

To Fund Or Not To Fund: Determinants of Loan Fundability in the Prosper.com Marketplace

**By Joe Ryan, Katya Reuk, and Charles Wang
The Stanford Graduate School of Business**



**“Only when the tide goes out do you discover who’s been swimming naked”
–Warren Buffett**

Executive Summary

The following is a summary of our key findings and recommendations, grouped into two categories of fundability determinants: personal/social and financial. The remainder of the report provides a detailed presentation of all findings and analyses.

1. Personal/Social Determinants: Endorsements, Listing Descriptions and Profile Descriptions

- A Group Leader Endorsement strongly increases both the percent funded (+33.8%) and the number of bids (+18.85) on a listing. Since an endorsement is either present or not, this effect is binary (e.g., either +33.8% or +0%).
 - ✓ Prosper should actively encourage members to join groups and interact with other group members. New features around group recommendations and other “profiling” techniques might also be worth considering.
 - ✓ The data supports the theory that the social networking aspects of the Prosper marketplace are a significant factor in loan fundability.
 - ✓ Our analysis only considered whether an endorsement was present or not; we could not qualitatively evaluate the contents of that endorsement, and we chose not to consider the length of the endorsement (if present). A qualitative analysis of endorsements may provide a deeper understanding of their effects.
- Listing Descriptions and Profile Descriptions were analyzed based on their length; we were unable to perform a qualitative analysis of this data. Longer Listing Descriptions have a positive effect on both funding percentage and the number of bids, while longer Profile Descriptions have a negative effect. However, both of these effects are considerably small. We had expected a much stronger effect for each, and we suspect that evaluating these variables based on the length of the description is not a good metric. Therefore, we strongly suggest that a qualitative analysis be performed; it’s likely that a large amount of the outcome behavior our models cannot explain is due to the qualitative effects of this data.

2. Financial Determinants: Credit Grades, Verified Bank Accounts, Debt-To-Income, Homeownership, and “Flash” Funding

- As expected, the borrower’s Credit Grade has a very significant effect on both the percent funded (+15.78% per credit grade increase) and the number of bids (+8.24 bids per credit grade increase). (We define a credit grade increase as a one-unit increase, e.g., from “C” to “B” credit). For example, borrowers with “B” credit (a numeric value of 5 on the credit grade scale of 0-7, where 0 = High Risk and 7 = AA credit) should attribute $(5 * 15.78\%) = 78.9\%$ of their total funding percentage to the effect of their credit grade.

- ✓ Having a Group Leader Endorsement (+33.8%) has roughly the same effect as *two* credit grade increases (+31.56%) on percent funded.
- Having a Verified Bank Account also has a significant effect on both the percent funded (+16.64%) and the number of bids (+6.44).
 - ✓ Having a Verified Bank Account has roughly the same effect as a single credit grade increase (+15.78%) on percent funded.
- The Debt-To-Income (DTI) ratio had a very small effect on the percent funded, and a statistically insignificant effect on the number of bids.
 - ✓ This result may be incorrect due to problems with the DTI data. (See the section on “Analysis and Validation of the Models” for more information).
- Surprisingly, Homeownership did not have a statistically significant effect on the percent funded or the number of bids.
- Automatic (“Flash”) Funding has a negative effect on both the percent funded (-4.03%) and the number of bids (-2.61). This is likely because selecting flash funding prevents a listing from being overfunded and also because it may create a market signal that the borrower is “desperate” to get funded as soon as possible; some lenders may avoid such borrowers.

(Please see the Conclusion & Recommendations section for additional information).

Study Synopsis

Who is Prosper.com?

Prosper is an online marketplace for people-to-people lending. Borrowers and lenders each create personal profiles and become members of the Prosper community. Borrowers can request loans through the marketplace by creating listings for a specific amount and maximum interest rate; these listings include a current credit rating & debt-to-income ratio, as verified by the system and made public to potential lenders. Lenders bid for the privilege of supplying all or part of the requested loans and specify the interest rate at which they are willing to lend. When the listing ends, the bids with the lowest rates “win” and are combined into a single loan to the borrower, with Prosper acting as the broker between the lender(s) and the borrower. In addition to personal profiles, borrowers and

lenders who share common backgrounds or interests can form groups. These smaller communities within the marketplace help bring together those who would be more willing to lend and borrow from each other, which often translates into lower rates for borrowers and lower risk of defaults to lenders.

Analysis Questions

A significant portion of Prosper's revenue stream comes from the origination and servicing fees charged on loans made through their site. Consequently, the success of listings in the Prosper marketplace is a primary concern for the company. Listings which reach full funding (i.e., at least 100% funded) are considered "fundable" and, barring any administrative issues, translate into an active loan. However, listings which do not reach full funding are considered unsuccessful ("not fundable") and no loan is created. Our analysis of the data showed that the loan conversion rate is roughly 9% in the Prosper marketplace. Therefore, a greater understanding of the determinants of both loan funding percentage and the number of bids a listing receives will enable Prosper to potentially increase its revenues from the marketplace. There are a number of features about the listing which are user-selected, most of which are decided upon prior to the start of the funding auction. Our analysis concentrates on this set of "input" variables and seeks to determine their effects on the "output" variables (funding percentage and number of bids).

The Data

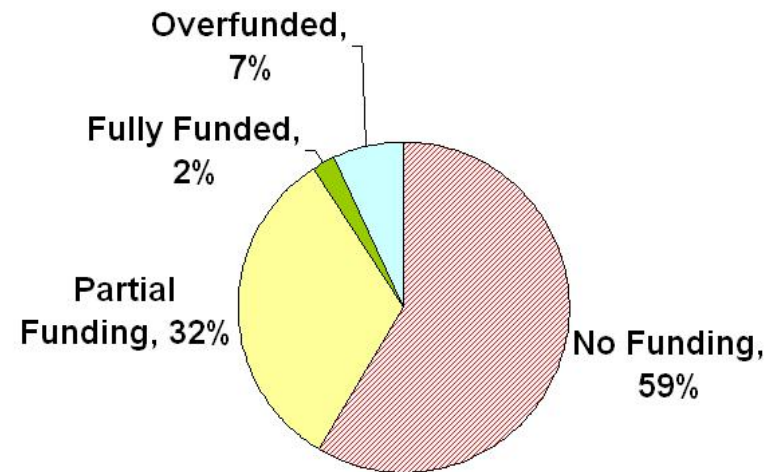
Prosper provided us with data for over 52,000 listings. We randomized this data to remove any ordering bias and extracted a sample of 10,000 listings for our study (90% of the sample was used to construct the regression models, and the remaining 10% was used as a "holdout sample" to independently test the models). Each listing represents an attempt by a borrower to request a loan from the marketplace, and this data constitutes listing activity for the seven-month period between April 19, 2006 and November 19, 2006.

Each data point encompasses a "before and after" snapshot of a listing. A number of data columns represent independent (X) variables that are determined prior to the start of the funding auction, such as the amount of money requested, the maximum rate the borrower is willing to pay, the number of days to keep the listing active, the borrower's credit rating, etc. We were also given outcome (Y) variable data for each listing, such as the final funding percentage, the final interest rate, the total number of bids on the listing, the actual duration of the auction (from start to close), etc. For a detailed explanation of all data columns, please see the Appendix.

Loan Fundability

The primary focus of our analysis was to study the determinants of loan fundability. We created a pie chart (right) showing the distribution of loan funding in our sample of 10,000 listings. Listings that received no funding constitute more than half of the sample. Conversely, listings that were either fully funded (100%) or “overfunded” (greater than 100%) together constituted 9% of the sample. Therefore, the data seems to indicate that listings in the Prosper marketplace have a 9% loan conversion rate. Increasing this rate would increase Prosper’s gross margins and is a motivation for conducting this study.

It is also interesting to note that a loan is over *three* times more likely to be overfunded than not. This yields lower rates for borrowers and a more competitive marketplace overall. It may also indicate a surplus of investment funds that could be directed toward converting more listings into loans, if listings could be made more attractive. It is hoped that our study will be helpful in this regard.



Distribution of Loan Funding

The Regression Models

Before running any regressions, we first formed a theory about what might be the key determinants of loan fundability. This gave us an initial set of independent (X) variables to consider. We then performed a correlation matrix analysis to illuminate the relationships between *all* of the variables in our data (see the Appendix). This analysis allowed us to refine our selection of X variables and to get a feeling for which variables might be significant in the final models. (Note: We also faced a limit of sixteen independent variables, due to constraints imposed on us by Excel’s regression package).

Model #1 – Regression of Percent Funded

We regressed Percent Funded on the fourteen variables listed in the following table. Below are the summarized results of the regression along with our interpretations. (See the Appendix for a complete listing of the model, as well as analyses of the residuals and the holdout sample).

Variable	Impact on Percent Funded (95% Confidence Interval)	Interpretation & Comments
Positive Effects: (Most to Least Significant)		
Group Leader Endorsement	33.8 ± 7.19 %	Having a group leader endorsement increases funding by 33.8%
Verified Bank Account	16.64 ± 2.23 %	Having a verified bank account increases funding by 16.64%
Credit Grade	15.78 ± 0.84 %	For each unit increase in the borrower's credit grade (e.g. C to B), funding increases by 15.78%
Group Membership	2.73 ± 2.25 %	Membership in a group increases funding by 2.73%
Number of Recent Listings	2.40 ± 0.80 %	For each additional recent listing a borrower creates, funding in subsequent auctions is increased by 2.40%. (A recent listing is defined as being created within the previous 30 days).

Borrower Rate	$1.32 \pm 0.15 \%$	For every 1% increase in the borrower rate, there is a 1.32% increase in funding
Listing Duration	$0.64 \pm 0.49 \%$	For every additional day the listing is active, funding increases by 0.64%. (Note: Listing Duration refers to the borrower-selected number of days to keep the listing active. <i>Auction</i> Duration is a separate variable describing how long the auction actually lasted, in fractional days (end date/time – start date/time). We considered Auction Duration an outcome (Y) variable and did not analyze it in this study; we selected Listing Duration because the user selects this value <i>prior</i> to the start of the auction).
Listing Description Length	$0.004 \pm 0.001 \%$	For every additional character in the listing description, funding increases by 0.004%. This implies that typing an additional 1,000 characters would increase funding by 4%. However, our analysis did not have the sophistication to evaluate the <i>qualitative</i> aspects of the description (i.e., spelling and grammar), and therefore we do not think this variable is the best predictor of how the Listing Description affects funding percentage.
Debt-To-Income Ratio	$9.72 \times 10^{-6} \pm 7.07 \times 10^{-6} \%$	A one unit increase (e.g., 0 to 1, or 1 to 2) in the debt-to-income (DTI) ratio increases funding by $9.72 \times 10^{-6} \%$. Since the DTI for most borrowers falls between 0 and 1, this implies that having a DTI of 1 (vs. 0) would lead to a slight increase in funding. Intuitively, we expected DTI increases to have a decreasing effect on funding percentage, and in fact, we found a slight negative correlation between these variables (see the Correlation Matrix). There were some data values in our sample that may have significantly skewed the effect of the DTI ratio in this model (see the section on Model Validity).

**Negative Effects:
(Most to Least
Significant)**

Flash Funding Selected	$-4.03 \pm 2.07 \%$	Selecting Automatic (“Flash”) Funding decreases funding by 4.03%. Users who select this option are effectively creating a market signal that they want the loan to originate as soon as possible. They are willing to trade off a longer auction period (and potentially a lower interest rate) in order to receive the funds more quickly. It’s possible that some smart lenders are detecting this signal and choosing not to bid. However, it seems more
------------------------	---------------------	---

likely that this negative effect simply indicates that a listing with Flash Funding selected can never become “overfunded”.

Number of Total Listings	$-2.04 \pm 0.30 \%$	For each additional listing the borrower creates over his/her entire history as a Prosper user, funding in subsequent auctions is reduced by 2.04%. There appears to be an interplay between the Number of Recent Listings (which has a positive effect on funding) and the Number of Total Listings (which has a negative effect). The net effect is approximately $2.40\% - 2.04\% = +0.36\%$. In general, it would seem reasonable that a borrower with a large number of total listings would be negatively perceived as someone who is often in need of financial assistance. However, coupled with a very positive loan repayment history, lenders may decide that the borrower is, in fact, very trustworthy and choose to ignore the large number of total listings. (It would be interesting to look for this effect once Prosper has a sizeable population of borrowers who have received several loans).
Profile Description Length	$-0.002 \pm 0.002 \%$	For every additional character in the profile description, funding decreases by 0.002%. Thus, typing an additional 1,000 characters would decrease funding by 2%. Interestingly, it is more advisable for a borrower to invest time writing a lengthy (and presumably high-quality) Listing Description instead of investing this energy into their Profile Description. Without detailed qualitative analysis of both the Listing and Profile Descriptions, it is not clear why Profile Description Length is negatively correlated with funding percentage.
Amount Requested	$-1.25 \times 10^{-3} \pm 2.07 \times 10^{-4} \%$	For each additional \$1 requested, funding decreases by $-1.25 \times 10^{-3} \%$. Thus, for every additional \$1,000 requested, funding decreases by 1.25%. Amount Requested has a positive correlation of 0.25 with Credit Grade, so we presume that higher Credit Grades will likely counterbalance the negative effects of higher loan amounts on percent funded.

Exceptions:

Homeownership	<i>Not statistically significant</i>	Surprisingly, we found that homeownership had no statistically significant effect on funding percentage. Other financial status variables, such as Credit Grade and having a Verified Bank Account, are far more significant in this model. This doesn't necessarily
---------------	--------------------------------------	--

mean that Homeownership should be dismissed. In fact, Homeownership has a 0.29 correlation with Credit Grade, so its effect may be partly accounted for in the Credit Grade effects on loan funding.

Model #2 – Regression of Number of Bids

We regressed the (total) Number of Bids on the fourteen variables listed in the following table. Below are the summarized results of the regression along with our interpretations. (See the Appendix for a complete listing of the model, as well as analyses of the residuals and the holdout sample).

Variable	Impact on Number of Bids (95% Confidence Interval)	Interpretation & Comments
----------	--	---------------------------

Positive Effects: (Most to Least Significant)

Group Leader Endorsement	18.85 ± 4.09 bids	Having a group leader endorsement increases the number of bids by 18.85.
Credit Grade	8.24 ± 0.48 bids	For each unit increase in the borrower's credit grade (e.g. C to B), number of bids increases by 8.24.
Verified Bank Account	6.44 ± 1.27 bids	Having a verified bank account increases the number of bids by 6.44.
Group Membership	1.88 ± 1.28 bids	Membership in a group increases the number of bids by 1.88.

Number of Recent Listings	1.07 ± 0.45 bids	For each additional recent listing a borrower creates, the number of bids in subsequent auctions is increased by 1.07.
Borrower Rate	0.62 ± 0.08 bids	For each 1% increase in the borrower rate, the number of bids increases by 0.62. This result is surprising: we expected the borrower rate to have a much stronger, perhaps even non-linear, effect on the number of bids. Presumably, a borrower offering to make interest payments at 25% should attract significantly more bids than a borrower offering 10%.
Listing Duration	0.45 ± 0.28 bids	For every additional day the listing is active, the number of bids increases by 0.45. Again, this effect is surprisingly low.
Listing Description Length	0.002 ± 0.001 bids	For every additional character in the listing description, the number of bids increases by 0.002. This implies that typing an additional 1,000 characters would generate 2 more bids. As mentioned in Model #1, we feel a qualitative analysis of the listing (and profile) description content might uncover more significant effects.
Amount Requested	$2.52 \times 10^{-4} \pm 1.18 \times 10^{-4}$ bids	For each additional \$1 requested, the number of bids increases by 2.52×10^{-4} . Because lenders can bid on any portion (large or small) of the borrower's requested amount, it is not too surprising that the amount requested has little effect on the number of bids. A risk-averse lender could simply offer to fund \$50 on a \$5,000 loan. The slightly positive effect of amount requested on bids might also be explained as: larger requested amounts provide a larger number of investment opportunities in the market, and thus encourage more bidders.

**Negative Effects:
(Most to Least
Significant)**

Flash Funding Selected	-2.61 ± 1.18 bids	Selecting Automatic ("Flash") Funding decreases the number of bids by 2.61.
------------------------	-----------------------	---

Number of Total Listings	-0.85 ± 0.17 bids	For each additional listing the borrower creates over his/her entire history as a Prosper user, the number of bids in subsequent auctions is reduced by 0.85.
--------------------------	-------------------	---

Exceptions:

Homeownership	<i>Not statistically significant</i>	(See the interpretation of this variable in Model #1 for more information).
Debt-To-Income Ratio	<i>Not statistically significant</i>	We feel that the DTI ratio <i>should</i> have been statistically significant. This result may be due to problems with the DTI data – see the Appendix for more details.
Profile Description Length	<i>Not statistically significant</i>	It may be true that bidders do not consider the borrower’s profile description in choosing to place a bid on the listing; however, we feel that there should be an effect. We had theorized that bidders will read the Listing Description and then, at least part of the time, read the borrower’s Profile Description as well. As we’ve discussed earlier in the report, length is probably not a useful metric for evaluating the effect of textual descriptions on both funding percentage and the number of bids. If data from a <i>qualitative</i> analysis of these descriptions were included in the model, we predict that these variables would have a significant effect.

Conclusion & Recommendations

The most interesting finding in this study is the significant effect of Group Leader Endorsements on both funding percentage (+33.8%) and the number of bids (+8.24). An Endorsement is also equal to the effect of *two* credit grade increases on funding percentage. We feel that this strongly supports the theory that social capital translates into economic capital in the Prosper marketplace; the trustworthiness and reputation of borrowers has a tangible economic value, and one that can generate revenues for Prosper. This theme should be a central part of Prosper’s marketing efforts, along the lines of: “Banks don’t care about you, but our members do” and “Don’t just be a borrower, be an opportunity and a friend.” Actively promoting group membership and interaction with other group members (even “real world” social events) could generate more successful listings; in general, site features that work to build trust and reputation among members should increase the quantity of endorsements and other positive market signals.

Also, as mentioned below in the “Next Steps” section, we feel that the qualitative aspects of the Listing Description and the borrower’s Profile Description should have a significant effect on funding percentage and the number of bids. Because we could not perform a qualitative analysis of this data, we cannot quantify any such effects nor include them in our models. Common sense would suggest that since human beings are making (perhaps fairly emotional) decisions about who to invest in, these descriptions are likely to weigh heavily in lenders’ decision-making. The R-Squared values for both of our regression models fall between 21-24%, suggesting (in a simplistic interpretation) that roughly 75% of the behavior we’re seeing in the market cannot be explained by the variables included in our models. We feel that qualitative variables related to the Listing and Profile Description would greatly improve the “fit” of our models.

(Please see the Executive Summary (at the beginning of this report) for a list of additional findings and conclusions).

Recommended Next Steps:

- Active promotion of group membership (potentially including a recommendation system similar to how Amazon.com recommends books or the Apple iTunes Music Store recommends related songs), increased group interaction, and other social networking features that build trust and reputation. Both eBay and Facebook.com offer good examples.
- A qualitative analysis (e.g., spelling, grammar, text analytics) of Listing and Profile Descriptions, to better understand how these variables effect funding percentage and the number of bids.
- Analysis of the determinants of Auction Duration (end time – start time). We noted that the user-selected Listing Duration is often longer than the Auction Duration. There were also cases where the auction lasted longer than the listing. The length of the auction (and the related Listing Duration), similar to the effects of “days on market” in real estate, are worth studying further. A bid-by-bid (discrete) analysis of the auction would be extremely interesting; we could only consider the outcome of the auction in its entirety.
- (Customized?) online educational materials for borrowers to help improve the fundability of their listings, given their credit grade and other contextual factors.

Appendix – Details of the Analysis

Correlation Matrix Analysis

Before beginning any analysis, we first theorized about which variables (columns) in our data should be part of a reasonable model of both funding percentage and the number of bids a listing receives in the Prosper marketplace. Some of these variables we classified as output (LHS) variables, while others seemed more likely to be input (RHS) variables. We used a correlation matrix to illuminate the interrelationships among all the variables in our data. From this analysis we learned: (1) which variables are strongly correlated with our chosen output variables, and therefore should not be eliminated from any model, (2) which variables are almost perfectly positively correlated (i.e., r is greater than 0.95) and might therefore be redundant in the model, and (3) which variables have a low correlation, and thus, do not seem to have a significant influence on our chosen output variables. (We did not strictly eliminate variables based on low correlations with output variables, but we did use this information as a basis for selecting which variables we ultimately included in our models).

CORRELATION MATRIX

	Percent Funded	Num Bids	Listing Duration (Days)	Flash Funding Dummy	Amount Req.	Monthly Pmt	Lender Rate	Bank Draft Fee	Sponsor Rate	Borrower Rate	Listing Desc Length	Credit Grade	Debt/ Income Ratio	GrpLdr Endorse Dummy	Home owner Dummy	Has Verified BnkAcct Dummy	Recent Listings	Total Listings	In Group	Now Del	Past Del	Public Records	Trade Lines	First Credit Line Yr	Inqs Last 6 Mos
PercentFunded	1.00																								
Num Bids	0.77	1.00																							
Listing Duration (Days)	0.05	0.06	1.00																						
FlashFunding Dummy	-0.09	-0.11	-0.10	1.00																					
AmountRequested	-0.03	0.13	0.06	-0.09	1.00																				
MonthlyPmt	-0.02	0.14	0.07	-0.08	0.95	1.00																			
LenderRate	0.13	0.09	0.03	0.08	-0.08	0.02	1.00																		
BankDraftFee	0.02	0.04	-0.04	-0.01	0.05	0.04	0.12	1.00																	
SponsorRate	-0.06	-0.05	0.08	-0.02	-0.02	-0.01	-0.06	-0.09	1.00																
BorrowerRate	0.12	0.08	0.05	0.08	-0.09	0.02	0.99	0.11	0.09	1.00															
Listing Desc Length	0.12	0.11	0.15	-0.02	0.05	0.07	0.18	0.06	0.07	0.19	1.00														
CreditGrade	0.37	0.38	0.00	-0.18	0.25	0.21	-0.21	-0.03	-0.12	-0.22	-0.02	1.00													
Debt/Income Ratio	-0.02	-0.03	-0.01	0.00	-0.01	-0.01	0.00	0.03	-0.03	-0.01	0.01	-0.11	1.00												
GrpLdrEndorse	0.13	0.13	0.04	-0.01	0.01	0.01	0.07	0.11	0.03	0.07	0.17	0.02	-0.01	1.00											
Homeowner Dummy	0.09	0.11	-0.04	-0.04	0.16	0.15	-0.07	-0.02	-0.04	-0.08	-0.03	0.29	-0.07	0.00	1.00										
HasVerifiedBnkAcct	0.24	0.18	0.11	-0.02	-0.10	-0.07	0.11	-0.46	0.07	0.12	0.13	0.18	-0.02	0.06	0.04	1.00									
RecentListings	0.03	0.01	0.05	0.05	-0.13	-0.11	0.16	0.05	0.13	0.18	0.23	-0.09	0.00	0.10	-0.05	0.16	1.00								
TotalListings	-0.09	-0.08	0.05	0.04	-0.08	-0.06	0.11	0.03	0.12	0.13	0.14	-0.13	-0.01	0.03	-0.05	0.24	0.62	1.00							
InGroup	0.08	0.08	0.14	-0.06	-0.01	0.01	0.08	0.04	0.44	0.15	0.27	0.00	-0.02	0.13	-0.01	0.14	0.31	0.22	1.00						
NowDel	-0.16	-0.16	-0.01	0.10	-0.16	-0.14	0.13	0.00	0.06	0.14	0.01	-0.37	-0.02	0.00	-0.11	-0.05	0.10	0.15	0.01	1.00					
PastDel	-0.11	-0.10	-0.01	0.07	-0.09	-0.08	0.05	-0.02	0.05	0.06	0.03	-0.19	-0.02	-0.01	-0.02	-0.02	0.06	0.09	0.02	0.36	1.00				
PublicRecords	-0.08	-0.07	0.00	0.07	-0.04	-0.03	0.02	0.00	0.06	0.03	0.01	-0.17	-0.01	-0.02	-0.02	-0.02	0.06	0.08	0.05	0.15	0.12	1.00			
TradeLines	0.03	0.05	-0.03	-0.01	0.07	0.06	0.01	0.01	0.00	0.01	0.00	0.19	-0.04	0.02	0.27	0.02	0.03	0.03	0.01	0.34	0.30	0.03	1.00		
FirstCreditLineYr	-0.05	-0.07	0.03	0.00	-0.09	-0.09	0.04	0.06	-0.02	0.04	0.00	-0.24	0.02	0.00	-0.25	-0.06	-0.01	-0.02	-0.03	-0.01	-0.19	-0.12	-0.40	1.00	
InqsLast6Mos	-0.06	-0.05	-0.02	0.06	0.04	0.05	0.08	0.08	0.03	0.09	-0.02	-0.12	-0.03	0.02	0.08	-0.04	0.09	0.10	0.03	0.04	0.01	0.02	0.10	0.01	1.00
Profile Desc Length	0.01	0.01	0.12	-0.03	-0.01	0.00	0.08	0.01	0.11	0.10	0.36	-0.05	0.01	0.07	-0.02	0.12	0.17	0.19	0.26	0.04	0.03	0.03	0.02	-0.02	0.00

We have further interpreted these results below:

Variables

*Percent Funded and
Number of Bids*

Observations & Interpretation

Our two chosen output variables, Percent Funded and Number of Bids, have a correlation of 0.77, which indicates that these variables are capturing a common effect: the success of a listing. For a listing to be considered successful, it must receive enough bids to reach at least 100% funding. We've chosen to model each of these variables separately, hence the two regression models.

<i>Amount Requested, Monthly Payment, and Credit Grade</i>	Amount Requested and Monthly Payment are almost perfectly positively correlated (0.95), because they essentially describe the same factor - the size of the loan. Therefore, having both of these variables in the model is probably redundant and we have chosen to use only Amount Requested. It is interesting that loan size (both in terms of the amount requested and monthly payments) has a substantial correlation of 0.25 and 0.21 (respectively) with the borrower's Credit Grade. Borrowers with higher credit grades are more likely to request larger sums of money (for presumably longer-term uses), as compared to borrower's with lower credit grades who borrow smaller amounts due to more short-term financing interests and due to their lack of confidence in receiving funding on larger sums.
<i>Borrower Rate and Lender Rate</i>	These variables have an almost perfectly positive correlation of 0.99 and each has a similar correlation pattern with other key variables, such as Credit Grade. Therefore, we decided that Lender Rate was redundant and have focused instead on the Borrower Rate in our models.
<i>Credit Grade</i>	Credit Grade has correlations with many other variables and, most importantly, has the highest correlations with our two LHS variables, Percent Funded and Number of Bids. It is likely that lenders rely heavily on Credit Grade as one of the key factors in making their bidding decisions. The negative correlation with Borrower Rate and Lender Rate shows that Credit Grade is a significant factor to the lender with respect to risk; therefore, higher credit grades should command lower interest rates. A substantial positive correlation with Homeownership might imply that people with higher credit grades generally have better access to mortgages, and as a result, are more likely to own a home. The negative correlation with First Credit Line Year indicates that the older a person's credit history is, the more likely it is that they will have built up a good credit rating.
<i>Verified Bank Account and Bank Draft Fee</i>	When a borrower chooses not to receive the loan as an electronic deposit to his/her bank account, Prosper charges a Bank Draft Fee. We assume that not having a Verified Bank Account signals that the borrower will incur the Bank Draft Fee, and this explains the significant negative correlation between the two variables. Since the Bank Draft Fee variable does not provide us with additional information that explains loan fundability, it has been excluded from the final model. However, having a Verified Bank Account could be an important factor in loan fundability, given its correlation of 0.24 with Percent Funded.
<i>Sponsor Rate and</i>	Group leaders receive a 0.25% reward for endorsing a loan that gets funded; however, the borrower has

<i>Group Membership (InGroup)</i>	to be in a group for this to be possible. Hence, there is a high positive correlation between the variables. We decided that the Sponsor Rate variable does not provide us with any additional information that could explain the variability in loan funding, so we have decided to exclude it from our models.
<i>Listing Description Length</i>	Listing Description Length is positively correlated with Number of Recent Listings, Group Membership (InGroup), and Profile Description Length. Lengthy Listing Descriptions for group members may indicate that they feel more accountable (within the group) for what they write and therefore invest more time in describing their reasons for borrowing money. Correlation between Listing Description Length and Number of Recent Listings may indicate a “learning effect” on the borrower; borrowers get better at positioning/describing their loans with each additional listing.
<i>Trade Lines and Home Ownership</i>	Trade Lines is positively correlated with Homeownership possibly because the more trade lines the borrower has opened, the more the likely the borrower has an active mortgage. We feel that the effect of the Trade Lines variable is summarized within the Credit Grade variable and/or the Homeownership variable, so we have excluded Trade Lines from the final model.
<i>Delinquencies</i>	Current and Past Delinquencies don't have significant correlations, possibly because the data these variables provide is summarized in the Credit Grade variable. Thus, we have excluded current and past delinquencies from our final model.
<i>Recent/Total Listings and Group Membership (InGroup)</i>	We cannot offer an explanation for the significant correlation between being in a group and having more recent and total listings.
<i>Trade Lines and First Year Line of Credit</i>	The negative correlation between these two variables makes sense, since the more recent the first line of credit is for the borrower, the fewer open credit lines he or she is likely to have.

Variables excluded from the regression models: Monthly Pmt, Now Del, Lender Rate, Past Del, Bank Draft Fee, Public Records, Sponsor Rate, Trade Lines, Inqs Last 6 Mos, and First Credit Line Yr.

The Regression Models

The following shows the complete output for our two regression models. The independent (X) variables highlighted in yellow are those with the largest positive (or negative) effects on the dependent (Y) variable. Variables highlighted in gray are statistically insignificant.

Output for Model #1 (Regression of Percent Funded):

Regression Statistics	
Multiple R	0.494575422
R Squared	0.244604848
Adjusted R Squared	0.243427303
Standard Error	48.78025%
Observations	8996

(The R-Squared value of 0.24 could be interpreted as meaning that this model explains approximately 24% of the behavior we observed in our data sample. It is also commonly interpreted as a measure of the “fit” of the regression model to the data).

Regression of Percent Funded On:	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
GrpLdrEndorse Dummy	33.79088364%	3.67015888%	9.206926668	4.11311E-20	0.26596535	0.409852322
HasVerifiedBnkAcct Dummy	16.64418681%	1.13806703%	14.62496179	6.9002E-48	0.144133158	0.188750578
CreditGrade	15.77592536%	0.43018245%	36.67263853	1.7066E-274	0.149326696	0.166191811
InGroup	2.73077083%	1.14972615%	2.375148933	0.017562742	0.004770453	0.049844963
RecentListings	2.39536173%	0.40691078%	5.886700134	4.08116E-09	0.015977238	0.031929997
BorrowerRate	1.31772980%	0.07555407%	17.44088436	5.09714E-67	0.011696266	0.01465833
Listing Duration (Days)	0.63849797%	0.24930787%	2.561082241	0.010450856	0.001497977	0.011271983
Listing Desc Length	0.00381555%	0.00063099%	6.046879593	1.5357E-09	2.57866E-05	5.05244E-05
Debt/Income Ratio	0.00000972%	0.00000361%	2.696790187	0.007014198	2.65544E-08	1.67894E-07
AmountRequested	-0.00124726%	0.00010555%	-11.81711158	5.49168E-32	-1.4541E-05	-1.04036E-05
Profile Desc Length	-0.00199181%	0.00091822%	-2.169208471	0.030093002	-3.7917E-05	-1.9189E-06
Homeowner Dummy	-0.11138520%	1.28900297%	-0.086411905	0.93114091	-0.02638125	0.024153547
TotalListings	-2.04099355%	0.15410488%	-13.24418545	1.14655E-39	-0.02343074	-0.01738912
FlashFunding Dummy	-4.03144149%	1.05641625%	-3.816148692	0.000136461	-0.06102258	-0.01960624
(Intercept)	-37.7111363%	2.56794013%	-14.68536428	2.89513E-48	-0.42744884	-0.32677387

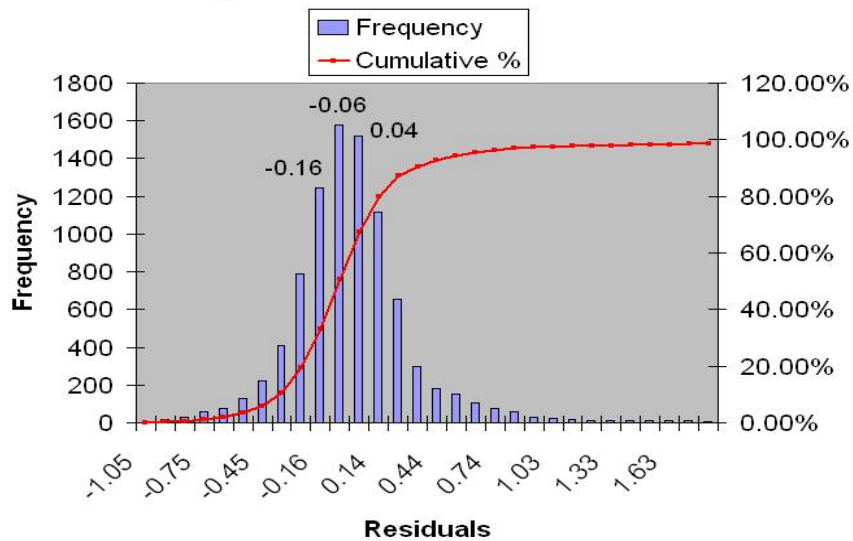
Output for Model #2 (Regression on Number of Bids):

Regression Statistics	
Multiple R	0.460343636
R Squared	0.211916263
Adjusted R Squared	0.210687762
Standard Error	27.72186 bids
Observations	8996

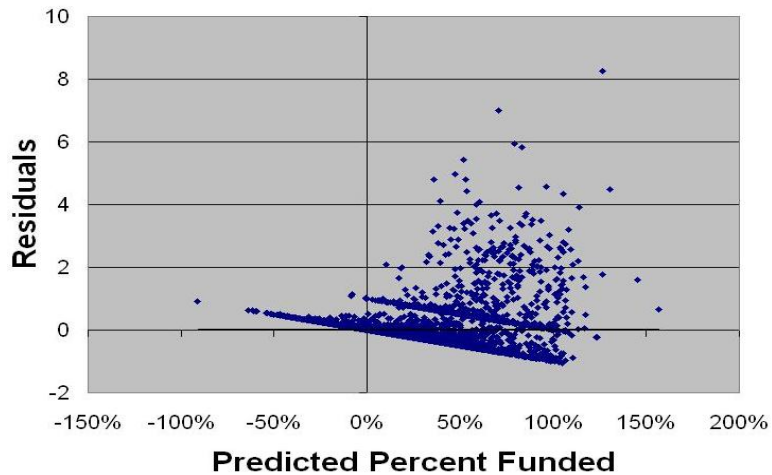
(The R-Squared value of 0.21 could be interpreted as meaning that this model explains approximately 21% of the behavior we observed in our data sample. It is also commonly interpreted as a measure of the “fit” of the regression model to the data).

Regression of Number of Bids On:	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
GrpLdrEndorse Dummy	18.85274351	2.085754796	9.038811057	1.91372E-19	14.76418833	22.94129869
CreditGrade	8.237912608	0.244473095	33.69660207	1.6726E-234	7.758689575	8.717135641
HasVerifiedBnkAcct Dummy	6.444515819	0.64676458	9.964237405	2.88548E-23	5.17670971	7.712321927
InGroup	1.876192533	0.653390468	2.871472152	0.004095206	0.595398172	3.156986893
RecentListings	1.069056778	0.231247784	4.622992526	3.83523E-06	0.615758371	1.522355184
BorrowerRate	0.620897439	0.04293745	14.46051021	7.21715E-47	0.536730242	0.705064636
Homeowner Dummy	0.562693226	0.732541621	0.768138232	0.442425261	-0.87325545	1.998641901
Listing Duration (Days)	0.451429788	0.141681903	3.186220535	0.00144638	0.17370094	0.729158636
Listing Desc Length	0.001867795	0.000358595	5.208648699	1.94439E-07	0.001164867	0.002570722
AmountRequested	0.000252141	5.99823E-05	4.203586785	2.65264E-05	0.000134562	0.00036972
Debt/Income Ratio	4.00682E-06	2.04883E-06	1.955662325	0.050535901	-9.3545E-09	8.02299E-06
Profile Desc Length	-0.000974455	0.000521825	-1.867397008	0.061878709	-0.00199735	4.84416E-05
TotalListings	-0.85227849	0.087577948	-9.731656302	2.84619E-22	-1.02395124	-0.68060573
FlashFunding Dummy	-2.612177123	0.600362364	-4.351000795	1.37016E-05	-3.78902430	-1.43532994
(Intercept)	-25.3449452	1.459362831	-17.36713082	1.77162E-66	-28.2056292	-22.4842611

Histogram of Residuals - % Funded

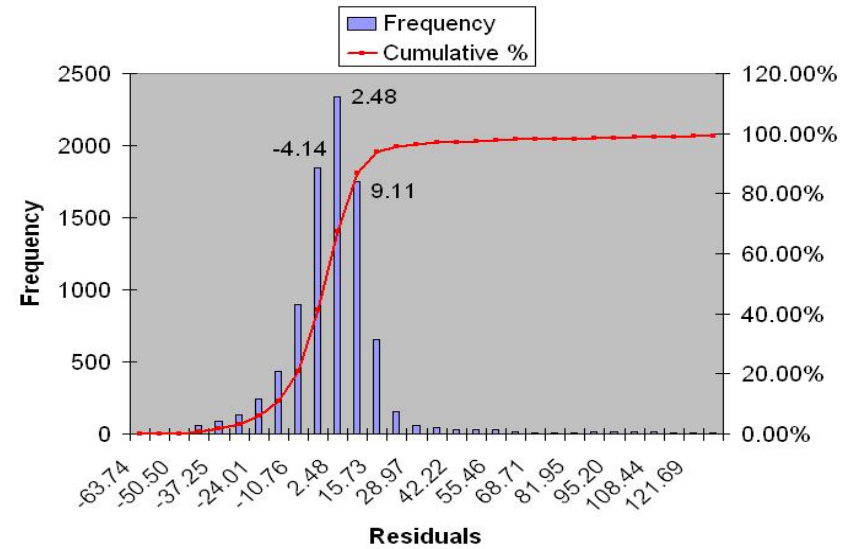


(a)

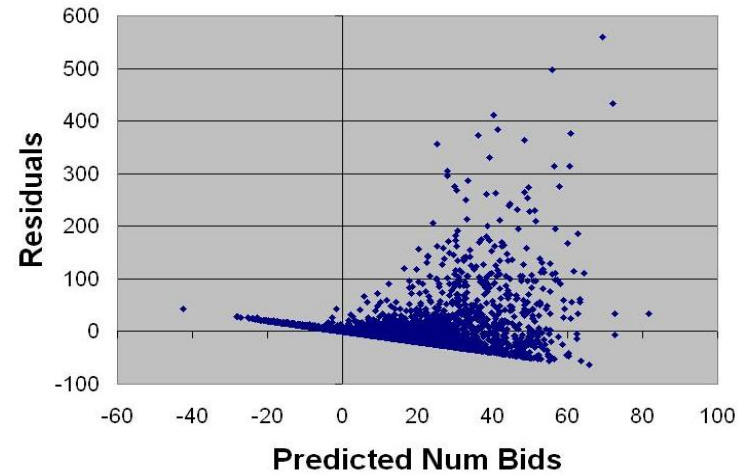


(c)

Histogram of Residuals - Num Bids



(b)



(d)

Exhibit 1 – (a-b) Histograms of Residuals, (c-d) Plots of Model Predictions vs. Residuals (Error)

Residuals Analysis

For each model, we plotted predicted values against residuals, as well as produced a histogram of residuals. (See Exhibit 1 on the previous page).

In the residual plots (Exhibit 1-c & d), we noticed that our graphs did not have the typical “cloud” or “football” shape, but were instead slightly cone-shaped. We believe that this is the result of the sample being heavily influenced by a large percentage (59%) of listings that received no funding. We investigated the residuals more closely and verified that in most cases, for each predicted value, the residuals are averaging to zero. However, because it is not possible to have negative percent funding or a negative number of bids, there is a “projection” effect that is stacking up what would’ve been negative predicted values. This can be seen in the thick, downward sloping line at the bottom of both plots.

In order to try to correct the cone shape, we could’ve transformed our Y values (percent funded and number of bids) using the logarithm function, i.e., $\log(Y)$, and then re-run the regressions on the transformed Y values. Since the “coning” of the scatter plots seems to be relatively minor, we opted to leave the regression model as is; changing the model to a logarithmic regression may potentially introduce other errors, and make the residuals, and hence the regression model, worse.

The Holdout Sample

From a sample of 10,000 listings, we used 90% of the sample to produce our models and withheld the remaining 10% for use in later testing of the models. Using the derived linear equation from each regression model (i.e., $\text{Predicted-Y} = \text{Intercept} + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \dots + \beta_{14} \cdot X_{14}$) and data from the holdout sample, we calculated the model-predicted values and compared them to the actual values in each case. (See Exhibit 2 on the next page for a plot of Predicted vs. Actual values).

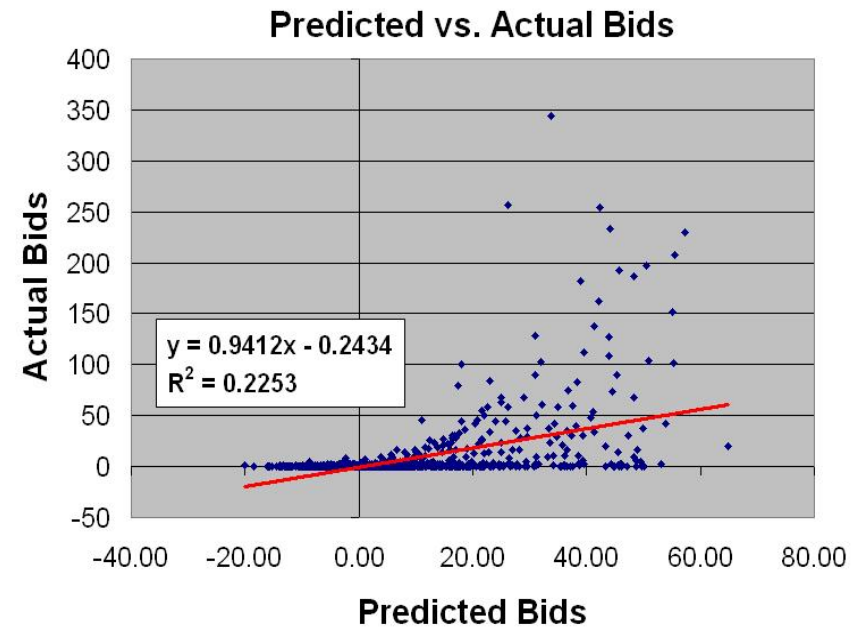
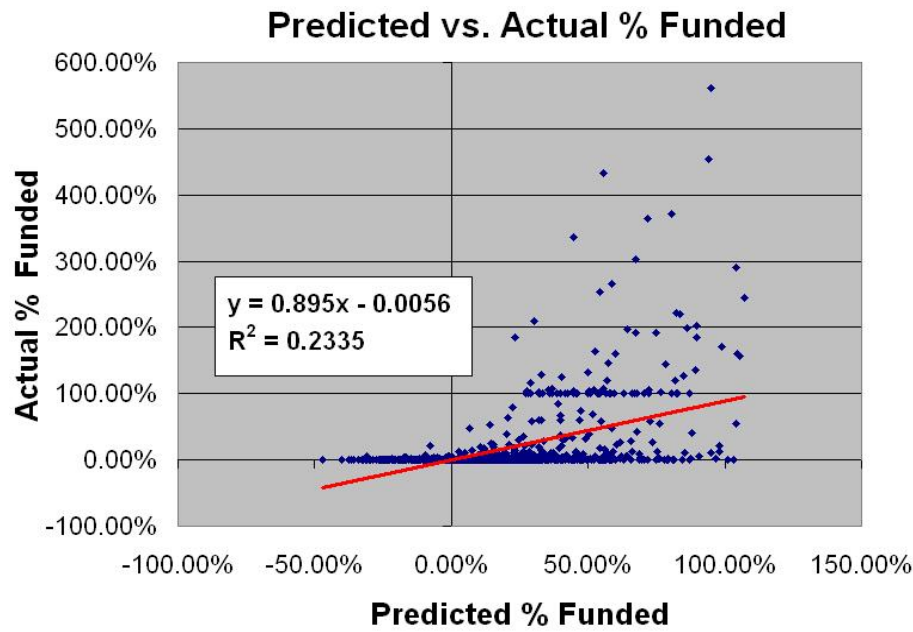


Exhibit 2 – Model Predictions vs. Actual Values using the Holdout Sample

As a result of the holdout sample testing, we found that our model for **percent funded had an average error of approximately 25%**, and our model for the **number of bids had an average error of approximately 12 bids**. These error rates are roughly twice as small as the error rates predicted by the regression models (49% and 28 bids, respectively). We feel that these prediction errors are reasonable given that our goal for this study was to provide a ranking of the relative importance of various fundability determinants, and not to provide predictive models. Further, the absence of variables that reflect the qualitative nature of both the Listing and Profile Descriptions are likely to account for a significant portion of this error; in general, these are not complete models.

We also attribute some of this error to the fact that 59% of the data sample contained listings that received no funding. We did not want to remove any of these data points since that would severely bias the sample, but we recognize that these data points are likely skewing the models: fundability determinants that are normally effective may be viewed as less significant by our models because so many listings (where such determinants were positively applied) still did not receive any funding at all. This would unfairly “water down” the positive effects of the determinants.

The Data

Data for each listing included the following columns (variables):

- **Percent Funded.** Percentage of the loan amount that is funded is the sum of the amounts of bids over the total amount requested by the borrower. The numbers in our data sample range from 0% to 952%. There is a high concentration of data at both 0% and 100%. In fact, more than half of the sample of listings received 0% funding. The use of the “Flash Funding” feature (described below) creates a high chance that loan funding will be exactly 100%.
- **Number of Bids.** Total number of bids on the loan includes multiple bids by the same lender. Our data range is 0 – 628.
- **Listing Duration.** Borrowers can select the number of days their listing is active, typically in the range of 3 to 10 days.
- **Flash Funding.** Borrowers have the option of selecting Flash Funding, which automatically ends the auction (and initiates the loan) when 100% funding has been achieved. This data is represented in our model as a dummy variable (“1” = flash funding selected, “0” = flash funding not selected).
- **Amount Requested.** This is the amount of money requested by the borrower and is measured in US Dollars. Our data range is from \$1000 (minimum listing amount allowed) to \$25,000 (maximum listing amount allowed).
- **Monthly Payment.** Monthly payments by the borrower include interest. Repayment occurs in 36 monthly installments.
- **Lender Rate and Borrower Rate.** Lender Rate is the rate all winning bidders will receive from the borrower. Borrower Rate is the maximum rate the borrower is willing to accept.
- **Listing Description Length.** Borrowers write a description of why they are seeking a loan. This can be up to 4,000 characters in total length.
- **Credit Grade.** Prosper assigns a credit grade based on the borrower’s score from Experian. Credit grade ranges from 0 to 7, where 0 = High Risk (no credit information available) and 7 = “AA” (the highest).
- **Debt/Income Ratio.** The Debt-to-Income ratio (or DTI) is a measurement of the borrower's ability to take on additional debt. This number measures the borrower's monthly debt payments (excluding housing payments) relative to their monthly income. It includes the monthly payment on the requested loan, as if the loan were already active. Generally, a DTI of 20% is at the upper end of “normal” when excluding housing debt.
- **In Group.** The use of borrower groups is one of the unique aspects of the Prosper model. Membership in a group is represented as a dummy variable in our data (“1” – borrower is in a group; “0” – borrower is not in a group).
- **Group Leader Endorsement.** Group leaders have the option to endorse the listing of a borrower in their group. This is represented as a dummy variable in our data (“1” – endorsed; “0” – not endorsed).
- **Homeowner.** Whether or not the borrower is a homeowner is another dummy variable in our data (“1” – homeowner; “0” – not a homeowner).

- **Has Verified Bank Account.** Indicates whether or not the borrower has a bank account as verified by Prosper (“1” – has a verified bank account; “0” – does not have a verified bank account).
- **Recent and Total Listings.** Recent Listings indicates the number of listings posted by the same borrower in the past 30 days, while Total Listings indicates the total number of listings posted by the borrower over the life of their Prosper profile.
- **Profile Description Length.** Similar to Listing Description Length, this measures the length of the borrower’s member profile description in characters.
- **Sponsor Rate.** Group leaders receive compensation for endorsing loans that subsequently get funded. Sponsor Rate is the percentage of the loan received by group leaders.
- **Bank Draft Fee.** Borrowers without a verified bank account are charged a bank draft fee by Prosper for each transaction.
- **Now Delinquent.** This data point indicates the number of accounts on which the borrower is currently late on a payment. This includes any unpaid charge-offs or other derogatory balances.
- **Past Delinquencies (in last 7 years).** The number of 90+ days past due delinquencies on the borrower's credit report in the last 7 years.
- **Public Records (in last 10 years).** The number of negative public records on the borrower's credit report over the last 10 years. Negative public records include, among other things, bankruptcies, liens, and judgments.
- **First Credit Line Year.** The month and year the borrower's first recorded credit line was opened. Credit lines may include, among other things, revolving, installment, and mortgage credit. The data is presented numerically beginning with the year, so that the highest number indicates the most recent first credit line.
- **Trade lines.** The total number of credit lines appearing on the credit report. Credit lines can be open or closed and may include, among other things, revolving, installment, and mortgage credit.
- **Inquiries in last 6 months.** The number of inquiries made by creditors to the borrower's credit report in the last six months.

END OF REPORT
