Contents

The Folly of the Crowd:
Market Inefficiencies in Peer to Peer Lending

Samuel Shleifer

Yale University

March 27, 2015

Abstract

This paper documents violations of market efficiency in Peer to Peer Lending, using data from Prosper.com, America's first Peer-to-Peer Lending Marketplace. Peer to Peer lending is the practice of individual and institutional investors lending money to borrowers over the internet, without the assistance of a traditional financial intermediary. Theories of Market Efficiency predict that investors price assets to earn an expected return commensurate with risk taken, and that asset "prices should fully reflect publicly available information." Contrary to these, and other, efficient market hypothesis predictions:

1. Investors in risky Prosper loans earn negative average returns during the first few years of its history.

2. Investors earn lower returns on riskier loans during the first few years of Prosper's history.

3. Strategies using publicly available information can earn superior risk-adjusted returns.

4. Poor loan performance during the first few years of Prosper's history is a consequence of poor, and not just an unfortunate time period. Had investors better used the information at their disposal, they would have earned better returns.

These violations of market efficiency are discussed in the midst of two market design changes by Prosper that guided and then forced investors towards setting more efficient interest rates. Such guidance and eventual price control, as opposed to unconstrained auction rate setting, improved loan performance.

1 Introduction

Historically, there are two major perspectives of efficient market prices. Fama ((1970)'s semi-strong efficient markets hypothesis ("EMH") predicts that a strategy using publicly available information should not be able to earn excess risk-adjusted returns. Risk models, such as CAPM, imply that risky assets should have positive expected returns and that riskier assets should have higher returns than less risky assets. The information aggregation perspective associated with (Hayek, 1945) and (Stiglitz and Grossman, 1980), further predicts that many individuals, each using their own information, collectively arrive at a more efficient price than a central planner who sets prices, since the central planner cannot access all the information available to the crowd.[1][2]

In the P2P Lending context, the market's price is the interest rate associated with a loan. This paper studies the question of whether interest rates on Prosper.com are set efficiently. (Fama, 1970)'s semi-strong EMH predicts that if interest rates were set efficiently, then...

1. Loans would have positive expected returns. Investors would be unwilling to lend to Borrowers if they expected to lose money. [3]

2. Riskier loans would have higher expected returns.[4]

3. It is not possible to do better than the average investor using publicly available information.[5]

I test hypothesis 1 by measuring the historical returns of loans up to a certain date, and then asking whether loans with negative expected returns get funded at that date. I test hypothesis 2 by comparing the Returns of high risk loans to low risk loans. I test hypothesis 3 by using publicly available information to select loans with a higher than average expected return.

Prosper.com data provide some additional valuable insights into investor behavior. As it happens, over the period I study, there are three interest rate determination regimes:

1. Early Auction Regime (April 19, 2006-October 15, 2008): Rates are set by investor auction. Investors have large amounts of data, but Prosper does not guide them meaningfully.

2. Late Auction Regime: (July 1, 2009-December 20, 2010): Rates are set by investor auction, with Prosper offering valuable guidance in the form of more complex credit grading.

3. Prosper-rate Regime: (December 21, 2010 - August 3, 2011): Prosper sets rates using its more complex credit grading algorithm.

---

[1]Sanford J. Grossman and Joseph E. Stiglitz. "On the Impossibility of Informationally Efficient Markets". In: Reprint Series 121 (Dec. 1980). URL: http://www.nber.org/papers/r0121.

[2]Friedrich A. Hayek. *The Use of Knowledge in Society*. Library of Economics and Liberty, 1945. URL: http://www.econlib.org/library/Essays/hykKnw1.html.

[3]Assuming a Risk-Free Rate of 0

[4]William F Sharpe. "Capital asset prices: A theory of market equilibrium under conditions of risk*". In: *The journal of finance* 19.3 (1964), pp. 425–442.

[5]Eugene F. Fama. "Efficient Capital Markets: A Review of Theory and Empirical Work". In: *The Journal of Finance* 25.2 (1970), pp. 383–417. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1970.tb00518.x. URL: http://dx.doi.org/10.1111/j.1540-6261.1970.tb00518.x.

The early auction regime is closest to Hayek's vision of the free market, since interest rates are determined freely by investors through auctions. In this case, the EMH predictions of no negative average returns, a positive relationship between risk and return, and no improvement in performance from using public information are most applicable. Yet, as I show, it is precisely in this regime that the three predictions of the EMH are most dramatically violated. In the two later regimes, when Prosper first provides much more guidance to investors, and then actually sets interest rates, investor returns improve dramatically. The evidence across regimes is broadly inconsistent with the view that local knowledge and the wisdom of the crowd improves price discovery. To the contrary, it appears that a "centralized' Prosper mechanism of either guiding or setting interest rates actually leads to fewer predictable mispricings: as Prosper becomes more involved in the rate-setting process, the excess returns available using publicly available information approach 0, while aggregate returns exhibit higher mean and lower variance.

A major confound of this analysis is the time period. Perhaps 2006-2008, the early auction period, is just a bad time to invest in consumer credit, whereas 2009-2011 is much better. To address this concern, I estimate the importance of time effects by comparing default rates of loans to similar borrowers in different periods. I then show that time period is not the only explanation of the various changes to the marketplace by estimating what loan performance would have been if pre-shutdown borrowers borrowed at the higher Prosper-set interest rates.

1.1 Background

Peer-to-peer lending ("P2PL") is the practice of lending money to borrowers over the internet, without going through a traditional financial intermediary. [6] From April 2006 until December 2010, interest rates were set by lenders, who competed for the lowest rate in a reverse auction model.

In October 2008, Prosper was shut down by the SEC, and re-opened after a 9 month quiet period with more stringent under-writing standards, and greater guidance for lenders . Importantly, all the information about borrower creditworthiness that is used in this paper was available to investors when they bid on loans. Since December 2010, the beginning of *Prosper-rate* regime, rates have been pre-set by Prosper on the basis of a Borrower's Credit history.

Under all three regimes, lenders could choose to whom to lend and how much to lend, allowing them to diversify across many loans, thereby decreasing the variance of their returns. In exchange for verifying borrowers' identities and managing servicing, Prosper charges a 0.5-5 percentage point origination fee on each loan depending on the borrower's risk profile and loan duration. This fee is usually 1%. Fees are added to the borrower's monthly payment.[7] This rate is lower than that charged by most traditional financial institutions, which have the higher costs of physical branches, and pay for human analysis of credit history.[8] The P2PL model relies on lenders being able to infer the riskiness of borrowers: since Prosper gets fees from monthly payments, they too are incentivized to maximize investor returns, both for short term profit, and long term reputation.

---

[6]en.Wikipedia.org/wiki/Peer-to-peer_lending

[7]There are no prepayment penalties.

[8]Samuel G Hanson et al. *Banks as Patient Fixed-Income Investors*. Tech. rep. National Bureau of Economic Research, 2014.

In my sample, Prosper's origination volume, as shown in Figure 1, peaked in the months before the shutdown, and then slowly resumed after the quiet period. Since 2011, when the sample ends, origination volume has eclipsed its 2008 peak.
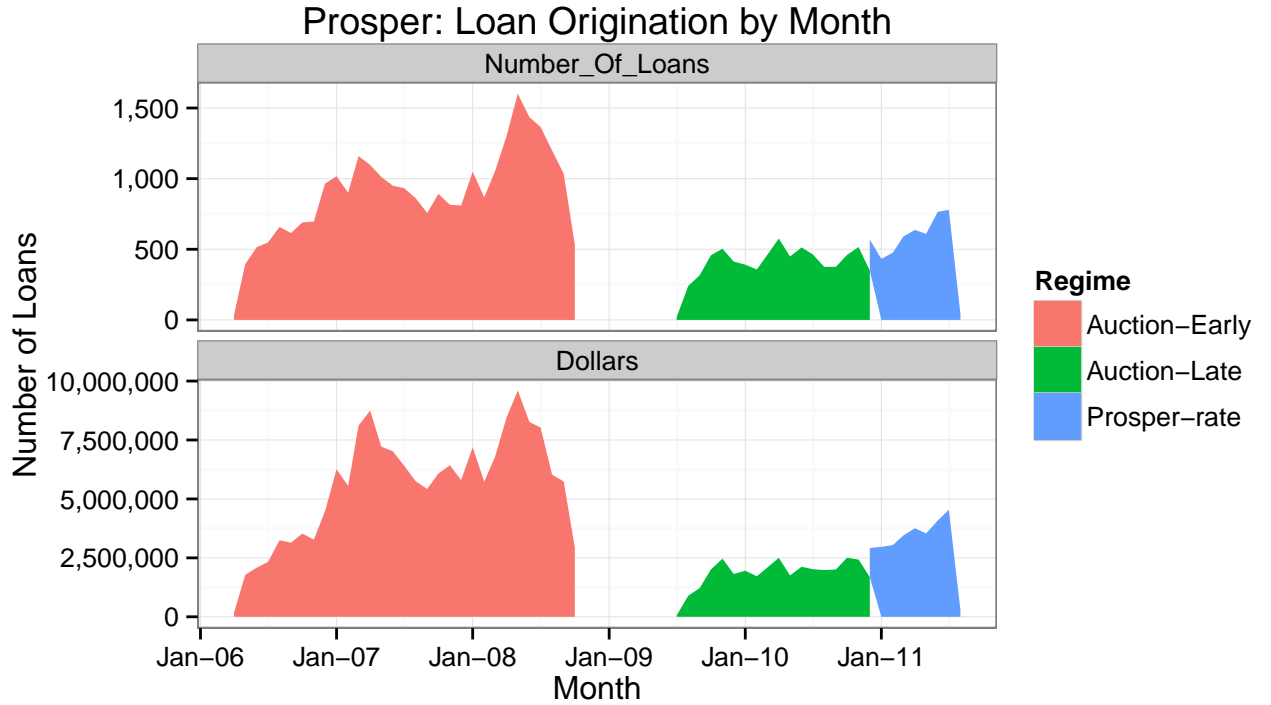
Figure 1: This chart shows two panels. The first panel shows the number of Loans funded on Prosper in each month of the sample. The second panel shows the total dollars lent to borrowers in each month of the sample.

	purpose	n	Frequency
1	Debt Consolidation	10531	0.446
2	Other	4167	0.176
3	Business	3016	0.128
4	Personal	2365	0.100
5	Home Improvement	1783	0.075
6	Auto	1017	0.043
7	Student Use	748	0.032

Table 1: Stated Reasons for Borrowing on Prosper (NAs removed)

Reasons for borrowing on Prosper, as shown in Table 1, have changed relatively little over

the site's 8 year history: the plurality of borrowers aim to consolidate credit, while fewer use the funds to start a business, or pay for a wedding or a car. Morse (2015) finds that "Borrowers are characterized as debt-laden, middle-to-high income, individuals who are consolidating credit cards and other debt." The average P2P loan face value comprises 20.5% of annual income, and payments absorb 7.5% of monthly income.[9]

1.2 Auction Design

From April 19, 2006 until December 20, 2010. Prosper operated a *variable rate* model, where lenders and borrowers determined interest rates using a Dutch auction-like system, where each investor bid consisted of a (dollar amount, interest rate) pair. In these auctions, borrowers would declare a maximum reservation interest rate, where bidding would start. Over the course of a 7-10 day bidding period, that rate would be bid down by investors, with the loan's interest rate closing at the lowest rate $R_{clearing}$ such that enough investors were willing to fully fund the loan at or below the $R_{clearing}$. All investors who bid at or below the clearing rate would receive the clearing rate. If a borrower could not reach 100% funding at his reservation rate, the loan would not be funded.

As an instructional example, suppose Bill the borrower needed to borrow $100, and set his maximum rate at 20%. Let Chris, David and Ellen be investors, with Chris bidding ($50, 15%), David bidding ($50, 18%) and Ellen bidding ($30,19%). In this scenario, Chris and David would win the auction, and each lend lend Bill $50, and receive an 18% interest rate. Bill would pay 18% interest before fees.

1.3 Early Auction Regime

In the early auction regime, loans originated between April 19, 2006 and October 15, 2008, auction participants had access to considerable information from Prosper, including over 60 Borrower characteristics, such as residence, employment status and a FICO score bucket, but few intelligent risk estimates. Additionally, Prosper did not regulate who could apply for a loan on the website. Besides basic fraud prevention, any applicant who received 100% funding from investors was eligible, for Prosper's first year of operation. On March 1, 2007, amidst poor returns, Prosper announced that listings with credit scores below 520 would no longer be offered to investors. This intervention is not discussed for lack of space, but shows that after a year, more or less, Prosper felt the need to nudge a fully decentralized investment process in the right direction.

In an October 15, 2008 email, Prosper announced that it would temporarily stop accepting new loan applications to register with the SEC, which subsequently released a cease and desist order characterizing Prosper as a "seller of securities," rather than a marketplace merely matching lenders and borrowers, as the site had argued. [10][11]

---

[9]Adair Morse. "Peer-to-Peer Crowdfunding: Information and the Potential for Disruption in Consumer Lending". In: Working Paper Series 20899 (Jan. 2015). DOI: 10.3386/w20899. URL: http://www.nber.org/papers/w20899.

[10]techcrunch.com/2008/11/26/sec-outlines-its-reasoning-for-shutting-down-p2p-lender-prosper/

[11]SEC LETTER

1.4 Late Auction Regime

After a 9 month quiet period, Prosper re-opened, using the same Auction model, with three significant market interventions. First, under-writing standards again became more stringent: new borrowers were required to have a FICO score above 640, and repeat borrowers were required to have a score above 600.

Second, whereas previously Prosper used only credit score to determine a borrower's Credit Grade (eg. AA), the site began offering a more advanced, multi variable formula to determine each borrower's credit grade for investors consideration. Finally, the site also partnered with FolioFN, a secondary market where lenders could liquidate their positions. Crucially, Prosper still had no direct role in setting interest rates, investors still used an auction, but with the extra information credit grades that more closely matched default probabilities. Besides credit score, Prosper began using (and still uses) the following "Key variables" to determine its revamped credit grades: [12]

- TotalInquiries: The number of times potential lenders requested the borrowers' credit portfolio from the credit bureau, which is usually Experian or a competitor.

- CurrentDelinquencies: The number of delinquent loans associated with the borrower.

- Credit card utilization: how much of available credit the borrower uses on a monthly basis.

- Number of open credit card accounts

- Debt to income ratio

- Loan payment performance on prior Prosper loans.

This Auction rate-setting mechanism continued for more than a year after the shutdown, until December 20, 2010.

1.5 Prosper-rate Regime

Under the current *Prosper-rate* model, enacted December 21, 2010, lenders no longer determine the loan rate in an auction. Instead, Prosper uses the intelligent credit grades given to investors in the late auction regime to set interest rates on loans, with slightly less stringent under-writing standards to determine eligibility. The change was motivated by a desire to accelerate the listing process, so that investors could reinvest money more quickly. Moreover, towards the end of the late auction regime, auctions were arriving at interest rates that were almost completely determined by the credit score. Under the fixed-rate regime, investors choose whether or not to invest in a given loan, at the interest rate set by Prosper, so bids only consist of a dollar amount.

---

[12]CITE = Estimated Loss Page

2 Literature

2.1 Efficient Markets Hypothesis

In layman's terms, the efficient markets hypothesis (EMH), states that 'it is impossible to beat the market.' More specifically, Fama explains that market prices should always fully reflect publicly available information. Therefore, EMH rules out the possibility of strategies that earn excess returns using only publicly available information.[13] In the P2P lending context, this market price is the interest rate determined by investors (or accepted by investors in the fixed-rate regime), and paid by the borrower, and returns are the cash flows received from borrowers divided by the amount invested. (Fama, 1970) notes the dependence of most tests of market efficiency on a model of risk and return. The models, including most famously, Sharpe's capital asset pricing model ('CAPM'), make two predictions: First, since an investor can always hold cash, the average return on risky investments must be positive. Second, securities with higher systematic risk earn higher average returns.[14]

The EMH rests on three assumptions. "First, investors are assumed to be rational and hence to value securities rationally. Second, to the extent that some investors are irrational, their decisions are random and uncorrelated. Third, to the extent that some investors are irrational in similar ways, they are met in the market by rational arbitrageurs, who eliminate their influence on prices."[15]

In a reverse interest rate auction, such as Prosper's market design, however, arbitrageurs cannot short a borrower and increase his interest rate. In the situation where irrational investors under estimate a borrower's riskiness, they will outbid rational bidders, and set the market price. To earn excess returns, then, an investor must select loans for which other investors have demanded too high a rate, so that he can participate in funding.

2.2 Soft Information and Credit Screening

The novel idea of P2P is a large group of investors may be better equipped to assess risk than traditional financial institutions. "Traditionally, the credit score provided by rating agencies has been the main tool banks use to screen smaller borrowers."[16] As such, the initial strand of literature focuses largely on the novel sources of information utilized by P2P investors. (Gao and Lin, 2012) analyze the text written by the borrower in the "description field" of the application, and find that descriptions that are easy to read have 2.3% lower default rates than average, while complex narratives are associated with a 3.6% higher default rate, but do not analyze return on investment.[17] Pope and Sydnor (2008) find that investors can reduce their defaults by statistically discriminating

---

[13]Fama, "Efficient Capital Markets: A Review of Theory and Empirical Work".

[14]Sharpe, "Capital asset prices: A theory of market equilibrium under conditions of risk*".

[15]Andrei Shleifer. *Inefficient Markets*. 2000.

[16]Rajkamal Iyer et al. "Screening in New Credit Markets: Can Individual Lenders Infer Borrower Creditworthiness in Peer-to-Peer Lending?" In: rwp09-031 (Aug. 2009). URL: http://ideas.repec.org/p/ecl/harjfk/rwp09-031.html.

[17]Qiang Gao and Mingfeng Lin. "Linguistic Features and Peer-to-Peer Loan Quality: A Machine Learning Approach". In: *Available at SSRN 2446114* (2013).

against racial minorities, but do not analyze return on investment. (Iyer et. al,2009) find that within the credit score buckets provided by Prosper, investors use soft information to further sort borrowers within buckets, reducing defaults.[18] These papers echo Hayek (1945)'s basic argument for market efficiency: that many individuals, each using their own information, collectively arrive at one efficient price. The right price "might have been arrived at by one single mind possessing all the information which is in fact dispersed among all the people involved in the process."[19]

This essay contributes to this literature along two dimensions. First, I use ROI, rather than a simple default rate, to measure the improvement of investment strategies over a benchmark. Second, and more importantly, my paper pushes back against the idea that the soft information used by borrowers is more useful for credit screening; I find that borrowers do not do a very good job aggregating the various hard credit information at their disposal, and synthesizing this into an interest rate that reflects the riskiness of the borrower. When Prosper, as a central planner, assists them, they set much more efficient prices.

3   Data

My sample consists of 39,555 three year, unsecured (no collateral) Loans, worth $231.7 Million, originated from 2006-2011, downloaded off Prosper.com at various points by various people. Borrowers apply for loans by creating a public listing on the Prosper.com website, and can request anywhere from $1,000 to $35,000.[20]. Since our analysis requires measurement of loan performance, which can only take place 3 years after origination(given 3 year maturity), we must end the sample in 2011, as we do not know what will end up happening to more recently originated loans. As shown in Table 2, this leaves us with many more early auction regime loans than the post shutdown regimes.

	regime	n	freq.
1	Auction-Early	27764	0.7020
2	Auction-Late	7238	0.1830
3	Prosper-rate	4550	0.1150

Table 2: Observations per regime. Each observation is one loan.

To help investors better assess the creditworthiness of borrowers, for every loan originated in the site's history, Prosper posts data regarding loan performance, such as payments made, interest rate, and principal lost. Borrower characteristics, including credit variables, which are usually obtained directly from a consumer credit agency like FICO, and verification of income and employment, are also posted on the website in the borrower's *listing* for prospective investors to consider as they decide who to lend to, and available for download.

---

[18]Iyer et al., "Screening in New Credit Markets: Can Individual Lenders Infer Borrower Creditworthiness in Peer-to-Peer Lending?"

[19]Hayek, *The Use of Knowledge in Society*.

[20]The minimum increased to $2,000 effective Nov. 1, 2011

3.1 Variable Definitions for the Prosper Data

	Auction-Early	Auction-Late	Prosper-rate
roi.mean	-0.03754	0.15381	0.13493
roi.sd	0.4152	0.3226	0.3756
default.rate.mean	0.3700	0.1585	0.2112
default.rate.sd	0.4828	0.3652	0.4082
LenderYield.mean	0.1730	0.1939	0.2205
LenderYield.sd	0.07266	0.09655	0.08532
BorrowerRate.mean	0.1833	0.2040	0.2305
BorrowerRate.sd	0.07391	0.09654	0.08531
LoanOriginalAmount.mean	6166	4759	5956
LoanOriginalAmount.sd	5562	3820	3613
Investors.mean	118.13	144.17	86.91
Investors.sd	117.74	125.89	74.38
CreditScoreEst.mean	653.2	714.2	705.1
CreditScoreEst.sd	88.71	58.85	53.29
DebtToIncomeRatio.mean	0.3284	0.2292	0.2458
DebtToIncomeRatio.sd	0.9423	0.2681	0.3769
InquiriesLast6Months.mean	2.879	1.031	1.132
InquiriesLast6Months.sd	3.911	1.532	1.640

Table 3: Means and Standard Deviations for Relevant Variables by Regime.

For each loan, Prosper offers 63 interesting variables, I now define an important subset.

- BorrowerRate: Annual interest rate paid by the borrower

- TotalCreditLinespast7years: Number of credit lines, including cards and loans, used by the borrower in the seven years before the prosper loan was originated.

- Inquirieslast6Months: Number of times a lender inquired about the Borrower's credit report, in the 6 months before the loan was originated. Traditional lenders make inquiries when a borrower applies for a loan, so this serves as a proxy for the number of times someone has borrowed in the past 6 months.

- CurrentDelinquencies: Number of loans on which the borrower is delinquent at the origination date.

- DelinquenciesLast7Years: The number of loans which the borrower has not repaid fully over the 7 years before the origination date.

- MonthlyLoanPayment: The borrower's scheduled monthly payment on the Prosper loan.

- CreditScoreRangeLower: The lower bound of the credit score range (which bounds his FICO score) , that bounds a borrower's credit grade, as provided by FICO.

- CreditScoreEst: The average of the lower and upper bounds of the credit score range.

- PublicRecordsLast10Years: Number of public records of borrowing activity in the past 10 years at the origination date.

- DebtToIncomeRatio: The debt to income ratio of the borrower at the origination date. This value is capped at 10.01 (any debt to income ratio larger than 1000% will be coded as 1001%).

- TotalProsperPaymentsBilled: Number of on time payments the borrower made on Prosper loans at the time they created the listing. This value is 0 if the borrower had no prior loans.

- LenderYield: The lender's Yield on the loan. Equal to the Interest rate minus servicing fee.

- BorrowerRate: The borrower's interest rate for the loan.

- Investors: The number of investors that funded the loan.

- *.sc: The Variable * was scaled to have 0 mean and unit variance. (found only in regression results)

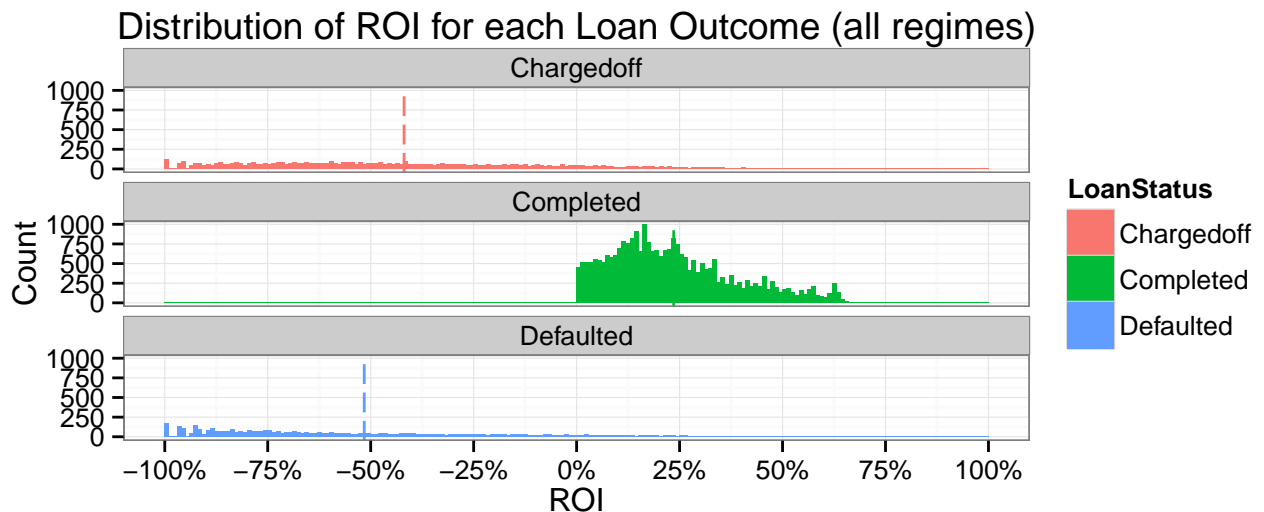## 3.2   Measuring Loan Performance with ROI

Figure 2: This chart shows the distribution of ROI for completed, defaulted and chargedoff loans over all 3 regimes. A continuous ROI variable offers more information about the quality of an investment than a binary default indicator. Dotted lines indicate group means.

In order to capture the actual impact of a loan on an investor's portfolio, we measure the ROI for a given loan as the sum of the borrower's payments / amount borrowed -1.

$$ROI = \frac{amount.repaid}{amount.borrowed} - 1 \tag{1}$$

Where amount.repaid is the sum of the borrower's monthly payments. As shown in Figure 2, a return on investment variable offers much more granularity than a binary default variable for both defaulted and repaid loans. In many observations, a borrower paid back half his principal before defaulting. A binary indicator would treat this loan, which only cost the investor half his money, as equal to a loan that repaid him nothing. Additionally, a default indicator fails to account for the higher interest payments associated with riskier loans. Still, measuring ROI and default lead to similar conclusions, as the correlation between ROI and default is -79.7% over the full sample.

I do not discount or annualize ROIs, because in the period of low interest rates, the difference between default and repayment is large enough that a few percentage points don't matter: loans that default average -46% ROI, loans that repay average + 24% ROI.[21] [22]

3.3 Prosper's Charged-Off vs. Default Terminology

After the shutdown, Prosper began using the term "Charged-Off" to mean that a loan is delinquent past 120 days. We observe more defaults than charge-offs before the shutdown, but more charge-offs than defaults after, likely because of the shift in terminology. According to Prosper's website:

> In general, a debt or account is considered charged off when it is unlikely that further payments will be received. Debts are usually charged off after they remain unpaid for a period of time (e.g., 90 to 180 days). Prosper uses the 120 days as the charge off threshold because loans that become over 120 days past due are eligible for sale to a debt buyer, and we have found that there is a steep drop-off in likelihood of further payments after 120 days of delinquency. [23]

4 Performance Across Regimes

This section begins with an examination of ROIs over time and and across borrower credit buckets. I then discuss regime differences in default rates and interest rates, the two features that largely determine returns. The discussion of interest rates includes measurement of investors' varying sensitivity to risk across regimes. The section closes by examing the impact of Prosper's interventions, which reduce the valid borrower pool to a less-risky subset and assign borrowers more informative credit grades, on investor returns.

---

[21]I will probably use discounting in the final draft, but it doesn't change results, and it feels like a low priority, at this stage.

[22]Elaborate

[23]http://blog.prosper.com/2008/09/16/charge-offs-explained/

I find the following: (1) the mean ROI of early auction regime loans is negative, (2) in the early auction regime, average ROI is lower for higher risk loans, (3) interest rates are insufficiently sensitive to risk in the early auction period and (4) investors made sub-optimal choices, given the information they had at their disposal, and (5) Prosper's interventions improved investor returns.

Finding (1) is inconsistent with the EMH prediction that risky assets like Prosper loans should offer investors a positive expected return.[24] Finding (2) is inconsistent with the prediction that riskier loans should have higher expected returns to compensate investors for taking risks. Finding (3) describes the sub-optimal investor decisions that led to (1) and (2), in response to arguments that investors chose an unlucky time period. Finding (4) is inconsistent with Hayek's notion that a free market will arrive at accurate prices without the meddling of a central planner.
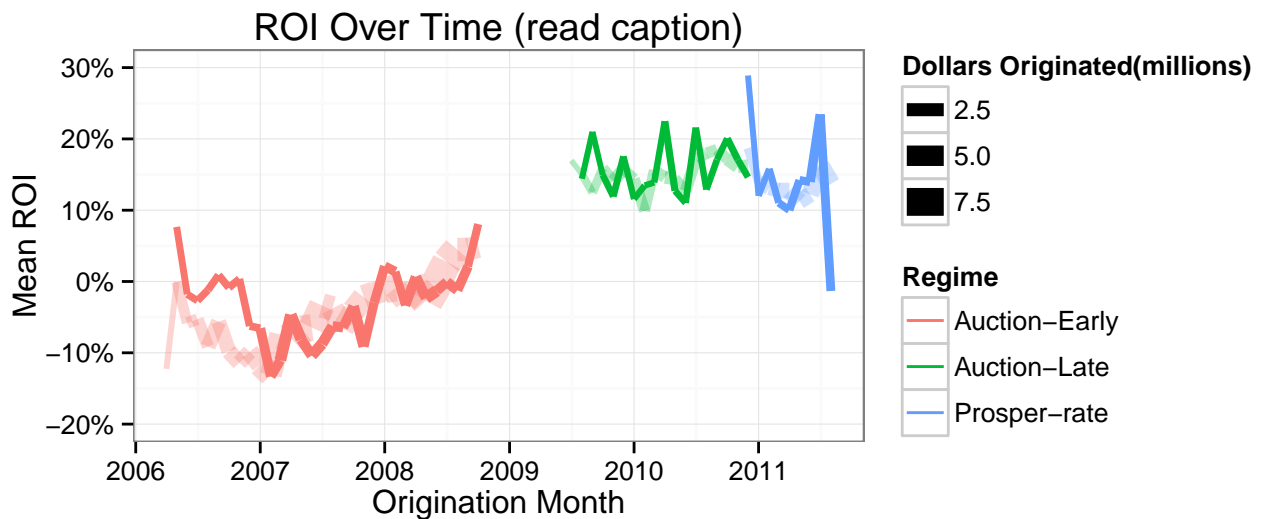
4.1   ROI Across Regimes

Figure 3: This chart shows the mean ROI of loans originated each month in our sample.The blurred line is the average ROI accross all loans. the solid line is the average ROI for loans to borrowers in the 640-660 credit score bucket. Loans funded after the shutdown are associated with higher ROI, even controlling for Borrower Creditworthiness.

Figure 3 shows the performance of loans originated in each month of the sample. The blurred lines represent the mean return across all loans for each month, while the solid lines represents the mean return to borrowers in the 640-659 credit score bucket, which was chosen because it is the most common in the data. The average ROI of all loans increases from -3.8343% in the early auction period to 15.2047% in the late auction period. The average return for loans originated in each of the site's first 27 months is negative, rejecting the EMH prediction of positive expected returns for

---

[24]Otherwise, investors would just hold cash.

risky assets. Of course, it is possible that the time period is unusual. I address this possibility in section 6.
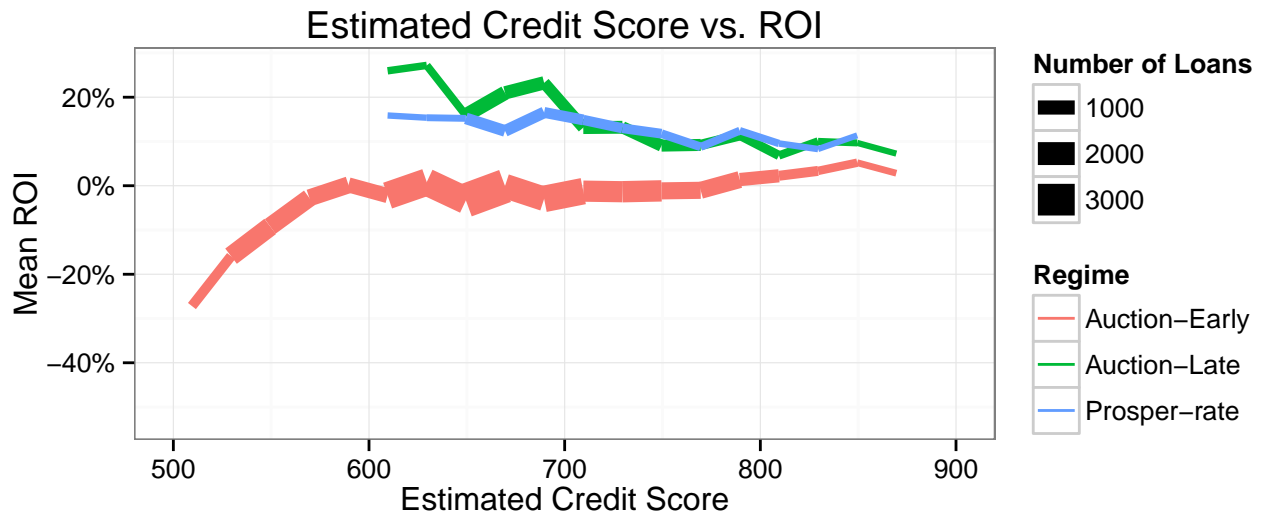
Figure 4: This chart shows the mean ROI by borrower credit bucket, under each regime. In the early auction regime, higher risk loans earn lower returns.

Figure 4 shows that higher risk loans earn lower returns in the early auction regime, inconsistent with the EMH prediction of a risk, reward tradeoff. The expected pattern is shown in later two regimes' curves: riskier loans earn higher average returns. Figure 4 also reveals early auction regime returns are worse than the other two regimes' returns for each bucket, but that the gap is much greater for riskier loans. Again, time period may be a confounding factor, as riskier investments often perform worse in bad economic times, a possibility I discuss later. We now examine default rates and interest rates, the two features that explain the variation in returns.

4.2 Interest Rates and Defaults Across regimes

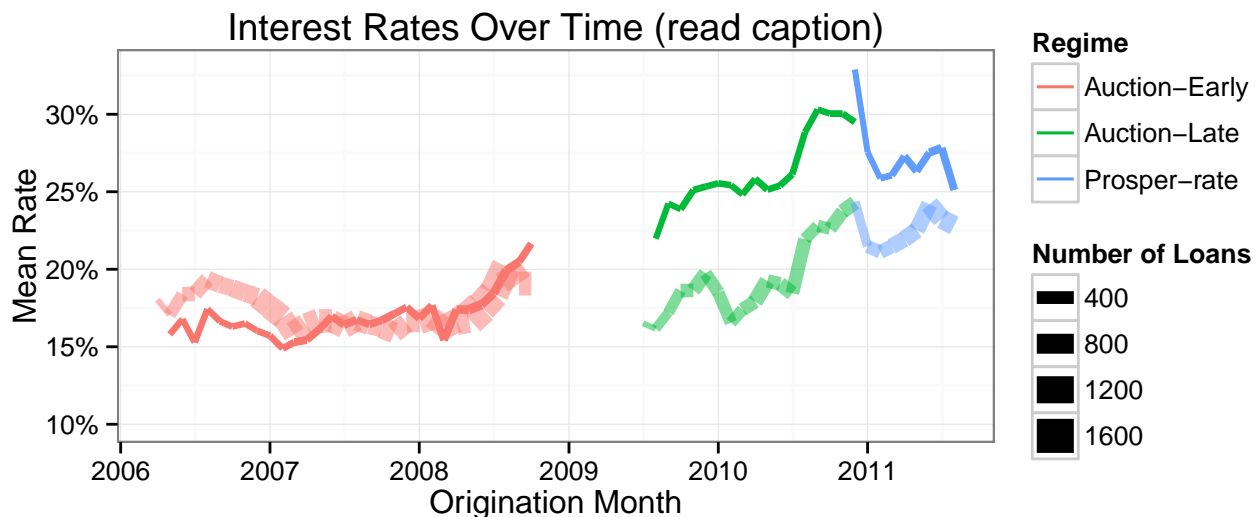Figure 5: The blurred line shows is the median interest rate of loans originated each month. The solid line represents the average rate for loans to borrowers in the 640-660 credit score bucket for each month. Loans funded after the shutdown are associated with higher rates after controlling for borrower credit score bucket.

Figure 5 presents the median interest rate [25] for each month of the sample. The solid line again represents borrowers in the 640-660 credit score bucket, while the blurred line represents all borrowers. Overall (blurred line), the median interest rate paid by borrowers increases from 17.4% to 19.6% to 22.5% over the early auction, late auction and Prosper-rate regimes. When we control for borrower selection by only considering the 640-660 credit score bucket,however, the climb in median interest rate,from 17.1% in the early auction period to 26.4% in the late auction period is more dramatic. Interestingly, rates for the 640-660 bucket are almost identical directly before and right after the shutdown but the upwards trend, from from 15% at the beginning of 2007 to 30% at the end of 2010 is clear.[26]

---

[25]Before Fees.

[26]WHY THE BIG 2010 RUNUP
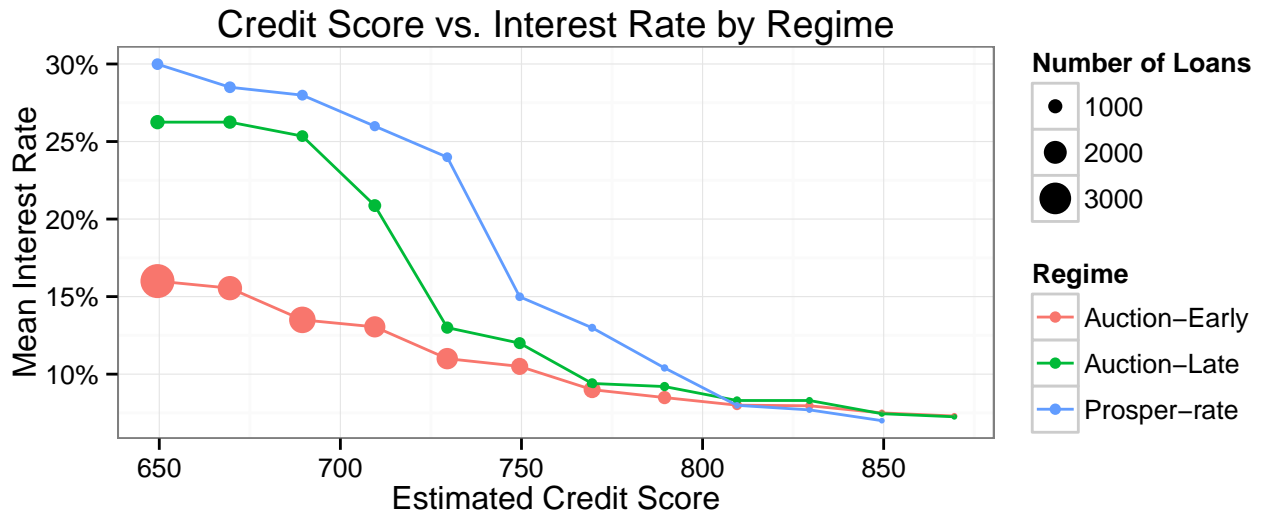
Figure 6: This Chart shows the median interest rate borrowers from each credit score bucket paid during each period. Under the Early Auction Regime, interest rates were lower, especially for risky borrowers.

Figure 6 shows that majority of the difference in interest rates between regimes can be attributed to the rates paid by borrowers in risky credit buckets, with credit scores in the 620-700 range. Investors across regimes demand similar interest rates from low-risk (high credit) borrowers. For the 760-780 bucket, early auction regime investors charged 9.0%, late auction regime investors charged 9.4%. However, for higher risk buckets, the prices are extremely different; for the 640-660 bucket, early auction regime investors charged 16.0%, while late auction investors charged 26.3%.

Figure 7: This chart shows the default rate of the borrowers in each credit score bucket.

16

More generally, the slope of the curves in Figure 6 show that late auction regime investors were much more sensitive to changes in credit score than early auction regime investors. In the early auction regime, a 100 point increase in credit score is associated with a 4% reduction in interest rate. In the late auction regime, a 100 point increase in credit score is associated with an 11% reduction. Early auction regime investors were 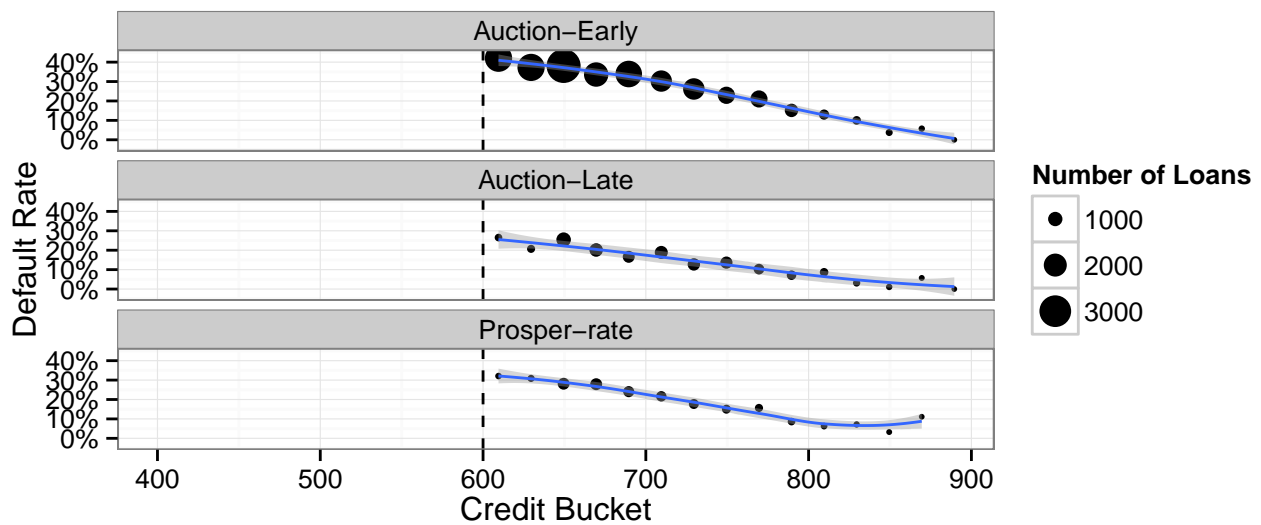especially insensitive to changes in credit score below the 600-620 buckets, as the four buckets directly below each have mean rate around 25%. Early auction regime investors' insensitivity to variance in credit score among risky borrowers works well in the context of Iyer et. al (2009)'s finding that investors relied on "soft" information when evaluating borrowers with low credit scores. This reliance on soft information, however, is not associated, in my findings, with higher returns. Figure 7 presents the relationship between credit bucket and default rate[27], and shows that the probability of default continues to increase, and may even accelerate, with reductions in credit score below the dotted line at 600.

Early auction regime investors, then, were not sufficiently sensitive to the hard credit variables that are predictive of default, especially when evaluating risky, low-credit borrowers, suggesting that the poor returns of the early auction regime were not caused only by the time period, but also by poor investor decisions.[28] We now look at the more informative grading algorithm Prosper began using after the shutdown, which made investors more sensitive to hard credit variables.

---

[27]Other Assets?

[28]NEED MORE HERE

4.2.1 More Informative Grading, Higher Interest Rates
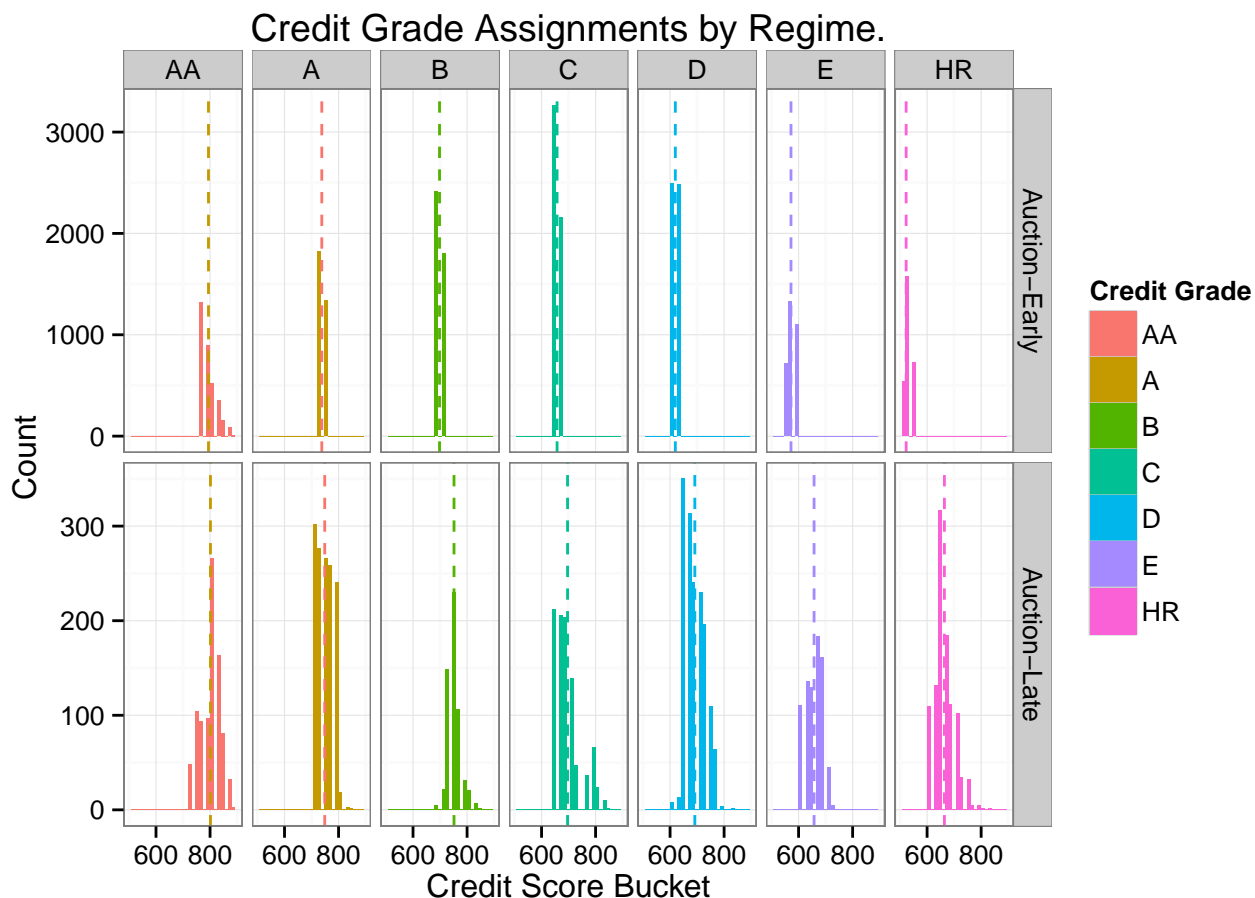
Figure 8: Another way to visualize the credit scores of borrowers in each grade. Dotted Lines indicate mean credit bucket for each grade.

In Figure 8, each $(row, column)$ entry is the distribution of borrower credit scores that were assigned grade $column$ during the row regime. For example the top row, leftmost entry presents the credit score buckets of the borrowers who were assigned grade AA in the early auction regime. In the early auction regime, all borrowers in a given credit bucket were assigned the same grade. Since all borrowers in the 640-680 buckets were assigned grade C, we can see two adjacent spikes, one for the 640-660 bucket, and another for the 660-680 bucket, in the top row of the grade C column of Figure 8.No other credit grades in the top row have any borrowers from the 640-680 buckets.

In the late auction regime, Prosper began using multiple variables to form its credit grade, and borrowers in each credit score bucket are spread across grades based on other credit variables. Correspondingly, in the bottom row of the grade C column of Figure 8, we can see that late auction regime, grade C borrowers came from 11 different credit buckets. When Prosper began aggregating hard credit variables into more informative grades, investors began using these grades to determine interest rates.

	Regime	Fit	Number of Loans
1	Auction-Early	0.48	27764
2	Auction-Late	0.87	7238
3	Prosper-rate	0.97	4550

Table 4: This table documents the predictability of interest rates in each of the three regimes in the Fit column. Predictability is measured as the Adjusted R squared of a regression with interest rate as the dependent variable and standard credit variables as independent variables, using every loan in the relevant regime.

	Regime	Fit	Number of Loans
1	Auction-Early	0.40	27764
2	Auction-Late	0.48	7238
3	Prosper-rate	0.39	4550

Table 5: This table documents how much of the variance in interest rates can be explained by borrowers' credit score buckets. Fit is measured as the Adjusted R squared with interest rate as the dependent variable and credit score bucket as the independent variable, using every loan in the relevant regime.

The tension between the importance of the credit score in the credit grade and the importance of the credit score in the interest rate is resolved by the fact that credit grades are much more predictive of interest rates in the later two regimes. As shown in Table 4, whose *Fit* column shows the fraction of the variance in interest rate that can be described by variance in hard credit variables, as measured by the adjusted r^2 of a regression with interest rate as the dependent variable and credit score bucket, current delinquencies and credit grade as independent variables. This regression yields adjusted r^2 of 48%, 88% and 97% for the early auction, late auction and Prosper-rate regimes, respectively, suggesting that Prosper's more informative grades are immediately used by borrowers to determine interest rates. [29]

Adversaries may argue that since we do not know Propser's exact grading formula, it may include information that was not previously available to borrowers.Table 5 addresses this argument, by running the same procedure as Table 4, but with only credit score bucket as an independent variable. The direction of the results is the same: Prosper's change to more informative grading did not only lead to a larger role for those grades in investors' interest rates, but increased the weight of credit score bucket, a variable that investors had access to in the same format in both periods. If the crowd of investors were as wise as (Hayek, 1945) theorizes, Prosper's repackaging of multiple variables into a more informative credit grade should not affect the weight of hard information in investors' risk estimates. Furthermore, Figure 7 makes is clear that investors would

---

[29]The full regression results are available in the appendix.

have experienced fewer defaults, and thus earned higher returns, if they had used credit buckets more heavily in their risk assessments, given the poor returns associated with low credit borrowers. We can thus interpret some of the poor returns of the early auction regime to a lower level of reliance on hard information.

In the next sub section, I examine Prosper's other interventions, which ban low-credit borrowers, and show that these interventions also similarly improved investor returns.

4.3   Impact of Low-Credit Borrower Ban
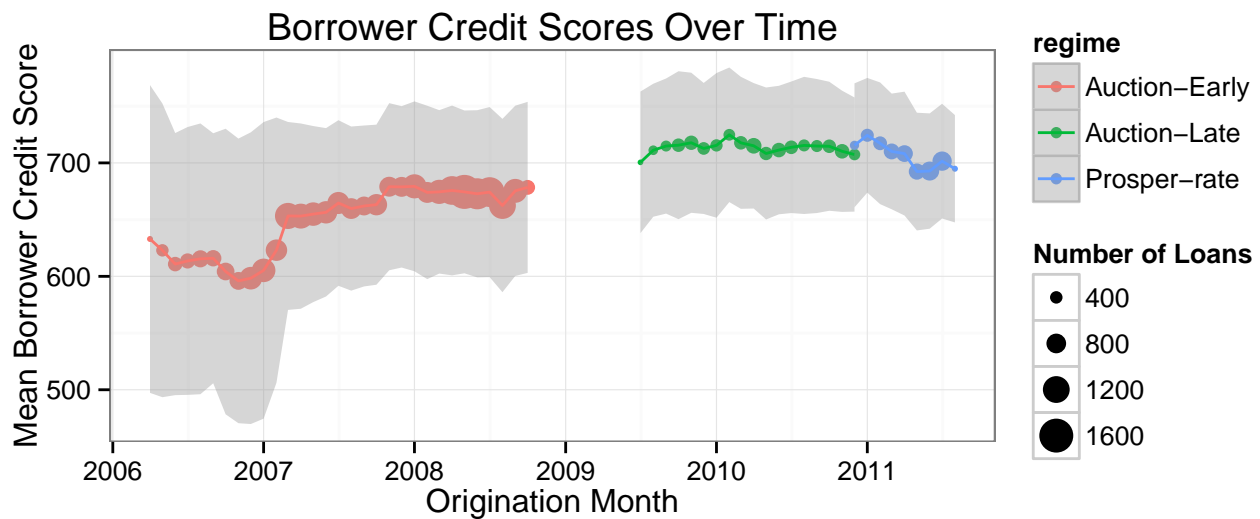
Figure 9: This figure presents average estimated credit score of loans originated each month. The average credit score increases from 653 in the early auction regime to 714 in the late auction regime, before returning to 705 in the Prosper-rate regime. Grey clouds indicate one standard deviation.

We now turn to Prosper's policy changes, which restrict the types of borrowers who investors are allowed to lend to. An auction perspective predicts that the crowd of lenders should be able to effectively screen all candidates, not just low risk candidates. On multiple occasions, however, Prosper has intervened to restrict the borrower pool to high credit borrowers, and thereby improved investor returns. Figure 9 shows the mean borrower credit score for each month in the sample, with the grey clouds showing one standard deviation bounds. Overall, for the three samples in full, the average borrower credit scores increases from 653.2 before the shutdown to 714.2 after the shutdown, then dips to 705 for the Prosper-rate period. For the first 10 months of Prosper's history, investors choose to lend to borrowers with an average estimated credit score of 609. Prosper's policy change, on March 1, 2007, that prevented loans to "low-credit borrowers," defined as those with credit score below 520, showed an immediate positive impact on investor decision making and returns.

Prosper's reasoning for the change was to protect investors: by March 1, 2007, 14.9% of loans to low-credit borrowers had already defaulted, more than triple the 4.6% of loans to other borrowers

that had already defaulted. [30] Theoretically, however, it is more difficult to understand why Prosper needed to step in. From the EMH perspective, theory predicts that investors would have reacted to the news of the defaults on low credit loans by decreasing their exposure to low credit loans, or increasing the interest rates they bid on them. However, the median interest rate on low-credit, actually decreased from 24.23% for the month of July, 2006 to 23.4% for February, 2007, the month before the intervention. Similarly, exposure to low credit loans did not substantially decrease. On December 1, 2006, 13.3% of these bad loans had already, after an average of only 3 monthly payments, gone into default, as compared to 3.7% of other loans. Yet over the month of December 2006, loans to borrowers with credit score below 520 accounted for 22% of orginations, the highest proportion in the site's history. The wisdom of the crowd theory expects market participants to adjust their beliefs to new information more rapidly and flexibly than a central planner. The data, however, show that investors continued to make avoidably bad decisions until Prosper disallowed them.

After the intervention, returns improved substantially. For the loans originated in February 2007, the month before the change, the mean borrower credit score was 620, with standard deviation 117, while the mean ROI on a loan was -13.1%. For loans originated in March 2007, the month after the change, the mean borrower credit score was 650, with standard deviation of 83, and mean ROI -7.4%. The 30 point month to month increase in mean borrower credit score is the largest in the sample, and the associated ROI change is not caused by time differences: the default rate to borrowers in the 640-660 bucket, a representative example, was 42.7% in February, and actually rose to 44.8% in March. Had Prosper enforced this rule from the site's inception, the default rate on loans originated before the rule change would have been reduced by 6%, and mean roi would have increased from -9.1% to -4%, a savings of roughly $2 million for investors.

A similar line of analysis can be applied to examine the impact of Prosper's second intervention. Upon re-opening after the shutdown, borrowers were required to have a credit score $\geq$ 640 and returning borrowers were required to have a credit score $\geq$ 600. This policy change is associated with a similar shift in average credit score from 678 in October, 2008 to 701 for July, 2009. The contemporaneous decrease in default rates makes it more difficult to explain the ROI jump exclusively through changes to the borrower population. But like in the previous intervention, low-credit loans continued until they were no longer allowed. By September 2008, the last full month before the shutdown, 19.8% of loans to borrowers in the 520-640 buckets had already defaulted, compared with 8.1% of loans to more creditworthy borrowers, yet 32% of originated loans for that month were to borrowers in the 520-640 range. August, 2009, the first full month after the rule change, only 7.5% of loans were to borrowers in that range, all of whom had borrowed with Prosper before. To reiterate, even with access to public information that showed that their lending strategies were not profitable, investors did not adjust. In the next section, we show that this lack of adjustment allows for the possibility of using public information to earn excess returns during the early auction period.

---

[30]the 10 % gap in default rates is even more striking when the average age of the two groups of loans is considered. On March 1, 2007, the loans to low-credit borrowers averaged 100 days since origination, while loans to other borrowers averaged 122 days since origination. 14.9% of low credit borrowers defaulted over an average time since origination of 100 days, whereas only 4.6% of other borrowers defaulted over an average time since origination of 122 days.

5 Excess Returns

5.1 Puzzle: Bad Returns in Early Auction regime

The key inconsistency with EMH found in the early auction regime data is that a simple loan selection strategy can earn excess returns using publicly available information. This section explains the methodological approach used to create such a strategy, including discussion of cross-validation and avoidance of look-ahead bias, and closes with a discussion of the excess returns earned by the strategy.

Although Figure 6 shows that investors charge higher rates to riskier borrowers, the high risk, low-reward data pattern of the early auction regime, shown in Figure 4 suggests that this preference for more credit-worthy borrowers is not extreme enough, since the extra interest charged to riskier borrowers does not, in aggregate, make up for the difference in default likelihood. This suggests that the variables given to investors by Prosper could be better used to assess borrower creditworthiness, and select loans with the lower likelihood of default. In this section, I use a regression model (that a naive investor might be able to use to choose loans) to show that a simple model to assess borrower creditworthiness can select loans, out of sample, that earn higher returns than the loans funded by investors. First I carefully discuss the methodology, which sets up the linear regressions that (1) can be tested by the mean ROI of the loans they choose out of sample, and thus will be penalized for overfitting, (2) are trained on data that would not have been available at the time of the decision (3) uses automated feature selection, so that I, with the benefit of hindsight, don't choose which variables will be included.

5.2 Methodology

5.2.1 Cross Validation Procedure

	Sample	Loans	Start	End	Days
1	Auction-Early.train	13841	2006-04-21	2007-10-01	528
2	Auction-Early.test	13923	2007-10-02	2008-10-16	380
3	Auction-Late-.train	3579	2009-07-20	2010-04-26	280
4	Auction-Late.test	3659	2010-04-27	2010-12-17	234
5	Prosper-rate.train	2265	2010-12-20	2011-04-28	129
6	Prosper-rate.test	2285	2011-04-29	2011-08-03	96

Table 6: This table documents number of loans and dates of the 6 samples used for training and testing regression strategies.

The regime discontinuity splits the data into early auction, late auction and Prosper-rate regimes. To test our predictions out of sample, we split each of these thirds in half. For those familiar with

cross validation, regressions are trained on the first half of loans originated under each regime, and tested on the second half. Table 6 describes the size and periods of the six sub samples. For those less familiar with cross validation, I now describe the six sub samples in words.

1. 13841 loans originated between April 21, 2006 and October 2nd, 2007, are used to generate coefficients for the early auction regime regressions.

2. 13923 loans originated after October 2nd, 2007, but before October 16, 2008, are used to test the predictions of the early auction regime regressions. [31]

3. 3579 loans originated between July 20, 2009 and April 26, 2010, are used to generate coefficients for the late auction regime regressions.

4. 3659 loans originated after April 26, 2010, but before December 17, 2010, are used to test the predictions of the late auction regime regressions.

5. 2265 loans originated between Dec 20,2010 and April 28,2011 are used to generate coefficients for the Prosper-rate regime regressions.

6. 2285 loans originated between April 29,2010 and August 3rd 2011 are used to generate coefficients for the Prosper-rate regime regressions.

5.2.2 Avoiding Look-ahead Bias with Known-Default

Unfortunately, since our time series is so short, the challenge of training models on the limited information a lender would have is considerable; the early Auction regime starts in April 2006, and ends in October 2008, and all loans last 36 months, so in October 2007, an investor only knows 18 months of payment results of the earliest origination. To address this challenge, we develop a binary Known-Default variable. For a given loan, Known-Default equals 1 if the ex-post sum of the payments of a loan is less than the percentage of 36 months that had elapsed by the end of the training period.

As an instructive example, suppose Bill borrows $100 in April 2006, and that the training period ends December 2007. If Bill ends up paying back $60 to investors over the full 36 months, his Known-default = 0 (even though he ended up defaulting) because the training period ends after 55.6% of his three years have elapsed, but he paid back 60% of his loan. If, however, Bill only ended up paying back $50 to investors, Known Default=1, because Bill MUST have already missed at least one payment by the end of the training sample.

Using Known-Default [32] as our left hand side variable instead of observed defaults or ROI allows us to restrict the information available to models to the same set of information that investors had at their disposal at the end of the training sample, namely borrower characteristics of each loan that has already been originated, and whether it has *already* gone into default. The downside to this approach is that for each period, many loans were originated in the month before the end of the training period, and almost none of these investors miss their first monthly payment. To avoid

---

[31]Between October 10, 2008 and July 20, 2009, no loans are originated, because Prosper is in the SEC-imposed Quiet Period

[32]It would be difficult to create known-ROI because there is no data on the magnitude of each monthly payment

these young loans, about whose performance almost nothing is known, from exerting a large impact on the model's coefficients, I weight observations by *observation maturity*, the amount of time that has elapsed between a loan's origination and the end of the training sample, in a weighted linear regression, using the training sample data.

5.2.3   Out of Sample Loan Selection

After training linear models to predict Known-Default, we use the coefficients to predict known-default as a proxy for real default, for each possible loan in the testing sample, and chooses the 10% of loans from each credit grade with the lowest predicted default likelihood. [33]

5.2.4   Summary of Procedure

For each regime . . .

1. divide each regime into training and testing samples of equal size, with all loans in the training sample originated before the first loan in the test sample was originated.

2. create variable Known-Default = (0||1) for each loan, depending on whether an investor would have known that the loan defaulted at the end of the training sample.

3. Start with 45 Features, select 8 that best predict Known-default.[34][35]

4. Run Logistic Regression to predict Known-Default using training period data and selected variables

5. Using trained coefficients, generate predicted Known-Default values for loans in the test period.

6. Select the 10% of Loans from each credit grade with the lowest predicted Known-default probability.

7. Compare ROI of selected loans against the benchmark of a strategy that chooses every loan.

5.2.5   Assumptions

Two assumptions underly the out-of sample loan selection technique which selects from the pool of loans that were funded by Prosper's investors, at the rate they received. First, in the auction regime, by sucessfully bidding on a loan, investors may decrease the interest rate. This analysis assumes that a strategy does not move the interest rates of the loans it selects. Second, a strategy never chooses to fund loans that will end up unfunded, since we cannot observe their ROIs.

---

[33]Loans are selected from each credit grade to ensure that strategies in the later periods can actually earn positive returns. Grade A loans in later periods return less than average.

[34]Thomas Lumley using Fortran code by Alan Miller. *leaps: regression subset selection*. R package version 2.9. 2009. URL: http://CRAN.R-project.org/package=leaps.

[35]Feature selection was performed by selecting the best combination of 8 features out of 45, based on in-sample performance (scoring with Bayesian Information Criterion).

5.3   In Sample Results

The Linear Regression results shown in Table 7 suggest that returns are much more predictable before the shutdown than after: the Adjusted R^2 for each regime's regression fall from 13.3171% for the early auction regime to 1.1529% for the late auction regime and 3.4981% for the prosper rate regime. The coefficients for all regressions are so close to zero because we are predicting Known-Default, which occurred in 11% of Auction Regime, training sample observations, but only 1.5% in the late auction regime training data and 1.0% of observations in the Prosper-rate regime data. The discrepancy in Known-default rates is caused by both a shorter period of training data and a lower actual default rate for the later regimes.

	Auction-Early	Auction-Late	Prosper-rate
(Intercept)	**0.0869*****	**0.0162*****	0.0059
	(0.0042)	(0.0046)	(0.0040)
BorrowerRate.sc	**0.0511*****		
	(0.0049)		
TotalCreditLinespast7years.sc	**−0.0187*****	**−0.0070***	
	(0.0037)	(0.0029)	
OpenRevolvingAccounts.sc	**−0.0189*****	−0.0014	−0.0028
	(0.0037)	(0.0031)	(0.0023)
InquiriesLast6Months	**0.0063*****		**0.0161*****
	(0.0007)		(0.0019)
CurrentDelinquencies.sc	**0.0455*****		
	(0.0025)		
DelinquenciesLast7Years.sc	**−0.0160*****		−0.0045
	(0.0027)		(0.0032)
MonthlyLoanPayment.sc	**0.0185*****		
	(0.0030)		
CreditScoreEst.sc	**−0.0347*****	−0.0068	−0.0045
	(0.0036)	(0.0038)	(0.0037)
TotalInquiries		**0.0022*****	**−0.0020***
		(0.0005)	(0.0008)
PublicRecordsLast10Years.sc		−0.0029	−0.0033
		(0.0039)	(0.0028)
DebtToIncomeRatio.sc		**0.0324****	
		(0.0119)	
TotalProsperPaymentsBilled.sc		**−0.0068****	
		(0.0024)	
ProsperPrincipalBorrowed.sc		0.0031	
		(0.0021)	
TotalProsperLoans.sc			−0.0027
			(0.0016)
StatedMonthlyIncome.sc			−0.0011
			(0.0016)
R^2	0.1337	0.0140	0.0384
Adj. R^2	0.1332	0.0115	0.0350
Num. obs.	13458	3189	2265

Y = Known-Default

Table 7: Coefficients extracted from Least Squares Regression to predict Known-Default in early auction, late auction, and Prosper-rate regime training data.

	Regime	BM.mean	BM.sd	BM.default	LM.mean	LM.sd	LM.default	Excess
1	Auction-Early	-0.001	0.031	0.323	0.029	0.032	0.211	0.030
2	Auction-Late	0.159	0.024	0.130	0.216	0.033	0.078	0.058
3	Prosper-rate	0.133	0.024	0.181	0.123	0.037	0.206	-0.009

Table 8: The mean return and standard deviation of return of the loans selected by the Linear Model (LM) compared with all loans (BM for benchmark). BM.default and LM.default are the default rates of the relevant selected loans. Excess returns equals LM.mean - BM.mean
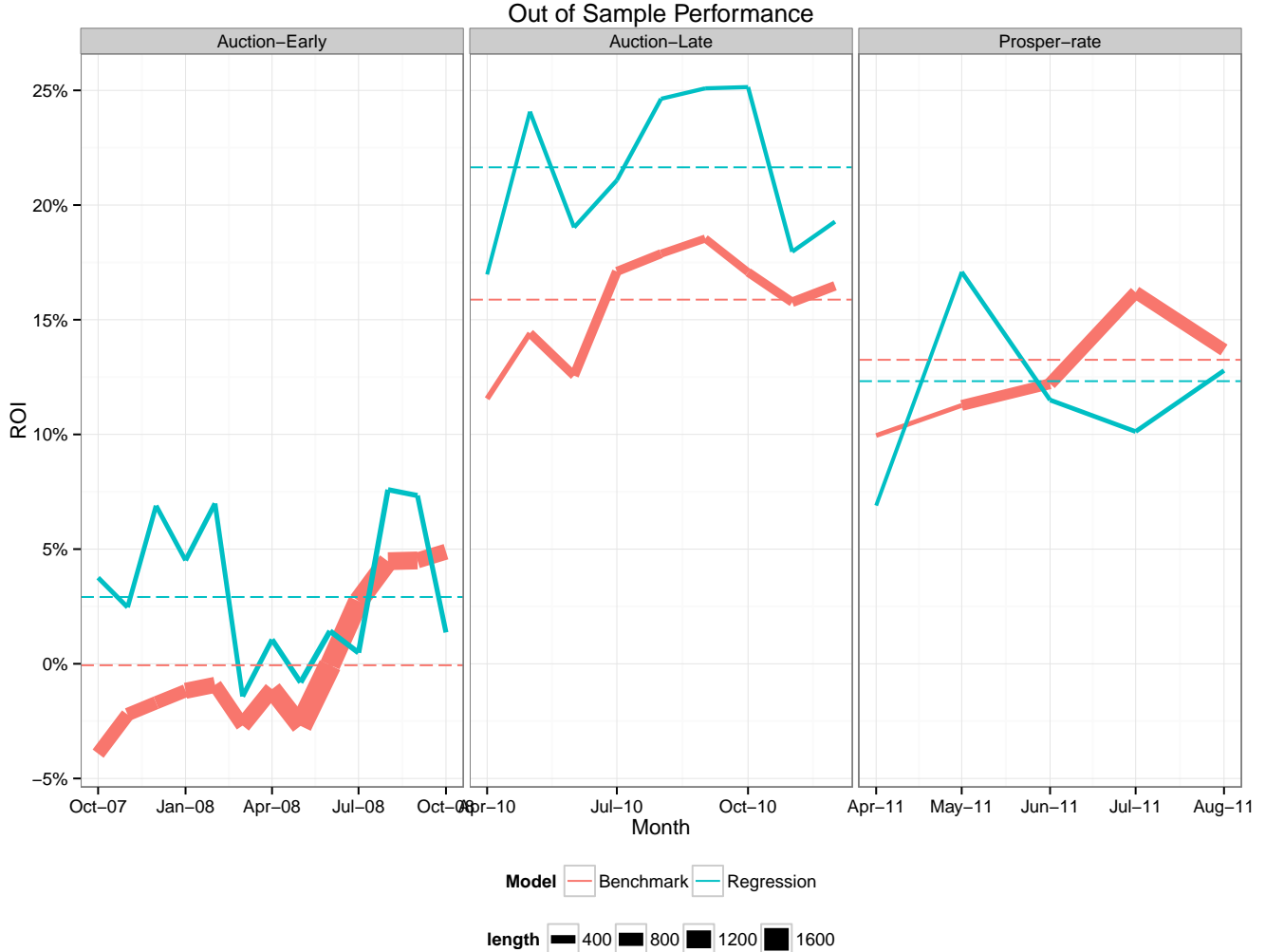
5.4 Out of Sample Results

Figure 10: Strategy's Out Of Sample Performance for Each Month. Loan Selection Strategies described in Methodology section.

Table 8 shows the mean and standard deviation of the returns of the loans selected by the regression strategy. For the early auction regime, when investors focused less on hard variables, set low rates, and after observing high default rates among similar borrowers, the model beats the benchmark by 3%, with similar variance[36]. Moreover, a T test comparing the 1434 loans selected by the early auction regime model and the 13,923 loans it could have selected, rejects the Null hypothesis that the two groups have the same mean ROI at the .1% significance level. This finding is inconsistent with the EMH prediction that it should be impossible to earn excess returns using publicly available information. As discussed in the methodology section, the model is trained on the observed defaults in the training sample, and chooses the loans its predicts to have the lowest probability of default. As such it ends up minimizing the riskiness of loans it selects, rather than trying to optimize a risk reward calculation. Early auction period loans selected by the model have a default rate of 21% , compared to the 32% default rate for all loans. Figure 10, which compares the various strategies out-of-sample performance over time, reveals that most of the model's outperformance is during the initial 6 months of testing.

For the late auction regime, when investors largely used the ratings prosper gave them to form interest rates, the model beats the benchmark by 5.8%, with slightly higher standard deviation. A T test comparing the 360 loans selected by the late auction regime model and the sample of 3569 loans it chose from rejects the hypothesis that the two groups have the same mean ROI at the 1% significance level. It is again clear that the superior returns of the loans chosen by the model were generated by a lower default rate: 7.8% compared to 13.0% for the full basket of loans. This result should interpreted carefully, however, given the noticeably higher variance and small number of selected loans.

For the Prosper-rate regime, when Prosper set interest rates, the model is unable to beat the benchmark: it chooses loans with worse ROI, higher default rate and identical returns.

In summary, over the inital two regimes where investors are involved in the process of setting interest rates, via auction, it is possible to use publicly available information to earn excess returns. In the final regime, where the central planner sets the interest rate, it is much more difficult, if not impossible, to earn excess returns using publicly available information.

5.5   Methodological Expansions

An expansion of this project could retrain the model every month with the benefit of newly observed defaults and newly originated loans, and expect performance improvements.

Engineering a known-roi dependent variable that equals the percentage of expected repayment of principal an investor received from each borrower up to the relevant date, would allow the model to make risk-reward calculations when deciding between loans, as opposed to just choosing the least risky loans.

---

[36]Variance is calculated as the standard deviation of monthly returns

6 Controlling for Time Period

The first two inconsistencies with EMH predictions: that early auction regime loans have negative average ROI, and that riskier early auction regime loans earn investors negative ROI, are both vulnerable to time period. Specifically, maybe the early period just happened to be bad for loan repayments. In this section, I aim to show that the decisions made by early investors, namely to lend to low credit borrowers and to demand far lower interest rates than hard variables would suggest, partly explain the difference in ROI. Thus, the question guiding this section is: **what would have happened to investor returns in the early auction regime if investors only lent to borrowers that would have qualified for loans in the late auction regime, and demanded the rates charged by investors in the late auction period?** This will help to understand whether the change in returns between the early and late auction regimes was merely a consequence of an unpredictably higher default rate.

6.1 Methodology

In order to estimate the hypothetical returns in an early auction regime where investors charged late auction regime rates to borrowers who would have met the late auction regime criteria, we must reduce early auction borrowers to the subset that would have been acceptable in the late auction regime, estimate what interest rate late auction regime investors would have charged them, and then use this estimated rate to predict a new ROI.

The first step, restricting the borrower pool, is simple. In the late auction regime, Prosper required that new borrowers have a credit score greater than 640, and that returning borrowers have a credit score greater than 600 in order to apply for a loan. Applying this rule reduces the number of early auction regime loans whose rate we will adjust by 40.4%, from 27,764 to 16,545. [37]

The second step of the method involves finding the hypothetical interest rate an early auction regime borrower would have paid in the late auction regime. For simplicity, this *adjusted rate* is computed by finding the median interest rate paid by late auction regime borrowers in the same credit score bucket, *EG* 700-720.

I then predict what the loan's ROI would have been if the borrower had paid that adjusted rate by multiplying the payments the borrower made under the old regime by the ratio of $\frac{(1+adjustedrate)}{(1+oldrate)}$.

As an instructional example, suppose Bill is an early auction regime borrower in the 700-720 credit score bucket. Under the early-auction regime, Bill borrowed \$100 from investors at a 13% interest rate, and returned \$113, over the one year course of his loan. [38]

In the late auction regime, the median borrower 700-720 borrower paid 23% interest, so, Bill's Adjusted-Rate is 23%. To calculate Bill's Adjusted ROI, we multiply Bill's Actual ROI by the ratio of the interest rates.

$$ADJ.ROI = \frac{Payments}{Amount.Borrowed} * \frac{ADJ.Rate}{Actual.Rate} \tag{2}$$

$$ADJ.ROI = \frac{113}{100} * \frac{1.23}{1.13} - 1 = 23\% \tag{3}$$

---

[37]This may not go far enough, as late auction regime investors refused to fund many borrowers applications.

[38]Neither maturity nor principal matter, for our purposes

Suppose, alternatively, Bill had repaid only $45. So Bill's actual ROI was -55%. In this second case, Bill's Adjusted ROI $= \frac{45}{100} * \frac{1.23}{1.13} - 1 = -51\%$.

6.1.1 Assumptions

Three assumptions facilitate this methodology, but also cause bias. Most importantly, the methodology ignores the idea of cost-induced default, as discussed by (Stiglitz and Weiss, 1976): first, forcing the borrower to pay more per month might push him over the default tipping point.[39] Empirical validation of this theory can be found in (Wei and Lin, 2013)'s examination of Prosper's transition from the late auction regime to the Prosper-rate regime, which shows that when borrowers were assigned higher interest rates, they defaulted more frequently. With this in mind, my adjusted ROI serves as an **upper bound** for the possible ROI because it assumes that each borrower makes the same number of payments, even though each payment is higher. More realistically, we might expect Bill to make fewer payments on his loan if he found it difficult to afford his interest rate.

Two other assumptions work in the other direction, however. First, we have assumed that all applications that satisfied Prosper's minimum borrower standards would have been funded by investors. In reality, many applications met the criteria but were not funded. Secondly, we have assumed that in our hypothetical world early auction regime borrowers would pay the same rate as late auction regime borrowers in their credit bucket. In reality, however, they would probably pay higher rates, because controlling for credit score bucket, they have worse average delinquencies, credit card utilization and other hard variables that would be used for grading.

To summarize, the total.ROI reported using the adjusted borrower population is probably a little lower than it should be, and the total.ROI reported using adjusted interest rates is probably higher than it should be.

---

[39]Joseph E Stiglitz and Andrew Weiss. "Credit rationing in markets with imperfect information". In: *The American economic review* (1981), pp. 393–410.
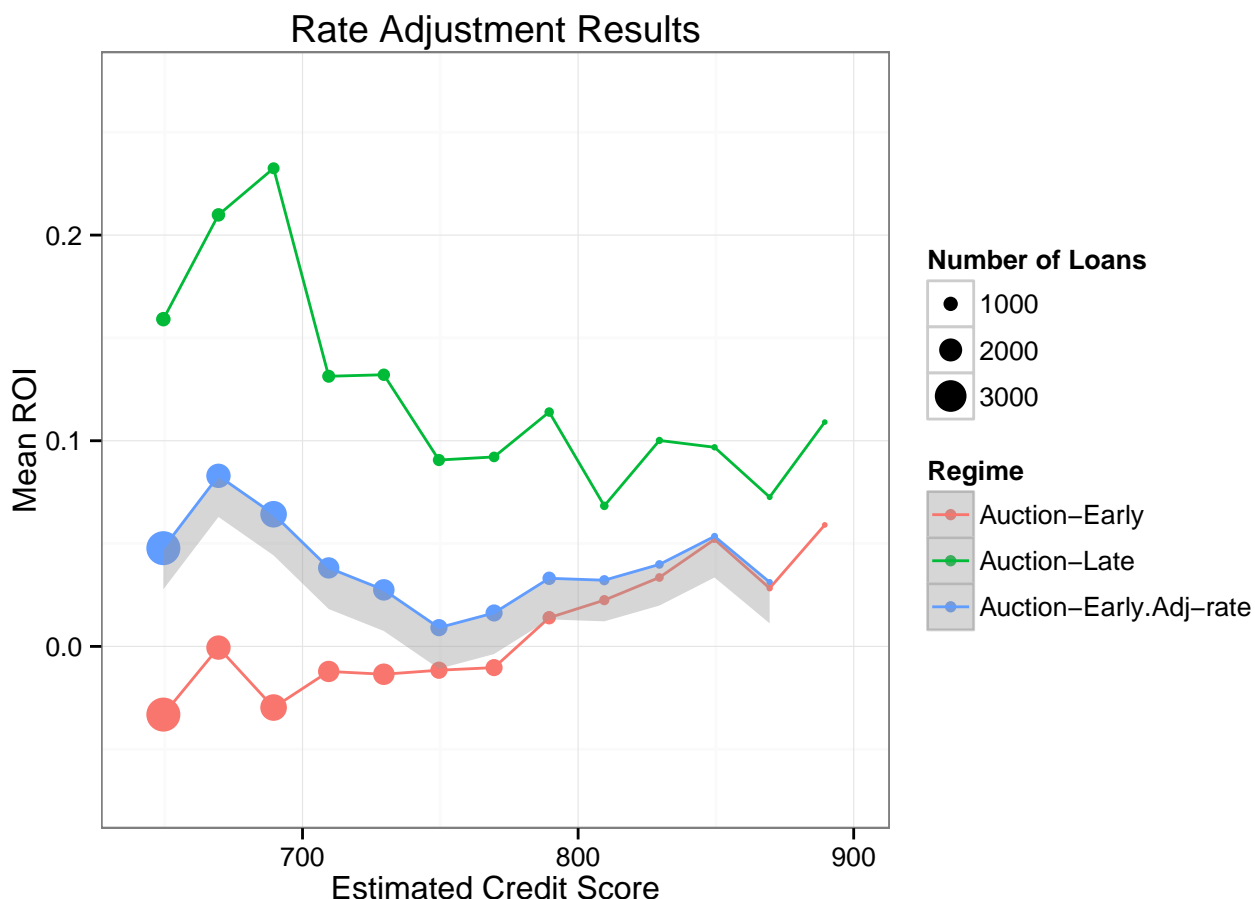
6.2   Results & Discussion

Figure 11: This chart displays the results of our rate-adjustment experiment. The blue line is the result of our rate-adjustment experiment, and displays the hypothetical ROI if late auction regime borrowers had paid early auction interest rates.

Figure 11 documents the relationship between credit score bucket and returns for early auction regime, late auction regime, and late auction regime borrowers with adjusted rates, where the rate adjustment and subsequent ROI adjustment was computed using the *upper bound* methodology described above. At low credit scores, where the difference in interest rates between the two regimes is largest, the gap in ROI is also the largest.[40] With adjusted rates, the inconsistency with EMH, where riskier assets had lower average return, is reversed, as well as the negative average return inconsistency. With adjusted interest rates, early auction regime borrowers earn higher average returnson riskier loans, and positive average returns overall.

---

[40]The grey cloud is 2% error, in the right direction. I could improve on this by trying to quantify the effect of a 1% increase in rate for similar borrowers on default.

	regime	total.ROI
1	Auction-Early	-0.0383
2	Auction-Early.Adj-borrowers	-0.0136
3	Auction-Early.Adj-rate	0.0439
4	Auction-Late	0.1520

Table 9: The aggregated results of the Rate Adjustment Experiment.

At each credit score, an upper bound estimated of the fraction of the ROI gap that would be accounted for by adjusted interest rates can be measured as:

$$rate.factor = \frac{mean(adj.roi) - mean(early.roi)}{mean(late.roi) - mean(early.roi)} \tag{4}$$

[41] This *rate factor* is, by construction, heavily correlated with the difference in interest rates for each regime. At the lowest included estimated credit score, 649.5, the difference in average interest rate is 9.3%, and that difference explains 42.1% of the 19.2% difference in average ROI. At the higher buckets, the difference in interest rates for borrowers in the 800-820 credit score bucket was 0.3%, and this difference explained only 3% (not half) of the 6% difference in average ROI.

The aggregated returns of the regimes, as shown in Table 9, indicate that if early auction regime borrowers had self enforced the minimum standards Prosper introduced in the second period, investor returns would have improved by at least 2.8%. If investors had charged these borrowers the rates similar borrowers were charged in the late auction regime, returns might have improved by another 5.75%.

These results suggest that although much of the difference in performance between regimes is due to an especially bad time period, roughly $\frac{1}{3}$ of the difference in returns between the early auction and late auction regimes can be explained by poor investor decisions.

7   Conclusion (unfinished)

EMH predicts that (1) risky assets should have positive expected returns, (2) riskier assets should have higher expected return than less risky assets and (3) it should be impossible to use publicly available information to earn excess returns. The previous sections document evidence that rejects all three of these predictions, as well as evidence that the inconsitencies with EMH were diminished and then removed as Prosper took more control of the rate-setting process. I conclude by offering alternative framings of the central planner as a market guiding, rather than disrupting force.

---

[41]Visually, think of it as if we are walking up the plot from the red line to green line, what percentage of the way would we make it before we encountered the blue line.

References

[1] Friedrich A. Hayek. *The Use of Knowledge in Society*. Library of Economics and Liberty, 1945. URL: http://www.econlib.org/library/Essays/hykKnw1.html.

[2] William F Sharpe. "Capital asset prices: A theory of market equilibrium under conditions of risk*". In: *The journal of finance* 19.3 (1964), pp. 425–442.

[3] Eugene F. Fama. "Efficient Capital Markets: A Review of Theory and Empirical Work". In: *The Journal of Finance* 25.2 (1970), pp. 383–417. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1970.tb00518.x. URL: http://dx.doi.org/10.1111/j.1540-6261.1970.tb00518.x.

[4] Sanford J. Grossman and Joseph E. Stiglitz. "On the Impossibility of Informationally Efficient Markets". In: Reprint Series 121 (Dec. 1980). URL: http://www.nber.org/papers/r0121.

[5] Joseph E Stiglitz and Andrew Weiss. "Credit rationing in markets with imperfect information". In: *The American economic review* (1981), pp. 393–410.

[6] Andrei Shleifer. *Inefficient Markets*. 2000.

[7] Thomas Lumley using Fortran code by Alan Miller. *leaps: regression subset selection*. R package version 2.9. 2009. URL: http://CRAN.R-project.org/package=leaps.

[8] Rajkamal Iyer et al. "Screening in New Credit Markets: Can Individual Lenders Infer Borrower Creditworthiness in Peer-to-Peer Lending?" In: rwp09-031 (Aug. 2009). URL: http://ideas.repec.org/p/ecl/harjfk/rwp09-031.html.

[9] Qiang Gao and Mingfeng Lin. "Linguistic Features and Peer-to-Peer Loan Quality: A Machine Learning Approach". In: *Available at SSRN 2446114* (2013).

[10] Samuel G Hanson et al. *Banks as Patient Fixed-Income Investors*. Tech. rep. National Bureau of Economic Research, 2014.

[11] Adair Morse. "Peer-to-Peer Crowdfunding: Information and the Potential for Disruption in Consumer Lending". In: Working Paper Series 20899 (Jan. 2015). DOI: 10.3386/w20899. URL: http://www.nber.org/papers/w20899.