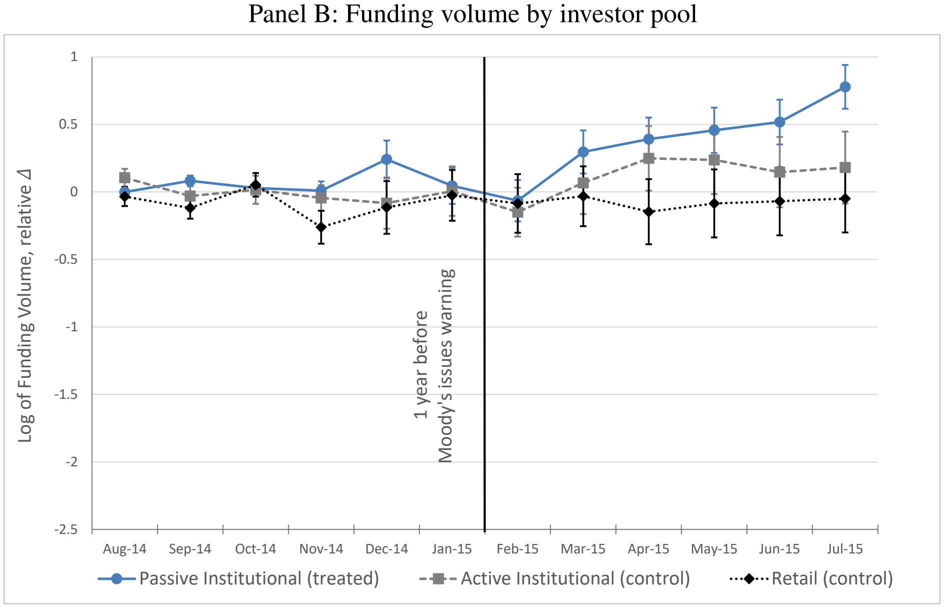
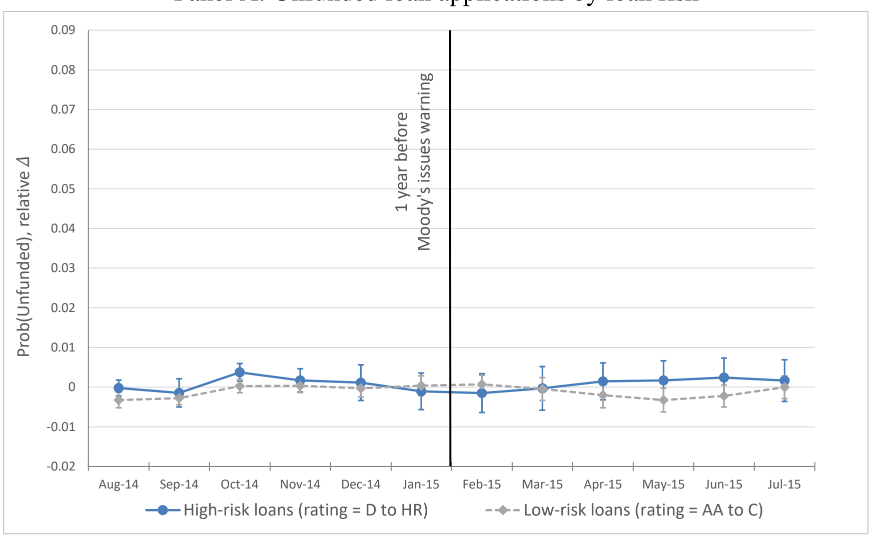


FIGURE F.2

Placebo tests: Heterogeneity analyses around Moody's downgrade warning

This graph plots regression coefficients of the month fixed effects from heterogeneity analyses one year before 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 2014. The vertical line corresponds to one year before Moody's downgrade warning on February 11, 2016. Each regression includes ELR fixed effects and controls for the interest rate. Standard errors are clustered at the ELR level.

Panel A: Unfunded loan applications by loan risk



Panel B: Unfunded loan applications by loan maturity

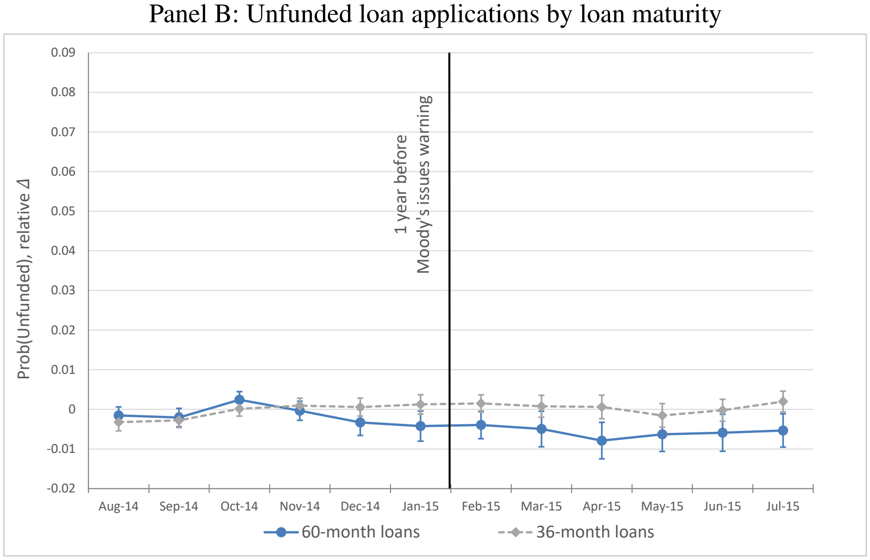


TABLE F.1

Investor response to Moody's warning: Placebo DiD analyses

This table reports the results of the placebo DiD analyses one year before Moody's downgrade warning. The dependent variable in Columns 1–2 is the loan-level probability of not receiving investor funding, and the analyses are at the loan level. The dependent variables in Columns 3–6 is the natural logarithm of monthly loan funding volume by different types of investors in each ELR bin, and the analyses are at the ELR-month level. See Section I 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			
Event Window:	(-2m; +2m)	(-6m; +6m)	(-2m; +2m)		(-6m; +6m)	
	1	2	3	4	5	6
POST	-0.0011	-0.00051	-0.12**		0.49***	
	(-0.91)	(-0.65)	(-2.44)		(10.42)	
POST×PASSIVE_INST				0.30***		0.49***
				(8.27)		(18.42)
POST×ACTIVE_INST				0.31***		0.38***
				(5.95)		(7.70)
PASSIVE_INST				1.96***		1.97***
				(43.55)		(43.27)
ACTIVE_INST				0.76***		0.79***
				(12.33)		(13.15)
INTEREST_RATE	-0.41**	-0.15*	-95.5***	-59.4***	8.20	-11.9
	(-2.20)	(-1.99)	(-10.06)	(-5.75)	(1.37)	(-1.65)
ELR FEs	YES	YES	YES	YES	YES	YES
Month FEs	NO	NO	NO	YES	NO	YES
N	85,111	288,238	260	775	667	1,993
Adj. R^2	0.004	0.002	0.887	0.957	0.908	0.948

TABLE F.2

Investor response to Moody's warning: Placebo heterogeneity analyses

This table reports the results of the placebo heterogeneity analyses around Moody's downgrade warning on February 11, 2016. The dependent variable is the loan-level probability of not receiving investor funding. See Section I 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.

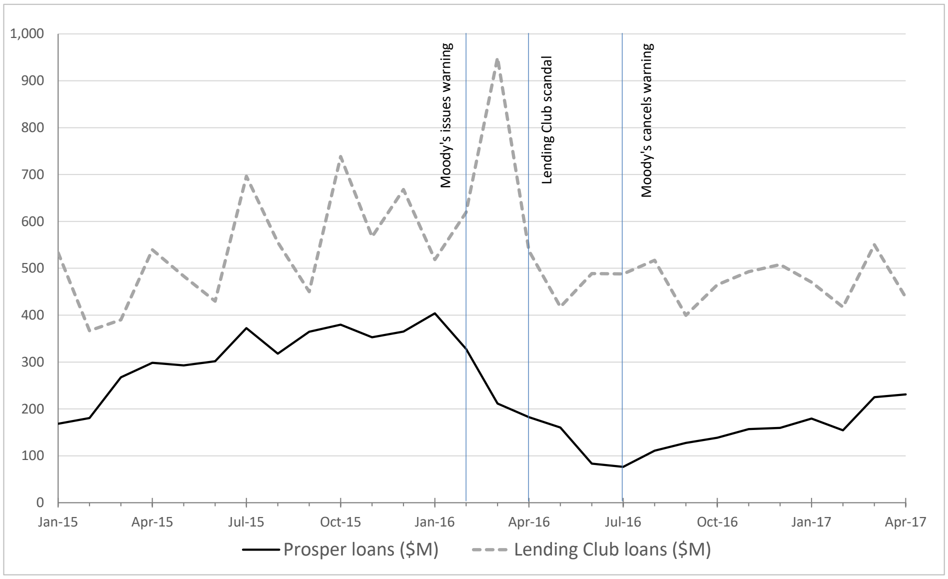
	Unfunded			
Event Window:	(-2m; +2m)		(-6m; +6m)	
	1	2	3	4
$\text{POST}_{pl} \times$HIGH_RISK_LOAN	0.0022		0.00037	
	(1.53)		(0.38)	
$\text{POST}_{pl} \times$LONG_TERM_LOAN		0.0025*		0.00071
		(1.97)		(0.94)
LONG_TERM_LOAN		0.0020**		0.0020***
		(2.11)		(3.41)
INTEREST_RATE	-0.71***	-0.63***	-0.13	-0.13
	(-3.15)	(-3.00)	(-1.11)	(-1.13)
ELR FEs	YES	YES	YES	YES
Month FEs	YES	YES	YES	YES
N	85,111	85,111	288,238	288,238
Adj. R^2	0.004	0.004	0.002	0.002

These results hold in formal regression analyses. The coefficients of the post-February 2015 dummy ($Post_{pl}$) in Columns 1–2 of Table F.1 are small and insignificant, as expected. We observe significant coefficients in the remaining columns (Columns 3–6). However, these coefficients are either substantially smaller than in our main DiD analyses in Table **??** of the main text or they have

FIGURE F.3

Market growth around the 2016 crisis

This graph shows the volume of new loans originated monthly through Prosper and LendingClub. The vertical lines mark major events during this time period.



the opposite sign. Table F.2 reports the results of the placebo heterogeneity analyses. We are interested in the interaction terms between the post-February 2015 dummy and indicators for riskier and long-term loans. All coefficients of these interaction terms are small in size and are either statistically insignificant or marginally significant. Therefore, our findings around the Moody's downgrade warning are not driven by differential seasonal trends across different loan types.

We also investigate a concern that the drop in Prosper's loan volume in early 2016 can be due to a market-wide trend in P2P lending rather than due to the effect of Moody's downgrade warning. It is also possible that the effect is confounded by other events that may have negative spillover effects on Prosper, such as governance problems faced by LC (its largest competitor) in April 2016. To rule out these concerns, we closely examine the monthly volume of loans originated by Prosper and LC during this period in Figure F.3. Despite parallel trends in loan volume for LC and Prosper before February 2016, we observe a large divergence in loan originations by the two P2P lenders in February and March 2016. This divergence is due to a sharp decrease in Prosper's originations. The fact that LC's loan volume dropped later than Prosper's (because of a governance scandal at LC) mitigates the concern that a market-wide event caused the decrease in Prosper's loan volume or that developments at LC confounded the effects of Moody's downgrade warning in the first two months after the announcement.

G. DiD analysis of Hurricane Irma

In this section, we test whether active investors perform incremental screening (above and beyond the platform's screening) of loans requested in areas affected by natural disasters. This test allows us to ascertain that active investors are indeed active loan pickers while passive investors are indeed passive, given than Prosper does not price loans based on local market characteristics and "disaster loans" are risky in expectation.

Our identification relies on a difference-in-differences (DiD) approach around Hurricane Irma, which occurred in September 2017.[6] We focus on a 12-month window around the event. Our

[6]We rely on borrowers' states for geographical location because Prosper's data set does not contain borrower county information and because borrower city names are noisy in the data. In view of these issues, we impose several requirements on natural disasters for this test. First, we focus on hurricanes because the extent of damage in these events is higher than for other disasters (e.g., extensive floods, droughts, inclement winter weather). We restrict our attention to major hurricanes. Second, we require that the disaster affect at least one entire state as opposed to having a localized effect (as with Hurricane Harvey). Third, we require that the hurricane occur toward the later part of our sample because of statistical power issues due low monthly loan origination volume before April 2013.

treatment group is disaster loan applications, which we define as applications from Florida and South Carolina. Our control group comprises untreated applications from all other states, except Texas and Louisiana.[7]

We first employ the event study methodology to examine the dynamics of loan rejection rates by investors after Hurricane Irma. Figure **??** in the main text plots the coefficients of month fixed effects from four distinct regressions of the probability of the application remaining unfunded, namely for treated applications in the active investor pools, untreated applications in the active pools, treated applications in the passive pool, and untreated applications in the passive pool.[8] Note that the coefficient estimates plotted in Figure **??** are not levels. The coefficients are estimated as changes relative to the respective rejection rates in February 2017, the month preceding the 12-month window. Thus, one should interpret this graph in terms of the relative trends in the evolution of rejection rates for the treated and untreated applications rather than in terms of the relative levels of the coefficients.

Figure **??** of the main text shows that investors reject disaster applications after the hurricane at a higher rate. By contrast, rejection rates for untreated applications stay generally flat during the 12-month window around the hurricane. It is noteworthy that rejections of treated applications increase abruptly exactly in the month when the hurricane hit and did so only for applications

[7]We exclude these states because of potential confounding effects from Hurricane Harvey. Including these states in the control group, however, yields similar results both qualitatively and quantitatively. The results are similar, although slightly weaker, when we include Georgia in the treatment group rather than the control group, as can be expected based on the hurricane's path and related damages.

[8]We conduct our analyses within ELR bins to compare applications of the same estimated risk by Prosper. We control for the interest rate to account for any changes in the ELR-rate mapping during this time.

TABLE G.1

DiD analyses of investor funding around Hurricane Irma

This table reports the results of the DiD analyses of loan-level probability of not receiving investor funding for the 12-month window around Hurricane Irma (-6m; +6m). The *treatment group* (TREATED=1) is loan applications by residents of Florida and South Carolina. The *control group* is loan applications by residents of all other states, except for Texas and Louisiana. The reference period is March 2017. POST is an indicator for the loan application listed for investor funding after the hurricane. ACTIVE is an indicator for the loan funded in the active (i.e., active institutional or retail) loan 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)			
	All	Active	Passive	Pooled
	1	2	3	4
POST×TREATED	0.0070***	0.042***	0.0011	0.0013
	(5.03)	(4.94)	(1.40)	(1.56)
POST×TREATED×ACTIVE				0.040***
				(4.74)
POST×ACTIVE				-0.018***
				(-7.22)
TREATED×ACTIVE				0.0042
				(0.81)
ACTIVE				0.040***
				(18.00)
INTEREST_RATE	-0.19	0.21	0.069	0.043
	(-1.05)	(0.28)	(0.58)	(0.28)
ELR FEs	YES	YES	YES	YES
Month FEs	YES	YES	YES	YES
State FEs	YES	YES	YES	YES
N	249,278	36,996	212,281	249,278
Adj. R^2	0.003	0.011	0.001	0.021

allocated to the active pools, which are open to investor screening. In addition to showing a relative increase in rejection rates for disaster applications after the hurricane, Figure **??** shows that the parallel trends assumption for our DiD analyses likely holds. The trends in rejection rates are similar for treated and untreated loan applications within active and passive pools before the event. We find similar results in pooled regression analysis where we interact month fixed effects with the indicator of a treated application (not reported).

We find consistent evidence of active investor screening in formal DiD regression analyses reported in Table G.1. The probability of not receiving investor finding is higher for disaster applications after the hurricane, an increase of 0.71% or 82.6% relative to the mean probability of not receiving funding (Column 1). The latter number, however, can be interpreted as a decrease in the investor funding rate of only 0.7 pp in view of loan funding rates in the upper 90 pp during this time period. Also consistent with graphical evidence, rejection rates go up only for applications in active investor pools (Column 2), but not in the passive pool (Column 3). This coefficient in Column 2 implies an increase in the rate of disaster applications remaining unfunded by 4.2 pp, which is almost a five-fold increase in investor rejection rate in the active investor pools, or a decrease in the funding rate of 4.24% relative to the mean funding rate over this time window. The respective coefficient in Column 3 is 38 times smaller and is statistically insignificant. These results hold in the pooled triple-difference analyses in Column 4.

H. Randomization across investor pools

Prosper claims that it is randomly allocating loan applications across the three investor pools: (1) Whole Active (i.e., active institutional), (2) Whole Passive (i.e., passive institutional), and (3) Fractional (i.e., retail). We test for such random allocation. The presence of three, rather than two, pools requires joint tests of randomization across the pools, which makes the traditional regression analyses with a dummy for each investor pool as the dependent variable inappropriate.

Therefore, we adopt the methodology developed in the peer effects literature, specifically in **?** with the bias adjustment from **?**. This methodology appropriately accounts for the presence of multiple pools and cohorts.

Another complication of our empirical setting is the conditional character of randomization. The allocation among pools is conditional on Prosper rating because Prosper bases the weights of each investor pool within a rating on its estimate of the relative investor demand for loans assigned this particular rating. While the exact Prosper estimates of loan demand are unobservable, we assume that Prosper conducts its demand estimation on a monthly basis due to the monthly nature of loan orders by passive institutional investors that we anecdotally observe from loan purchase contracts.

Table H.1 shows the results of a randomization test. We examine ELR in Column 1 and a subset of other borrower characteristics in Columns 2–9. We focus on characteristics that are most correlated with loan default risk and are most significant in regressions of default probability (not reported). Each dependent variable, X, is a characteristic (e.g., ELR) measured at the loan application level which is regressed on the average characteristic of all other applications in the same pool and month-rating cohort, $AVG(X)_{-i,mrp}$ (also called leave-me-out characteristic). We include the average value of X for all other loan applications in the same month and with the same rating (i.e., excluding the value for the application itself) as an additional control. This adjustment is necessary to correct for the bias in the randomization test, as per **?**. The coefficient on the bias correction term is approximately equal to $-(N-1)$, where N is the average number of observations in a month–rating–pool category, and can thus expected to be large in our setting (see **?**, p. 12). The coefficient is statistically significant in the presence of a bias.

TABLE H.1

Randomization test

This table reports the results of the randomization test for the allocation of loan applications on Prosper across three investor pools: (1) Whole Active (i.e., active institutional), (2) Whole Passive (i.e., passive institutional), and (3) Fractional (i.e., retail). The sample period is November 2013 (introduction of Passive Institutional pool) to March 2017. Dependent variables are defined in Columns 1–9 and are measured at the loan application level. See Section I for the definitions of variables. Variable AVG(X)$_{-i,mrp}$ is the average characteristic of all other applications in the same month–rating–pool category. Variable AVG(X)$_{-i,mr}$ is the bias correction term, which is the average characteristic of all other applications in the same month–rating category. The estimates of the intercept and fixed effects are omitted for brevity. 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.

	Dependent variable, X =								
	ELR	Log of income	Debt-to income	Self-employed	Trades last 6 m.	Inquiries last 6 m.	Prior loan	Long-term loan	SCOREX score
	1	2	3	4	5	6	7	8	9
AVG(X)$_{-i,mrp}$	0.078	0.26*	0.086	-0.051	0.11	0.037	0.089	0.12	0.036
	(1.01)	(1.96)	(1.03)	(-0.45)	(1.32)	(0.27)	(0.99)	(1.59)	(0.28)
AVG(X)$_{-i,mr}$	-1,125***	-918***	-858***	-1,086***	-810***	-817***	-1,123***	-2,199***	-787***
	(-11.4)	(-7.69)	(-8.55)	(-9.18)	(-7.67)	(-6.89)	(-9.69)	(-15.9)	(-7.54)
Month-rating FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	837,998	837,998	836,667	837,998	836,648	836,667	837,998	837,998	835,182
Adj. R^2	0.976	0.323	0.335	0.369	0.302	0.325	0.442	0.654	0.464

The results in Table H.1 show that Prosper's assignment of loan applications to investor pools is likely as good as random. If the assignment is indeed random or if idiosyncrasies in the assignment can be ignored (e.g., when ex post nonrandom assignment is due to pure chance of an outlier allocated to one of the pools), the coefficient of $AVG(X)_{-i,mrp}$ should be close to zero and insignificant. If the assortative matching within pools is perfect (e.g., when institutional investors get only applications from self-employed borrowers and retail investors only get applications from employed borrowers), then the absolute value of the coefficient should be close to one (and statistically significant). As expected, most coefficients in Table H.1 are small and insignificant. Although the coefficient of *Log of income* is statistically significant at 10%, the significance becomes much lower after winsorization and goes beyond the conventional levels of significance. Thus, we cannot rule out the hypothesis that Prosper's allocation is random.

I. Definitions of variables

Variable	Definition
ACTIVE	Indicator for the loan funded in the active institutional or retail pool
ACTIVE_CREDIT_CARDS	Number of open credit cards with activity in the last 6 months
ACTIVE_INST	Indicator for the loan funded in the active institutional loan pool
CANCELLATION_RATE	Rate at which the platform cancels loan applications
CONSOLIDATION_PURPOSE	Indicator of the loan requested for debt repayment or debt consolidation
CREDIT_CARD_BALANCE	Average credit line of open credit card trades verified in past 12 months
DEBT_TO_INCOME	Ratio of the monthly debt of the borrower to their stated monthly income, including the Prosper loan
DEFAULT	Indicator of a loan becoming delinquent over the life of the loan
EXPECTED_RETURN	Annual expected return on a loan calculated by discounting expected monthly cash flows on the loan given the default probability in each month based on a hazard model of default; a detailed description of the procedure is available from the authors upon request
EXPECTED_DEFAULT	Expected probability of loan becoming delinquent over the life of the loan calculated using a hazard model of default
EMPLOYED	Indicator for the borrower being employed at the time of application
ESTIMATED_LOSS_RATE (ELR)	Amount of principal that would be lost due to defaults and charge-offs on the loan estimated by Prosper
FICO_SCORE	Midpoint value of the FICO credit score, as binned by Prosper
FRACTION	Fraction of all loan applications allocated to the respective investor pool
FUNDING_RATE	Rate at which loan applications allocated to the respective investor pool are funded by investors
HIGH_RISK_LOAN	Indicator for a loan rated as D, E, or HR by Prosper

Variable	Definition
INTEREST_RATE	Interest rate on the loan set by the platform
LOAN_AMOUNT	Loan amount requested by the borrower
LOAN_MATURITY	Number of months over which the loan amortizes
Log of Application Volume	Natural logarithm of monthly application volume in each ELR bin
Log of Funding Volume	Natural logarithm of monthly loan application volume within each ELR bin funded by the respective pool of investors
LOG_OF_INCOME	Natural logarithms of the borrower's monthly income
LONG_TERM_LOAN	Indicator of the 60-month loan
MATCH	Indicator that a loan listing can be matched to the loan data
MONTHLY_DEFAULT_RATE	Annualized default rate on all outstanding loans on Prosper in a given month
MONTHLY_INCOME	Borrower's monthly income
MORTGAGE	Indicator for the positive balance on real estate trades of the borrower
PAST_BANKRUPTCIES	Number of public record bankruptcies
TRADES_EVER_DELINQUENT	Percentage of trades ever delinquent
PASSIVE_INST	Indicator for the loan funded in the passive institutional pool
POST	Indicator for the loan application listed for investor funding after Moody's downgrade warning in February 2016
POST_{pl}	Indicator for the loan application listed for investor funding after February 2015 (placebo Moody's downgrade warning)
PRIOR_LOAN	Indicator for the borrower having at least one prior Prosper loan
Prob(Default)	Probability of a loan becoming delinquent over the life of the loan
Prob(Unfunded)	Probability of a loan not receiving enough commitments from investors for the loan to originate

Variable	Definition
RATING	Proprietary rating developed by Prosper allowing to analyze an application's level of risk that ranges from 1 (rating HR) to 7 (rating AA), 7 having the lowest risk
REALIZED_DEFAULT_RATE	Annualized probability of loan becoming delinquent, averaged over the life of the loan
REALIZED_RETURN	Realized return on a loan calculated as the annualized geometric average of monthly total returns on the loan based on the value of the loan balance in a given month, the loan payment received in that month, and he reinvested value of payments received in prior months; a detailed description of the procedure is available from the authors upon request
RECYCLED_RETAIL	Indicator for the loan application that was unfunded in an institutional pool and transferred to the retail pool
RESUBMITTED	Indicator for an application resubmitted after loan cancellation
SCOREX_SCORE	Midpoint value of the Scorex Plus credit score binned by Prosper
SELF_EMPLOYED	Indicator for the borrower being self-employed at the time of application
TRADES_LAST_6M	Total number of trades opened within 6 months of the profile date
UNFUNDED	Indicator of a loan not receiving enough commitments from investors for the loan to originate
YEARS_EMPLOYED	Length of employment with the current employer, measured in years