MANAGEMENT SCIENCE
Vol. 62, No. 6, June 2016, pp. 1554–1577
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2015.2181

Screening Peers Softly: Inferring the Quality of Small Borrowers

Rajkamal Iyer
MIT Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, riyer@mit.edu

Asim Ijaz Khwaja
Harvard Kennedy School, Harvard University, Cambridge, Massachusetts 02138; and National Bureau of Economic Research, Cambridge, Massachusetts 02138, akhwaja@hks.harvard.edu

Erzo F. P. Luttmer
Dartmouth College, Hanover, New Hampshire 03755; and National Bureau of Economic Research, Cambridge, Massachusetts 02138, erzo.fp.luttmer@dartmouth.edu

Kelly Shue
Booth School of Business, University of Chicago, Chicago, Illinois 60601; and National Bureau of Economic Research, Cambridge, Massachusetts 02138, kelly.shue@chicagobooth.edu

This paper examines the performance of new online lending markets that rely on nonexpert individuals to screen their peers' creditworthiness. We find that these peer lenders predict an individual's likelihood of defaulting on a loan with 45% greater accuracy than the borrower's exact credit score (unobserved by the lenders, who only see a credit category). Moreover, peer lenders achieve 87% of the predictive power of an econometrician who observes all standard financial information about borrowers. Screening through soft or nonstandard information is relatively more important when evaluating lower-quality borrowers. Our results highlight how aggregating over the views of peers and leveraging nonstandard information can enhance lending efficiency.

Keywords: peer-to-peer credit markets; market-based lending; crowd sourcing; screening; market inference; information and hierarchies; soft information

History: Received July 15, 2014; accepted February 2, 2015, by Amit Seru, finance. Published online in *Articles in Advance* August 12, 2015.

1. Introduction

An important challenge for financial markets is understanding how to best screen borrowers when allocating credit. Predicting borrower creditworthiness is understandably hard. Whether a person defaults on a loan is not just the result of a mechanical financial calculus but is also driven by the complexities and idiosyncrasies of human behavior. Even when facing an identical financial situation, individuals may vary in their ability and willingness to meet their financial obligations.

Traditionally, this screening process has been performed by banks and financial experts using summary statistics of hard information, such as credit scores, compiled through sophisticated models based on the payment history of the borrower along with other verified information.[1] Recently, technological advances have allowed for new peer-to-peer online lending platforms that provide an opportunity for individuals to assess the creditworthiness of their peers. In addition to standard financial variables, these markets also provide nonstandard or "soft" information about borrowers. The promise of peer-to-peer markets is that lending decisions are based on the collective assessment of many individual lenders, and that these lenders not only have access to standard financial variables but can also base their assessment on nonstandard or soft information, drawing on their own experiences and understanding of human behavior. The downside, however, is that lenders in such markets typically have limited experience and no formal training in judging borrower creditworthiness. Further, the nonstandard information is self-reported by borrowers and not readily verifiable. Given the growing importance of these types of markets in the lending industry, a better understanding of the functioning of these markets is important.

In this paper, we examine screening in one such peer-to-peer lending market (Prosper.com). We investigate

[1] Although soft information can also be used by banks, doing so is time consuming. Thus, often credit scores are used in the initial phase of screening, especially for smaller borrowers. For instance, Petersen (2004) argues that technological advances and competitive pressures have pushed banks more toward using hard information.

how screening in this market compares with the screening achievable based on the borrower's exact credit score (which is not observable by Prosper lenders) or the best possible screening achievable using all hard information typically available to banks. We examine the role of soft or nonstandard information in screening borrowers, and how the role of such information varies with borrower quality.

The Prosper marketplace forms an excellent setting to quantify collective inference by nonexpert market participants and to examine how inference depends on different sources of information, such as hard and soft/nonstandard information. Borrowers post loan listings on Prosper.com and then multiple individual lenders bid to fund a portion of the loan at a desired interest rate. Lenders have access to standard hard financial information commonly used by banks, such as the borrower's income and number of past delinquencies. In addition, lenders can view nonstandard information, such as the maximum interest rate the borrower is willing to pay as well as softer and less quantifiable information, such as the borrower's picture and a textual description of his or her reasons for the loan application.[2] The interest rate for a funded loan is determined through sequential bidding and reflects the lenders' collective perception of the quality and, hence, the creditworthiness of the borrower.[3]

We exploit a unique feature of the Prosper marketplace in our proprietary data: whereas lenders only see the borrower's aggregate credit category, we as researchers observe a borrower's exact credit score—a much finer measure of the borrower's underlying creditworthiness. We first examine the efficiency of screening in these markets by comparing the power of the interest rate (proxying for lenders' inference) set by market participants in predicting default against the default predictability obtained using the borrower's exact credit score. In theory, the credit score should be the best available aggregator of standard hard financial data in terms of predicting default. The credit score in question—the Experian ScoreX Plus credit score—is estimated using a sophisticated prediction model that uses a superset of the hard financial variables available to Prosper lenders. It is designed specifically to improve screening of low-quality small borrowers, such as those in the Prosper marketplace. However, individual Prosper lenders may be able to improve upon the predictive power of this credit score because they make use of nonstandard/soft borrower information in addition to standard financial variables. At the same time, the interest rate set by lenders could be less predictive of default because lenders tend to be nonexpert individuals, may be driven by personal biases, or lack access to the larger pool of data on which the credit score is based.

We find that lenders in this market substantially outperform the credit score in terms of predicting default. We first show, by examining the R^2, that the market interest rate on loans explains more variation in ex post default than the credit score can explain. We then present a more formal comparison using tools from signal detection that are common in credit scoring. Specifically, we construct "receiver operating characteristic" curves and show that the "area under the curve" (AUC)—a simple metric used to judge the screening power of a screening score—is both large in an absolute sense and also significantly higher for the market interest rate than for the borrower's credit score. In particular, the interest rate set by lenders predicts default 45% more accurately than the borrower's credit score. We note that the median number of lenders (bidders) per funded loan is 60 (115), which implies that a relatively small number of individuals effectively predict loan outcomes.

We present the credit score as one natural benchmark by which to measure the quality of inference. However, a traditional lender, even one that relies exclusively on standard hard financial information, may improve upon the credit score by optimally weighting observable hard information in ways that are tailored to predict default within its particular borrower base. In addition, such lenders may incorporate very recent hard information on borrower employment that is not reflected in the credit score. As an even more demanding benchmark, we use the best possible default predictor that an econometrician could construct using all available standard financial information, including updated information on employment. This benchmark is particularly demanding because the econometrician holds an unfair advantage: to estimate the econometrician's predictions, we use in-sample data on default realizations, something that is unavailable to any bank or lender at the time of loan origination. We find that the market interest rate still achieves a remarkable 87% of the AUC of the econometrician.

Next, we examine whether the extent of lender inference varies by borrower quality. We find that inference is greater in the higher credit categories (better borrowers) than in the lower ones. We also explore

[2] Some of the additional information in the Prosper marketplace is soft in a strict sense of the word (not codeable) and some of it is merely nonstandard (e.g., the borrower's reservation interest rate, which can be coded). We use the term "soft/nonstandard" to capture both of these sources of information.

[3] The loan is funded only if the total amount bid equals or exceeds the amount requested by the borrower, and the final interest rate is determined by the highest reservation interest rate among the set of lenders that bid successfully.

how lenders weight standard financial versus nonstandard/soft information in forming their predictions of default. We find that both sources of information are important in screening, but that inference from soft/nonstandard information appears relatively more important when assessing worse borrowers. These results suggest that crowd sourcing of nonstandard or softer borrower information may be particularly helpful in terms of screening lower-quality borrowers.

The results presented so far explore the ability of lenders to use listing information, particularly nontraditional and soft information, to predict *default*. In supplementary analysis, we also explore the extent to which lenders infer the information content of the missing credit score itself. We caveat this portion of the analysis by acknowledging that, by construction, the credit score only measures the dimension of creditworthiness that can be captured by hard financial information. However, an advantage of this supplementary analysis is that it allows an exact decomposition of inference by source of information.

We find that, within a given credit category (spanning 40 points in the credit score), lenders infer a third of the difference in creditworthiness that is captured by a borrower's exact credit score. This effect is economically meaningful because such a degree of inference allows lenders to offer a rate that is 140 basis points lower for borrowers at the top of a typical credit category than for borrowers at the bottom of that category. Given that the credit score is computed based on proprietary formulas and not all variables that go into the computation are available to lenders, it was by no means obvious that lenders would be able to piece together the information provided in the listing and infer a third of the true credit score.[4] Moreover, we estimate that lenders infer 69% of what they could have potentially extracted from the information provided on the Prosper website.

We find that, along the credit-score dimension, lenders base most of their inference on standard financial variables. Yet, soft/nonstandard variables also contribute to inference. Of the soft/nonstandard variables, we find that lenders draw the most inference from the maximum interest rate that a borrower posts she is willing to pay for the loan. This rate likely serves as a credible signal that satisfies the single-crossing property because (i) borrowers posting too low a rate risk not having the loan funded and (ii) it is costlier for lower-quality borrowers to risk not having the loan funded as they have fewer alternate funding options. Our results suggest that, consistent with models of cheap talk, individuals pay greater attention to the more credible signals sent (Crawford and Sobel 1982, Farrell and Rabin 1996).[5] Consistent with our primary analysis, we find a high degree of inference from the noncoded component of the listing, especially among the lower credit categories. In general, coding soft information is challenging because it is difficult to quantify the information content of pictures or lengthy personal text descriptions. We can nevertheless measure the inference drawn from uncoded information by computing this inference as a "residual," that is, the variation of interest rates with the exact credit score that remains after controlling for a very flexible functional form of all coded information.

Our paper contributes to the literature that examines how nonexpert individuals aggregate information, such as also occurs in prediction markets (Wolfers and Zitzewitz 2004, Arrow et al. 2008). Consistent with the results from prediction markets, we find that individual lenders perform quite well in aggregating information. A related strand of literature examines information aggregation, inference, and learning in markets more generally. There are several theoretical papers that focus on information aggregation through prices (Grossman 1976; Townsend 1978; Grossman and Stiglitz 1980; Vives 1993, 1995). Another strand of literature focuses on learning in decentralized markets (Wolinsky 1990, Duffie and Manso 2007, Duffie et al. 2009). On the empirical front, Biais et al. (1999) and Davies (2003) examine learning in the preopening period in equity markets. There are also several experimental papers that examine price formation in asset markets (Plott and Sunder 1988, Forsythe and Lundholm 1990, Bronfman et al. 1996, Cao et al. 2000, Hanson et al. 2006).

Our work complements the recent literature that specifically examines lending in peer-to-peer markets, summarized in Morse (2015). Pope and Sydnor (2011), Ravina (2012), and Theseira (2009) examine whether these markets display discrimination based on personal attributes, such as race and physical appearance. Peer-to-peer markets may also make better use of social network information. Although Freedman and Jin (2015) find evidence of adverse selection due to informational problems faced by lenders in Prosper, they also find that social networks (endorsements by friends) can help alleviate these problems.[6] In a

[4] The R-squareds of regressions of credit scores within each credit category on a flexible specification of all hard information variables are low (average R^2 of around 0.3).

[5] The borrower maximum rate also censors our observations when the interest rate that the market requires to fund a listing exceeds the borrower maximum rate. As we explain in more detail in the methodology section, our estimation strategy corrects for this mechanical censoring effect.

[6] On the other hand, Hildebrand et al. (2014) provide evidence that individuals misinterpret signals of group leaders.

similar spirit, Lin et al. (2013) find that stronger and more verifiable relational networks help reduce the adverse selection problems in Prosper. Butler et al. (2014) find that consumers residing in counties with a larger supply of traditional sources of finance seek loans at lower interest rates from an alternative source of finance (an online person-to-person consumer lending intermediary) than do similar borrowers residing in counties with poor access to finance. In contrast to the papers that document lending behavior and biases in peer-to-peer markets, our focus is on evaluating the screening ability of these markets and decomposing the extent of inference along different information sources.

Our paper also contributes to the literature that examines the importance of soft information in screening. Liberti and Mian (2009), Agarwal and Hauswald (2010), and Rajan et al. (2015) find that greater hierarchical distance discourages the use of subjective and more abstract information in banks. Our findings of effective use of soft/nonstandard information in online credit markets, where the hierarchical distance between the borrower and lender is small, is consistent with these papers. Our paper adds to this literature by quantifying the increase in the accuracy in assessing borrower creditworthiness that arises from the use of soft/nonstandard information. The role of soft information also received considerable attention in research on the mortgage market and housing crisis (see, e.g., Keys et al. 2010, Agarwal et al. 2011, and Jiang et al. 2014). By decomposing the extent of inference from different sources of information, our paper sheds light on the relevance of different types of information in markets (Crawford and Sobel 1982, Farrell and Rabin 1996, Berger and Udell 2002, Berger et al. 2002, and Petersen 2004).

Although we examine one particular market, we believe our results are informative about alternative screening mechanisms more broadly. Our results suggest that collective inference by nonexpert participants can be remarkably accurate, and that participants successfully draw inference from soft/nonstandard information. One form of soft/nonstandard information that lenders found particularly informative was the signal that borrowers sent by their choice of reservation interest rate, which suggests that similar signals may also improve screening in other contexts. Overall, our results highlight the importance of soft/nonstandard information in screening smaller borrowers and emphasize the screening ability of alternative markets.

2. Context and Data

2.1. Context

Peer-to-peer lending on the Internet enables individual lenders to locate individual borrowers, and vice versa. In the United States, there are around 12 active, online peer-to-peer lending websites. Furthermore, online peer-to-peer lending markets are on the rise (Samaad 2014). We analyze data from Prosper.com, an online peer-to-peer lending marketplace that was founded in February 2006. Prosper focuses on U.S. clients and intermediates capital mostly between individual lenders and small borrowers. It has funded over $440 million in loans and currently has 1,590,000 members.

All Prosper loans are personal, three-year, fixed-rate, unsecured loans. Borrowers request loans by creating public listings on the Prosper.com website. They choose the amount of money to request (up to $25,000) and the duration of the loan listing (3, 5, 7, or 10 days). The online listing consists of three components: pictures, listing text, and credit information. The pictures and text contain unverified soft information provided voluntarily by the borrower. Often, borrowers describe why they need a loan, why they are good credit risks, and their income and expenditure flows. Some borrowers also post optional pictures of themselves or of themes related to their loan purpose. The third listing component, credit information, contains verified hard information obtained by Prosper through a credit check. The credit information section contains information on each borrower's delinquencies, credit lines, home ownership status, debt, inquiries, and public records. A sample listing is provided in Online Appendix D (online appendices available as supplemental material at http://dx.doi.org/10.1287/mnsc.2015.2181).

The credit information also contains the borrower's credit category. Prosper assigns each borrower to one of seven credit categories based on the borrower's Experian ScoreX PLUS credit score. Prosper lenders and borrowers observe only these credit categories, not the exact credit score. The relationship between credit scores and credit categories is shown below.[7]

Category:	HR	E	D	C
Score:	520–559	560–599	600–639	640–679
Category:	B	A	AA	
Score:	680–719	720–759	760–900	

[7] The above credit-category chart reflects the Prosper classification at the end of our sample period. A major change in credit-category criteria occurred on February 12, 2007. Prior to the credit criteria change, the credit categories were set such that HR(0-539), E(540-600). After February 12, 2007, credit scores below 520 were disqualified and the credit-category stratification was finalized to the numbers described in the chart. For consistency of results, we restrict our sample to the period after February 12, 2007. However, results are robust to using the sample from before February 12, 2007 (see Table 4 in §4).

In addition, borrowers can join borrower groups led by "group leaders." The ratings and financial rewards of group leaders depend on the payment profiles of the group's members. Group leaders can write public messages endorsing the borrower or pledging to exert pressure on the borrower to repay and can bid on group members' loans. In addition, borrowers can friend other registered Prosper users, who then can add friend endorsement texts to listings and can cast friend bids on listings.

After listings are posted, lenders can browse through Prosper's website for listings to bid on. Multiple lenders can bid on and fund each listing. Lenders can bid on portions of listings ($50 minimum) and set the lowest interest rate at which they are willing to fund the listing. The bidding begins at the maximum interest rate the borrower is willing to pay. The listing is funded only if the total amount of money bid by lenders matches or exceeds the loan amount requested by the borrower. In the latter case, the interest rate is bid down. Lenders with lower reservation interest rates are given priority in the bidding hierarchy. The final interest rate is determined by the highest reservation interest rate among the set of lenders that successfully bid for the loan.[8]

After the listing is funded and approved by the borrower, the borrower begins to make monthly payments that are divided across lenders in proportion to each lender's winning bid size. The borrower never directly interacts with the lenders, and all payments are routed via Prosper. If a borrower is late in making payments or defaults on the loan, his behavior is reported to the major credit agencies and the borrower's credit rating suffers. If the borrower is late for four or more months, Prosper sells the loan to a collection agency and splits the proceeds among the lenders.

2.2. Data

Our data set contains all credit information variables displayed on a borrower's loan listing, as well as the text of the listing and the complete history of each borrower's loan repayment stream. In addition, our data include the exact credit score (unobserved by lenders and borrowers) for each borrower.[9] Our sample contains all listings posted between February 12, 2007, and October 16, 2008.[10] The sample covers 194,033 listings, 17,212 of which were funded.

Table 1 provides summary statistics both for the universe of listings and for the set of funded listings (listings that resulted in loans). We divide the set of variables into standard financial variables and soft/nonstandard variables. The standard financial variables include hard information from the borrower's credit report that is typically used by traditional banks. As expected, funded listings tend to have borrowers with better credit scores—in particular, funded listings tend to have far fewer "high risk" borrowers (those in the lowest credit categories). Still, 30.6% of the funded listings default at some point in the three-year duration of the loan, with the default rate ranging from 14.7% in credit category AA to 51.6% in credit category HR. Because defaults often occur after some of the principal has already been repaid, and because some of the principal gets recovered when a defaulted loan is sold off to a collection agency, the fraction of the principal repaid at the end of the loan term is higher than one minus the default rate. The fraction repaid is 79.7% on average, and ranges from 91.0% in credit category AA to 62.5% in credit category HR. Among the universe of listings, the average loan amount requested is $8,015 and the maximum interest rate borrowers are willing to pay is 21% on average.

The soft/nonstandard variables capture soft information that may be difficult to fully quantify, as well as information that is quantifiable but not typically used by banks, i.e., nonstandard variables that represent borrower choices. Borrower choice variables include the maximum interest rate the borrower is willing to pay, the listing duration (number of days the listing remains public), and listing category (e.g., debt consolidation or student loan). We also code basic proxies for soft information, such as whether the borrower posts a picture or the number of words used in the listing text descriptions. We code the soft infor-

[8] Recently, Prosper.com moved to a system where they create portfolios of loans for individuals based on the desired credit categories for which individuals are interested in lending. In addition, because of competition from other lending platforms such as LendingClub.com and regulatory oversight from the SEC following the financial crisis, Prosper moved to a system of preset rates (based on Prosper's internal risk rating) within narrow categories of loans and no longer follows an auction-based approach. Lenders screen by choosing whether to accept the interest rate.

[9] Note that even borrowers do not have access to the exact Experian ScoreX PLUS credit score obtained from the credit rating agency because it is not available for purchase by borrowers. We are able to work with this data under a nondisclosure agreement that safeguards the confidential and proprietary nature of some of the variables in the data set.

[10] Prosper entered a "quiet period" in October 2008, during which it ceased making new loans in anticipation of an SEC cease-and-desist procedure. Prosper emerged from the quiet period in July 2009 using a new system of classifying prospective borrowers into credit categories. We therefore do not use data on loans originating after October 2008. We also use data from May 2006 to February 12, 2007 as part of a robustness check. However, we exclude data from this period in our baseline sample because the credit-category boundaries changed on February 12, 2007. See §2.1 for more details.

Table 1 Summary Statistics

	All listings		Funded listings	
	Mean	S.D.	Mean	S.D.
Loan outcomes				
Annual lender interest rate			0.166	0.068
Default dummy			0.306	
Credit category HR			0.516	
Credit category E			0.424	
Credit category D			0.356	
Credit category C			0.318	
Credit category B			0.305	
Credit category A			0.234	
Credit category AA			0.147	
Fraction of loan repaid			0.797	0.334
Credit category HR			0.625	0.406
Credit category E			0.708	0.377
Credit category D			0.762	0.352
Credit category C			0.793	0.334
Credit category B			0.798	0.329
Credit category A			0.852	0.292
Credit category AA			0.910	0.235
Credit scores and categories				
Credit score	609.5	73.8	676.0	74.5
Credit category dummies				
Credit category HR	0.343		0.068	
Credit category E	0.164		0.074	
Credit category D	0.178		0.173	
Credit category C	0.136		0.211	
Credit category B	0.082		0.183	
Credit category A	0.055		0.140	
Credit category AA	0.044		0.152	
Standard financial variables				
Amount requested ($)	8,015	6,577	6,761	5,788
Number of current delinquencies	2.89	4.54	0.77	2.28
Number of delinquencies, last seven years	9.68	15.78	4.30	10.52
Number of public record requests, last 10 years	0.57	1.20	0.33	0.83
Total number of credit lines	25.61	14.57	24.30	14.29
Number of credit score inquiries, last six months	3.71	4.45	2.38	3.35
Amount delinquent ($)	3,191	12,662	855	4,504
Bank card utilization (total balances/total limits)	0.63	0.42	0.54	0.37
Number of public records, last 12 months	0.07	0.34	0.03	0.22
Number of current credit lines	8.52	6.08	9.70	5.89
Number of open credit lines	7.51	5.41	8.34	5.22
Revolving credit balance ($)	13,446	33,874	16,773	38,030
Debt-to-income ratio	0.54	1.37	0.33	0.90
Fraction homeowners	0.37		0.48	
Credit history age (years)	13.3	7.1	13.4	7.2
Length of current employment status (months)	20.91	51.90	22.73	53.52
Personal annual income dummies				
N/A or unable to verify	0.053		0.025	
Not employed	0.021		0.007	
$1 – $24,999	0.163		0.120	
$25,000 – $49,999	0.402		0.372	
$50,000 – $74,999	0.211		0.253	
$75,000 – $99,999	0.078		0.117	
$100,000+	0.064		0.101	
Employment status dummies				
Full time	0.812		0.859	
Part time	0.041		0.040	
Self-employed	0.096		0.074	
Retired	0.028		0.020	
Not employed	0.023		0.008	

Table 1 (Continued)

	All listings		Funded listings	
	Mean	S.D.	Mean	S.D.
Soft/nonstandard variables				
Borrower maximum interest rate	0.21	0.09	0.21	0.08
Duration of loan listing dummies				
3 Days	0.044		0.037	
5 Days	0.046		0.055	
7 Days	0.693		0.661	
10 Days	0.218		0.247	
Listing category dummies				
Not available	0.386		0.380	
Debt consolidation	0.281		0.262	
Home improvement loan	0.024		0.033	
Business loan	0.098		0.100	
Personal loan	0.114		0.121	
Student loan	0.025		0.024	
Auto loan	0.017		0.017	
Other	0.056		0.063	
Bank draft annual fee dummy	0.010		0.007	
Borrower lists city of residence dummy	0.11		0.16	
Borrower provides image dummy	0.54		0.69	
Characteristics of listing text				
HTML character number	283	271	309	350
Text character number	963	716	1,106	806
Average word length	4.63	0.58	4.59	0.55
Average sentence length	122.75	97.14	106.96	68.62
Number of numerics	13.03	11.31	14.49	14.32
Percent of words misspelled (%)	0.03	0.03	0.03	0.04
Number of dollar signs	8.98	5.78	8.49	7.25
Percent of listing as signs (%)	0.23	0.88	0.46	1.26
Number of characters in listing title	30.76	13.74	32.36	13.54
Member of group dummy	0.18		0.30	
Group leader reward rate dummies				
0%	0.916		0.867	
0.25%	0.002		0.010	
0.50%	0.015		0.046	
0.75%	0.001		0.002	
1.00%	0.034		0.047	
1.50%	0.004		0.007	
2.00%	0.019		0.017	
3.00%	0.006		0.003	
4.00%	0.003		0.001	
Number of friend endorsements	0.324	0.769	0.519	0.973
Observations	194,033		17,212	

Notes. For brevity, we do not summarize 66 occupation and 52 state dummies (including District of Columbia and Puerto Rico). These are included as controls in the relevant specifications in Tables 5 and 7, and the online appendix tables. *Default dummy* equals one if the loan is three or more months late at the end of the three-year loan term. *Percent of listings as signs* refers to the percentage of the listing composed of non-alpha-numeric characters. *HTML character number* refers to the number of HTML formatting characters in the listing and reflects the extent to which borrowers formatted their listings. *Public records* includes bankruptcies, judgments, tax liens, court records, and overdue child support. *Bank draft annual fee dummy* equals one if the borrower elected to pay a 1% annual fee for not using the electronic funds transfer option.

mation to roughly estimate the relative importance of pictures, listing text, friend endorsements, etc., for lender inference. However, we do not attempt to fully quantify the large selection of soft information available in Prosper listings. Rather, as we explain in the next section, we develop a methodology to measure how much inference is drawn from *uncoded* sources of listing content.

To assess the extent to which the findings from this setting (peer-to-peer markets) are applicable to a more general population, we compare our distribution of credit scores to the distribution of the credit scores (for the general population) reported by Keys et al. (2010). We find that distribution of potential applicants in the peer-to-peer market is more tilted toward the lower credit scores. The average score of a loan applicant is 610 in the peer-to-peer setting as against an average for the general population of approximately 680. Thus, the peer-to-peer market caters more to borrowers of lower credit quality as measured by credit scores (where the screening problems are more severe).[11] We report many results separately for low-quality borrowers and high-quality borrowers, and we suspect that one should place more weight on our results for high-quality borrowers if one wants to apply our results to a more general population.

3. Methodology

3.1. Estimating Screening Performance

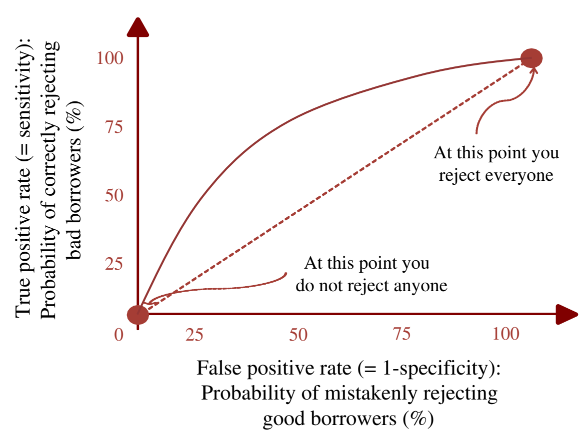
We begin by examining the ability of the Prosper marketplace to infer borrower quality as proxied by ex post loan performance. Under the assumption that the objective of marginal lenders on the Prosper marketplace is to maximize the returns on their portfolios, the interest rate is the market's best predictor of loan performance. We measure the quality of a screening method as (a) the simple goodness-of-fit (R^2) from a linear regression of ex post loan performance on the predictor of loan performance used by that screening method (Online Appendix A details why this, and not the regression coefficient, is the appropriate statistic) and, more formally, as (b) the area under a receiver operating characteristic (ROC) curve, a technique that is commonplace in commercial financial banking markets. We prefer using the ROC curve over the R^2 statistic, as it provides a more interpretable estimate of inference. Because ROC curves require a binary outcome measure, we use them for default as the measure of ex post loan performance but not for the fraction of loan repaid as the outcome measure.

We compare the quality of inference by Prosper lenders to two benchmarks. Our first benchmark is the most common instrument used by banks to screen borrowers: the credit score. In theory, the credit score should be the best available aggregator of standard hard financial data in terms of predicting default. Second, we develop an even more challenging benchmark based on what an econometrician would do if he had access to all standard financial information or even all the coded information available in the Prosper marketplace. This benchmark is particularly demanding because the econometrician's prediction model uses in-sample data on ex post default, something that is not available to banks at the time of loan origination.

3.1.1. Receiver Operating Characteristic Curves. A receiver operating characteristic curve measures the quality of a screening test in a setting with a continuous predictor for a binary outcome. Consider a threshold value for the continuous predictor (e.g., interest rate) above which the screening test predicts the outcome to occur (e.g., default). If we set the threshold very high, the test predicts default for relatively few individuals. As we decrease the threshold value, this increases the number of correctly predicted defaults as a fraction of all default realizations. This fraction is called the "sensitivity" or true positive rate (TPR). At the same time, reducing the threshold value increases the number of incorrectly predicted defaults as a fraction of individuals that did not end up defaulting. This fraction is referred to as the false positive rate (FPR), which equals one minus the "specificity." Thus, the choice of this threshold value allows us to trade off the sensitivity and the specificity of the screening test.

The ROC curve illustrates this trade-off by plotting the true positive rate on the y axis against the false positive rate as we reduce the threshold value. Figure 1 illustrates this for a hypothetical screening procedure, with each point on the curve showing the TPR and FPR for a particular threshold value. The ROC curve always starts at the origin and ends at the upper right-hand corner as we move the threshold value from the highest possible value to its lowest possible value. A "perfect" screening tool is represented by a single point on the top left-hand corner (TPR = 100% and FPR = 0%), whereas the worst possible screening tool would be one that is

Figure 1 (Color online) Stylized Receiver Operator Curve (ROC)



Notes. This figure shows a stylized receiver operator curve that we use to summarize the predictive power of various credit screening measures, such as the credit score or the interest rate. It plots the fraction of true positives out of the positives (TPR = true positive rate) vs. the fraction of false positives out of the negatives (FPR = false positive rate), at various threshold settings. ROC curves with greater area under the curve (AUC) represent superior predictors for default. See §3 for more details.

[11] As a caveat, we note that we are unable to assess how the quality of borrowers on Prosper compares to the general population on dimensions of creditworthiness that are not captured by the credit score, i.e., whether conditional on credit score Prosper borrowers are less or more likely to default than borrowers in the general population.

essentially random, causing the curve to lie on the 45° line.[12]

More specifically, the cardinal measure used to quantify the accuracy of a screening tool is the AUC, which ranges from 1 (perfect system) to 0.5 (purely random predictions). This corresponds closely to a Gini coefficient (Gini $= 2 * \mathrm{AUC} - 1$), and the literature in credit scoring uses both the AUC and the Gini as a way of quantifying the quality of a screening tool. We will present results using the AUCs.

Although there are no objective benchmarks for levels of AUC since its values depend on the usage context, it is the most common metric used in the credit-scoring industry. As a rough rule of thumb, an 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.

We estimate the AUCs (and show the ROC curves) using the market interest rate for a given loan as determined in the Prosper marketplace, since this acts as a simple metric by which the market judges the quality of a loan (i.e., higher interest rate loans are deemed to be of lower quality). We first compare the AUC of the market interest rate to the AUC using the credit score of the borrower. Second, we compare the market interest rate AUC to the AUC of the best possible score that an econometrician could create based on a regression of ex post default on the credit score as well as all standard financial information in the borrower listing. Finally, we compare to the AUC of an econometrician who uses coded nonstandard information in addition to credit score and all standard financial information.[13] These econometrician AUCs represent upper bounds for the predictive power of the best possible screening tool because the prediction model uses data on in-sample ex post default, which is never available to banks or lenders at the time of loan origination.[14]

We estimate the best possible score using a split-sample approach in order to avoid issues associated with overfitting. Specifically, we regress ex post default on all observed borrower characteristics for a randomly chosen half of the sample to obtain the regression coefficients. We then combine these coefficients with the observed borrower characteristics of the other half of the sample to create the best possible score.

To test whether the AUCs for the market interest rate and the other benchmark screening tools are statistically different from one another, we use the nonparametric approach developed in DeLong et al. (1988) and implemented in the routine "roccomp" by the STATA statistical software program.

3.2. Estimating Sources of Inference

In addition to estimating how well the market is able to screen, we investigate how lenders weight different sources of information in forming their screening measure (the market interest rate). Of particular interest is the relative importance placed on standard hard financial variables (such as debt-to-income ratio and past delinquencies) that banks typically use in addition to the credit score, and soft/nonstandard variables that banks typically would not use but are more common in peer-to-peer marketplaces.

We estimate the marginal contribution to inference from each of three sets of information as described in §2: standard financial variables (including the credit-category bins that lenders observe), nonstandard variables, and uncoded listing content (which represents soft information that is difficult to quantify). We define the marginal contribution as the extent to which inference would improve if lenders were able to observe one additional set of information, conditional on having already observed the other two sets of information. If the information content of one set of variables is strongly correlated with the information content of another set of variables, the marginal contribution from both will be low (because it is already contained

[12] A perfect screening test would correctly rank all borrowers. Hence, reducing the threshold would at first only predict default among those who end up defaulting (move along the y axis while the x axis stays at 0). Only after reducing the threshold to the point that all of the actual defaults are correctly predicted (the top left-hand corner of the graph), would further tightening start predicting defaults among those who do not end up defaulting, moving the curve along the y axis until all individuals are predicted to default. Conversely, the ROC curve for a screening tool with no predictive power would not distinguish those who default from those that do not, causing them to be evenly distributed across all scores/values of the screening tool. Thus, starting with the worst value of the score and tightening the threshold would lead to a prediction of default among those who end up defaulting and among those who do not in equal proportion (to their population), and the resulting ROC curve would be along the diagonal line. Note that this diagonal is the worst a screening tool could do. If the ROC curve is strictly below the diagonal then one can simply invert the score and come up with a better (than random) screening mechanism.

[13] We flexibly control for coded standard and nonstandard variables as quadratics, with *amount delinquent* and *revolving credit balance* measured in log form. We also include dummy variables for each of the following variables taking on a value of zero: *number of current delinquencies*, *number of delinquencies in last 7 years*, *number of public record requests in last 10 years*, *number of public records in last 12 months*, *revolving credit balance*, *amount delinquent*, and *revolving credit balance*. We further include dummy variables for *amount delinquent* and *revolving credit balance* less than 100 USD. All other variables are as described in Table 2 (see §4.1). To allow for additional flexibility, we interact all coded variables with a dummy variable for high-quality (*credit category B* or above) borrower.

[14] Our estimate of the best possible screening tool uses only coded information. In §4.2 and §4.3, we develop a methodology to measure inference from uncoded listing content. For now, we note that it is theoretically possible to outperform our "best possible" score by using uncoded information, such as pictures or listing text.

in the other set of information). As our results in §4.2 will show, this is indeed the case for inference from higher credit categories.

We measure the information from uncoded listing content contained in the interest rate as the residual from a regression of the interest rate on a flexible functional form of all the coded listing content (standard financial variables and nonstandard variables). This residual thus allows us to quantify the contribution of uncoded listing content, and we refer to it as the "deduced measure for uncoded information."

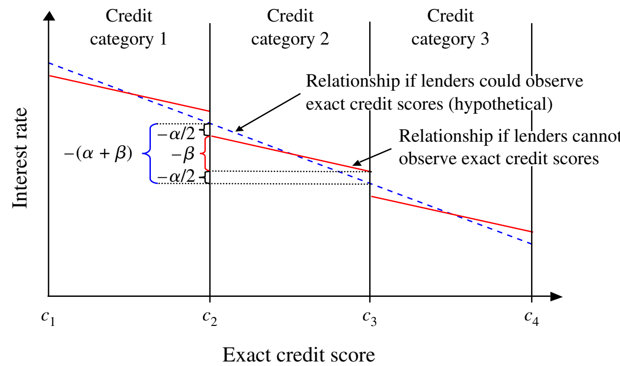
To estimate the marginal inference drawn from standard financial information, we estimate a first-stage regression of the interest rate for the loan on a flexible functional form of nonstandard variables and the deduced measure for uncoded information. The residual from this first-stage regression only contains variation in the interest rate that comes from standard financial variables that cannot be explained by the other two sets of information. We then estimate an ROC curve using the residual from this first-stage estimation. The ROC curve's corresponding AUC represents the marginal inference from standard financial variables. Alternatively, we can measure marginal inference using R-squareds instead of AUC. In a second-stage specification, we regress default on the residual from the first-stage regression. The R^2 from the second-stage estimation measures the marginal inference drawn from standard financial variables.

We follow the same procedure to measure the marginal contribution of inference from nonstandard variables, except that now the first-stage regression has controls for standard financial variables and the deduced measure for uncoded content. Similarly, to measure inference from uncoded listing content, we use a first-stage regression that has standard financial variables and nonstandard variables as controls.

3.3. Inferring the Unobserved Credit Score

The previous two sections describe methods that evaluate the ability of Prosper lenders to predict overall borrower creditworthiness, including along dimensions *not* captured by the credit score. Inference beyond the credit score is important because the credit score is primarily based on hard information (e.g., past repayment history) and may miss other valuable predictors of borrower quality. In this section, we describe a complementary methodology that measures how well lenders infer creditworthiness along the dimension that is directly captured by credit score (unobserved by lenders). We perform this analysis, despite the drawback of only examining one dimension of creditworthiness, because it provides an exact decomposition of inference by source of information.

Figure 2 **(Color online) Stylized Relationship Between Interest Rate and Credit Score**



Note. This figure shows the stylized hypothesized relationship between a borrower's credit score and the market interest rate on her (funded) loan.

The main idea is straightforward. Lenders observe credit categories but not the exact credit score. Consequently, if they offer loans at lower interest rates to borrowers who have better credit scores *within* a given credit category, then lenders must have correctly inferred from other information that these borrowers are more creditworthy than others in the same credit category. We can precisely quantify lenders' inference of creditworthiness along the dimension captured by the credit score by comparing the degree to which the interest rate declines with the exact credit score within credit categories to the overall decline in the interest rate across credit categories.[15]

We illustrate our empirical methodology with a stylized graph of the relationship between the exact credit score and the market interest rate. The x axis of Figure 2 plots the borrower's exact credit score, which is a proxy for one dimension of creditworthiness. Since the repayment probability is higher for more creditworthy people, the market interest rate should fall monotonically with the credit score if lenders could observe the true score (as shown by the dashed blue line). If the credit-score categories were the only information that lenders observed, the interest rate would be constant within categories and would only jump at the category borders. Thus, if we observe that the interest rate falls *within* credit-score categories, it must be the case that lenders are able to infer information about borrowers' creditworthiness from information other than the categorical credit-score variable (as illustrated by the discontinuous, downward-sloping, solid red line).

[15] Although the context is different, our method of using information not available to Prosper lenders to measure inference is similar to Farber and Gibbons (1996) and Altonji and Pierret (2001), who estimate employer inference of worker quality using AFQT scores, which are observed by the econometrician but not by the economic agents.

The degree to which lenders infer creditworthiness from other information is given by the amount by which the interest rate falls within credit-score categories relative to the total drop in interest rates, both within and between credit-score categories. In the figure, the interest rate drops by an amount β within each credit-score category and drops discontinuously by an amount α at each credit-score boundary. Hence, the total drop over one credit category (including one boundary) equals $\alpha+\beta$. Of this total drop, the interest rate falls by β because of the change in creditworthiness that lenders inferred from information other than credit category. We denote the fraction of information learned from all sources other than credit category by the symbol $\gamma \equiv \beta/(\alpha+\beta)$, and refer to γ as the amount of "inference" made by lenders along the credit-score dimension.

In this stylized setup, the following regression yields parameter estimates α and β:

$$InterestRate_i = \mu + \alpha Cat(CreditScore_i) + \beta CreditScore_i/CatSize + \varepsilon_i, \quad (1)$$

where $InterestRate_i$ is the interest rate charged on loan i, $CreditScore_i$ is the exact credit score of the borrower of loan i, and $Cat(\cdot)$ is a scalar that denotes the category of the credit score; because there are seven credit-score categories, $Cat(\cdot)$ takes on integer values 1–7. *CatSize* is a constant that is equal to the range of credit scores that each credit category spans. Finally, ε_i denotes the error term, and the remaining Greek symbols are parameters to be estimated.

In Online Appendix B, we present a more formal analysis of the case (which we implement empirically) where we allow for the underlying relationship between interest rate and exact credit score to be nonlinear and the bin sizes of credit categories to be of unequal size. We also detail how this method allows us to decompose inference along the credit-score dimension by source of information. For this decomposition, we include in Equation (1) a flexible functional form in variables for each source of information. If an information source contributes to inference, it will lower the coefficient on the credit score. The standard omitted variable bias formula then determines the exact amount by which each included control variable lowers the coefficient on the credit score, which corresponds to that control variable's contribution to inference. Although it may seem challenging to quantify or code qualitative data (such as pictures and other personal details), an advantage of our methodology is that we can still derive the contribution of such information: the contribution of uncoded information is inferred from the remaining relation between the exact credit score and interest rate within credit categories while controlling for a flexible functional form of all coded information.

4. Results

We first examine the extent to which the interest charged by lenders predicts loan performance. We compare the predictive power of the interest rate set by market participants against the predictive power of two benchmarks: the exact credit score of the borrower and the best predictors achievable by an econometrician. Next, we explore how the accuracy of lender inference differs across borrower quality, and examine the weight placed by lenders on different sources of information (standard financial versus soft/nonstandard information) for inference. Finally, we focus on inference along the credit-score dimension of creditworthiness because, along that dimension, we can provide an exact decomposition of inference by information source.

4.1. How Well Does the Market Screen?

Table 2 first examines the predictive relationship between interest rates and loan performance. Our performance measures are the default rate (an indicator for whether the loan is over three months late) and the fraction of the principal repaid after the end of its term. The functional form of our specification is derived in Online Appendix A and uses $1/(1+r)$ as our independent variable, where r denotes the interest rate. In panel A, we find that the interest rate is indeed a significant predictor of default and fraction repaid. The adjusted R^2 of the regressions are 8% and 9%, respectively. To provide a benchmark for the performance of the interest rate, we examine default predictability using exact credit score in panel B. We find that the credit score also significantly predicts default and fraction repaid. However, the adjusted R^2 obtained using credit score (4% and 5%) is only half of that obtained using the interest rate.[16] We examined the robustness of the results in panels A and B by double clustering the standard errors either by borrower and time (year $*$ month) or by borrower and largest lender. The standard errors increase but we still find significant results at 1% significance levels (see Online Appendix Tables C.2 and C.3).

These initial regressions show that the interest rate set by market participants is a better predictor of creditworthiness than the exact credit score. To rule out the possibility that the interest rate merely performed better because it happened to have a better fitting functional form than the credit score, we reran the credit score regressions with a very flexible functional form of the credit score (a seven-part spline with break points at the credit-category boundaries). This did not meaningfully increase the R^2.

[16] See also Adams et al. (2009), Einav et al. (2013), and Keys et al. (2010), who find that credit score is a good predictor of default.

Table 2 Interest Rates and Loan Performance

	Default—three or more months late	Fraction repaid
Panel A: OLS—Do interest rates predict loan performance?		
1/(1 + Interest rate)	−1.525*** (0.038)	1.173*** (0.029)
N	17,212	17,212
Adjusted R^2	0.0814	0.0914
Panel B: OLS—Do credit scores predict loan performance?		
Exact credit score/100	−0.129*** (0.004)	0.096*** (0.003)
N	17,212	17,212
Adjusted R^2	0.0432	0.0456
Adjusted R^2 using seven-part spline in credit score	0.0442	0.0476
Panel C: IV—Do interest rates causally affect loan performance?		
1/(1 + Interest rate)	0.166 (0.415)	0.061 (0.298)
N	17,212	17,212
First stage *F*-statistic	58.45	58.45

Notes. Default is a dummy for whether the loan is three or more months late as of three years after the loan is initiated (all loans have three-year maturities). Fraction Repaid measures the fraction of the principal that has been repaid after three years, not including missed interest payments. Formally, fraction repaid is defined as (principal—netdefaults)/principal where netdefaults is the principal balance minus loan sale proceeds and forfeited group rewards. Specifications in panel A regress these measures of loan performance on $1/(1+r)$, where r is the three-year interest rate that lenders receive on the loan. Panel B regresses loan performance measures on exact credit score. We also report the *r*-squared from a second specification allowing for a seven-part spline in exact credit score. Panel C examines whether interest rate has a causal effect on loan performance using credit-category dummies as the excluded instruments. We report the second-stage results from a two-stage least squares regression of default or fraction repaid on $1/(1+r)$ with controls for a spline in credit score (kinks in the spine are set at credit category boundaries); $1/(1+r)$ is instrumented with credit category dummies (the excluded instruments). Standard errors are allowed to be clustered by borrower (some borrowers hold more than one loan) and are in parentheses.

***Significant at 1%.

To investigate the concern that higher interest rates could lead to default by increasing the burden on borrowers (as in Stiglitz and Weiss 1981), we estimate the causal effect of the interest rate on default using credit-category borders as instruments. The intuition for the instrument is that there is a sharp jump in the interest rate at the exogenously defined borders (AA, A, B, etc.), but that underlying creditworthiness should be smooth at the borders because the credit score does not change discontinuously there. Panel C shows no evidence of a causal effect of the interest rate on loan performance, and we conclude that our results are not driven by reverse causation.

A more formal way to compare the predictive power of two screening mechanisms is through ROC curves. The first two panels of Figure 3 show the ROC curves for the interest rate and credit score. We find that the area under the ROC curve (AUC) for the interest rate is 0.6824, and the AUC using the exact credit score is lower at 0.6254. Note that a completely uninformative screening system has an AUC of 0.5 and that a 0.01 improvement in AUC is considered a noteworthy gain in the credit scoring industry. By the AUC metric, lenders predict default with 45% greater accuracy than what is possible by using just the borrower's credit score, an improvement that is significant at the 1% level.[17] This is particularly noteworthy since the Experian ScoreX PLUS credit score is designed by Experian as a special score (not even available to borrowers themselves) that is supposed to be better tailored to the types of borrowers on Prosper.[18]

Next, we construct ROC curves based on the best possible score that an econometrician would construct if he used (i) all observable standard financial variables, including the exact credit score and recent employment data, or (ii) all available coded information, i.e., both standard financial variables and coded nonstandard information. These scoring systems offer an AUC of 0.7103 and 0.7140, respectively. In comparison to these more demanding benchmarks, Prosper lenders still perform fairly well. They achieve 87% of the predictive power of an econometrician who observes all standard financial information and 85% of the predictive power of an econometrician who observes all coded information, standard and soft/nonstandard.[19] To examine the contribution of the borrower maximum interest rate in the inference drawn from coded nonstandard information, we calculate the AUC using standard financial variables and borrower maximum interest rate. This AUC is 0.7135, which implies that the borrower maximum rate can achieve 86% ($=(0.7135-0.7103)/(0.7140-0.7103)$) of improvement in inference coming from coded nonstandard information.[20]

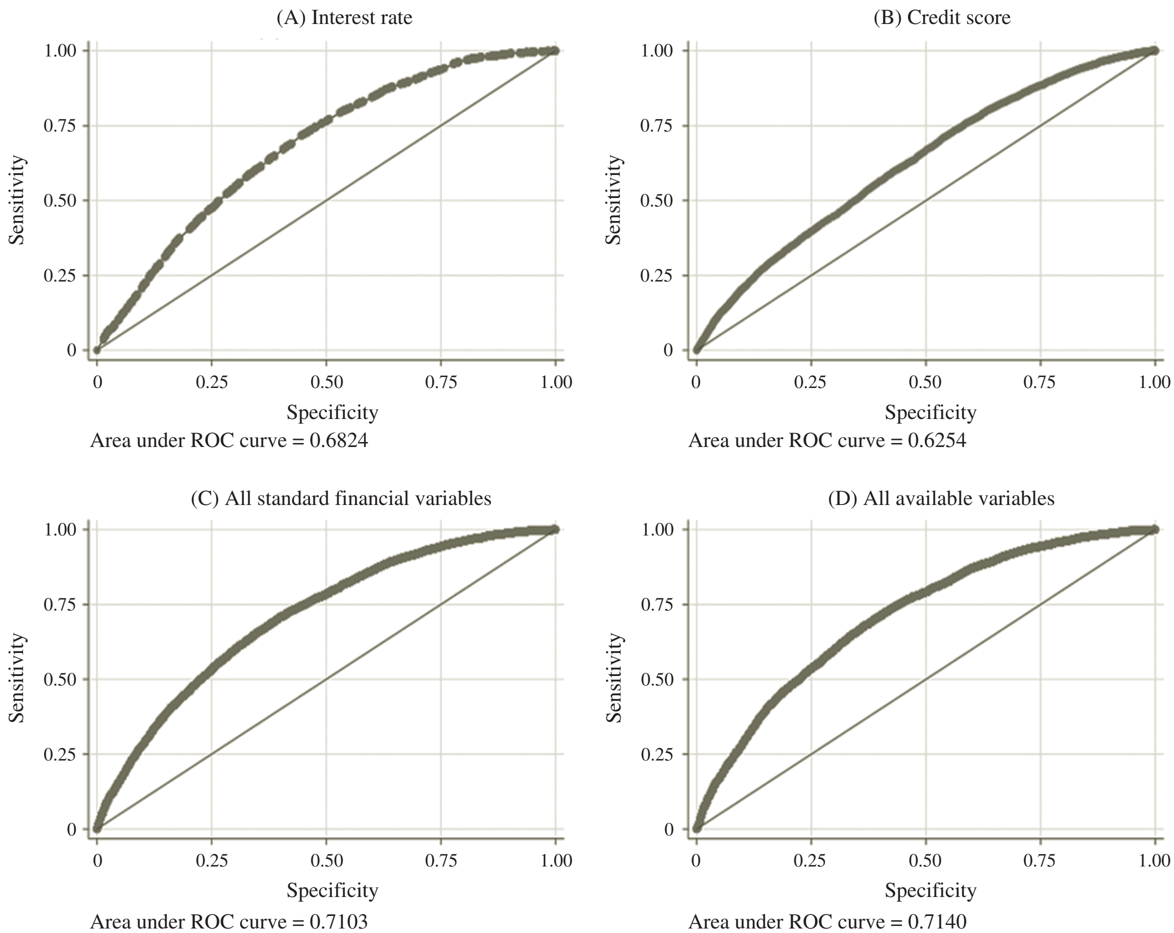
[17] We calculate the percentage improvement as $(0.6824-0.5)/(0.6254-0.5)=1.45$, where 0.5 is subtracted from both the interest rate and credit score AUC because 0.5 is the AUC under a noninformative (random) scoring system.

[18] Experian claims that the ScoreX PLUS provides "a stronger separation of good and bad accounts and classifies more bad accounts into the worst-scoring ranges." They further note that "traditional risk models typically are unable to score a significant percentage of consumers due to insufficient credit information," but that "with ScoreX PLUS, almost all consumers can be effectively scored to rank order risk, thereby reducing the need for manual review" and that "in multiple market validations with traditional risk models, ScoreX PLUS performed better in over 90% of the head-to-head comparisons" (see Experian 2004).

[19] We again adjust both AUCs by 0.5, which corresponds to zero inference, and estimate that $(0.6824-0.5)/(0.7103-0.5)=0.87$ and $(0.6824-0.5)/(0.7140-0.5)=0.85$.

[20] Because the borrower maximum rate and the other coded nonstandard variables capture largely the same information, the other coded nonstandard variables (excluding the borrower maximum rate) can also account for a large fraction of improvement in inference coming from coded nonstandard information. In particular, the AUC of standard financial variables and coded nonstandard information other than the borrower maximum rate is 0.7133.

Figure 3 ROC Curves—Full Sample



Notes. Panels (A) and (B) present the ROC curves for the interest rate and credit score, respectively. Panel (C) presents the ROC curve using the best possible score that an econometrician would construct if he used all observable standard financial variables including exact credit score. Panel (D) presents the ROC curve using the best possible score that an econometrician would construct if he used all observable standard financial variables as well as coded nonstandard or softer variables. See §3.1 for a detailed description of the creation of the curves in panels (C) and (D). We can reject equality between the interest rate AUC in panel (A) and each of the AUCs in panels (B), (C), and (D) with p-values of less than 0.001.

The econometrician's AUC is a particularly demanding benchmark because we allow the econometrician to use data on in-sample, ex post default. Perhaps a fairer and more realistic benchmark is the AUC achieved by the econometrician where we split the sample in time (rather than randomly) to give us an "out-of-sample, out-of-time prediction." We use the first half of the time period of the sample to estimate the econometrician's model and then use that model to predict default for loans originating in the second time period. Lenders now come closer to the econometrician's benchmark, achieving 93% (91%) of the econometrician's AUC from all standard financial information (all coded information).

A caveat to this analysis is that we estimate inference only within the selected sample of listings that were funded because our outcome variable (default or fraction repaid) is available only for funded listings. However, we make the comparisons between the market interest rate, the credit score, and the econometrician's predictors all within the same sample, and have no reason to believe that the selected sample would skew the relative performance of these predictors. When we examine inference along the credit-score dimension (in §4.2), we can explicitly correct for the selective funding decision because credit scores are observed both for funded and unfunded listings. In that setting, we estimate that our inference estimate drops from 39% to 33% when we correct for sample selection. This suggests that the absolute magnitude of inference is biased up if inference is estimated within the selected sample of funded listings.

4.1.1. Screening by Borrower Quality. The results thus far show that, on average, the interest rate is a substantially better predictor of loan performance than credit score and comes reasonably close to an upper bound achievable by the econometrician. We now explore how the quality of screening varies across different ranges of borrower quality. Table 3 presents

Table 3 Variation in Inference by Borrower Quality

	Default (1)	Fraction repaid (2)
Panel A: Low credit categories (HR − C)		
1/(1 + Interest rate)	−1.220*** (0.067)	1.045*** (0.051)
N	9,041	9,041
Adjusted R^2	0.0349	0.0460
Panel B: High credit categories (B − AA)		
1/(1 + Interest rate)	−2.201*** (0.069)	1.602*** (0.052)
N	8,171	8,171
Adjusted R^2	0.1124	0.1229
Panel C: Low credit categories (HR − C)		
Exact credit score/100	−0.157*** (0.012)	0.132*** (0.009)
N	9,041	9,041
Adjusted R^2	0.0191	0.0243
Adjusted R^2 using seven-part spline in credit score	0.0192	0.0249
Panel D: High credit categories (B − AA)		
Exact credit score/100	−0.164*** (0.010)	0.113*** (0.007)
N	8,171	8,171
Adjusted R^2	0.0282	0.0277
Adjusted R^2 using seven-part spline in credit score	0.0282	0.0278

Notes. Panels A and B estimate the relationship between $1/(1+r)$ and loan performance separately for the sample of low-quality borrowers (credit categories HR − C) and high-quality borrowers (credit categories B − AA). Panels C and D estimate the relationship between exact credit score and loan performance separately for low- and high-quality borrowers. All variables are as defined in Table 2. Standard errors are allowed to be clustered by borrower (some borrowers hold more than one loan) and are in parentheses.

***Significant at 1%.

the analogous regressions to Table 2 separately for low-quality borrowers (credit categories HR, E, D, and C) and high-quality borrowers (credit categories B, A, and AA). Panels A and B show that the interest rate explains a higher fraction of the variation in default among high-quality borrowers than among low-quality borrowers. The adjusted R^2 from the regression of default on interest rate for low-quality borrowers is 3.5%, whereas that for high-quality borrowers is 11%. Similarly, the adjusted R^2 for fraction repaid on interest rate for low-quality borrowers is 4.6%, whereas that for high-quality borrowers is 12%. Panels C and D examine the differences in inference using credit score and find similar patterns. The adjusted R^2 for default predictability for low-quality borrowers using credit score is 1.9%, whereas that for high-quality borrowers is 2.8%. The adjusted R^2 for fraction repaid using credit score for low-quality borrowers is 2.5%, whereas that for high-quality borrowers is 2.8%. Altogether, the results show that overall predictability is higher for high-quality borrowers and that the level of predictability obtained using interest rates is in each case higher than that obtained using credit score.

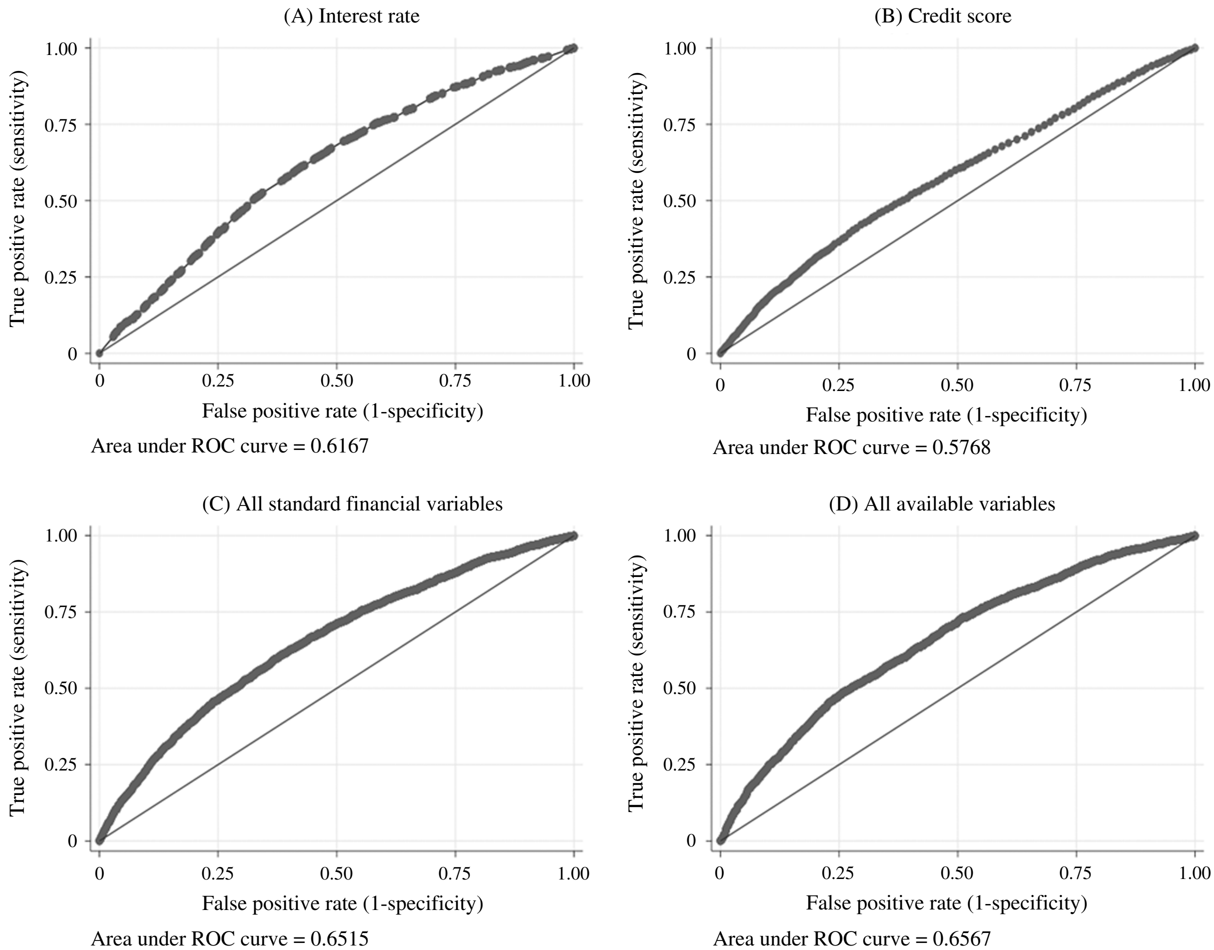
Figures 4 and 5 show the ROC curves for low- and high-quality borrowers. In line with the earlier results, we find that the AUC using the interest rate is larger than the AUC using the credit score for both low- and high-quality borrowers. The differences are large—the interest rate outperforms the credit score by 52% for low-quality borrowers and by 100% for high-quality borrowers. As before, we also present the more demanding benchmark of the best prediction an econometrician would produce. We find that the market performs well against this benchmark for both the low- and high-quality borrowers. Prosper lenders achieve 77% and 92% of the predictive power of an econometrician who observes all standard financial information for low- and high-quality borrowers, respectively. Lenders also achieve 74% and 91% of the predictive power of an econometrician who observes all coded information for low- and high-quality borrowers, respectively.

4.1.2. Robustness. One might be concerned that credit categories incorporate some of the information in credit scores and, hence, the interest rate (which may include information on credit categories) will have an edge in predicting default over the credit score. To address this concern, we examine the AUC *within* each credit category. Rows (1)–(7) of Table 4 show that within each credit category, the interest rate outperforms the credit score in predicting default. We also present the AUC of the econometrician who observes ex post default as well as all coded standard financial information or all coded information. The interest rate performs close to this benchmark as well.

Row (9) addresses the concern that borrowers may directly inform lenders of their credit score.[21] We find that the interest rate also significantly outperforms the credit score in the sample period when there was no facility for questions and answers. Another potential concern is that Prosper introduced policy changes over the sample period that may affect our inference estimates and interpretation. Suggested ranges for the borrower maximum rate might impact the extent of inference. The introduction of portfolio plans could have guided lenders. In rows (10) and (11), we find similar AUC estimates before and after these changes. Another concern is that borrowers in some states are

[21] We think this channel is unlikely because Prosper strongly discourages borrowers from revealing detailed personal information and a text search through all listing text does not reveal any self-reported credit scores. Prosper allows borrowers to post information in the listing and also has a facility for questions and answers. However, because all of this information is unverified, borrowers would have an incentive to report the highest possible credit score within a credit category. In communications with Prosper staff, we were told that great care was taken by Prosper to purge any personal references such as credit score or Social Security number.

Figure 4 ROC Curves—Low Credit Categories



Notes. This figure plots the same ROC curves as described in Figure 3, but with the sample restricted to credit categories HR – C (lower-quality borrowers). We can reject equality between the interest rate AUC in panel (A) and each of the AUCs in panels (B), (C), and (D) with *p*-values of less than 0.001.

subject to usury laws (Rigbi 2013), which create a ceiling on interest rates that could impact inference. In row (12), we find similar results in the period without usury law restrictions.

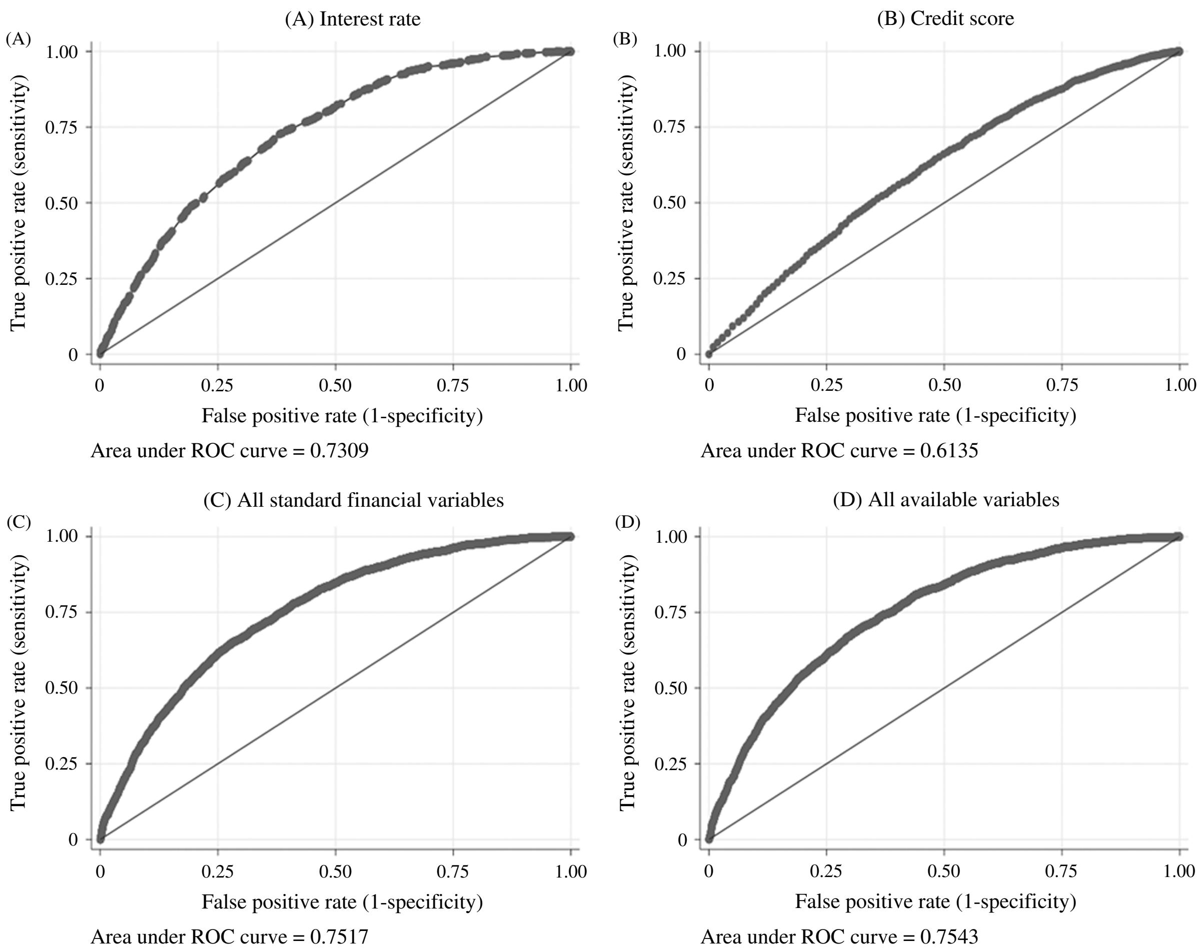
To address the concern that members of Prosper groups might share personal information with one another, we restrict the sample to borrowers who are not affiliated with any group. To ensure that inference is not driven by learning about borrowers from their previous listings, we restrict the sample to first-time listings (row (14)) and first-time loans (row (15)). We again find similar results.

4.2. Sources of Inference

In this section, we examine how lenders use different types of information in setting the interest rate. We group information into two broad categories of interest: standard financial variables and soft/nonstandard information. Standard financial variables are hard, readily coded, and tend to be verifiable. We provide the details and summary statistics of variables included in this category in Table 1.[22] Meanwhile, soft/nonstandard information tends to be subjective, nonfinancial, potentially harder to verify, and more likely to behave like signals. Of particular interest are the various "softer" pieces of information such as pictures, individual background, description, and online exchanges, which are readily identifiable, but much harder to code in a way that is suitable for empirical analysis. For example, one may be able to code whether a listing has a picture or even attributes about the picture, but it is impossible to code all aspects of the picture from which lenders might draw information. However, we can infer how much uncoded information is incorporated into the interest rate: we create the deduced measure of uncoded listing content as the residual of a regression of the

[22] For the sake of brevity, Table 1 does not provide summary statistics for 66 borrower occupation dummies and 52 borrower state-of-residence dummies (50 states, District of Columbia, and Puerto Rico). However, these variables are included as controls in the relevant specifications in Tables 5, 7, and the online appendix tables.

Figure 5 ROC Curves—High Credit Categories



Notes. This figure plots the same ROC curves as described in Figure 3, but with the sample restricted to credit categories B – AA (higher-quality borrowers). We can reject equality between the interest rate AUC in panel (A) and each of the AUCs in panels (B), (C), and (D) with *p*-values of less than 0.001.

interest rate on a flexible functional form of all sources of coded information. We thus distinguish between the coded content of soft/nonstandard information and the uncoded content, which leaves us with three sources of information: standard financial variables (which are coded), soft/nonstandard *variables* (which are coded), and the deduced measure of uncoded listing content (which also reflects soft/nonstandard information). One caveat of the deduced measure of the uncoded listing content is that it may pick up information from the coded variables if our functional form is misspecified or not sufficiently flexible.

As described in §3, our estimates measure marginal contributions for each type of information, assuming that lenders already see the other two types of information. Panel A of Table 5 presents the fraction of inference on the default rate. Columns (1)–(3) report the results for low-quality borrowers and columns (4)–(6) report the results for high-quality borrowers.

In row (1), we focus on the inference using the AUC measure for the marginal contribution of each source. We find that all three types of information contribute toward inference. For low-quality borrowers, the AUC from standard financial variables is 0.573, whereas the corresponding figure for high-quality borrowers is 0.591. Thus, the point estimate for inference from standard financial variables is 0.019 lower for low-quality borrowers than for high-quality borrowers. Note that a drop of 0.019 is considered fairly substantial for AUCs. For low-quality borrowers, the AUCs for soft/nonstandard variables and uncoded listing content are 0.542 and 0.538, and the corresponding figures for high-quality borrowers are similar at 0.545 and 0.543. Because the marginal contribution to inference from soft/nonstandard sources remains basically stable across different credit categories, whereas that of hard information clearly drops in lower credit categories, the results indicate that the marginal contribution of soft/nonstandard information for inference is *relatively* more important for lower-quality borrowers than for higher-quality borrowers. When we compare the overall inference using

Table 4 Robustness of Measure of Inference

Estimation sample	AUC using interest rate	AUC using credit score	Econometrician's AUC using standard financial variables	Econometrician's AUC using all coded variables	*N*
(1) Credit category HR	0.578	0.514	0.634	0.649	1,169
	(0.017)	(0.017)	(0.016)	(0.016)	
p-value test of equality		0.009	0.017	0.002	
(2) Credit category E	0.548	0.517	0.587	0.591	1,274
	(0.016)	(0.016)	(0.016)	(0.016)	
p-value test of equality		0.167	0.097	0.055	
(3) Credit category D	0.592	0.534	0.631	0.632	2,971
	(0.011)	(0.011)	(0.011)	(0.011)	
p-value test of equality		0.000	0.005	0.003	
(4) Credit category C	0.602	0.515	0.650	0.661	3,627
	(0.010)	(0.010)	(0.010)	(0.010)	
p-value test of equality		0.000	0.000	0.000	
(5) Credit category B	0.659	0.535	0.703	0.709	3,149
	(0.010)	(0.011)	(0.010)	(0.010)	
p-value test of equality		0.000	0.000	0.000	
(6) Credit category A	0.734	0.511	0.741	0.742	2,414
	(0.012)	(0.014)	(0.012)	(0.012)	
p-value test of equality		0.000	0.553	0.453	
(7) Credit category AA	0.780	0.591	0.786	0.788	2,608
	(0.012)	(0.015)	(0.012)	(0.012)	
p-value test of equality		0.000	0.578	0.493	
(8) Baseline sample	0.682	0.625	0.710	0.714	17,212
(all listings 2/12/2007–10/16/2008)	(0.004)	(0.005)	(0.004)	(0.004)	
p-value test of equality		0.000	0.000	0.000	
(9) Period without question	0.739	0.678	0.753	0.761	767
and answers (pre 2/12/2007)	(0.019)	(0.021)	(0.019)	(0.019)	
p-value test of equality		0.000	0.444	0.203	
(10) Period before suggested	0.728	0.663	0.739	0.741	5,852
borrower maximum rate and	(0.007)	(0.007)	(0.007)	(0.007)	
portfolio plans (pre 10/30/2007)					
p-value test of equality		0.000	0.121	0.044	
(11) Period after suggested	0.664	0.602	0.692	0.696	11,360
borrower maximum rate and	(0.005)	(0.006)	(0.005)	(0.005)	
portfolio plans (post 10/30/2007)					
p-value test of equality		0.000	0.000	0.000	
(12) Period without state usury law	0.658	0.600	0.676	0.681	6,420
restrictions on interest rates	(0.007)	(0.008)	(0.007)	(0.007)	
(Post 4/15/2008, excl. TX and SD)					
p-value test of equality		0.000	0.021	0.002	
(13) Sample restricted to listings	0.676	0.608	0.696	0.703	12,061
with no group affiliation	(0.005)	(0.006)	(0.005)	(0.005)	
p-value test of equality		0.000	0.000	0.000	
(14) Sample restricted to listings	0.695	0.615	0.712	0.719	6,350
posted by borrowers with no	(0.007)	(0.008)	(0.007)	(0.007)	
previous prosper listings					
p-value test of equality		0.000	0.025	0.001	

AUCs for interest rates, across low-quality and high-quality borrowers, we find that the AUCs are 0.6167 and 0.7309, respectively, and that the difference is significant at the 1% level. Similarly, we find that the AUC based on credit score is significantly higher for high-quality borrowers than low-quality borrowers. Thus, we find that inference in general is higher for higher-quality borrowers.

In row (2), we measure the marginal inference using R-squareds. We find similar results to those obtained using AUC measures. Inference is important for all three sources of information, but soft/nonstandard information appears relatively more important for lower-quality borrowers.

In panel B, we report the marginal inference using R-squareds for fraction repaid as the outcome variable. As in panel A, we find that all three sources of information are important for inference. Moreover, as before, the point estimate of the marginal contribution to inference from standard financial variables is

Table 4 (Continued)

Estimation sample	AUC using interest rate	AUC using credit score	Econometrician's AUC using standard financial variables	Econometrician's AUC using all coded variables	*N*
(15) Sample restricted to listings posted by borrowers with no previous prosper loans	0.682 (0.004)	0.623 (0.005)	0.708 (0.004)	0.713 (0.004)	15,303
p-value test of equality		0.000	0.000	0.000	

Notes. This table examines the robustness of our inference estimates from Table 2, using default as our outcome variable; *p*-values test whether the AUC in each column is equal to the interest rate AUC. All variables are as described in Table 2. Rows (1) through (7) restrict our sample to borrowers within each of the seven credit categories, and row (8) shows estimates using the full baseline sample. Row (9) restricts our sample to the period before public and private questions were allowed between borrowers and lenders (pre February 12, 2007). This ensures that inference is measured from lender inference rather than from possible direct exchanges of credit score information between borrowers and lenders. Note that our baseline sample excludes the pre February 12, 2007 period because credit-category cutoffs changed on February 12, 2007. Rows (10) and (11) restrict our sample to the periods before and after Prosper added (a) a Web application to suggest borrower maximum rates to borrowers and (b) an application allowing automatic bids on loans through lender portfolio plans (pre and post October 30, 2007). Representatives from Prosper have confirmed that Prosper does not use exact credit score in its calculations of suggested borrower maximum rate or its implementation of lender portfolio plans. Row (12) restricts our sample to the period after Prosper became exempt from most state usury laws that capped the maximum interest rate (post April 15, 2008) and excludes the two states, Texas and South Dakota, for which usury laws are still enforced. Row (13) restricts the sample to listings posted by borrowers with no group affiliations. Rows (14) and (15) restrict the sample to listings posted by borrowers with no previous Prosper listing or loan (funded listing), respectively. These tests confirm that our measurements of inference do not depend on information about the past repayment and listings history of borrowers who apply for more than one loan. Standard errors are allowed to be clustered by borrower (some borrowers apply for more than one loan) and are in parentheses.

smaller for the low credit categories (an adjusted R^2 of 0.014 for the low versus 0.018 for the high credit categories) while, if anything, inference from the uncoded listing content is *higher* in the low credit categories (an adjusted R^2 of 0.006 for the low versus 0.004 for the high credit categories). However, in contrast to the results with default as an outcome, the inference from soft/nonstandard variables now exhibits a clear drop in the low credit categories (an adjusted R^2 of 0.008 in the low versus 0.011 in the high credit categories).

The broad message from these results is that both standard financial and soft/nonstandard information

Table 5 Sources of Inference

	Low credit categories			High credit categories		
	(1)	(2)	(3)	(4)	(5)	(6)
	Marginal contribution to inference by			Marginal contribution to inference by		
	Standard financial variables	Soft/nonstandard variables	Uncoded listing content	Standard financial variables	Soft/nonstandard variables	Uncoded listing content
	Panel A: Predicting default					
AUC	0.5726	0.5421	0.5381	0.5916	0.5452	0.5345
Adjusted R^2	0.0140	0.0049	0.0038	0.0196	0.0057	0.0040
	Panel B: Predicting fraction repaid					
Adjusted R^2	0.0140	0.0077	0.0055	0.0184	0.0109	0.0042

Notes. Panel A estimates the marginal contribution to inference about borrower creditworthiness (as measured by ex post default) from three sets of information: standard financial variables, soft/nonstandard variables (which are coded), and uncoded listing content (which reflects soft/nonstandard information that cannot be easily coded). We define the marginal contribution to inference as the extent to which inference will improve if lenders are able to observe one additional set of information, conditional on already observing the other two sets of information. For example, column (1) measures the increase in total inference if investors are able to observe standard financial variables, conditional on already observing soft/nonstandard variables and uncoded listing content. To estimate the marginal inference from standard financial information, we estimate a first-stage regression of interest rate on the other two sets of information (in this case, soft/nonstandard variables and uncoded listing content, which is estimated as the residual of $1/(1+r)$ regressed on all coded information). We then estimate an ROC curve using the residual from this first-stage estimation. The ROC curve's corresponding AUC represents the marginal inference from standard financial variables. Alternatively, we can measure marginal inference using r-squareds instead of the AUC. In a second-stage specification, we regress default on the residual from the first-stage estimation. The r-squared from the second-stage estimation represents the marginal contribution to inference from standard financial variables. Panel B repeats the exercise using fraction repaid as the outcome variable. Note that because fraction repaid is a continuous variable, we present the r-squared as our measure of inference instead of the AUC. We control for coded standard and nonstandard variables as quadratics, with Amount Delinquent and Revolving Credit Balance measured in log form. We also include dummy variables for each of the following variables taking on a value of zero: *number of current delinquencies*, *number of delinquencies in last 7 years*, *number of public record requests in last 10 years*, *number of public records in last 12 months*, *revolving credit balance*, *amount delinquent*, and *revolving credit balance*. We further include dummy variables for *amount delinquent* and *revolving credit balance* less than 100 USD. All other variables are as described in Table 2.

contribute on the margin toward inference (and therefore add value above each other). In addition, our estimates suggest that standard financial variables are less predictive of loan performance for lower-quality borrowers than for higher-quality borrowers, both absolutely and relative to soft/nonstandard information. These results highlight the value of complementing hard information with soft/nonstandard information, especially when screening low-quality borrowers.

4.3. Quantifying Inference Along the Credit-Score Dimension

The results presented so far explore the ability of lenders to use listing information, particularly nontraditional and soft information, to predict *default*. We now explore the extent to which lenders infer the information content of the missing credit score itself. Estimating inference along the credit-score dimension provides a precise decomposition of inference that arises from different sources of information, rather than bounds as presented in the previous sections.

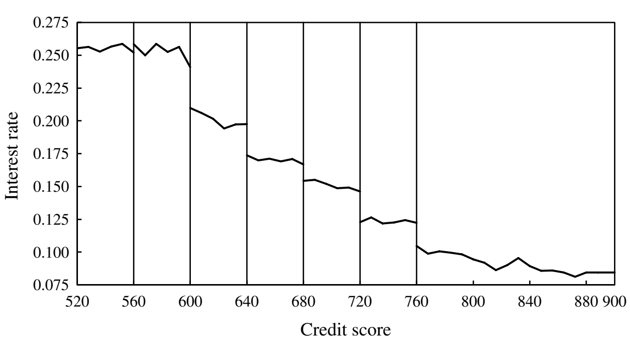
4.3.1. Estimating Overall Inference Magnitude. We begin by presenting the empirical analogue to Figure 2. In Figure 6, we plot raw market interest rates against credit score. The average interest rate declines by 18 percentage points as we move from the lowest to highest credit scores. Importantly, the figure shows that the interest rate also declines with credit score *within* credit categories, suggesting that lenders are able to infer creditworthiness along the credit-score dimension from other listing information. In addition, the interest rate jumps at the credit-category boundaries, consistent with inference being incomplete.

Column (1) of Table 6 presents the simple ordinary least squares regression of the market interest rate on credit score and credit category, as specified by Equation (1). The coefficient on credit score/40 shows that the interest rate falls by 0.55 percentage points within the typical credit category, which has a width of 40 points in the credit score.[23] The constant term shows that the interest rate falls 2.18 percentage points at the typical credit-category border. Of the 18.3 percentage point fall in the interest rate from the lowest to the highest credit score, 13.1 percentage points ($= 6 \times 2.18$) occurs at the category borders, and the remaining 5.2 percentage points occur within credit categories. Hence, a first take on the magnitude of inference would be that lenders infer $5.2/18.3 = 28\%$ of the variation in creditworthiness (along the dimension of credit score) from other listing information.

There are two reasons why the analysis so far is only suggestive. First, the regression in column (1) of Table 6 has a rigid functional form that imposes a constant slope of interest rate with respect to credit score and a constant size of jumps in interest rate at the credit-category boundaries. To relax these restrictions, we will estimate the more flexible model specified in Equation (B.1) in Online Appendix B. Second, the market interest rate is a censored variable: it is only observed when the interest rate at which lenders are willing to lend is lower than the maximum interest rate that the borrower has specified. Hence, the market interest rate could mechanically fall within a credit category if borrowers with higher credit scores within a credit category specify lower borrower maximum rates and if the rate at which lenders are willing to lend has a random component. To capture lender inference, we need to estimate how the offer rate—i.e., the uncensored interest rate at which lenders are willing to lend—varies with credit score within credit categories. We estimate a censored regression, where the censoring takes place at the listing-specific borrower maximum rate.

Column (2) of Table 6 reports that a censored regression with a flexible functional form estimates that lenders infer a third (0.33) of the difference in creditworthiness (along the dimension measured by credit score) within a credit category. A benchmark for our estimate is the amount of inference that could have been attained if lenders had optimally used all coded information from the listing content. This benchmark, estimated using the method described in Online Appendix B, is 0.42. Because the benchmark

Figure 6 **Market Interest Rate and Credit Scores**



Notes. This figure shows the "raw" relationship between a borrower's credit score and the one-year interest rate on her funded loan. Each point in the graph plots the average interest rate over an eight-point range in credit scores. Solid lines separate the seven credit categories. Starting from left to right, the categories are HR, E, D, C, B, A, AA. Lenders observe the borrower's credit category but do not observe the borrower's exact credit score.

[23] To examine whether the results are driven by the AA category having a greater dispersion in underlying credit scores relative to the other categories, we estimate the results in column (1) of Table 6, excluding category AA. The results remain similar: the coefficient on credit score/40 becomes -0.0061^{***} (0.0015) and the coefficient on credit category changes to -0.0219^{***} (0.0015). In column (2) of Table 6, our combined estimate of gamma is 0.256^{***} (0.043) if we exclude category AA.

Table 6 Inferring Creditworthiness Along the Credit-Score Dimension

	(1)		(2)	
Dependent variable: Interest rate	OLS		Censored regression	
Estimate	Coefficient	(S.E.)	Coefficient	(S.E.)
Combined γ: Inference			0.330***	(0.033)
Regression coefficients				
Credit score/40	−0.0055***	(0.0008)		
Credit category	−0.0218***	(0.0009)		
α_2: Change between categories HR and E			−0.038***	(0.005)
α_3: Change between categories E and D			−0.059***	(0.005)
α_4: Change between categories D and C			−0.049***	(0.004)
α_5: Change between categories C and B			−0.051***	(0.005)
α_6: Change between categories B and A			−0.031***	(0.005)
α_7: Change between categories A and AA			−0.042***	(0.005)
β_1: Change within category HR			−0.011*	(0.006)
β_2: Change within category E			−0.011*	(0.007)
β_3: Change within category D			−0.027***	(0.005)
β_4: Change within category C			0.000	(0.005)
β_5: Change within category B			−0.014**	(0.006)
β_6: Change within category A			−0.005	(0.007)
β_7: Change within category AA			−0.052***	(0.008)
N	17,212		194,033	
R^2	0.492		0.431	
Implied coefficients and tests				
$\gamma_1 = \beta_1/\delta_1$: Inference in credit category HR			0.229*	(0.120)
$\gamma_2 = \beta_2/\delta_2$: Inference in credit category E			0.189*	(0.099)
$\gamma_3 = \beta_3/\delta_3$: Inference in credit category D			0.332***	(0.056)
$\gamma_4 = \beta_4/\delta_4$: Inference in credit category C			−0.006	(0.107)
$\gamma_5 = \beta_5/\delta_5$: Inference in credit category B			0.253***	(0.092)
$\gamma_6 = \beta_6/\delta_6$: Inference in credit category A			0.165	(0.192)
$\gamma_7 = \beta_7/\delta_7$: Inference in credit category AA			0.450***	(0.055)
p-value: $\gamma_i = \gamma$			0.002	
p-value: $\gamma_i = 0$			0.000	

Notes. This table examines the ability of lenders to infer borrower creditworthiness along the credit-score dimension. Column (1) takes a simple approach and asks whether, conditional on the observable credit category, credit score predicts the interest rate (measured as the one-year interest rate). It estimates an OLS specification in which the sample is restricted to funded listings. Column (2) implements a more flexible specification described in Equation (B.1) in Online Appendix B, and estimates the extent of inference that takes place using the full baseline sample, including unfunded listings. In column (2), all estimates are based on a censored normal regression with the interest rate as the dependent variable. Standard errors are allowed to be clustered by borrower (some borrowers apply for more than one loan) and are in parentheses.
*Significant at 10%; **significant at 5%; and ***significant at 1%.

is only for coded listing content and our inference estimate may be partly based on uncoded listing content, a fairer comparison is to relate the inference based solely on coded content with the benchmark. As we will see later (Table 7), we estimate inference of 0.04 from uncoded content, so $0.33 - 0.04 = 0.29$ is inferred from coded content. Thus, lenders inferred $0.29/0.42 = 69\%$ of what was attainable. This is a significant achievement: lenders infer creditworthiness beyond what is captured by credit score and also infer more than two-thirds of the information available that is directly captured by the credit score. To understand the economic significance of this result, note that the mean offered interest rate falls by 411 basis points within each 40-point credit category.[24] $\gamma =$ 0.330 implies that lenders offer an interest rate that is 137 $(= 0.330 \times 411)$ basis points lower to the borrowers with the highest credit score within a credit category relative to the borrowers with the lowest credit score in that same category.

If we relax our assumption that lenders exclusively try to maximize returns by allowing for charitable motives, gamma remains an unbiased estimate as long as charitable motives do not vary systematically with credit score within a credit-category bin. If charitable motives systematically decrease with credit

[24] We calculate the 411 basis-point decline by summing the αs and βs and scaling this sum by the width of a single credit category. This decline in the offer rate is greater than the decline in the market interest rate because censoring is much more severe in the lowest credit categories than in the highest credit categories. In particular, only 1.8% of listings are funded in the lowest credit category, whereas 30.9% of listings are funded in the highest credit category.

Table 7 Decomposing Inference Along the Credit-Score Dimension

	(1) All credit categories		(2) Low credit categories (HR – C)		(3) High credit categories (B – AA)		(4) Low = High *p*-value	
All listing content (γ)	0.328*** (0.027)		0.244*** (0.044)		0.417*** (0.028)		0.001	
Decomposition of γ								
1. Standard financial variables	0.312*** (0.020)		0.210*** (0.020)		0.421*** (0.034)		0.000	
1.1. Number of current delinquencies		0.079*** (0.006)		0.110*** (0.010)		0.045*** (0.007)		0.000
1.2. Number of credit inquiries, last six months		0.054*** (0.003)		0.073*** (0.004)		0.034*** (0.003)		0.000
1.3. Amount delinquent		0.051*** (0.006)		0.085*** (0.010)		0.015*** (0.006)		0.000
1.4. Debt-to-income ratio		0.048*** (0.007)		0.001 (0.008)		0.099*** (0.011)		0.000
1.5. Amount requested		−0.005 (0.005)		−0.124*** (0.006)		0.122*** (0.009)		0.000
1.6. All other standard financial variables		0.085*** (0.016)		0.065*** (0.017)		0.106*** (0.028)		0.226
2. Soft/nonstandard information	0.016 (0.032)		0.034 (0.045)		−0.004 (0.044)		0.557	
2.1. Borrower maximum rate		0.064*** (0.004)		0.083*** (0.005)		0.043*** (0.007)		0.000
2.2. Listing category		−0.026*** (0.003)		−0.048*** (0.005)		−0.002 (0.005)		0.000
2.3. Member of group		−0.016*** (0.002)		−0.028*** (0.004)		−0.003*** (0.001)		0.000
2.4. Group leader reward rate		−0.015*** (0.002)		−0.028*** (0.004)		−0.002 (0.002)		0.000
2.5. All other nonstandard variables		−0.031*** (0.005)		−0.042*** (0.008)		−0.019*** (0.006)		0.025
2.6. Uncoded listing content		0.040 (0.032)		0.096** (0.045)		−0.020 (0.044)		0.066

Notes. This table decomposes our estimate of inference of creditworthiness (along the dimension captured by credit score) presented in Table 6, column (2) into sources of inference. The decomposition is based upon the baseline censored normal specification with the addition of 216 control variables, each interacted with seven credit-category dummies, such that the coefficient on each control variable is allowed to vary by credit category. Column (1) presents results for the entire sample, and columns (2) and (3), present the combined gamma separately for the lower credit categories (C, D, E, and HR) and the higher credit categories (AA, A, and B). Column (4) presents the *p*-value from a test of whether the estimates for the lower and higher credit categories are equal. The top row presents our estimate of gamma. The rows below decompose the gamma in the top row into two main groups: 1. standard financial variables and 2. soft/nonstandard information. Each of these two main groups are broken down further into subgroups 1.1–1.6 and 2.1–2.6, respectively. Refer to Online Appendix C, Table C.1 for the full decomposition results and variable definitions. Standard errors are allowed to be clustered by borrower (some borrowers apply for more than one loan) and are in parentheses.

*Significant at 10%; **significant at 5%; ***significant at 1%.

score within a bin, the interest rate would fall less within a bin than would be the case absent charitable motives, and gamma would be an underestimate of true inference.

Although we focus on the combined gamma, we note that there is variation across credit categories. However, we caution against making too much of comparisons between the categories, since each individual estimate is not precise given the smaller sample sizes. Our preferred approach is to compare high and low credit categories by grouping individual ones, and we will do so later.

Incomplete inference ($\gamma < 1$) implies that borrowers just below a category boundary pay a higher interest rate than borrowers just above the boundary. One might expect that Prosper disproportionally attracts listings by individuals with credit scores in the lower ranges of each category. Freedman and Jin (2015) present evidence consistent with such adverse selection. Adverse selection, however, does not bias our estimates since we observe exact credit score and our estimator does not depend on the density of observations by credit score within a category.

4.3.2. An Exact Decomposition of Inference. Table 7 decomposes inference along the dimension of creditworthiness captured by the credit score. The first three columns present the results from a single regression (Equation (B.4) in Online Appendix B) that decomposes the total combined gamma into components that are explained by specific variables in the listing. The last column presents the *p*-value from a test of whether the combined gamma is equal across the low and high categories.

We start by presenting analogous results from our baseline specification in Table 6 (column (2)). As before, the total combined gamma is 0.33.[25] The gamma for the lower and higher credit categories are 0.24 and 0.42, respectively, consistent with earlier results in Table 3 showing that lenders better inferred borrower creditworthiness for high-quality borrowers. The next rows present the contributions that the standard and nonstandard/soft variables make to the total combined gamma. We also report the gammas for the variables within each subcategory that show the largest inference. Table C.1 in Online Appendix C presents individual gammas for all the variables separately.

We take away four main points from Table 7. First, along the credit-score dimension, lenders learn more from standard financial variables. This is not surprising, as we measure inference of creditworthiness only along the dimension of credit score, which is a summary statistic of hard information. As shown in previous sections, soft/nonstandard information is an important contributor to inference of aspects of creditworthiness that are not captured by the credit score.

Second, among standard financial variables, most of the inference is drawn from the number of current delinquencies, the number of credit inquiries, the amount delinquent, and the debt-to-income ratio. Inference for the first three variables is greater in the lower credit categories. For the debt-to-income ratio, there is greater relative inference in the higher credit categories.

Third, soft/nonstandard variables appear to be relatively more important for lower credit categories. Among the coded soft/nonstandard variables, inference content is highest for the borrower maximum rate (the maximum interest rate the borrower is willing to pay to get the loan funded). In fact, it is the second most important inference variable among the 40 (including standard financial variables) that we examined. The average inference from the borrower maximum rate is 0.064 (or 19% of total inference) across all credit categories and is greater for lower credit categories (33.9%) than for higher credit categories (10.2%). We suspect that lenders draw much inference from the borrower maximum rate because it serves as a credible signal of creditworthiness. We find that borrowers who choose a lower borrower maximum rate have a lower probability of their listing being funded, even conditional on credit score (results not reported). Since more creditworthy borrowers likely have better outside borrowing options, it is less costly for them to post a lower borrower maximum rate. Although establishing this as a separating equilibrium requires further assumptions that we do not have the data to test for, it does suggest that such a single crossing property may in fact be generated in equilibrium.

The fourth main finding concerns the importance of inference from uncoded information, which is soft/nonstandard by definition. Whereas the gamma on uncoded listing content is insignificant for the whole sample, we estimate a statistically significant gamma of 0.096 (39% of total inference) from uncoded sources in the lower credit categories. This result is consistent with our earlier finding from Table 5 that the marginal contribution to inference from uncoded listing content is most important for the lower credit categories.[26] Thus, uncoded listing content plays an important role for inference for weaker borrowers.

We finally note that several variables contribute to negative inference. For some variables, like amount requested, this negative inference could reflect correct inference along other dimensions of creditworthiness not captured by credit score. This would be the case if, for a given credit score, larger loan amounts increase default likelihood. For other variables such as posting a picture, this negative inference may be indicative of mistakes. Alternatively, lenders may know that a borrower is more likely to default but offer a better interest rate because of charitable motives.

5. Conclusion

Our results show that lenders in peer-to-peer markets infer borrowers' creditworthiness using the rich information set that these markets provide. Lenders predict default with 45% greater accuracy than the credit score and achieve 87% of the predictive power of an econometrician who observes all standard financial information as well as in-sample future default realizations. We further find that lenders rely on nonstandard or soft sources of information in their screening process and that such information appears to be

[25] In the first line of Table 7, we report the sum of all the components of γ. As noted in the methodology section in Online Appendix B, the decomposition of gamma into its components only holds in expectation in the case of a censored regression. As a result, the estimate of the sum of the components, 0.328 from Equations (B.4) and (B.5), is close but not identical to the direct estimate of gamma, 0.330 from Equation (B.1), that we presented in Table 6.

[26] We find similar (not reported) inference from uncoded listing content for subsamples where softer information may be more important: listings with images, listings where the borrower has at least one delinquency, and listings where the number of characters exceeds 900 (median). We find similar inference from uncoded listing content using linear or cubic controls for our x variables, suggesting that this estimate is robust to the functional form of the control variables.

relatively more important when screening borrowers of lower quality. In addition, the use of credible signals (like borrower maximum rate) that are nonstandard in banking contexts suggests that enhancing the opportunity for borrowers to post credible signals can further facilitate the screening process.

Our results highlight that even markets with nonexpert individuals can effectively screen for borrower creditworthiness. Individuals collectively perform well in solving a problem that is generally thought to be best left to experts with access to "hard verified data." Ex post, this may not be surprising—after all, whether a person defaults is not determined by a mechanical financial formula but a complex human calculus. One's peers may be able to glean additional information by reflecting upon their own experiences and understanding of human behavior. In effect, given the nuances of human behavior, peers likely have an advantage in interpreting nonstandard information. Our paper highlights the value of harnessing peer-evaluation mechanisms, and especially those that use soft/nonstandard information. Given peer-to-peer markets' ability to effectively screen borrowers, and given their noncollateral-based lending structure, such markets can offer a potential capital source for small borrowers who may otherwise be limited to more costly sources of finance, such as payday lenders and credit-card debt (Morse 2010).

Our results also underscore the need to design better mechanisms to incorporate soft or nonstandard information in banking systems that rely on more rule-based lending.[27] For instance, banks could ask borrowers to provide their reservation rate in loan applications. With more information being generated on individuals than ever before and with technology drastically reducing peer-to-peer transaction costs, such mechanisms hold major potential in enhancing the effectiveness of financial markets.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2015.2181.

Acknowledgments

The authors are extremely grateful to Prosper.com for being so generous with their time in answering queries. The authors thank Raymond Fisman, Matthew Gentzkow, Lawrence Katz, Atif Mian, Enrichetta Ravina, David Scharfstein, Jesse Shapiro, and Jeremy Stein, as well as seminar participants at Harvard University, the National Bureau of Economic Research finance meetings, and UCLA Anderson for helpful comments. The authors thank David Robinson for excellent research assistance. The views expressed in this paper are solely those of the authors.

[27] In banks, loan officers typically acquire soft information for larger clients during the screening process. However, this process is time consuming and is often bypassed when screening smaller borrowers and in the automated underwriting process.

References

Adams W, Einav L, Levin J (2009) Liquidity constraints and imperfect information in subprime lending. *Amer. Econom. Rev.* 99(1): 49–84.

Agarwal S, Hauswald R (2010) Distance and private information in lending. *Rev. Financial Stud.* 23(7):2757–2788.

Agarwal S, Ambrose BW, Chomsisengphet S, Liu C (2011) The role of soft information in a dynamic contract setting: Evidence from the home equity credit market. *J. Money, Credit Banking* 43(4):633–655.

Altonji JG, Pierret CR (2001) Employer learning and statistical discrimination. *Quart. J. Econom.* 116(1):313–350.

Arrow KJ, Forsythe R, Gorham M, Hahn R, Hanson R, Ledyard JO, Levmore S, et al. (2008) The promise of prediction markets. *Science* 320(5878):877–878.

Berger AN, Udell GF (2002) Small business credit availability and relationship lending: The importance of organizational structure. *Econom. J.* 112(477):F32–F53.

Berger AN, Miller NH, Petersen MA, Rajan RG, Stein JC (2002) Does function follow organizational form? Evidence from the lending practices of large and small banks. *J. Financial Econom.* 76(2):237–269.

Biais B, Hillion P, Spatt C (1999) Price discovery and learning during the preopening period in the Paris Bourse. *J. Political Econom.* 107(6):1218–1248.

Bronfman C, McCabe K, Porter D, Rassenti S, Smith V (1996) An experimental examination of the Walrasian "Tatonnement" mechanism. *RAND J. Econom.* 27(4):681–699.

Butler AW, Cornaggia J, Gurun UG (2014) Do local capital market conditions affect consumers' borrowing decisions? Working paper, Rice University, Houston.

Cao C, Ghysels E, Hatheway F (2000) Price discovery without trading: Evidence from NASDAQ preopening. *J. Finance* 56(3):1339–1365.

Crawford V, Sobel J (1982) Strategic information transmission. *Econometrica* 50(6):1431–1451.

Davies RJ (2003) The Toronto Stock Exchange preopening session. *J. Financial Markets* 6(4):491–516.

Delong ER, Delong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* 44(3): 837–845.

Duffie D, Manso G (2007) Information percolation in large markets. *Amer. Econom. Rev. Papers Proc.* 97(2):203–209.

Duffie D, Malamud S, Manso G (2009) Information percolation with equilibrium search dynamics. *Econometrica* 77(5):1513–1574.

Einav L, Jenkins M, Levin J (2013) The impact of credit scoring on consumer lending. *RAND J. Econom.* 44(2):249–274.

Experian (2004) ScoreX PLUS. Accessed July 13, 2015, http://annualcreditreport.experian.com/products/pdf/scorex_plus.pdf.

Farber HS, Gibbons R (1996) Learning and wage dynamics. *Quart. J. Econom.* 111(4):1007–1047.

Farrell J, Rabin M (1996) Cheap talk. *J. Econom. Perspect.* 10(3): 103–118.

Forsythe R, Lundholm R (1990) Information aggregation in an experimental market. *Econometrica* 58(2):309–347.

Freedman SM, Jin GZ (2015) Learning by doing with asymmetric information: Evidence from prosper.com. NBER Working Paper 16855, National Bureau of Economics, Cambridge, MA.

Grossman S (1976) On the efficiency of competitive stock markets where traders have diverse information. *J. Finance* 31(2): 573–585.

Grossman S, Stiglitz J (1980) On the impossibility of informationally efficient markets. *Amer. Econom. Rev.* 70(3):393–408.

Hanson R, Oprea R, Porter D (2006) Information aggregation and manipulation in an experimental market. *J. Econom. Behav. Organ.* 60(4):449–459.

Hildebrand T, Puri M, Rocholl J (2014) Adverse incentives in crowdfunding. Working paper, Duke University, Durham, NC. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1615483.

Jiang W, Nelson AA, Vytlacil E (2014) Liar's loan? Effects of origination channel and information falsification on mortgage delinquency. *Rev. Econom. Statist.* 96(1):1–18.

Keys BJ, Mukherjee T, Seru A, Vig V (2010) Did securitization lead to lax screening? Evidence from subprime loans. *Quart. J. Econom.* 125(1):307–362.

Liberti JM, Mian AR (2009) Estimating the effect of hierarchies on information use. *Rev. Financial Stud.* 22(10):4057–4090.

Lin M, Viswanathan S, Prabhala NR (2013) Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. *Management Sci.* 59(1):17–35.

Morse A (2010) Payday lenders: Heroes or villains. *J. Financial Econom.* 102(1):28–44.

Morse A (2015) Peer-to-peer crowdfunding: Information and the potential for disruption in consumer lending. Working paper.

Petersen M (2004) Information: Hard and soft. Working paper, Northwestern University, Evanston, IL.

Plott CR, Sunder S (1988) Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometrica* 56(5):1085–1118.

Pope DG, Sydnor J (2011) What's in a picture? Evidence of discrimination from prosper.com. *J. Human Resources* 46(1):53–92.

Rajan U, Seru A, Vig V (2015) The failure of models that predict failure: Distance, incentives, and defaults. *J. Financial Econom.* 115(2):237–260.

Ravina E (2012) Love and loans: The effect of beauty and personal characteristics in credit markets. Working paper, Columbia Business School, New York.

Rigbi O (2013) The effect of usury laws: Evidence from the online loan market. *Rev. Econom. Statist.* 95(4):1238–1248.

Samaad MA (2014) Peer-to-peer lending poised for more growth: Fed. *Credit Union Times* (August 14), http://www.cutimes.com/2014/08/14/peer-to-peer-lending-poised-for-more-growth-fed.

Stiglitz JE, Weiss AM (1981) Credit rationing in markets with imperfect information. *Amer. Econom. Rev.* 71(3):393–410.

Theseira W (2009) Competition to default: Racial discrimination in the market for online peer-to-peer lending. Dissertation, The Wharton School, University of Pennsylvania, Philadelphia.

Townsend RM (1978) Market anticipations, rational expectations, and bayesian analysis. *Internat. Econom. Rev.* 19(2):481–494.

Vives X (1993) How fast do rational agents learn? *Rev. Econom. Stud.* 60(2):329–347.

Vives X (1995) The speed of information revelation in a financial market mechanism. *J. Econom. Theory* 67(1):178–204.

Wolfers J, Zitzewitz E (2004) Prediction markets. *J. Econom. Perspect.* 18(2):107–126.

Wolinsky A (1990) Information revelation in a market with pairwise meetings. *Econometrica* 58(1):1–23.