Journal of Business and Economics, ISSN 2155-7950, USA
December 2010, Volume 1, No. 1

http://www.academicstar.us

The Prosper Credit Risk Rating System: Does It Improve Market Decision Making?

Christina J. Bradbury
(College of Business Administration, Plymouth State University, Plymouth 03264-1595, New Hampshire, USA)

Abstract: Anonymous online interaction presents a new challenge in peer-to-peer consumer credit markets: effectively predicting and screening risk. Prosper has implemented new policies to minimize its disadvantage in information access. These are improved information transparency and the introduction of the "Prosper Credit Rating System". Therefore, the objective of this research endeavor is to examine the credit risk rating system newly introduced into Prosper's peer-to-peer loan marketplace to see if it leads to improved market decision making as it relates to pricing. The proprietary credit rating system was specifically designed to evaluate and rank borrower credit risk associated with their loan request listings so as to support lenders in their pricing decision. It is untested, only introduced in July 2009. The sample covers 2,525 funded loan listings. An analysis of covariance was employed to test whether the newly created Prosper credit rating system has an impact on the outcome of interest rate after removing the variance for which other variables (dollar amount of loan request, Fair Isaac credit score range, homeownership status, and debt-to-income ratio) may account. Research findings clearly indicate that six of the seven Prosper credit ratings have a statistically significant impact on the dependent variable, individual loan interest rates.

Key words: credit risk evaluation; information access; peer-to-peer loan marketplace; prosper.com
JEL Codes: D81, G21

1. Introduction

Prosper Marketplace, Inc., known as Prosper, matches people who need small loans with others who have cash to lend. An online financial marketplace, Prosper enables individual lenders to locate individual borrowers and vice-versa. Disintermediation having found its way to the unsecured consumer loan industry, Prosper seeks to act as the middleman for all transactions between lenders and borrowers, taking a small fee for its efforts. Prosper's responsibilities range from collecting money and distributing it pro rata to each lender on a given loan, to spot information verification before money is disbursed on a loan, to sending delinquent loans to collection agencies. It differs from other online peer-to-peer financial conduits, such as Lending Club, in its operational style as it offers its customers a blend of ebay-like peer-to-peer loan auctions supplemented with dozens of high-traffic discussion forums (Andrews, Dholakia, Herzenstein, & Lyandres, 2008).

Screening borrowers and efficient allocation of credit based on creditworthiness of the borrower is an important function of the credit market (Iyer, Khwaja, Lettmer & Shue, 2009). Peer-to-peer (P2P) online lending

Christina J. Bradbury, Ph.D., College of Business Administration, Plymouth State University; research areas: finance accounting. E-mail: cjbradbury@plymouth.edu.

is a new platform for the consumer credit market. The uncollateralized nature of lending in these online markets makes it particularly attractive for small borrowers who might otherwise turn to payday lenders or credit card debt, often at exorbitant rates. The first P2P lending network in the US, going live to the public in February 2006, Prosper.com has originated just over 36,000 loans for a total in excess of $210 million. The quick expansion of Prosper coincided with a number of similar new P2P lending sites in the US such as Kiva and Lending Club. Nevertheless, Prosper remains the leader in the United States P2P lending business with about 80% of total loan volume (Bogoslaw, 2009).

2. Research design

2.1 Theoretical framework

The P2P lending platform is not yet five years old and as such there is limited empirical research specific to the peer-to-peer credit market. While Prosper lenders face the traditional information imperfection in assessing borrower risk, anonymous online interaction presents new challenges. For instance, individual lenders, by definition smaller and less professional than financial institutions, may not have the expertise to predict and screen risks. Using transaction data from June 2006 to July 2008, Freedman and Jin (2008) present empirical evidence that suggests the presence of adverse selection due to information problems on Prosper.

Aware of these issues, Prosper has implemented new policies to minimize its disadvantage in information access. The first was improved information transparency. Prosper posts all of its up-to-date activities; from listing to loan performance, on its website. The second, and the focus of this paper, was the introduction of the “Prosper Credit Rating System”, a proprietary credit risk rating system specifically designed to evaluate and rank borrower credit risk associated with their loan request listings. In July 2009, following an eight month moratorium associated with obtaining SEC approval for the sale of loans in the secondary market, Prosper rolled out its new proprietary credit rating system. Now, each loan listing on Prosper is assigned a “Prosper Rating”, a proprietary credit rating that according to their prospectus “allows one to easily analyze a listing’s level of risk because the rating represents an estimated average annualized loss rate range” (US SEC S-1 Amendment 6, 07/13/2009).

The Prosper Rating system consists of seven ratings: AA, A, B, C, D, E and HR corresponding to lowest to highest credit risk. According to their prospectus, this allows Prosper to “maintain consistency when assigning a rating to the listing” (US SEC S-1 Amendment 6). The underlying determinates consist of two scores. The first is the credit score from the Experian credit reporting agency. The minimum threshold for new a borrower is currently 640 on a FICO scoring scale of roughly 400 to 900. The second is a Prosper Score of which is based upon a custom risk model using Prosper data to predict the probability of a loan going “bad” where “bad” is 61 days or more past due. The specifics behind this in-house scorecard have not been released, making it difficult to analyze. It is, however, disclosed on their website that loans booked from April, 2007 through October, 2008 were used to build the model, with the performance measured for the following fifteen months. Moreover, it is also revealed that key variables in the scorecard are number of trades, number of delinquent accounts, number of inquires, number of recently opened trades, amount of credit available on bankcards and bankcard utilization. The Prosper score was built specifically on the Prosper population, so it incorporates behavior that is unique and inherent to this population (US SEC S-1 Amendment 6).

The FICO credit score obtained from a credit reporting agency is based on a much broader population, of which Prosper borrowers are just a small subset. As such, Prosper asserts that “the credit reporting agency score should, and does, rank order risk on the Prosper population, but is not as discriminating as a custom score" (Prosper.com). The Prosper Rating encompasses both the custom score and the credit reporting agency score together to assess the borrower's level of risk and determine estimated loss rates. It is contended that the Prosper rating is more powerful than using just the FICO score.

Given the purpose for which the Prosper risk rating system was created and its very recent introduction into the Prosper market, the purpose of this research endeavor will be to examine the Prosper credit risk rating system to see if it leads to improved market decision making as it relates to supporting lenders in their pricing decision. It is untested, having only been introduced in July of 2009. It is reasonable to discern that average loan rates may be influenced by the Prosper rating should the market deem that this risk rating system adds noteworthy value; providing additional information beyond that which was already available.

Credit risk scoring overall has become widely used in consumer lending, ranging from issuing credit cards, to making automobile loans and home equity financing. It is a method of evaluating the credit risk of loan applications. Using historical data and statistical techniques, credit scoring attempts to isolate the effects of various applicant characteristics on delinquencies and defaults (Mester, 1997). The method produces a score or a rating if you will that can be used to rank loan applicants or borrowers in terms of risk.

A credit risk scoring system serves to support the attainment of optimal credit making decisions. According to Blochlinger and Leippold (2005), the basic utilization of a credit risk scoring system is to provide guidance for lending cutoffs, supporting an "accept" or "reject" decision, and/or to support risk-adjusted pricing on the loan. Risk-adjusted pricing may be interpreted as the pricing that is based on the obligator's creditworthiness. In a pricing regime, the bank sets the price of a loan according to the credit score or rating.

Prosper does not use its credit risk rating system to accept or reject loan listings. That decision is reached via the mechanism of the online P2P bidding platform where lenders and prospective borrowers bid openly. This is an operating style unique to Prosper. Rather, Prosper uses its credit risk rating system to support lenders in their pricing decision or otherwise said to support lenders to make sound final bids on loan listings.

2.2 Hypothesis and research variables

Given the aforementioned, the focus of this research is to determine whether the Prosper credit rating system, consisting of seven specific ratings, provides any value to the lender with respect to interest rate pricing. To test this, it is the writer's null hypothesis that the Prosper credit ratings will have no impact on individual Prosper loan interest rates, more specifically that not one of the seven credit ratings has any impact on interest rates. Otherwise written, Ho: $\tau_{\text{ProsperRating}\mathbf{AA}} = \tau_{\text{ProsperRating}\mathbf{A}} = \tau_{\text{ProsperRating}\mathbf{B}} = \tau_{\text{ProsperRating}\mathbf{C}} = \tau_{\text{ProsperRating}\mathbf{D}} = \tau_{\text{ProsperRating}\mathbf{E}} = \tau_{\text{ProsperRating}\mathbf{HR}} = 0$. It is the writer's alternative hypothesis that at least one of the seven Prosper credit ratings will significantly impact interest rates on Prosper loans; Ha: at least one ($\tau_{\text{ProsperRating}\mathbf{AA}}$, $\tau_{\text{ProsperRating}\mathbf{A}}$, $\tau_{\text{ProsperRating}\mathbf{B}}$, $\tau_{\text{ProsperRating}\mathbf{C}}$, $\tau_{\text{ProsperRating}\mathbf{D}}$, $\tau_{\text{ProsperRating}\mathbf{E}}$, $\tau_{\text{ProsperRating}\mathbf{HR}}$) $\neq 0$.

The dependent variable being measured in this framework is the stated interest rate on loans receiving funding, denoted as $Y_{\text{interestrate}}$. Explanatory variables include the dollar amount of the loan request ($X_{\text{loanamount}}$), Fair Isaac credit score range ($\gamma_{\text{FICOrange}}$), borrower's homeownership status ($\alpha_{\text{homeowner}}$), borrower's debt-to-income ratio ($X_{\text{debt-to-income}}$) and Prosper credit risk rating ($\tau_{\text{ProsperRating}}$). Of the six explanatory variables, it is acknowledged that two are numeric, $X_{\text{loanamount}}$ & $X_{\text{debt-to-income}}$, one is nominal, $\alpha_{\text{homeowner}}$, and two are ordinal, $\gamma_{\text{FICOrange}}$ & $\tau_{\text{ProsperRating}}$. The "X" designation represents scale data while "α, γ and τ" represents categorical. The dependent variable, interest rate, is expressed as a decimal such that 0.20 equates to a 20% interest rate. Loan amount is expressed as a dollar figure, ranging from $1,000 as a low upward to $25,000 at a maximum. Debt-to-income is represented as a decimal taken to the hundredths place, ranging from a low of 0.00 to a high of 10.10. The variable measuring for homeownership is either true or false. The Fair Isaac credit scores include the following five ordinal categories ranging from best to worst: 770-900, 730-769, 700-729, 680-699 and 640-679. Prosper prohibits prospective borrowers with a FICO credit score below 640 from listing their request. Lastly and aforementioned, the Prosper Rating system consists of seven ratings: AA, A, B, C, D, E and HR corresponding to lowest ("best") to highest ("worst") credit risk.

3. Research methodology

3.1 Data employed

The data employed in this study came directly from Prosper itself. Prosper prides itself on making its market fully transparent and freely available. One can access Prosper's public site data in a downloadable format on the developer tools and data mining resources page. This is found on their website, www.prosper.com. Historical data can be downloaded in a raw XML format which can then be imported into a data management program or converted into a CSV file and read on Excel. The data provided includes but is not limited to listings, bids, user profiles, groups and all loans ever created on Prosper. This snapshot is updated regularly.

As the proprietary Prosper credit rating system was introduced in July 2009, the dataset spans July 20, 2009 through January 28, 2010. The data was downloaded on January 29, 2010. Cumulative data is being collected for an extended timeframe under which the Prosper system will be reexamined at a future date.

The dataset contains all variables displayed on a borrower's fully funded loan listing as per the publically available Prosper data export file with exception of the borrower's loan repayment status. The reason for the omission of loan repayment status as a variable is the short and recent nature of time period examined. By definition for a loan to be considered "in default", it takes 120 days of non-payment from the loan origination date (US SEC S-1 Amendment 6). In other words, it takes four months for a loan to season to where it could even be possibly observed as a default. Within the six-month dataset observed only the months of July and August would satisfy this requirement.

Other limitations include the fact that Prosper does not provide for all extended credit information on the public data download, including but not limited to borrower income. Prosper was contacted for this additional data but was informed, as is indicated on the website, that disclosure thereof is "explicitly forbidden" on the individual member (Prosper.com). Be that as it may, specific to borrower income, the writer considers the integrity of this information to be questionable. This is for reason that it is self-reported and no income verification is performed on the part of Prosper given that no tax return, W-2 or bank statement requirements exist. Furthermore, obtaining a precise income figure would not be possible as the borrower is directed to report their income in increments of $25,000, options including the likes of $0 - $24,999 or $25,000-$49,999, etcetera.

The sample initially covered 2,527 funded loan listings. Table 1 provides summary statistics of variables used in our analysis, the dependant variable, $Y_{interestrate}$, and five independent explanatory variables: $X_{loanamount}$, $X_{debt\text{-}to\text{-}income}$, $\alpha_{homeowner}$, $\gamma_{FICOrange}$ and $\tau_{ProsperRating}$.

3.2 Methodology

An analysis of covariance (ANCOVA) was employed to test whether the newly created Prosper credit risk rating system, $\tau_{ProsperRating}$, has an impact on the outcome of $Y_{interestrate}$ after removing the variance for which the other four variables denoted as $X_{loanamount}$, $X_{debt\text{-}to\text{-}income}$, $\alpha_{homeowner}$ and $\gamma_{FICOrange}$ account. The ANCOVA model was selected for the analysis for reason that the Prosper rating is an ordinal, categorical variable. A merger of an analysis of variance and a general linear regression model for continuous variables, the analysis of covariance can provide for an increase in statistical power because it accounts for some of the variability (Babin, et al., 2006). Alternate tests for robustness were further incorporated, evaluating the change in model predicting power when excluding and including the Prosper rating variable as well as repeating some of the main analysis exclusive of outliers outside the 0.01 significance level.

Inferential statistics, specifically correlation analyses, F-test, t-test and partial eta squared statistics were reviewed. These statistics provide information on which to evaluate statistical significance and strength of the linear dependence (between dependent and explanatory variables). Providing for a high degree of assurance, a 99% confidence level was used on individual results. Descriptive statistics including tests of normality were further reviewed.

Table 1 Summary statistics

Variable		Measurement	mean	Standard deviation	Valid values	Count	Percent (%)
Loan interest rate	$Y_{interestrate}$	Scale	0.189848	0.0897005			
Loan amount	$X_{loanamount}$	Scale	4900.28	4412.663			
Debt-to-income ratio	$X_{debt\text{-}to\text{-}income}$	Scale	0.22795	0.31203			
Homeownership status	$\alpha_{homeowner}$	Nominal					
					False	1123	44.5
					True	1402	55.5
FICO range	$\gamma_{FICOrange}$	Ordinal					
					770-900	720	28.6
					730-769	468	18.5
					700-729	533	21.1
					680-699	626	24.8
					640-679	178	7.0
Prosper rating	$\tau_{ProsperRating}$	Ordinal					
					AA	413	16.3
					A	555	21.8
					B	143	5.6
					C	547	21.5
					D	442	17.4
					E	210	8.4
					HR	215	9.0

4. Results

Results initially examined included graphical descriptive illustrations of the variables examined herein. The measure of association between the independent explanatory variables and the dependent were further analyzed.

A logarithmic transformation was made to the numerical independent variable $X_{loanamount.}$ Squeezing together larger values in the data set and stretching out smaller values, $X_{log(loanamount)}$ provided for a much more normal distribution given demonstrated data skewing to the right.

Review of the data of the other numerical independent variable, $X_{debt\text{-}to\text{-}income,}$ indicates two loans booked with a staggering reported debt-to-income ratio of 1010%! While not disclosed on Prosper's website or within their publicly available data download just how this ratio is arrived at, it is probable that these two outliers are likely representative of input error. This conclusion is supported by the summary statistical data indicating a mean debt-to-income ratio of 22.8%. For this reason, these two funded loan listings were excluded from the dataset. The final sample encompasses 2,525 funded loan listings.

There is a statistically significant lack of independence between $\gamma_{FICOrange}$ and $Y_{interestrate.}$ The greatest number of observations was found for the middle FICO range, lessening on both ends where the FICO ranges improved upward and deteriorated downward. Overall, the interest rate declined as the FICO credit score ranges improved. Providing a measure of the strength of dependence between two variables where the independent is nonparametric, Kendall's tau-b rank correlation of -0.517 indicates a moderately strong inverse relationship between $\gamma_{FICOrange}$ and $Y_{interestrate.}$ This statistical information can be found in Table 2. As for $\tau_{ProsperRating}$ and the dependant variable, much of the same conclusion can be reached only the relationship is undeniably stronger. The Kendall's tau-b rank statistic measuring association between the Prosper ratings explanatory variable and loan interest rates is -0.793 (Table 2). As the Prosper credit rating declines, moving down from "AA" to "A" to "B" to "C" and so forth, this means that the borrower is judged less creditworthy or otherwise said a greater credit risk. Therefore, as is intuitive, the interest rate pricing of the loan increases as the Prosper rating worsens. It is further useful to assess the significance of concordance for the two risk rating variables, FICO and the Prosper rating, to evaluate the level of redundancy. Given that the Prosper Rating encompasses both one's FICO score range as well as a custom Prosper score unique to its P2P population to assess a borrower's level of risk, a positive, strong relationship was expected. There is indeed a high degree of concordance between these two as exemplified by a +0.641 (Table 2), albeit not all information is redundant.

Table 2 Measure of association–Kendall's tau-b

	Kendall's tau-b	Approx. sig.
Interest rate and FICO range	-0.517	0.000
Interest rate and prosper rating	-0.793	0.000
FICO range and prosper rating	0.641	0.000

With a better understanding of the variables under examination and some adjustment made to help minimize error, a general linear regression model was run. The relationship is represented as follows: $Y_{interestrate} = ß_0 + ß_1 X_{log(loanamount)} + ß_2 X_{debt\text{-}to\text{-}income} + \alpha_i + \gamma_j + \tau_k + \varepsilon$, where $ß_0$ is the intercept, α_i is the i^{th} level of homeowner, γ_j is the j^{th} level of FICO range and τ_k is the k^{th} level of Prosper Rating and ε is the error term.

Adjusting for the number of explanatory terms in a model, the adjusted R-squared statistic was reported to be 0.864. (See Table 3) These results are exceptionally good. What is more, the $\tau_{ProsperRating}$ variable accounts for the greater part of the explanatory power of the dependant variable. Evidenced by the partial eta squared stat., 74.2% of the model's explanatory power came from this variable! Far more practically significant that any of the other variables, it is without a doubt the primary driver in this model to effectively determine interest rate.

Table 3 Analysis of covariance–original model

$Y_{interestrate} = ß_0 + ß_1 X_{log(loanamount)} + ß_2 X_{debt\text{-}to\text{-}income} + \alpha_i + \gamma_j + \tau_k + \varepsilon$

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial eta squared
Corrected model	17.532[a]	13	1.349	1235.377	0.000	0.865
Intercept	0.256	1	0.256	234.847	0.000	0.086
Prosper rating	7.877	6	1.313	1202.551	0.000	0.742
Homeownership status	0.008	1	0.008	6.946	0.008	0.003
FICO range	0.025	4	0.006	5.627	0.000	0.009
logLoanAmount	0.125	1	0.125	114.855	0.000	0.044
Debt-to-income	0.025	1	0.025	22.987	0.000	0.009
Error	2.741	2511	0.001			
Total	111.459	2525				
Corrected total	20.273	2524				

Note: [a]. R Squared = 0.865 (Adjusted R Squared = 0.864)

Parameter estimates–99% confidence interval (Dependent variable: Interest rate)

Parameter	B	Std. error	t	Sig.	99% Confidence interval		Partial eta squared
					Lower bound	Upper bound	
Intercept	0.235	0.009	25.894	0.000	0.212	0.259	0.211
[Prosper rating=1]	-0.236	0.004	-66.224	0.000	-0.245	-0.227	0.636
[Prosper rating=2]	-0.216	0.003	-68.843	0.000	-0.224	-0.208	0.654
[Prosper rating=3]	-0.178	0.004	-42.799	0.000	-0.189	-0.167	0.422
[Prosper rating=4]	-0.116	0.003	-41.207	0.000	-0.123	-0.109	0.403
[Prosper rating=5]	-0.059	0.003	-20.493	0.000	-0.067	-0.052	0.143
[Prosper rating=6]	0.009	0.003	2.641	0.008	0.000	-0.018	0.003
[Prosper rating=7]	0[a]	.	.	.	.	.	.
[Homeownership status=1]	-0.004	0.001	-2.635	0.008	-0.007	-8.091E-5	0.003
[Homeownership status=2]	0[a]	.	.	.	.	.	.
[FICO range=1]	-0.003	0.002	1.419	0.156	-0.003	0.009	0.001
[FICO range=2]	-0.002	0.003	0.690	0.490	-0.005	0.008	0.000
[FICO range=3]	0.001	0.003	-0.206	0.837	-0.008	0.007	0.000
[FICO range=4]	0.015	0.004	-3.396	0.001	-0.026	-0.004	0.005
[FICO range=5]	0[a]	.	.	.	.	.	.
logLoanAmount	0.010	0.001	10.717	0.000	0.008	0.013	0.044
Debt-to-income	0.022	0.005	4.794	0.000	0.010	0.033	0.009

Note: [a]. This parameter is set to zero because it is redundant.

Before further exploring the significance of the Prosper Rating, it would be prudent to examine the each of the variables under consideration to determine significance. As evidenced in Table 3 by p-values of 0.000 in all cases except one, $\alpha_{homeowner}$ where the p-value stands at just 0.008, there is no doubt that all variables in the model under consideration are statistically significant.

The significance of the Prosper Rating as the driving explanatory variable in the model is reinforced by an examination of the coefficients. Results can be found in Table 3. The relationship $Y_{interestrate} = ß_0 + ß_1 X_{log(loanamount)} + ß_2 X_{debt\text{-}to\text{-}income} + \alpha_i + \gamma_j + \tau_k + \varepsilon$ can now be represented as follows: $Y_{interestrate} = .235 + 0.010\, X_{log(loanamount)} + 0.022X_{debt\text{-}to\text{-}income} - 0.004\alpha_{homeowner\mathbf{FALSE}} - 0.003\gamma_{FICOrange\mathbf{730\text{-}769}} - 0.002\gamma_{FICOrange\mathbf{700\text{-}729}} + 0.001\gamma_{FICOrange\mathbf{680\text{-}699}} + 0.015\gamma_{FICOrange\mathbf{640\text{-}679}} - 0.236\tau_{ProsperRating\mathbf{AA}} - 0.216\tau_{ProsperRating\mathbf{A}} - 0.178\tau_{ProsperRating\mathbf{B}} - .116\tau_{ProsperRating\mathbf{C}} - 0.059\tau_{ProsperRating\mathbf{D}} + .009\tau_{ProsperRating\mathbf{E}}$.

A profile plot, found in Figure 1, provides a pictorial illustration of the estimated average interest rate on an average size loan (of $4,900) at each of the seven Prosper ratings, holding covariates constant (log(loan amount) at 8.1614 and debt-to-income at 22.039%). Pooled over FICO and Homeownership, mean interest rates by Prosper Rating are as follows: AA- 8.54%, A- 10.54%, B- 14.34%, C- 20.54%, D- 26.24%, E-33.04% and HR- 32.14% (albeit the HR rating proved insignificant). In sum, as borrower creditworthiness deteriorates, interest rates required by lenders on Prosper loans rise. The aforementioned observations are reflected and a fairly linear relationship reflected with exception of the last Prosper rating, "7" corresponding to "NR". This was not unexpected given the lack of significance found by the model for the final rating.

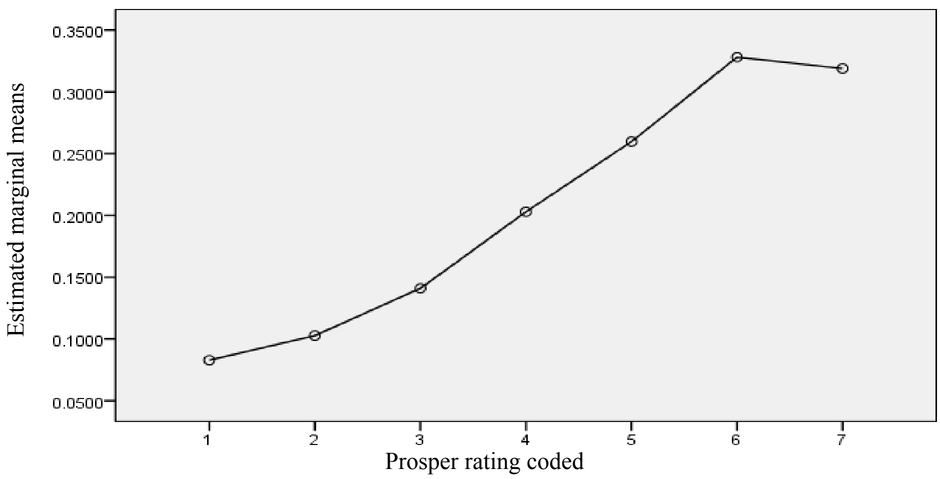

Figure 1 Profile plot–original model

Note: Covariates appearing in the model are evaluated at the following values: logLoanAmount=8.1614, Debt-to-income=0.22039

It is noted that the residuals of the subject model evidence normality, a mean of zero and overall independence, however some evidence of increasing volatility is reflected. Otherwise said, as the Prosper rating system moves from the end of the spectrum of strong ratings toward weaker, less favorable credit ratings, greater variance is shown. This may be interpreted to mean that lenders on Prosper in general have a more difficult time pricing the more risky borrowers.

Having examined the dataset and the results of the general linear regression model as a whole; we now turn to the results to definitely answer the research hypothesis. Do any of the seven Prosper credit ratings have a statistically material impact on individual Prosper loan interest rates?

All but one of the seven Prosper ratings is statistically significant with p-values below 0.01, namely AA, A, B, C, D & E. As evidenced in Table 3, “HR” is not statistically significant.

With that said; the lower and upper bound values of the confidence interval can be examined for those six Prosper ratings found to be significant. More specifically, the confidence interval is examined for overlap in values. If there is any overlap between any of the six ratings, question may be raised concerning autonomy of the impact of the Prosper Rating itself. Review of the confidence intervals is observed at the 99% level; as opposed to the more conventional 95% level so as to provide for better assurance of the research findings. It is noteworthy to point out that analysis of the confidence intervals for the Prosper ratings in a paired ordinal fashion leads to a reduction in the overall confidence level.

No overlap is observed. This observation provides for a reject conclusion to the null hypothesis. Results found in Table 3 provide the details leading to this conclusion. No overlap exists as evidenced by the following results: -0.245 to -0.227 for “AA”, -0.224 to -0.208 for “A”, -0.189 to -0.167 for “B”, -0.123 to -0.109 for “C”, -0.067 to -0.052 for “D” and .000 to -0.018 for “E”.

5. Robustness tests

Apart from the original ANCOVA model, three alternate models were considered to examine the robustness of the results. For one, the model was examined exclusive of the Prosper Rating variable, as well as run another time exclusive of the FICO score variable. This provided for a compare and contrast on how much more of an impact (or lack thereof) that the Prosper rating system provided relative to interest rate to that of the FICO rating system alone. Tables 4 and 5 reflect this analysis. Results indicate that much of the explanatory power for interest rate pricing, using publically available data, comes from the Prosper Rating. This is evidenced by an adjusted R-squared of 0.475 on the model excluding the Prosper Rating variable (Table 4) compared with that of an adjusted R-squared of 0.863 on the model including the Prosper variable and instead excluding the FICO score variable (Table 5). Additionally, the analysis suggests that the FICO score variable provides minimal added value given that the original model (Table 3) reflected an adjusted R-squared of 0.864. Finally, it is observed that the average interest rate rises for the most creditworthy borrowers to 11.15% and declines for the least creditworthy to 31.85% when the model is run without the Prosper Rating. For reference purposes, this compares to lower and upper bound values of 8.54% and 33.04% on the original model. This suggests that the presence of a AA Prosper rating provides additional assurances to lenders such that they are willing to accept a lower interest rate on a loan to this borrower than they would have without this Prosper rating. Conversely, the presence of an E Prosper rating warrants a slightly greater interest rate than would have been required otherwise.

An alternate test was also performed where the scale covariates, namely the log Loan Amount and the Debt-to-Income variables, were standardized within a range of -2.6 < □ < 2.6 so as to provide for a significance level of 0.01. While no “outliers” were found for the log Loan Amount variable, a handful of observations (nine in total) did indeed exist relative to the debt-to-income variable and as such the ANCOVA model was run once again without those observations. Table 6 reflects these results. In summary, exclusive of the nine “outlier” observations, adjusted R-squared improves only 0.02 to that of 0.866. Average interest rates, as are illustrated by a profile plot in Figure 2, reflect a reduction for the most creditworthy, down to a 7.44% for AA Prosper Rating, as well as for the least creditworthy, down to a 32.04% for E Prosper Rating.

Table 4 Model excluding prosper rating as an independent variable

$Y_{interestrate} = ß_0 + ß_1 X_{log(loanamount)} + ß_2 X_{debt\text{-}to\text{-}income} + \alpha_i + \gamma_j + \varepsilon$

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial eta squared
Corrected model	9.654[a]	7	1.379	326.932	0.000	0.476
Intercept	0.039	1	0.039	9.257	0.002	0.004
HomeownershipStatus	0.012	1	0.012	2.769	0.096	0.001
FICOrange	8.011	4	2.003	474.802	0.000	0.430
logLoanAmount	0.467	1	0.467	110.597	0.000	0.042
DebttoIncome	0.076	1	0.076	17.931	0.000	0.007
Error	10.617	2517	0.004			
Total	111.460	2525				
Corrected total	20.271	2524				

Note: [a]. R squared = 0.476 (Adjusted R squared = 0.475)

Parameter estimates–99% confidence interval (Dependent variable: Interest rate)

Parameter	B	Std. error	t	Sig.	99% Confidence interval		Partial eta squared
					Lower bound	Upper bound	
Intercept	0.155	0.015	10.393	0.000	0.116	0.193	0.041
[HomeownershipStatus=1]	-0.005	0.003	-1.664	0.096	-0.012	0.003	0.001
[HomeownershipStatus=2]	0[a]	.	.	.	.	.	.
[FICOrange=1]	-0.207	0.006	-34.534	0.000	-0.222	-0.192	0.321
[FICOrange=2]	-0.167	0.006	-28.371	0.000	-0.182	-0.152	0.242
[FICOrange=3]	-0.107	0.006	-18.789	0.000	-0.122	-0.092	0.123
[FICOrange=4]	-0.059	0.006	-10.596	0.000	-0.073	-0.044	0.043
[FICOrange=5]	0[a]	.	.	.	.	.	.
logLoanAmount	0.019	0.002	10.516	0.000	0.014	0.024	0.042
DebttoIncome	0.038	0.009	4.234	0.000	0.015	0.060	0.007

Note: [a]. This parameter is set to zero because it is redundant.

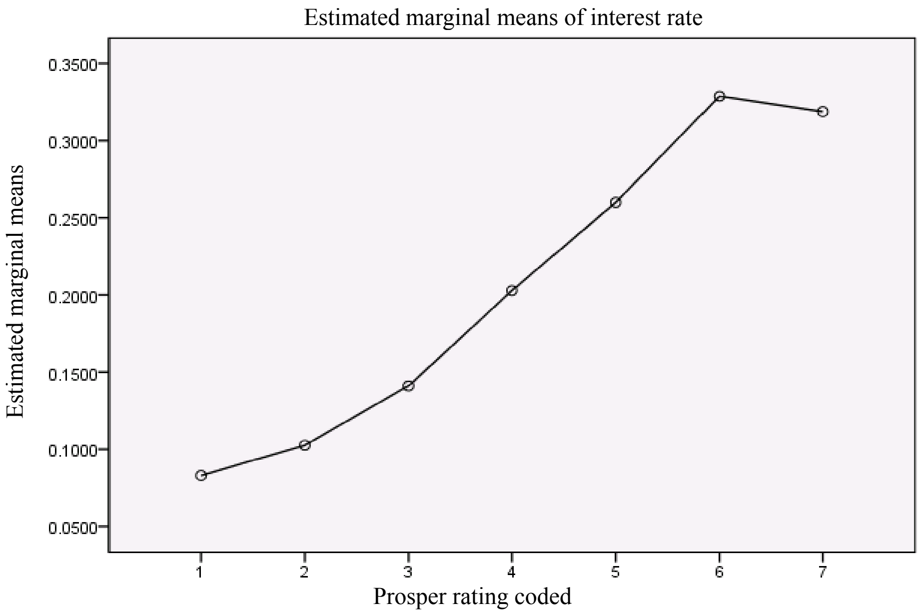

Figure 2 Profile plot–model exclusive of outliers outside the 0.01 significance level

Note: Covariates appearing in the model are evaluated at the following values: Debt-to-income=0.21621, logLoanAmount=8.1646,

Table 5 Model excluding FICO score as an independent variable

$Y_{interestrate} = ß_0 + ß_1 X_{log(loanamount)} + ß_2 X_{debt\text{-}to\text{-}income} + \alpha_i + \tau_k + \varepsilon$

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial eta squared
Corrected model	17.506[a]	9	1.945	1768.889	0.000	0.864
Intercept	0.317	1	0.317	287.861	0.000	0.103
HomeownershipStatus	0.009	1	0.009	8.235	0.004	0.003
Prosperrating	15.863	6	2.644	2404.418	0.000	0.852
logLoanAmount	0.153	1	0.153	139.527	0.000	0.053
DebttoIncome	0.026	1	0.026	23.475	0.000	0.009
Error	2.765	2515	0.001			
Total	111.460	2525				
Corrected total	20.271	2524				

Note: [a]. R squared = 0.864 (Adjusted R squared = 0.863)

Parameter estimates–99% confidence interval (Dependent variable: Interest rate)

Parameter	B	Std. error	t	Sig.	99% Confidence interval		Partial eta squared
					Lower bound	Upper bound	
Intercept	0.233	0.007	32.169	0.000	0.214	0.251	0.292
[HomeownershipStatus= 1]	-0.004	0.001	-2.870	0.004	-0.008	0.000	0.003
[HomeownershipStatus= 2]	0[a]	.	.	.	.	.	.
[Prosperrating= 1]	-0.232	0.003	-81.015	0.000	-0.240	-0.225	0.723
[Prosperrating= 2]	-0.211	0.003	-78.293	0.000	-0.218	-0.204	0.709
[Prosperrating= 3]	-0.172	0.004	-47.295	0.000	-0.181	-0.163	0.471
[Prosperrating= 4]	-0.112	0.003	-42.012	0.000	-0.119	-0.105	0.412
[Prosperrating= 5]	-0.055	0.003	-19.987	0.000	-0.062	-0.048	0.137
[Prosperrating= 6]	0.004	0.003	1.083	0.279	-0.005	0.012	0.000
[Prosperrating= 7]	0[a]	.	.	.	.	.	.
logLoanAmount	0.010	0.001	11.812	0.000	0.008	0.012	0.053
DebttoIncome	0.022	0.005	4.845	0.000	0.010	0.034	0.009

Note: [a]. This parameter is set to zero because it is redundant.

Table 6 Model excluding covariate outliers beyond the 0.01 significance level

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial eta squared
Corrected model	17.487[a]	13	1.345	1248.310	0.000	0.866
Intercept	0.260	1	0.260	241.680	0.000	0.088
FICOrating	0.022	4	0.006	5.217	0.000	0.008
HomeownershipStatus	0.007	1	0.007	6.628	0.010	0.003
Prosperrating	7.860	6	1.310	1215.692	0.000	0.745
DebttoIncome	0.038	1	0.038	34.909	0.000	0.014
logLoanAmount	0.116	1	0.116	107.241	0.000	0.041
Error	2.696	2502	0.001			
Total	110.840	2516				
Corrected total	20.183	2515				

Note: [a]. R squared = 0.866 (Adjusted R Squared = 0.866)

(to be continued)

Parameter estimates–99% confidence interval (Dependent variable: Interest rate)

Parameter	B	Std. error	t	Sig.	99% Confidence interval		Partial eta squared
					Lower bound	Upper bound	
Intercept	0.222	0.008	27.255	0.000	0.201	0.243	0.229
[FICO=1]	0.015	0.004	3.398	0.001	0.004	0.026	0.005
[FICO=2]	0.018	0.004	4.251	0.000	0.007	0.029	0.007
[FICO=3]	0.016	0.004	4.248	0.000	0.006	0.026	0.007
[FICO=4]	0.013	0.004	3.757	0.000	0.004	0.023	0.006
[FICO=5]	0[a]	.	.	.	.	.	.
[HomeownershipStatus=1]	-0.004	0.001	-2.574	0.010	-0.007	4.629E-6	0.003
[HomeownershipStatus=2]	0[a]	.	.	.	.	.	.
[Prosperrating= 1]	-0.236	0.004	-66.502	0.000	-0.245	-0.227	0.639
[Prosperrating= 2]	-0.216	0.003	-69.199	0.000	-0.224	-0.208	0.657
[Prosperrating= 3]	-0.178	0.004	-42.978	0.000	-0.188	-0.167	0.425
[Prosperrating= 4]	-0.116	0.003	-41.392	0.000	-0.123	-0.109	0.406
[Prosperrating= 5]	-0.059	0.003	-20.421	0.000	-0.066	-0.051	0.143
[Prosperrating= 6]	0.010	0.003	2.874	0.004	0.001	0.019	0.003
[Prosperrating= 7]	0[a]	.	.	.	.	.	.
DebttoIncome	0.031	0.005	5.908	0.000	0.017	0.044	0.014
logLoanAmount	0.010	0.001	10.356	0.000	0.007	0.012	0.041

Note: [a]. This parameter is set to zero because it is redundant.

6. Conclusion

Research findings clearly indicate that six of the seven Prosper credit ratings have a statistically significant impact on the dependent variable, individual loan interest rates. The null hypothesis is rejected. The Prosper credit risk rating system does improve market decision making as it provides support to lenders in their pricing decision.

Future areas for exploration may include examination of the market decision impact of the Prosper credit risk rating system from the standpoint of historical loan repayment performance. Given the newness of the Prosper rating system at the time of this writing, there had not been sufficient time that had elapsed to observe the performance of a portfolio of seasoned loans. It would prove interesting to see how well or lack thereof the Prosper rating system was at predicting default, late payments, and etcetera.

While Prosper prides itself on transparency and providing for a proprietary credit rating system to assist lenders to better assess borrower risk, it is acknowledged that not all of the data collected is available for each individual loan in the publically available data export download. Providing for this, on an anonymous basis certainly, would provide for a richer data set to mine and explore and may encourage additional research on this P2P market.

Finally, given the lack of significance and redundancy captured by the seventh and weakest Prosper rating, designated as “NR”, it may prove beneficial to collapse this rating into the next lowest rating, “E”. As it was the endeavor of this paper to examine all of the ratings for potential impact on individual loan interest rates, this was not carried out. However, future studies may benefit from the work performed herein in this regard. The author aims to add value to future research and exploration in the growing field of risk-based lending practices in the peer-to-peer marketplace.

Acknowledgement

I would like to thank Thomas Boucher, L. Michael Couvillion, Christian Ola and Kelly Rush for their most helpful comments and suggestions. All errors are my own.

References:

Andrews, R., Dholakia, U., Herzenstein, M. and Lyandres, E. (2008, February), "The democratization of personal consumer loans? Determinants of success in online peer-to-peer loan auctions", manuscript in preparation, Rice University: Houston, Texas.

Blochlinger, A. and Leippold, M. (2005), "Economic benefit of powerful credit scoring", working paper, Swiss Banking Institute, University of Zurich: Switzerland.

Bogoslaw, D. (2009), Peer-to-Peer Lending: Problems and Promise, *BusinessWeek Online*, 21.

Freedman, S. and Jin, G. Z. (2008, November), "Do social networks solve information problems for peer-to-peer lending?", NET Institute working paper, The Networks, Electronic Commerce and Telecommunications Institute, pp. 08-43.

Anderson, R., Babin, B., Black, W. Hair, J. and Tatham, R. (2006), *Multivariate Data Analysis,* Upper Saddle River, NJ: Pearson Prentice Hall.

Iyer, R., Khwaja, A. I., Luttmer, E. F. P. and Shue, K. (2009, August), "Screening in new credit markets: Can individual lenders infer borrower creditworthiness in peer-to-peer lending?", NBER working papers: 15242, National Bureau of Economic Research, Inc.

Mester, L. (1997, September/October), "What's the point of credit scoring?", *Business Review,* Federal Reserve Bank of Philadelphia, pp. 1-16.

Prosper.com. (n.d.). Retrieved December 14, 2010, available online at: http://www.Prosper.com.

United States Securities and Exchange Commission (2009), S-1 AMENDMENT 6 (Prospectus), filed 07/13/2009 (File No. 333-147019), available online at:
http://www.sec.gov/Archives/edgar/data/1416265/000141626509000033/ prosper_s-1a6d7d13d2009.htm.