STRATEGIC INFORMATION TRANSMISSION IN PEER-TO-PEER LENDING MARKETS

Fabio Caldieraro, Jonathan Zhang, Marcus Cunha Jr, Jeffrey D. Shulman*

*Fabio Caldieraro is an Associate Professor of Marketing at FGV/EBAPE, Rio de Janeiro, Brazil (fabio.caldieraro@fgv.br). Jonathan Zhang is an Assistant Professor of Marketing at Michael G. Foster School of Business, University of Washington, Seattle (zaozao@uw.edu).
Marcus Cunha Jr. is a Professor of Marketing at the Terry College of Business, University of Georgia, Athens, GA (cunhamv@uga.edu). Jeffrey D. Shulman is the Marion B. Ingersoll Associate Professor of Marketing at Michael G. Foster School of Business, University of Washington, Seattle (jshulman@uw.edu)

All authors contributed equally to this paper and are listed alphabetically.

Acknowledgement: Marcus Cunha Jr. gratefully acknowledges funding from a Terry-Sanford research grant from the University of Georgia for this research project. Jeffrey Shulman acknowledges the generous funding of the Marion B. Ingersoll professorship.The authors are thankful of José Fajardo, Marcelo Brogliato, Rafael Goldszmidt and the participants of the University of Washington marketing seminar, who have provided valuable comments on our research.

Abstract

Peer-to-peer (P2P) marketplaces, such as Uber, AirBnB, and Lending club, have been experiencing massive growth in recent years. They now constitute a significant portion of the world's economy and provide opportunities for people to transact directly with one another. However, such growth also challenges participants to cope with information asymmetry about the quality of the offerings in the marketplace. By conducting an analysis of a P2P lending market, the authors propose and test a theory in which countersignaling provides a mechanism to attenuate information asymmetry about financial products (loans) offered on the platform. Data from a P2P lending website reveal significant, non-monotonic relationships among the transmission of non-verifiable information, loan funding, and ex post loan quality, consistent with the proposed theory. The results provide insights for platform owners who seek to manage the level of information asymmetry in their P2P environments so as to create more balanced marketplaces, as well as for P2P participants interested in improving their ability to process information about the goods and services they seek to transact online.

Keywords: asymmetric information, consumer-to-consumer interactions, consumer financial decision making, electronic commerce, P2P platforms.

The Internet and information technology increasingly produce more disintermediated and democratized industries by connecting individual actors in unprecedented manners. Such development fostered the explosive growth of the peer-to-peer (P2P) economy and enabled the rise of a number of successful P2P platforms. Uber has quickly become the world's largest driving service; Alibaba is now the most valuable retailer; and Airbnb offers more rooms than any other hospitality service (*The New York Times* 2015). Similarly, Lending Club, a P2P lending platform is now the world's largest online marketplace connecting individual borrowers and investors.

Deservedly so, the P2P economy and its major societal impacts have attracted substantial research interest, as well as calls for more studies that apply decision-making perspectives to these consumer-to-consumer interactions (Kumar 2015; Yadav and Pavlou 2014). In response, a few recent papers in marketing and economics have studied peer-influenced consumer decisions in industries such as music (Sinha, Machado, and Sellman 2010), video games (Landsman and Stremersch 2011), used cars (Lewis 2011), lending (Lin, Prabhala and Viswanathan 2013), and retailing (Backus, Blake, and Tadelis 2015).

Even as P2P platforms expand in various industries, information asymmetry remains a challenge for both participants and P2P platform managers. Without the signals of brand power and other reputational heuristics that consumers often use as proxies for quality, participants in P2P platforms need to make decisions with limited information, causing transaction risks to be higher than those in traditional business settings. On the one hand, buyers need to decide how to interpret the information provided by sellers to infer quality and minimize risk; on the other hand, sellers can strategically reveal or withhold information about themselves to increase their chances of a favorable outcome. In turn, platform managers likely need to weigh the information provided by sellers and create mechanisms to reflect transactional risk accurately and engender more trust in the platform (Schlosser, White, and Lloyd 2006).

In this research, we examine the issue of information asymmetry in P2P markets by studying a social lending platform. We centered attention on the Lending Club platform since it is an exemplar P2P marketplace that is gaining substantive importance and in which the effects of asymmetric information lead to significant consumer losses. P2P lending is a multibillion dollar industry that has experienced staggering 100% annual growth since 2010 (*The Economist* 2014) and is expected to reach $150 billion in size by 2025 (PWC 2015). It has democratized capital markets by allowing people to bypass traditional banking roadblocks and enabling them to become customers and suppliers of their own financial products. Lending platforms provide verifiable information about borrowers but a considerable degree of information asymmetry does remain, causing lenders to bear significant risk because loans are unsecured. For instance, as of March of 2016, Lending Club reports that across all loans, 7.8 % of the amount issued to borrowers is charged off. Because lenders shoulder the default losses, it is optimal for them to minimize default risk by looking for the borrowers who are most likely to repay their loans.

We consider the Lending Club platform and examine whether the strategic transmission of non-verifiable information by a borrower, represented by the length of the description of the reasons for the loan, offers signaling content that complements other verifiable information, and helps lenders distinguish the likelihood of repayment of each loan. Any mechanism on a P2P lending platform that can further distinguish those borrowers more likely to repay a loan has the potential to improve the lending market for both borrowers and lenders. To the best of our knowledge, this is the first study to consider how the mere presence of a description and the length of the description might help borrowers strategically transmit information about their repayment prospects.

According to extant theories on cheap-talk, if the provision of non-verifiable information is costless, then it should not affect a buyer's decision. If the provision has costs in terms of effort, then such information carries a signal which might affect buyer's decision in a monotonic manner: the higher the effort, the stronger the signal. In contrast with these traditional perspectives, we recognize that non-verifiable communication in P2P platforms is a more complex phenomenon because there are multiple sources of information. Thus, we propose to study the P2P market under the lenses of a theory of *countersignaling* as the potential major force governing P2P interactions under information asymmetry.

Specifically, within the same credit-worthiness class, as measured by verifiable information, loan applicants who make the decision to provide no loan description (i.e., choose not to transmit non-verifiable information) are expected to have a higher likelihood of getting funded and a lower likelihood of delinquency, according to the countersignaling argument. However, when borrowers decide to write descriptions, applications featuring longer descriptions have a greater likelihood of getting funded and a lower likelihood of delinquency than those with short descriptions, a result that is consistent with an effort-as-a-signal argument. We contrast the predictions of countersignaling with those of competing theories. Using a data set of loan applications from the P2P platform Lending Club over 3 years, we find support for our theory: Lenders' funding decisions are influenced by strategic countersignaling by borrowers, and these decisions are confirmed by the borrower's subsequent likelihood of delinquency.

This research thus contributes to information transmission and consumer decision-making literature in several ways. Theoretically, we show that the countersignaling mechanism is present in the P2P transaction setting, and it helps resolve information asymmetry. This is a novel finding in light of competing theories based on signaling, cheap talk, persuasion, and psycholinguistics. Empirically, we provide evidence that is consistent with the countersignaling theory and inconsistent with the competing mechanisms. In particular, we show that individuals indeed strategically transmit information to other individuals in an online P2P environment through the effort they exert to write the loan description. This strategic transmission provides an informative signal about loan quality, as evidenced by subsequent loan performance. We show that in equilibrium, individuals on both sides of the platform are sophisticated actors, capable of sending and interpreting quality signals.

Our research also adds to literature on consumer lending decisions, a stream of research that has increasingly been receiving attention from marketing scholars. For instance, research has investigated lenders reaction to race and appearance of an applicant's uploaded photograph (Galak, Small, and Stephen 2011; Ravina 2012), the number and role of the members of an applicant's friendship group (Lin, Prabhala, and Viswanathan 2013), lender herding behaviors (Herzenstein, Dholakia, and Andrews 2011; Zhang and Liu 2012), and the impact of type of media on micro-lending (Stephen and Galak 2012). Our paper adds to the literature by investigating by showing how borrowers can use loan descriptions to signal quality to lenders.The findings have implications for P2P lending platform designers, who should consider countersignaling behavior when they seek to fine-tune their risk/return algorithms. The findings provide guidance to borrowers regarding when to countersignal; for lenders, they reveal how to weight information that goes beyond the verifiable information provided by the platform in order to better identify true risks. A growing industry of hedge funds and algorithm-based services (e.g., Lending Robot) promise that their risk assessments are more comprehensive than those from existing platforms' and select loans on the basis of a borrower's non-verifiable information to boost returns. Our finding reveals an area of information that these companies could productively exploit. Our work informs not only the growing number of P2P marketplaces, but also provides new insights into consumer financial decision making in the age of data prevalence.

More generally, our work has managerial implications for various platforms (such as eBay, Etsy, AirBnB, and Upwork) on which sellers may wish to communicate their quality credibly to buyers. Given the evidence that counter-signaling can indeed convey information about quality, managers of P2P platforms should consider opening this avenue of information exchange and incorporate it in their composite seller rating score presented to the buyer, to reduce information asymmetry and increase efficiency, which in turn would build trust among participants. As the P2P economy keeps on growing, information asymmetry and trust in the platform will continue to be notable issues. Platforms that can fine-tune their rating system to better reflect the risk and to resolve information asymmetry more effectively will instill more confidence among participants and thus gain advantages over their rivals. Sellers (in our case, borrowers) on P2P platforms can use the insights of this research to help them decide when to rely exclusively on verifiable information provided by the platform and when it is worth to engage in the production of non-verifiable information. Buyers in P2P marketplaces can learn how to aggregate platform-provided with participant-provided information about the products and services marketed on the platform. Buyers who are adept at picking up informational cues might achieve higher returns (e.g., buying high quality products at a lower price) while mitigating risks.

Literature Review

Our work brings together two research streams – asymmetric information in P2P platforms and P2P lending specifically, and the mechanism of counter-signaling. We now briefly discuss related research in each stream and our contributions.

Asymmetric Information in P2P Markets

It is well understood that for markets to work efficiently, buyers and sellers need to possess symmetric information. In the presence of information asymmetry, the market will not allocate resources efficiently and may even collapse (Akerlof 1970).

Thus, research on P2P platforms has largely focused on information disclosure and signaling to alleviate asymmetric information. For example, Lewis (2011) studied the P2P marketplace eBay Motors and found that the disclosure of some degree of verifiable information by a seller, such as pictures and text with specifications of the automobile, can serve to reduce adverse selection, provided the seller is contractually obligated to fulfill products that match the information they provide. Backus, Blake, and Tadelis (2015) identified how participants bargain on eBay's "Best Offer" listings can signal their level-of-impatience types by posting round-number prices. Li, Tadelis, and Zhou (2016) investigated how sellers can signal quality by offering incentives for consumers to leave feedback in the online P2P marketplace Taobao. Tadelis and Zettelmeyer (2015) used a field experiment to investigate how information disclosure about the quality of objects can improve the efficiency of markets. The authors find that the disclosure decreases search costs and thus it helps bidders better match their preferences with the quality of the products being offered in the marketplace.

All of the above studies found that sellers can alleviate asymmetric information by either voluntarily revealing verifiable information about quality types, or by sending a costly signal. Our work adds to the literature by recognizing that strategic information transmission in P2P platforms can be a more complex phenomenon in situations in which the P2P platform can serve as an additional source of information. In such cases, the sellers can resolve additional information asymmetry by 1) the voluntary disclosure of unverifiable information (even if unrelated to quality types), and the 2) the effort of providing lengthy disclosures. These two elements combine can result in a non-monotonic relationship between the degree of the seller's disclosure and the buyer's interpretation of quality.

Asymmetric Information in P2P Lending

Likewise, P2P lending platforms also experience asymmetric information. Research in this domain has primarily focused on investigating factors beyond borrower's creditworthiness can influence lender behavior.

Freedman and Jin (2011) identified that some of the asymmetric information and adverse selection can be reduced through a learning-by-doing process, in which the entire market learns about the risk level of the financial products being offered in the market and gradually excludes low-quality borrowers in favor of higher quality borrowers. The likelihood of a loan application getting funded can be affected by the race and appearance of an applicant's uploaded photograph (Duarte, Siegel, and Young 2012; Pope and Sydnor 2011; Ravina 2012), the number and role of the members of an applicant's friendship group (Lin, Prabhala, and Viswanathan 2013), and lender herding behaviors (Herzenstein, Dholakia, and Andrews 2011; Zhang and Liu 2012).

Kawai, Onishi, and Uetake (2014) also study the issue of adverse selection, using data from an earlier version of the Prosper P2P platform, where potential borrowers posted public reserve interest rates to signal their credit worthiness. In our framework and data, interest rates are set by the lending platform based according to the borrower's verifiable risk profiles, which is similar to the situation proposed by Milde and Riley (1988). As a result, our borrowers cannot use interest rates to signal their quality and instead rely on unverifiable information to signal and countersignal. When the borrower is required to provide a loan description, Sonenshein, Herzenstein, and Dholakia (2011) demonstrate that, using the perspectives of persuasion, borrower with poor credit history can improve funding likelihood by explaining and taking responsibility for their financial mistakes. Similarly, the number and the content of borrower identity claims influence lenders' decisions (Herzenstein, Sonenshein, and Dholakia 2011; Michels 2012).

Our current research differs from previous P2P Lending contributions in two dimensions. First, previous work has only looked at lender's funding decision. We take it a step further and look at whether these decisions are correct in the long run, as measured by loan performance. Second, whereas previous platforms *require* borrowers to provide description, Lending Club offers the option for borrowers not to do so. The proposed framework allows us to investigate the differential signaling values of the description length (and its related effort) as well as the value from the mere presence of (or the lack of) the description.

Therefore, we can corroborate prior findings that non-verifiable information such as the purpose and the content of the loan description affect loan funding. However, we uniquely provide theory and evidence suggesting that 1) the mere presence and 2) the length of the loan description are signaling mechanisms that can attenuate information asymmetry between the borrower and lender regarding the borrower's ability to repay the loan. Our framework contributes to the P2P lending literature by offering a unified theory of countersignaling, explaining both borrower and lender behaviors, which is likely sustainable in the long run. We show that in equilibrium, individuals on both sides of the platform are sophisticated actors, capable of sending and interpreting quality signals.

Countersignaling

In traditional signaling models, all the information originates from the sender, and the effect of the signaling instrument is monotonic. Research in signaling have examined a broad range of context from education choice (Spence 1973), to advertising decision (Milgrom and Roberts 1986) and pricing (Desai 2000).

The P2P lending context features a mix of verifiable information, screened and provided by the platform (e.g., credit score, debt level, public records), and non-verifiable information provided by the borrower (e.g., loan purpose, loan description). When borrowers prepare their applications, they do not know with certainty how their verifiable information (compiled by the P2P lending platform) will appeal to lenders, according to the platform's underwriting model. This reality more closely relates to the counter-signaling theory, in which there are two sources of information: one provided by the sender and the other provided by a trusted third party (Feltovich, Harbaugh and To 2002).

Prior research theorizes that when additional sources of information are available, high-quality senders countersignal by choosing not to provide information about course grades Feltovich et al. (2002), by using low-quality packaging (Clements 2010), or by spending less on advertising (Orzach, Overgaard, and Tauman 2002), or by engaging in either image or informative advertising (Mayzlin and Shin 2011). Counter-signaling creates a non-monotonic relationship between sender quality and signaling effort, and this relationship can hold even if there is heterogeneity in how consumers process information (a common phenomenon, as noted by Bart et al. 2005, and Zhu and Zhang 2010).

Our research is one of the very few studies to empirically investigate countersignaling. We do so in the growing field of P2P lending. We now describe our theory, followed by empirical test of the theory.

A Theory of Countersignaling in Peer-to-Peer Lending

In this section, we put forth a theory in which loan descriptions serve as an instrument for countersignaling, formalize hypotheses associated with this theory, and contrast its predictions with competing theories regarding the role that loan descriptions serve on peer-to-peer lending platforms. In the P2P marketplace, informational asymmetry exists regarding a borrower's type. Borrowers' inherent quality may not be expressed perfectly in the available verifiable information (e.g., credit scores), and borrowers have superior knowledge about their own likelihood to pay back a loan, relative to lenders. Lenders may attempt to infer the true quality of the loan from the information that borrowers provide on their applications. Thus, a borrower might try to use the loan request or description as a signaling instrument, to facilitate the exchange of information from the prospective borrower (sender) to the potential lenders (receivers). Countersignaling theory offers predictions about a lender's behavior in response to a borrower's communication of such non-verifiable information.

Intuition. Consider first a stylized P2P social lending situation in which there are three types of borrowers for a given asset class: high-quality, medium-quality, and low-quality, where quality indicates borrowers' unobservable likelihood to repay the loan. Each loan has a level of risk and an interest rate that compensates the lender for taking that risk. The lender's goal is to choose the asset classes that match his or her portfolio objectives and within each class identify the quality of the loans. The signaling mechanism (i.e., loan description in our case) does not differentiate applicants across verified credit grades; this classification already has been done by the credit grade itself. Instead, the loan description functions to differentiate among the loans within the same credit grade. All potential lenders know that a loan with credit grade A (best), priced at an interest rate of 5%, has a lower risk of default than a loan with credit grade G (worst), priced with an interest rate of 21%. The informational problem is the differentiation among loans within a range of similar credit grades and interest rate combinations. Thus, high-, medium-, and low-quality types refer to the types *within* the same credit grade (e.g., among three A-grade borrowers or three G-grade borrowers).

It is important to note that Lending Club loans are unsecured personal loans, so their credit worthiness is supported only by and is a direct function of the credit worthiness of the borrower. The credit grade, however, is not a completely deterministic measure of credit worthiness (e.g., 95% of A-grade borrowers honor their loans, but 5% do not), nor is it the only verifiable information reported by the Lending Club. The platform also reports additional verifiable information such as FICO scores and the number of previous hard credit inquires. Credit grades largely depend on FICO scores though, as we demonstrate empirically later in this article.

In our data set, Lending Club borrowers apply for loans and may write loan descriptions *before* the platform performs the formal credit worthiness assessment (e.g., checking credit scores, employment, public records) and reveals this third-party–verified information to lenders on the platform. At that point, borrowers know that the third-party–verified information correlates positively with their identified type, but they do not know the exact information with certainty.

Borrowers seek to maximize their payoff when choosing whether to write a loan description and how long to make it.[1] A borrower is willing to incur the cognitive cost (effort) of creating a

[1] Prior research reveals the cognitive cost of writing text. For example, Greiner and Wang (2010) report that prospective borrowers need to invest effort to write high-quality loan requests. Shavell (2010) also asserts that people

loan description if the benefits outweigh the costs. Thus, the loan description can facilitate a lender's inference of borrower quality. Medium-quality borrowers cannot be confident about whether the information provided by the third party will be viewed positively. Thus, they exert effort to send a signal to differentiate themselves from low-quality borrowers and write long descriptions. Low-quality borrowers recognize that the verifiable information is unlikely to benefit them, so it is unprofitable to attempt to overcome the negativity of this information by exerting a high level of effort by providing a lengthy loan description. They thus exert less effort describing the loan than medium quality borrowers do. High-quality borrowers recognize that the third-party information has a high probability of distinguishing them from a low-quality borrower. That is, they have a low probability of being confused with low-quality borrowers, and by providing no loan description, they can profitably draw a distinction from medium-quality borrowers.

Formal model for a continuum of borrower types. Assume that, in accordance to their previous portfolio allocation decisions,[2] lenders on the platform may wish to invest some of their money into a certain risk-reward category of loans offered in the platform. Because each risk-reward category has fixed interest rates, it is optimal for lenders to infer the quality of the loans within each asset class and to allocate higher shares of their budget to the loans that are less likely to default (i.e., have higher quality).[3]

usually experience some disutility for writing well-crafted written works, and Liebowitz and Margolis (2005) even suggest that the act of writing may be subject to opportunity costs.

[2] Fabozzi (2013 p. 464) states that the first decision that portfolio managers should do is the asset allocation decision, the decision as to how much to invest in each asset class. Brus (2010) provides and example stating that the Oklahoma Teachers Retirement System has an executive directive to invest 70% in the equity market and 30% in the bond market. Bodie, Kane, and Marcus (2009 p. 218) report that even sophisticated investment companies may adopt a multi-stage decentralized approach by first deciding between asset-class allocations and then performing security section within each class, because of the overwhelming complexity of optimizing an organization's entire portfolio decision in one stage. Our analysis is abstracting from higher-level asset-class allocation and concentrates on the security selection of loans within the same risk-reward category.

[3] A straightforward portfolio analysis in which loans have fixed interest rates and differ in their likelihood to default will conclude that lenders should be more likely to fund loans they believe to be less likely to default. Such an analysis is available from the authors upon request.

The quality of the loans within a risk-reward category is linked to the quality of the borrower, which can be represented by a parameter $\theta \in [\underline{\theta}, \bar{\theta}]$, with $0 < \theta < 1$. Higher values of θ represent higher quality borrowers in terms of the likelihood of repayment of the loan according on schedule. Both borrowers and lenders have common knowledge about the distribution of borrowers in the market, but ex-ante, only a particular borrower knows his own true quality.

The independent platform compiles verifiable information and sends a noisy signal, $x = \theta + \varepsilon$, where ε is a random variable distributed uniformly in the interval [-a, a]. This signal represents a measure of all the verifiable information provided by the lending platform, such as monthly income, delinquencies, credit inquiries, etc. Ex-ante, the borrower knows his own type θ, but not the realization of x. Without loss of generality, we assume that the interval [-a, a] is such that $\underline{\theta} - a \geq 0$ and $\bar{\theta} + a \leq 1$.[4]

Before x is revealed on the platform, borrowers can write descriptions of a length s (with $s > 0$) to send a signal to the market. To send the signal, borrowers experience a cognitive cost of effort s, where k is a cost parameter. We consider that a proportion λ of lenders are sophisticated and make decisions based on the verifiable information x and on their own inferences from the signal s. The likelihood these sophisticated lenders will fund a loan is equal to $\mu_s = \Phi(s, x)$, where $\Phi(s, x)$ represents the lenders' belief function as they rationalize both the verifiable and non-verifiable information. As discussed above, this assumption is consistent with the optimal portfolio allocation of rational lenders who consider a mean-variance tradeoff, as these lenders should allocate higher shares of their budget to higher quality loans, and consequently be more likely to fund higher quality loans.

We also allow for a proportion $1 - \lambda$ of lenders to be naïve lenders who myopically believe in the non-verifiable information. Our treatment of naïve individuals is similar to Inderst and Ot-

[4] The intervals for the values the borrower quality θ and the noise ε can be constructed from a simple normalization in which alternative parameters θ' and a' are members of $\mathbb{R}_+$ with $\underline{\theta'} > a'$. The normalization is formed by choosing a large enough number M, ($M = \bar{\theta}' + a'$ is sufficient) and making $\theta = \theta'/M$ and $a = a'/M$.

taviani (2012) who also allow for individuals who do not properly account for the strategic incentives behind the information they receive. For these lenders, their belief of borrower quality increases directly with the amount of non-verifiable information and with the quality of the platform's provided noisy signal (later we will discuss the ensuing outcomes when these types of lenders are absent from the market).

The likelihood naïve lenders will fund a loan is equal to $\mu_n = \frac{s}{1+s} x^\gamma$, where x^γ reflects the importance of verifiable information to the naïve lenders and captures the effect that borrowers who have good objective information have a better basis to write enticing descriptions (for instance, a borrower who has a prestigious job or high income can write a description that highlights these facts). The parameter γ allows for a non-linear interaction. Whereas this functional form nicely captures the possibility that naïve lenders are more easily swayed by non-verifiable information when the verifiable information is positive, we note the counter-signaling equilibrium result is robust to modifications to this function.[5] Our assumption about the likelihood a naïve lender funds a loan also goes in line with portfolio allocation decisions of lenders who consider a mean-variance tradeoff. In the same vain as sophisticated lenders, the naïve lenders should be more likely to fund the loans that they perceive to be of higher quality. The difference between these two types of lenders is the way they form beliefs about the quality of the loans.

Notice that the presence of naïve lenders may provide an incentive for some borrowers to convince these lenders via description length. As it will be seen later, in the equilibrium result section, for any type θ, both too little and too much effort investment in the description can be costly — invest too much and the borrower's disutility for writing descriptions overpowers the benefit of convincing borrowers; invest too little and the borrower leaves too much "money on the table" from naïve lenders.

[5] The counter-signaling equilibrium that arises in our model is robust to other specifications that preserve the standard single-crossing property. For instance, the effect of signaling on naïve borrowers can be independent of borrower type and the signaling cost can be a function of borrower type as in Feltovich, Harbaugh, and To (2002).

Assuming that borrowers get a value V from obtaining a loan,[6] and recalling that x is a function of θ, we can write the borrowers' expected utility as:

$$E[U(s,\theta)] = E\left[V\left(\lambda\Phi(s,x(\theta)) + (1-\lambda)\frac{s}{1+s}x(\theta)^{\gamma}\right) - k\,s\right]. \qquad [1]$$

Given that x is the only random variable, we can rewrite this expression as:

$$E[U(s,\theta)] = V\left(\lambda E[\Phi(s,x(\theta))] + (1-\lambda)\frac{s}{1+s}E[x(\theta)^{\gamma}]\right) - k\,s\,. \qquad [2]$$

Since we are interested in whether the amount of non-verifiable information can signal the quality of the borrower, we will be looking for an informational equilibrium that satisfies the following Perfect Bayesian Equilibrium conditions:

(*i*) $E[U(s^*,\theta)] \geq E[U(s',\theta)]$ for any $s' \in S$.

(*ii*) $E[\Phi(s^*,x(\theta))] = \theta,$ for all $\theta \in [\underline{\theta},\bar{\theta}]$.

The first condition states that each borrower type θ sends a signal s^* that maximizes her own utility. The second condition states that the sophisticated lenders' beliefs about the borrowers' types are confirmed in equilibrium. In other words, $E[\Phi(s^*,x(\theta))]$ will capture the derived belief supports our countersignaling equilibrium.

Because $E[\Phi(s,x(\theta))]$ is ultimately a function of s and θ, we can define $\Phi_x(s,\theta) \equiv E[\Phi(\mathrm{s},x(\theta))]$. In addition, we can compute the expectation $E[x(\theta)^{\gamma}]$ by integrating over the random noise ε: $E[x(\theta)^{\gamma}] = \int_{-a}^{a} x(\theta)^{\gamma}\frac{1}{2a}d\varepsilon = \frac{(\theta+a)^{\gamma+1}-(\theta-a)^{\gamma+1}}{2a(\gamma+1)}$. Hence, we define the function $\xi(\theta) \equiv E[x(\theta)^{\gamma}] = \frac{(\theta+a)^{\gamma+1}-(\theta-a)^{\gamma+1}}{2a(\gamma+1)}$ to be this expectation.

Following Condition (*i*), we maximize Expression [2] with respect to s. We take derivatives with respect to s and find the first order condition:

$$V\left(\lambda\Phi_x{}'(s^*,\theta) + (1-\lambda)\frac{\xi(\theta)}{(1+s)^2}\right) - k = 0\,, \qquad [3]$$

where s^* denotes the equilibrium signaling effort.

[6] In an extension of this model in the Web Appendix 1, we show that results are preserved even if the value V is dependent on the borrower`s type θ, provided the single-crossing property is preserved.

By solving the ordinary differential equation given by Expression [3], we find that $\Phi_x(s^*,\theta)$ can be expressed as:

$$\Phi_x(s^*,\theta) = \frac{k(1+s^*)}{V\lambda} + \frac{(1-\lambda)\,\xi(\theta)}{\lambda+\lambda s^*} + C\ , \tag{4}$$

where *C* is a constant to be determined by the appropriate boundary condition.

As in Milgrom and Roberts (1982) and Daughety and Reinganum (1995) we consider that a Pareto-efficient outcome requires that, in a separating equilibrium, the lowest quality borrower $\underline{\theta}$ has no incentive to distort her optimal amount of non-verifiable information. Hence, we rewrite the borrower's expected utility function in Expression [1] as:

$$E[U(s,\underline{\theta})] = V\left(\lambda\Phi_x(s,\underline{\theta}) + (1-\lambda)\frac{s}{1+s}\xi(\underline{\theta})\right) - k\,s\ . \tag{5}$$

By maximizing this expression with respect to *s* we find that the optimal (undistorted) amount of non-verifiable information sent by a $\underline{\theta}$ borrower is:

$$s^*(\underline{\theta}) = \frac{-k+\sqrt{(1-\lambda)Vk\,\xi(\underline{\theta})}}{k}\ . \tag{6}$$

By using Expression [6] as a boundary condition, and solving the equality $\underline{\theta} = \frac{k(1+s^*(\underline{\theta}))}{V\lambda} + \frac{(1-\lambda)\xi(\underline{\theta})}{\lambda+\lambda s^*(\underline{\theta})} + C$, we can determine the constant *C* and rewrite Expression [4] as:

$$\Phi_x(s^*,\theta) = \frac{k(1+s^*)}{V\lambda} + \frac{(1-\lambda)\xi(\theta)}{\lambda+\lambda s^*} + \frac{-2k(1-\lambda)\xi(\underline{\theta})+\lambda\underline{\theta}\sqrt{(1-\lambda)Vk\,\xi(\underline{\theta})}}{\sqrt{(1-\lambda)Vk\,\xi(\underline{\theta})}}\ . \tag{7}$$

By observing Condition (*ii*) in our signaling equilibrium that $\Phi(s^*,\theta)=\theta$, we solve $\theta = \frac{k(1+s^*)}{V\lambda} + \frac{(1-\lambda)\xi(\theta)}{\lambda+\lambda s^*} + \frac{-2k(1-\lambda)\xi(\underline{\theta})+\lambda\underline{\theta}\sqrt{(1-\lambda)Vk\,\xi(\underline{\theta})}}{\sqrt{(1-\lambda)Vk\,\xi(\underline{\theta})}}$ for the optimal signaling amount s^*. Inverting the equation, we find that the curve

$$s^*(\theta) = \frac{1}{2k}\left(\begin{array}{l}\lambda V(\theta-\underline{\theta}) - 2\mathrm{k} + 2\sqrt{(1-\lambda)Vk\,\xi(\underline{\theta})} + \\ \sqrt{V\left(4(1-\lambda)k\left(\xi(\underline{\theta})-\xi(\theta)\right)+\lambda(\theta-\underline{\theta})\left(V\lambda(\theta-\underline{\theta})+4\sqrt{(1-\lambda)Vk\,\xi(\underline{\theta})}\right)\right)}\end{array}\right) \tag{8}$$

can represent the optimal amount of non-verifiable information s^* for a borrower of type θ.

Lastly, we prove that a counter signaling equilibrium is possible in this model. By observing Condition (*i*) in our signaling equilibrium, we find that a counter signaling equilibrium is possible if the higher types find it optimal to switch from the function described in [8] to the curve $s = 0$ (i.e., to provide no description). Hence, we compute the higher-quality-type borrowers' expected outcome when following the belief function given by [7]:

$$U^{s*} = E[U(s^*(\theta),\theta)] = V\left(\lambda\theta + (1-\lambda)\frac{s^*(\theta)}{1+s^*(\theta)}\xi(\theta)\right) - ks^*(\theta). \tag{9}$$

On the other hand, the higher-quality-type borrowers' expected outcome when exerting zero signaling effort is simply:

$$U^{s0} = E[U(0,\theta)] = V\left(\lambda\theta + (1-\lambda)\frac{0}{1+0}\xi(\theta)\right) - 0k = V\lambda\theta\ . \tag{10}$$

The belief $\Phi_x(0,\theta) = \theta$ is rational only for types that prefer the outcomes given by expression [10] over those given by expression [9].

To verify existence, consider the difference in utilities $U^{s*} - U^{s0}$ (expressions [9] and [10]) has to be positive for the lowest quality borrower $\underline{\theta}$. This difference is:

$$U^{s*}|_{\theta=\underline{\theta}} - U^{s0}|_{\theta=\underline{\theta}} = (1-\lambda)V\xi(\underline{\theta})\frac{s^*(\underline{\theta})}{1+s^*(\underline{\theta})} - ks^*(\underline{\theta})\ .$$

Next, we consider the difference in utilities $U^{s*} - U^{s0}$ for the highest quality borrower $\bar{\theta}$ has to be negative. The difference is: $U^{s*}|_{\theta=\bar{\theta}} - U^{s0}|_{\theta=\bar{\theta}} = (1-\lambda)V\frac{s^*(\bar{\theta})}{1+s^*(\bar{\theta})}\xi(\bar{\theta}) - ks^*(\bar{\theta})$.

Countersignaling occurs when the two conditions above are satisfied. They can be combined in the condition:

$$\frac{(1-\lambda)V\xi(\bar{\theta})}{1+s^*(\bar{\theta})} < k < \frac{(1-\lambda)V\xi(\underline{\theta})}{1+s^*(\underline{\theta})}\ . \tag{11}$$

One can verify that, because $s^*(\theta) > \theta$, there are parameter values that satisfy Condition [11]. Furthermore, because both $\xi(\theta)$ and $s^*(\theta)$ are strictly increasing in θ, there exists a cut-off borrower quality level θ^* such that types $\theta > \theta^*$ find it better to countersignal by sending $s^* = 0$ because $U^{s*}|_{\theta>\theta^*} - U^{s0}|_{\theta>\theta^*} < 0$, whereas types $\theta < \theta^*$ find it unprofitable to do so because

$U^{s*}|_{\theta<\theta^*} - U^{s0}|_{\theta<\theta^*} > 0$ (the types $\theta < \theta^*$ thereby optimally choose to send a signal s^* according to Equation [8]). It is easy to see from Condition [11] that a countersignaling equilibrium requires that at least an infinitesimal proportion of lenders in the market are naïve. Otherwise, the inequalities cannot be satisfied.

Thus, it follows that the equilibrium amount of non-verifiable information as a function of borrower type is:

$$s^{*\text{counter\_signal}} = \begin{cases} s^*(\theta) & \text{for } \theta \in [\underline{\theta}, \theta^*] \\ 0 & \text{for } \theta \in [\theta^*, \overline{\theta}] \end{cases}. \quad [12]$$

Notice, however, that if there were no naïve consumers, and thus $(1-\lambda) = 0$, it is impossible to find any values for the parameters that satisfy Condition [11]. In such case, a countersignaling equilibrium could not be sustained, and the optimal for all players would be to write no description. The intuition is that if the benefit of writing a description cannot be rationalized and the cost of writing a description still exists, then the optimal is for all borrowers to simply avoid the cost of writing a description (thus setting $s^{*\text{no\_naives}} = 0$).

Figure 1 illustrates our theory's prediction about the relationship between the equilibrium amount of non-verifiable information sent by the borrowers, and the likelihood a loan will be funded. The figure shows loan requests with no-description are interpreted as being sent by higher quality borrowers and are rewarded by a high likelihood of being funded. For the remaining borrowers, the likelihood a loan request is funded is increasing in the amount of signaling effort.

(Figure 1 Goes Here)

In summary, in the P2P lending setting, where information asymmetry regarding the quality of the borrowers plays a significant role, if potential lenders are able to account for countersignaling when inferring the quality of a loan, the effect of loan description on funding will be non-monotonic. Specifically, the absence of a loan description should increase the likelihood that the loan gets funded relative to having a loan description. However, once a borrower provides a loan description, the likelihood of getting funded should be higher when the description is lengthy than when the description is short. Furthermore, no borrower has an incentive to pretend to be of a different quality by mimicking the strategy of other borrowers. The implications of the analytical model are synthetized in four testable hypotheses listed below:

H_1: The absence of a loan description has a positive effect on the likelihood of a loan being funded.

H_2: The absence of a loan description has a negative correlation with the ex post likelihood of a loan being delinquent in repayment.

H_3: Given that a loan has a borrower-provided description, the likelihood of a loan being funded increases with description length.

H_4: Given that a loan has a borrower-provided description, the ex post likelihood of a loan being delinquent in repayment correlates negatively with its description length.

At this point, it is useful to compare our hypotheses with competing theories. If providing a lengthy loan description were costless to borrowers, a model of cheap talk would predict that it would have no effect on lender behavior (e.g., Crawford and Sobel 1982). Thus, if loan descriptions are costless to borrowers, H_1–H_4 would not be supported, and we would not be able to reject the null hypotheses.

Literature on persuasion and compliance (Langer, Blank, and Chanowitz 1978) would predict that a lender is more likely to comply with a funding request when a reason is offered, whether that reason is legitimate or not. In contrast with H_1, this account would predict a negative effect of the absence of a loan description on loan funding. Moreover, we would not be able to reject the null version of H_2. Research that predicts an effect of the number of persuasive arguments (e.g., Petty and Caccioppo 1984) predicts that more arguments tend to improve persuasion. This prediction is consistent with H_3, but it implies the opposite of H_1. Moreover, we would not be able to reject the null hypotheses associated with H_2 and H_4.

Finally, we compare our theory's predictions with psycholinguistics theory, which predicts that in asynchronous computer communications, liars produce more words when lying than when telling the truth (e.g., Hancock et al. 2008). This theory then would predict ex post loan delinquency increases monotonically with description length, in conflict with H_4, and if lenders rationally infer this relationship, loan funding monotonically decreases with description length, in contrast with H_3.

In summary, in the P2P-lending context that we study, the counter-signaling model generates four testable hypotheses that capture the non-monotonic effects of loan descriptions on funding and ex post delinquency. Competing theories based on models of cheap talk, persuasion, and psycholinguistics generate one or more predictions that conflict with H_1–H_4. Thus we can test whether countersignaling governs how borrowers and lenders use the loan description or if this relationship instead can be explained better by one of the competing theories.

Institutional Details, Data, Empirical Model, and Results

In this section, we describe the institutional settings that govern the theory construction and the empirical model. We first discuss the data used to test our hypotheses and provide some model free evidence, then present the tests of the impact of borrowers' provision of non-verifiable loan information on attracting funding and the ensuing relationship between non-verifiable loan information and ex post loan performance.

Institutional Details of the P2P Lending Platform

We examine loan requests on Lending Club. The two dependent variables of interest are whether the loans are funded by lenders, and their subsequent performance. The loan application process is as follows: First, the borrower fills out an online form with his or her name, address, date of birth, annual income, and the loan amount that he or she would like to request. Second, Lending Club immediately makes an *initial* non-binding offer that includes information about the monthly payment. This offer is based on the information provided, as well as publically available information about the borrower; it is known as a soft credit pull. Borrowers are reminded to be truthful in providing the information, because false information will lead to denial of the loan. Third, once the borrower agrees with the initial terms, he or she fills out another online form, providing additional information that can be used for credit verification and displayed to lenders (e.g., employment, loan description, home ownership, social security number). At this point, the borrower also can write the loan description. All of this occurs before Lending Club performs the final credit worthiness verification. Fourth, after the borrower makes a decision about whether to write non-verifiable information in the form of a loan description, Lending Club conducts further identity verification by checking additional paperwork (e.g., recent tax returns), conducting phone call verifications, and making inquiries into the borrower's credit history. Finally, if the borrower passes Lending Club's underwriting review, the loan request will be posted online, along with the verified information (credit score, debt-to-income ratio, number of delinquencies) and any borrower-provided information about the purposes and description of the loan. Once the loan is funded, Lending Club deposits the money in the borrower's bank account.

Several details make Lending Club unique relative to other P2P platforms (e.g. Prosper.com), such that it offers a cleaner venue to study countersignaling. First, borrowers are anonymous and not able to communicate with lenders offline; lenders in turn have no capability to identify whether a particular borrower is a first-time or repeat borrower. In our data set, we observe the same information available to lenders when making their funding decision; it seems implausible that lenders could gain any private, unobserved information about borrowers. Second, some factors included in prior research, such as applicant photos, borrower groups, and borrower history, are not available to lenders on this platform. Lenders thus cannot be influenced by heuristics based on the borrower looks, group associations, or prior repayment history. Third, borrowers cannot set their own interest rates, and interest rates are set based solely at Lending Club's discretion. The platform's website states that interest rates are a result of Lending Club's

Base Rate plus an adjustment for risk and volatility, resulting from credit worthiness scores.[7] Borrowers thus cannot signal their quality using interest rates. Lending Club assigns credit grades (A to G) to indicate the risk of each loan (A is the safest, G is the most risky). As expected, the correlation between credit grade and interest rate in our data is .933, such that credit grades are determined almost entirely by verifiable credit risks.

The sequence of events in this loan application process may motivate borrowers to engage in signaling/countersignaling: Borrowers need to provide loan descriptions prior to Lending Club performing the verification and publishing the verifiable information. Borrowers have private information about their true quality, which they expect will correlate positively with Lending Club's published, verifiable information. However, they do not know this with certainty or a priori. Because the interaction is anonymous, and borrowers cannot communicate with lenders or entice lenders with higher interest rates, the only action they can perform is writing a loan description.

Data Description

The source of data is Lending Club archives, which feature all loan applications from 41 consecutive months from May 2007 to September 2010, providing a total of 26,314 applications. During this time period, the platform is open only to individual instead of institutional lenders[8]. To be considered for a loan, borrowers must have a valid bank account, valid social security number, a sufficiently high credit score (640 or above), and a debt-to-income ratio below 25% (excluding mortgage). After the verifiable information is authenticated, the loan request is listed on the site for two weeks or until it gets funded, whichever happens first.

The Lending Club platform provides verifiable credit history information collected from the major credit bureaus and reports. The following information thus is reported about each application: the borrower's FICO credit score range, number of open lines of credit, earliest credit

[7] Details are available from www.lendingclub.com/public/how-we-set-interest-rates.action.
[8] https://help.lendingclub.com/hc/en-us/articles/215437958-How-has-Lending-Club-s-investor-base-changed-

line, credit line utilization, revolving credit balance, number and timing of delinquencies, home ownership status, physical location of the applicant (state), number of credit inquiries in the last six months, and other relevant public records such as public records on file, months since last record, and months since last major derogatory report. Potential borrowers are also required to select a purpose of the loan (e.g. debt consolidation, home improvement etc.) and have the *option* to provide a loan description and state freely why lenders should lend them money (see Web Appendix 2 for samples of loan descriptions of varying lengths), which is non-verifiable information. Interested lenders can fund a portion (minimum of $25) or the entirety of the loan request. Any defaults are managed by collection agencies commissioned by the platform. The key dependent variable of interest is the Loan funding outcome, 1 if it gets fully funded, or 0 if it is not funded. [9]

The platform allows lenders to diversify across different loans. Paravisini et al. (2016), using both lending club data in its early days as well as private 3rd party data source on lender identification, observe diversification decisions and estimate risk-aversion parameter for different lenders on the platform. While we do not observe lender identification or lender's other investment vehicles and hence do not explicitly model the lenders' diversification decisions, our model does speak to the lenders' selection of loans to invest once a risk-balance allocation has been made and investors decided how much to allocate to each loan asset class.. As such, the lending decisions we model goes in line with the modern portfolio theory prescription of selecting loans such that the expected return is maximized for a given level of risk. A loan that is fully funded implies that a sufficient number of lenders deemed it of having low probability of default and therefore being worthy of funding (see the aforementioned discussion about loan selection).

As researchers, we observed exactly the same information that lenders did. In what follows,

[9] The majority (58.8%) of the loans is fully funded, and 34.7% of the loans received 0 funding. 6.5% received partial funding. For our empirical analysis, we have excluded the partially funded loans from our dataset. Comparing estimates with these loans included and coded as 1 if >0.5 funded and 0 if <0.5 funded yields the same substantive results. This results in 24594 observations.

we classify the information available to the lenders as verifiable or non-verifiable.

Verifiable Information. Lending Club collects information to verify borrowers' identity and assess credit worthiness. Using key indicators of credit worthiness, the lending platform classifies each potential borrower into seven grade categories, A to G, where A is the best and G the worst credit grade. These credit grades are determined by the lending platform as a function of the verifiable credit worthiness indicators. Interest rates are not set by the borrowers but instead are reflected by the credit grades; the relationship between the interest rate and risk is made salient by the lending platform. More specifically, much of the verifiable information (such as "FICO score range", "# of delinquencies in past 2 years", "# of credit inquiries in last 6 months", "revolving balance utilization", "debt-to-income ratio", "total credit lines", "# of derogatory public records", "# of public record bankruptcies", "zip code", "income", "length of employment", "employment title", etc.) is disclosed by the platform to potential lenders.

Non-Verifiable Information. While all borrowers are required to provide a "loan purpose" by selecting one of the categories offered by the platform, the thesis of our paper focuses on the provision of the *optional*, open-ended loan description. We hypothesize that the presence and length of the description serve as a proxy for the effort a borrower uses to signal. Accordingly, we create a "no_description" indicator variable that is equal to 1 when borrowers provide no description and 0 otherwise, and a "description length" variable that measures the number of words in the description. Of the 24,594 applications, 6372 have zero words (no loan description).

We create several other variables to control for variations in the content of the descriptions. With the classification and linguistic processing algorithms provided by SPSS software,[10] we identify the top concepts that appear in the loan description. Automatic classification helps control for the content coding biases that might arise with a researcher-developed classification

[10] We use the software package SPSS Text Analytics for Surveys 3.0, which is designed to extract and categorize free-text responses using natural language processing capabilities. For more information, see the IBM Software Business Analytics white papers, "Analyzing survey text: a brief overview" and "IBM SPSS Text Analytics for Surveys," available at www.ibm.com/software/analytics

scheme. The top six categories account for more than 95% of the observations, and account for both concreteness (e.g. budget) as well as attitudes of the description. Overall, the *budget* category represents the largest number of observations (65.7%). The most frequently used words in the descriptions in this category are “loan,” “pay,” “payment,” “paying,” “money,” and “rate.” Immediately following the budget category in magnitude is the *positive* category, classifying 36.4% of the observations (e.g., “excellent,” “good,” “timely”), followed by the *negative* category (e.g., “problem,” “bad,” “difficult”). For each of the top six categories, we create a dummy variable that indicates whether the observation falls in that category. A single loan description can belong to multiple categories. We use these 6 dummy variables to control for variations across the content of loan descriptions.

Finally, the data include the date the application was posted on the platform. We create a time trend variable to account for potential changes in loan funding behavior over time due to macroeconomic environments.

Model Free Evidence

We first show some model-free evidence that highlights the effects the optional description might have on the funding decision that is consistent with our theory. Table 1A shows, by each grade, the percentage of loans with no description, the overall loan funding percentage rate, and the funding percentage for those loans with no description. Table 1B shows for those loans with description, the description length mean, median and standard deviation for the funded loans and for the non-funded loans. Table 2 examines the overall funding percentages for loans with different description lengths.

(Table 1A, 1B and 2 Go Here)

The evidence shows that 1) loans with no descriptions are funded with higher probability than those with descriptions; 2) if the description is provided, the length improves the probability of funding, though the probability is less than those with no description. These nonparametric analyses point to a correlation between description length and loan funding. Lastly, we see that high credit-grade borrowers are more likely to provide no descriptions. This pattern is consistent with our analytical model, which states that unobservable loan quality drives the signaling effort and is also imperfectly yet positively correlated with the verifiable information.

Table 3 shows the summary statistics for our dataset. The length of the description is relevant to the countersignaling account, because it can serve as a proxy for signaling effort. For those loans with description, the average length is 54 words, with upper 90th percentile at 139 words.

(Table 3 Goes Here)

The Non-Randomness of Description Decisions

Before we present the empirical analysis, we first address the matter that description provisions are non-random decisions and can be influenced by the factors that also influence the verifiable information. Recall that in our analytical model, borrowers first make description decisions based not only on their knowledge about their type, on their belief about how their verifiable information will be presented to the potential lenders by the platform, and on the expectation about how lenders would react to their description efforts. Those that are confident about the later realization of their verifiable information (the high-types) will choose to counter-signal, and those that are less confident will want to bolster the chances by providing a description, with the length varying based the cost and benefit of their description-writing effort for the borrower (higher cost benefit and hence longer description for medium-types, lower cost benefit and hence shorter description for low-types). Subsequently, lenders will observe the information provided by the borrowers and the platform and make their investment decisions.

To model this two step process and to control the non-randomness of the description writing decisions, we use a two-stage regression approach. In the first stage, we use a zero-inflated

Poisson regression to model borrowers' description decision as a function of the most salient-to-the borrower creditworthiness information (i.e. their own FICO scores). Then in the second stage, we use the residuals from the first stage regression as a control for the lenders' funding decisions. This modeling approach not only provides us with empirical evidence as to whether salient credit information would impact description decisions as we have hypothesized, but also controls in the second stage for any potential biases arising from endogenous description decisions.

We model borrower's description decisions (number of words) as zero-inflated Poisson (Lambert 1992, Greene 1994, Bohning et al. 1999), which we believe is appropriate for the description decisions with a large number of zeros. Zero inflated Poisson (ZIP) is a two-component mixture model combining a point mass at zero with a count distribution. It assumes that the excessive zeros (i.e. not providing descriptions) are generated by a separate process from the count values and that the excess zeros are modeled separately. Therefore, it has two parts, a Poisson count model and the logit model for predicting excess zeros.[11] Specifically, the first stage contains the following explanatory variables.

$$\text{description decision}_i = f(\text{FICO score}_i\text{ , loan\_length}_i\text{ , loan\_amount}_i)$$

We use FICO score in our dataset as we believe it is the most popular credit information and should be the most salient to borrowers and hence would impact his/her decisions to provide descriptions, and exclude it in the second stage.[12] We use also loan amount and loan length as large amounts or length might drive borrows to offer explanations. We perform this analysis within each credit grade.

[11] For robustness check to assess dispersion, we have also ran a zero-inflated negative binominal regression. The results are similar but the fit is slightly worse, suggesting that over-dispersion is not an issue.
[12] We choose FICO as an instrument for the following institutional reasons: 1) it is the most salient credit information that forms the borrower's belief of his/her quality, hence impacting his/her description decision; 2) the econometrist point of view needs to consider the lender's decision. The platform assigns credit grades based largely on FICO score, so credit grade already soaks up most of the effect of FICO score. The residual information should then be captured by other verifiable credit related information, such as debt-to-income ratio, past delinquency etc. Thus, conditional on credit grades and other verifiable information, FICO score should have negligible residual effect.

The results, shown in Table 4, suggest for each of the credit grades, that borrower's with higher FICO scores are less likely to offer a description, but for those that do write a description, higher FICO score will result in longer description length. The result is consistent with our theory for borrowers.

(Table 4 Goes here)

Model Specification for Loan Funding Outcome

In the second stage of our empirical analysis we model the loan funding outcome and test for the existence of countersignaling from the lenders' perspective. Specifically, we wish to answer the following question: do lenders infer quality within a risk-reward class by taking into account the borrower's decision to exert effort to provide unverifiable loan descriptions? To answer this question, we include all information about the borrower that was available to the lender to avoid omitted variable biases. We performed a variance inflation factor (VIF) analysis to address potential multicollinearity. The resulting model specification exhibited VIFs of less than 5 for all variables, so this model was unlikely to suffer from multicollinearity concerns. Interest rates correlated highly with credit grade (.933), so we could only include one of these two variables. After testing, the specification that provided the best model fit was the one that used credit grade instead of interest rates. Both models share the same qualitative results.

In summary, for the i^{th} loan request, the model we test is:

$$\begin{aligned}\text{loan\_funding\_outcome}_i = f(&\text{loan\_amount}_i\ ,\ \text{loan\_length}_i\ ,\ \text{\#\_open\_credit\_lines}_i,\\ &\text{\#\_delinquencies\_past\_2\_years}_i, \text{\#\_total\_credit\_lines}_i,\ \text{revolving\_balance\_util}_i,\\ &\text{monthly\_income}_i, \text{debt-to-income\_ratio}_i, \text{home\_ownership\_status}_i, \text{state\_residence}_i\\ &\text{\#\_credit\_inquiries}_i,\ \text{"currency"}_i, \text{"buying"}_i\ ,\ \text{negative}_i\ ,\ \text{positive}_i\ , \text{"date"}_i\ , \text{"budget"}_i\ ,\\ &\text{credit\_grade}_i\ ,\ \text{description\_length}_i\ ,\ \text{description\_length}^2_i\ ,\\ &\text{"no\_description"}_i\ , \text{time trend}_i, \text{residual}_i)\end{aligned}$$

Estimation and Model Comparison

We estimate the model for loan funding outcome using logistic regression on funding, and we compare our proposed model against two benchmark models that vary in their degrees of signaling. The first benchmark model assumes that the funding outcomes are based solely on verifiable information (controlling for loan amount, loan length, and date of the loan request, as captured in the "time trend" variable). Then the second benchmark model accounts for the length of description but ignores the effect of counter-signaling, so it excludes the "no_description" variable. For ease of presentation, we first run an aggregate model (using credit grades as dummies) for model comparison to assess model fit with and without countersignaling and present the result in Web Appendix 3 (Table W1 – Model Comparison for Loan Funding), Then, as our model is within credit grade, we present the loan funding result for each individual credit grade in Table 5.

Table W1 shows the estimated parameters for the lending decision across the proposed and benchmark models. According to the Bayesian information criterion (BIC) which compares nested models and penalizes model complexity, the best fitting model is the proposed model which accounts both the effect of non-verifiable information as well as counter-signaling. All the parameters from the proposed model, as well as those of the two benchmark models, are in the expected direction and support our theory. For instance, borrowers with lower credit grades are less likely to be funded, and we find a residual impact of verifiable information after controlling for credit grades (e.g., borrowers with higher debt-to-income ratios, more past delinquencies, and more inquiries in the last six months are less likely to be funded). These results provide face validity for our analyses and show that lenders base their decisions on multiple pieces of information, in addition to the summary credit grade. Loans of larger amounts and longer terms also are less likely to be funded, but we find no impact of the borrower's state of residence on funding decisions. For time trend, we use the year indicator[13] and find that compared to 2007, all subsequent years result in lower funding likelihood, which reflects increasing lender caution as

[13] We also run a model using month/year indicators and quarter/year indicators, both models insignificant results for these indicators.

economy proceeded deeper into the recession after the bankruptcy of Lehman Brothers.

Several interesting insights stem from these model comparisons. First, the two models that account for the effects of non-verifiable information (proposed and benchmark model 2) fit the data better than the model that accounts only for verifiable information (benchmark model 1), Therefore, non-verifiable information influences loan funding decisions and is not viewed by lenders as uninformative cheap talk (Crawford and Sobel 1982). For instance, mentions of "budget" in the description might signal concreteness of financial planning and thus increases funding likelihood. Second, in our proposed model, the parameter estimate for "no_description" is positive and statistically significant. We show this effect to hold across all credit grades in Table 5, showing that within each credit grade and conditional on verifiable information, borrowers who do not provide a loan description are more likely to have their loans funded than those who provide a loan description, in support of H_1.

(Table 5 Goes here)

We also find a positive, statistically significant parameter value for "description_length" and a negative, statistically significant parameter value for "description_length^2." Once borrowers decide to provide a reason for the loan request, their chances of getting funded increase with the number of reasons, in a concave manner (i.e., decreasing returns to the number of words), in support of H_3.

We then run separate models for each credit grade, and the same patterns emerge for each credit grade, we present these results graphically in Figure 2, using the parameter estimates from the separate regressions. The patterns in Figure 2 are consistent with the prediction in our analytical model. Specifically, refraining from giving loan description is a countersignal of high quality and is rewarded by lenders with greater funding likelihood. When borrowers provide descriptions though, lengthier descriptions are perceived more favorably than shorter ones. Together, these results are consistent with the proposed countersignaling theory. The statistical significance in a number of interactions between description length and verifiable borrowers' creditworthiness information goes in line with our theory that some consumers naively use the combination of description length and verifiable creditworthiness information to assess the quality of the loan requests in the platform.

(Figure 2 Goes Here)

A critical element of our theory is that lenders correctly infer loan descriptions as a mechanism for strategic information transmission. We next consider ex post loan performance, to test H_2 and H_4.

Loan Description and Borrower Performance

The relationship between lending decisions and the optional borrower-provided non-verifiable information empirically supports countersignaling theory. In this section, we provide further support for countersignaling as an explanation of loan funding outcomes in P2P lending. An integral part of countersignaling theory is that the interpretation of the signal should be consistent with the strategy and type of sender. Otherwise, lenders would learn not to trust the signal, and the descriptions would become uninformative. We therefore turn our attention to testing the observed ex post performance of the loans (i.e., long-term performance, as measured by payment delinquency). With this analysis, we can determine whether, within a given credit grade, borrowers who countersignal are of higher quality, as indicated by a lower likelihood of delinquency.

Because high-quality borrowers should countersignal and refrain from providing a description, they also should be less likely to be delinquent on their loans. Countersignaling theory also suggests that once a description is provided, a longer description should be negatively correlated with loan delinquency. We describe the loan performance measurement and specify the model to test our loan performance prediction based on countersignaling theory. We perform the analysis using the funded applications from our dataset.

Loan Performance. 6% of the loans are non-current and exhibit the following four statuses of late (16–30 days), late (31–120 days), charged off, or default, which represent various stages of delinquency. We pool these four categories, due to the sparseness of these data points, and create a binary "delinquency" variable that is equal to 1 if the loan is delinquent and 0 otherwise. We present in Table 6 model free evidence showing delinquency % by grades and compare the % across various description length. Preliminary evidence suggests that having no description results in the lowest delinquency rate, followed by long descriptions, with short-descriptions having the highest rate.

(Table 6 Goes Here)

To examine the effect of countersignaling on the probability of delinquency, we model the delinquency of loan *i* with the same two step approach, with zero-inflated Poisson as the first stage. The second stage is a logistic regression with delinquency as the dependent variable:

$$\begin{aligned}\text{delinquency}_i = f(&\text{loan\_amount}_i, \text{loan\_length}_i, \text{\#\_open\_credit\_lines}_i,\\ &\text{\#\_delinquencies\_past\_2\_years}_i, \text{\#\_total\_credit\_lines}_i, \text{revolving\_balance\_util}_i,\\ &\text{monthly\_income}_i, \text{debt-to-income\_ratio}_i, \text{home\_ownership\_status}_i,\\ &\text{state\_residence}_i, \text{\#\_credit\_inquiries}_i, \text{"currency"}_i, \text{"buying"}_i, \text{negative}_i,\\ &\text{positive}_i, \text{"date"}_i, \text{"budget"}_i, \text{credit\_grade}_i, \text{description\_length}_i,\\ &\text{description\_length}^2_i, \text{"no\_description"}_i, \text{time\_trend}_i)\end{aligned}$$

Similar to the loan funding results, we present in Web Appendix 3 (Table W2 – Model Comparison for Loan Funding) the model comparisons results based on aggregate analysis using credit grade as dummies, then present our full results in Table 7, broken down by each credit grade.

Table W2 shows the estimated parameters for the likelihood of the loan being delinquent. We compare the proposed model with the same benchmark models used in the funding outcome analysis in Table W1 to examine the extent of the impact of countersignaling on loan performance. Overall, the results confirm the convergence between the funding decisions and delinquency rates. The proposed model provides the best fit for the data, based on the BIC criterion, in support of the predicted, non-monotonic relationship between the number of words in the loan description and the borrower's quality (in terms of delinquency).

As Table 7 shows, within each credit grade, the coefficient for the "no_description" variable is negative and statistically significant, providing evidence that borrowers who provide no description in their loan requests are less likely to be delinquent in their payment than are those providing short descriptions, in support of H_2. The coefficient for "description_length" is negative and statistically significant, which demonstrates that borrowers who provide lengthier descriptions are less likely to be delinquent than are borrowers who provide shorter descriptions, in supports of H_4.

(Table 7 Goes Here)

The graphic depiction of the results in Figure 3 uses the parameter estimates from Table 6. Combined with the loan funding results, these findings support the four hypotheses predicted by countersignaling theory with respect to how lenders correctly discriminate borrowers' types using the signal associated with the length of their loan description.

(Figure 3 Goes Here)

Recall our theoretical model assumes there is proportion of naïve lenders, and the analytical results hold even for a very small proportion of such lenders. Our empirical evidence supports our theory for the existence of some naïve lenders. In reality, even in the world of institutional lending, the fact that some funds perform better than others is an indication that there are variations in the processing of financial and creditworthiness information.

We now describe a series of robustness checks and discuss competing explanations.

Robustness Checks

To provide further evidence for the theory, we perform the following robustness checks:

1) Tested various specifications of "no_description" and "description_length".

2) Checked if the effect of "no_description" is unique to zero words.

3) Tested for the influence of repeat borrowers.

4) Tested whether there exists evidence of other signal mechanisms (e.g., loan terms and loan amounts).

5) Checked the potential effect of description content and quality.

First, to check 1), we tested various combinations of “no_description” and “description_length” in the model. The best fitting model is one that includes “no description”, “description length” and “description length^2”. As all three variables are significant, excluding one would results in a worse fit. The cubed term “description length^3” is not significant, which indicates that the probability will not invert and increase after certain number of words. Interaction terms of “no_description” and “description_length” with credit grades are not significant, suggesting that although no_description is important within each credit grade, its effects do not differ significantly across different credit grades. We also discretized “description_length” in 3 to 5 groups and re-run the model using these indicator variables (e.g. short, medium and long). We also teste an empirical model in which description lengths were classified as H, M, or L instead of considering a continuous classification. More specifically, lumped those descriptions less than the median length as L type, description longer than the median length as M type, and the absence of descriptions as H type. The substantive results hold with these discretization efforts (that is, M type results in higher likelihood of funding compared to L type), albeit with a slightly worse fit compared to using continuous number of words as in the proposed model. This result is expected, as we do not a priori know the cut-offs for the discretization, and the continuous type offers more flexibility. Finally, we included interaction between credit grades and the description variables, as well as interaction terms between description variables and other credit information of borrowers, both as continuous word length as well as discretized buckets, and found no effect of the interaction.

Next, to check 2), we investigated whether the proposed model correctly captured a unique zero-word effect of no loan description rather than an alternative effect caused by words fewer than some other threshold, we compared the proposed model with models including dummy variables that reflected whether the descriptions contained fewer than X number of words, where X = 3, 5, or 10 words, while holding everything else the same. We find that as "X" increases from 0, the effect size of the dummy variables "X_words_or_less" decreases and the model fit worsens. Therefore, providing no loan description is a unique signal. This phenomenon holds for both the loan funding decision and the ex post likelihood of loan delinquency. The result of this check is available in the Web Appendix 4.

As mentioned earlier, the lenders and borrowers are anonymous to each other, and the platform does not provide information about repeat borrowers or past performance indicator on the platform. Nevertheless, to check 3), we empirically investigated the presence of repeat borrowers by performing exact matches on time-invariant demographic and behavioral information such as income, credit grade, state of residence, earliest credit line opened, home ownership. We found 120 borrowers meet the criteria suggesting they may have applied twice, representing less than 0.5% of applications, and that there were no matches to indicate a borrower has applied three times or more. When we exclude these 120 observations from the data, the same results hold.

To check 4), we considered two situations. First, we investigated whether borrowers signal through loan amount or loan terms (durations) and whether certain lenders might treat loans of varying magnitudes differently (e.g., larger loans might signal a more responsible borrower, with strong repayment ability and consequently easy access to outside lending markets). To do so, we have divided the loan amounts into quintiles and created dummy variable for each of the following quintiles: amount ≤ \$500; \$500 < amount ≤ \$5000; \$5000 < amount ≤ \$9000; \$9000 < amount ≤ \$15000; \$15000 < amount ≤ \$25000. We included these indicator variables in addition to the amount, and found that relative to the baseline of the first quintile (amount ≤ \$500), the other four quintile dummy variables are non-significant. This indicates that, aside from the fact that larger loan amounts tend to decrease funding probability, there is no significant signaling values with loan amounts that would lead to lenders to behave differently. The same non-significant result also holds on the loan performance side, confirming that loan magnitude has no significant effect on the borrowing and lending behavior in the platform we study.[14]

In addition, to rule out potential signaling via loan length we run the model on a subset of the data from the time Lending Club did not allow borrowers to request 60 month loans, hence all loans were 36 months in duration. All of the substantive results are similar between the subset of the data and the full dataset.

Finally, 5) assessed whether shorter or longer descriptions contain specific words that drive decisions. We modeled the description length as a dependent variable on content dummies and found no significant differences between the effects of different content dummies. Correlations between description lengths and various content dummies ranged from .16 to .22; we observed no particular content variable that had a systematically higher impact on description_length.

To assess whether different description lengths exhibit different qualities of writing, we randomly selected 40 descriptions from each of the following groups: less than 33 percentile or 12 words (short group), between 33 and 66 percentile or 13 to 39 words (medium group), and greater than 66 percentile or 40+ words (long group), and conducted survey of quality with 153 Amazon Mechanical Turk respondents. We asked them to rate the quality of and the effort exhibited in the description. Correlation is 0.7 between description_length and quality, 0.88 between description_length and perceived effort, and 0.92 between effort and quality. Regressing perceived effort on description_length yields R-squared of 0.7, and regressing quality on description_length yields

[14] It is worth mention that even if some degree of signaling by amount was operating in the platform, undetected by our empirical analysis, such signaling effort would not invalidate our countersignaling results. As discussed in the literature (Feltovich, Harbaugh, and To 2002), both signaling and contersignaling can coexist in the same strategic interaction of economic agents.

R-squared of 0.5. This shows that longer descriptions are seeing as both more effortful and higher quality, and perceived that it takes more effort to write quality descriptions.[15]

In summary, the robustness analyses lend support for H_1–H_4. We now discuss how our counter-signaling theory and results compare with other prior theories.

General Discussion

Following this theoretical and empirical analysis of countersignaling in P2P lending, we discuss candidate explanations for the observed phenomena, and describe how these theories do not explain the totality of our results.

Persuasion. According to persuasion theory, the provision of descriptions might be interpreted as a compliance issue. Individuals and companies acting as persuaders often attempt to increase compliance with their requests by offering reasons why others should behave accordingly (Langer et al.1978; Petty and Caccioppo 1984). These predictions they could accommodate the behavior of naïve consumers in the analytical model, but cannot accommodate the effect of no_description. These predictions also cannot account for the observed relationship between loan description and ex post loan delinquency.

Preemptive behavior. In mortgage applications, if there is something negative in a borrower's verifiable information, the lender often asks the borrower to write an explanation for this potential blemish in the credit record. Borrowers in this setting might act preemptively and, without being asked, write to explain any negative verifiable information. However, if such preemptive behavior were successful in attracting lenders, it could explain the positive relationship between ex post loan defaults and the presence of a description, but it would not explain the negative relationship between funding and the presence of a description. Moreover,

[15] We also note that the respondents are all likely to have limited lending experience, so in absence of other cues, they make take the length of description a proxy for effort and quality of the writing, thus behaving in the way we theorized naïve lenders would do.

this explanation cannot address the non-monotonicity of both borrower and lender behavior.

Sufficiency of verifiable information. A final alternative explanation is that some borrowers do not provide explanations because their applications get funded without any further explanation. That is, similar to the countersignaling explanation, only borrowers who are truly confident that positive information about their profile would be revealed by the platform can afford not to send a description, so they forgo the opportunity to persuade some naïve lenders. However, in this explanation, the empirical model accounting for all other information would reveal no effect of description length as it has no signaling effect. Such an explanation is not consistent with the empirical findings.

In summary, none of these explanations can explain the totality of the non-monotonic pattern of behavior we observe and the ex post efficiency of borrowers' and lenders' decisions.

Theoretical Contributions

In this research, we find the strategic transmission of non-verifiable information by borrowers is an important influence on P2P lending platforms. Lenders make decisions on loan investments on the basis of both verifiable information and the optional non-verifiable information provided by borrowers, and such decisions are validated by subsequent loan performance.

We show that the optional loan description and its length both influence loan funding, so the provision of such information is not viewed as uninformative, cheap talk. Borrowers and lenders use non-verifiable information in a way that is consistent with the properties of a countersignaling equilibrium. Specifically, our empirical evidence suggests that lenders view those who provide no loan description as high-quality borrowers. Moreover, medium-quality borrowers can distinguish themselves from lower-quality borrowers by exerting greater effort to provide more non-verifiable information (i.e., longer descriptions) than lower-quality borrowers do. It is important to note that we are not suggesting the act of writing requires so much effort that someone might not undertake it. Rather, the marginal benefit of writing a description (or not) is the probabilistic improvement of obtaining funding, and this improvement depends on the quality of the borrower. Thus, the marginal effort cost of writing need not be so great as to offset the value of the loan completely; rather, borrowers weigh the however small effort costs relative to the associated marginal improvement in funding probability.

The evidence from loan funding and ex post loan performance suggest that countersignaling is a robust explanation of lending behavior in a P2P environment. Even after controlling for all the information that lenders encounter at the time they make funding decisions, loan applications with no loan description turn out ex post to be higher-quality loans within a given credit grade, as measured by their lower delinquency rates. Applications that provide longer descriptions are less likely to be of lower quality than applications with short descriptions. Thus, lenders correctly interpret the signaling effect of non-verifiable information when making loan funding decisions, because their funding decisions correspond with ex post loan quality.

The convergence between lenders' view of non-verifiable information and ex post loan performance suggests that non-verifiable information serves as a mechanism to attenuate information asymmetry regarding unobserved borrower quality. Borrowers who shrewdly recognize the signal/countersignal properties of non-verifiable information are rewarded, on average, by lenders. Lenders who correctly identify the signals will be rewarded with better performing loans. An essential characteristic of marketing is understanding the marketplace and the factors that encourage consumers to engage in it (Bradford 2015). We identify a robust factor that affects both consumer exchange behaviors and the makeup of the marketplace. As P2P lending becomes a big part of the current financial landscape, our work contributes to a growing research that provides empirical insights into the intricacies of consumer lending decisions, as the result of availability of large scale datasets (Galak, Small, and Stephen 2011; Stephen and Galak 2012), something that was difficult to do a few years ago.

Managerial Implications

Substantively, we provide several insights that can be broadly contextualized to the parties involved in P2P transaction platforms such as Lending Club, eBay, Upwork, and AirBnB. First and foremost, designers of the platforms should realize that verifiable information is not enough to eliminate asymmetric information and reveal the true nature of a seller. Although offline communication between buyer and seller could reduce the information asymmetry, the platform runs the risks of prolonging the transaction time and the risks of the two parties transacting outside of the platform, which would reduce platform revenue; thus, platforms always need to balance the benefits of allowing parties to transmit information with the risks of potentially losing platform revenue (e.g. Amazon and eBay both block email addresses in communication and cancel accounts that try to do business outside). As countersignaling behavior has quality content platforms that disallow communication (such as Lending Club) should improve their proprietary seller quality score by incorporating countersignaling behavior in their computations (e.g. those who countersignal could get a 3, long description a 2, short description a 1). This scoring algorithm should remain proprietary so participants could not game the system (i.e. akin to the review filtering algorithm on Yelp, or the "portfolio health meter" by Lending Robot). As P2P platforms become more important, platforms that can fine-tune their rating system to better reflect true risks and to resolve information asymmetry more effectively without communicating with each other offline would: 1) ensure the revenue stays on the platform, and 2) instill more confidence among participants.

For sellers who are reading this article, those who are confident in the quality of their profile as a provider should abstain from providing unsolicited information about their superior capabilities (such as repayment ability, reliability in terms of packaging quality and speed of delivery, or the quality of their work as a free-lance worker) when verifiable information about their quality as a provider is available. For these sellers, volunteering non-verifiable information can lower their probability of making a deal (e.g., getting a loan funded, renting their apartment, or making a sale). To illustrate, high-quality sellers on eBay should refrain from providing superfluous information and instead let their “power-seller” status do the talking; similarly, professionals who market their services in freelancing marketplaces (e.g., Upwork) should let their references, performance records, and professional portfolios speak for themselves, instead of providing non-verifiable descriptions about their skills or the quality of their services.

Limitations and Future Research

We note some limitations and suggest avenues for future investigation. During the time frame in our analysis, the platform was open only to individual, instead of institutional investors. Starting in 2012, the platform was opened to institutional investors and by December 2015 consists of 33% of the funding. One can examine how asymmetric information is resolved differently in P2P vs. P2B. Our current model suggests that for institutional investors, the percentage of sophisticated investors would perhaps be higher than those individual investors and the impact of countersignaling would be even stronger. Further, in light of the recent Lending Club scandal in the P2B domain,[16]. In light of the event, we have performed an analysis in which we exclude data from the month of December 2009 and observed that the results do not change (which is to be expected given the small proportion of loans relative to the entire transactions in the platform). Nevertheless, one can examine how the lender decision model and the weights they place on various pieces of information evolve. We suspect that in the future, when the scandal reveals that certain variables within the loan might be manipulated, lenders would deemphasize these variables and shift weight to other variables to form their decisions. Even so, we believe that as the public company and the industry as a whole conduct business in legitimate fashion and

[16] In April 2016, , an investigation indicated that the now ousted CEO Renaud Laplanche and three of his family members had taken 32 inside loans (totaling $720,000)   to inflate such a growth, a practice Silicon Valley insiders call “growth hacking”. As a result of this scandal, Lending Club shares plunged 51% the week following the reporting of the scandal as a result of institutional investors suspending debt purchases and the announcement of a probe by the U.S. Justice Department probe. (http://www.bloomberg.com/news/features/2016-08-18/how-lending-club-s-biggest-fanboy-uncovered-shady-loans)

provide values to both lenders and borrowers. During our analysis time frame, buyers and sellers cannot communicate off line, and the platform's disclosure rules have not changed. Further research can examine a period when there is change in platform's communication rules (e.g. change of sequence, opening or closing avenues of disclosure), which can shed light on how signaling mechanisms shift to alleviate asymmetric information. Finally, further research should examine whether characteristics specific to other P2P contexts moderate or replicate the signaling/countersignaling dynamics present in peer-to-peer lending.

References

Akerlof, George A. (1970). “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism,” *Quarterly Journal of Economics*, 84(3), 488-500.

Backus, Matthew, Tom Blake, and Steven Tadelis (2015), “Cheap Talk, Round Numbers, and the Economics of Negotiation,” *NBER Working Paper No. 21285*.

Bart, Iacov Y., Venkatesh Shankar, Fareena Sultan, and Glen L. Urban (2005), “Are the Drivers and Role of Online Trust the Same For all Web Sites and Consumers? A Large-scale Exploratory Empirical Study,” *Journal of Marketing*, 69(4), 133-152.

Bodie, Zvi, Alex Kane, and Alan J. Marcus (2009), *Investments* [8th ed], New York, NY: McGraw Hill/Irwin.

Böhning, Dankmar; Dietz, Ekkehart; Schlattmann, Peter; Mendonca, Lisette; Kirchner, Ursula (1999). “The zero-inflated Poisson model and the decayed, missing and filled teeth index in den-tal epidemiology,” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 162(2), 195-209.

Bradford, Tonya W. (2015), “Beyond Fungible: Transforming Money into Moral and Social Resources,” *Journal of Marketing*, 79(2), 79-97.

Brus, Brian (2010), “Oklahoma teachers retirement exec reflects on change,” *The Oklahoma City Journal Record*, July 13.

Clements, Matthew T. (2010), “Low Quality as a Signal of High quality,” *Economics: The Open-Access, Open-Assessment E-Journal* No. 2010-20.

Crawford, Vincent P. and Joel Sobel (1982), “Strategic Information Transmission,” *Econometrica*, 50(6), 1431-1451.

Daughety, A. F. and J. F. Reinganum (1995), Product Safety: Liability, RandD, and Signaling. *American Economic Review*, 85(5), 1187-1206.

Desai, Preyas (2000), “Multiple Messages to Retain Retailers: Signaling New Product Demand,” *Marketing Science*, 19 (4), 381-389.

Duarte, Jefferson, Stephan Siegel, and Lance Young (2012), Trust and Credit: The Role of Appearance in Peer-to-peer Lending. *Review of Financial Studies*, 25(8), 2455-84.

Fabozzi, Frank J. (2013), Bond markets, analysis and strategies [8th ed], Upper Saddle River, NJ: Pearson/Prentice Hall.

Feltovich, Nick, Richmond Harbaugh and Ted To (2002), “Too Cool for School? Signaling and Countersignaling,” *The Rand Journal of Economics*, 33(4), 630-649.

Freedman, Seth M. and Ginger Z. Jin (2011), “Learning by Doing with Asymmetric Information: Evidence from Prosper.com,” *NBER Working Paper 16855*.

Galak, Jeff, Deborah Small, Andrew T. Stephen (2011), “Microfinance Decision Making: A Field Study of Prosocial Lending,” *Journal of Marketing Research*, 48(SPL), S130-S137.

Greene, William H. (1994), “Some Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models,”. *Working Paper EC-94-10*: Department of Economics, New York University.

Greiner, Martina E., and Hui Wang (2010), “Building Consumer-to-consumer Trust in E-finance Marketplaces: an Empirical Analysis,” *International Journal of Electronic Commerce*, 15(2), 105-136.

Hancock, Jeffrey, Lauren Curry, Saurabh Goorha, and Michael Woodworth (2008), “On Lying and Being Lied To: A Linguistic Analysis of Deception in Computer-Mediated Communication,” *Discourse Processes*, 45(1), 1-23.

Herzenstein, Michal, Utpal M. Dholakia, and Rick L. Andrews (2011), “Strategic Herding Behavior in Peer-to-peer Loan Auctions,” *Journal of Interactive Marketing* 25(1), 27-36.

Herzenstein, Michal, Scott Sonenshein, and Utpal M. Dholakia. (2011), “Tell Me a Good Story and I May Lend You My Money: The Role of Narratives in Peer-to-peer Lending Decisions,” *Journal of Marketing Research*, 48 (November), S138-S149.

Inderst, Roman and Marco Ottaviani (2012), “How (Not) to Pay for Advice: a Framework for Consumer Financial Protection,” *Journal of Financial Economics*, 105(2), 393-411.

Kawai, Kei, Ken Onishi, and Kosuke Uetake (2014), “Signaling in Online Credit Markets,” *Available at SSRN 2188693*.

Kumar, V. (2015), “Evolution of Marketing as a Discipline: What Has Happened and What to Look Out For,” *Journal of Marketing*, 79(1), 1-9.

Lambert, Diane (1992). “Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing,” *Technometrics*, 34 (1), 1-14.

Landsman, Vardit and Stefan Stremersch (2011), “Multihoming in Two-Sided Markets: An Empirical Inquiry in the Video Game Console Industry,” *Journal of Marketing*, 75(6), 39-54.

Langer, Ellen, Arthur Blank, and Benzion Chanowitz (1978), “The Mindlessness of Ostensibly Thoughtful Action: the Role of "Placebic" Information in Interpersonal Interaction,” *Journal of Personality and Social Psychology*, 36(6), 635-642.

Lewis, Gregory (2011), “Asymmetric Information, Adverse Selection and Online Disclosure: The Case of eBay Motors,” *American Economic Review*, 101 (4), 1535-1546.

Liebowitz, Stan J. and Stephen E. Margolis (2005), “Seventeen Famous Economists Weigh in on Copyright: The Role of Theory, Empirics, and Network Effects,” *Harvard Journal of Law and Technology*, 18(2), 435-457.

Lin, Minfeng, N.R. Prabhala, and Siva Viswanathan (2013), “Judging Borrowers by the Company They Keep: Social Networks and Adverse Selection in Online Peer-to-Peer Lending,” *Management Science*, 59(1), 17-35.

Li, Lingfang (Ivy), Steven Tadelis, and Xiaolan Zhou (2016), “Buying Reputation as a Signal of Quality: Evidence from an Online Marketplace,” *NBER Working Paper 22584*.

Mayzlin, Dina and Jiwoong Shin (2011), “Uninformative Advertising as an Invitation to Search,” *Marketing Science*, 30(4), 666-685.

Michels, Jeremy (2012), “Do Unverifiable Disclosures Matter? Evidence from Peer-to-Peer Lending,” *The Accounting Review*, 87(4), 1385-1413.

Milde, Hellmuth and John G. Riley (1988), “Signaling in Credit Markets,” *Quarterly Journal of Economics*, 103(1), 101-129.

Milgrom, Paul and John Roberts (1986), “Price and Advertising Signals of Product Quality,” *The Journal of Political Economy*, 94(4), 796-821.

Orzach, Ram, Peter B. Overgaard, and Yair Tauman (2002), “Modest Advertising Signals Strength,” *The RAND Journal of Economics*, 33(2), 340-358.

Paravisini, Daniel, Veronica Rappoport, and Enrichetta Ravina (2016), “Risk aversion and wealth: Evidence from person-to-person lending portfolios,” *Management Science*, 63(2), 279-297.

Petersen, J. Andrew, Tarun Kushwaha, and V. Kumar (2015), “Marketing Communication Strategies and Consumer Financial Decision Making: The Role of National Culture,” *Journal of Marketing*, 79(1), 44-63.

Petty, Richard E. and John T. Cacioppo (1984), “The Effects of Involvement on Response to Argument Quantity and Quality: Central and Peripheral Routes to Persuasion,” *Journal of Personality and Social Psychology*, 46(1), 69-81.

PWC (2015), “Peer Pressure: How Peer-to-Peer Lending Platforms are Transforming the Consumer Lending Industry”, Feb 2015, (https://www.pwc.com/us/en/consumer-finance/publications/assets/peer-to-peer-lending.pdf).

Pope, Devin G. and Justin R. Sydnor (2011), “What's in a Picture? Evidence of Discrimination from Prosper.com,” *Journal of Human Resources*, 46(1), 53-92.

Rasmusen, Eric (2006), *Games and information: An introduction to game theory*. Oxford, UK: Blackwell Publishing.

Ravina, Enrichetta (2012), “Love & Loans: The Effect of Beauty and Personal Characteristics in Credit Markets,” *Working paper*. Columbia University.

Schlosser, Ann E., Tiffany B. White, and Susan M. Lloyd (2006), “Converting Web Site Visitors into Buyers: How Web Site Investment Increases Consumer Trusting Beliefs and Online Purchase Intentions,” *Journal of Marketing*, 70(2), 133-148.

Shavell, Steven (2010), “Should Copyright of Academic Works be Abolished?,” *Journal of Legal Analysis*, 2(1), 301-358.

Sinha, Rajiv K., Fernando S. Machado, and Collin Sellman (2010), "Don't Think Twice, It's All Right: Music Piracy and Pricing in a DRM-Free Environment," *Journal of Marketing*, 74(2), 40-54.

Sonenshein, Scott, Michal Herzenstein, and Utpal M. Dholakia (2011), "How Accounts Shape Lending Decisions through Fostering Perceived Trustworthiness," *Organizational Behavior and Human Decision Processes*, 115(1), 69-84.

Spence, Michael (1973), "Job Market Signaling," *The Quarterly Journal of Economics*, 87(3), 355-374.

Stephen, Andrew T., Jeff Galak (2012) The Effects of Traditional and Social Earned Media on Sales: A Study of a Microlending Marketplace. Journal of Marketing Research: October 2012, Vol. 49, No. 5, pp. 624-639.

Stiglitz, Joseph E. and Andrew Weiss (1981), "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 71(3), 393-410.

Tadelis, Steven, and Florian Zettelmeyer (2015), "Information Disclosure as a Matching Mechanism: Theory and Evidence from a Field Experiment," *American Economic Review*, 105(2), 886-905.

The Economist (2014), "Banking without Banks," *Economist*, (March 1st), 70-71.

The New York Times (2015), "Airbnb Grows to a Million Rooms, and Hotel Rivals are Quiet, for Now," *New York Times*, (May 11th).

Yadav Manjit S. and Paul A. Pavlou (2014), "Marketing in Computer-mediated Environments: Research Synthesis and New Directions," *Journal of Marketing*, 78(1), 20-40.

Zhang, J. and P. Liu (2012), "Rational Herding in Microloan Markets," *Management Science* 58(5), 892-912.

Zhu, Feng and Xiaoquan (Michael) Zhang (2010), "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing*, 74(2), 133-148.

Table 1A: Percentage of Loan Funded with and without Description

Credit Grade	% of loans with no description	Overall loan funding %	Funding % for loans with no description
A	29%	66%	71%
B	27%	57%	62%
C	25%	54%	63%
D	25%	53%	58%
E	23%	50%	56%
F	21%	48%	52%
G	16%	45%	51%

Table 1B: Mean and Median Description Length for Funded and Non-funded Loan Applications

Credit Grade	mean length for funded	median length for funded	mean length for non-funded	median length for non-funded	sd length for funded	sd length for non-funded
A	65	43	46	28	68	56
B	72	49	55	35	75	66
C	73	48	54	34	77	64
D	76	51	62	36	80	75
E	78	53	63	38	78	78
F	90	59	78	40	102	94
G	76	47	73	38	74	84

Table 2: Percentage of Loans Funded Varies with Description Length

Quintile	0 words	25%	50%	75%	90%
Length of description		6	28	72	139
% of loans funded	62.2	53.2	55.1	55.4	58.3

Table 3: Descriptive Statistics for Loan Applications

Total number of observations	26,314				
	Mean	**Std. Dev.**	**Lower 10%**	**Median**	**Upper 90%**
Loan amount applied ($)	10,541	6,755	3,000	9,000	21,000
Loan description length (words)	54	73	6	28	139

Percent funded	0% funded	35.0%
	100% funded	65.0%
Credit Grades	Grade A	16.8%
	Grade B	27.8%
	Grade C	23.9%
	Grade D	16.8%
	Grade E	9.1%
	Grade F	3.7%
	Grade G	2.0%
Loan Length	36 months	88.7%
	60 months	11.3%
Borrower Demographics	# of delinquencies in past 2 years	0.17
	# of credit inquiries in last 6 montl	1.59
	Revolving balance utilization	45.2%
	Monthly income ($)	6,017
	Debt-to-income ratio	12.1%
	Total credit lines	9
Home Ownership	None	1.2%
	Mortgage	39.3%
	Own	9.8%
	Rent	49.7%
Description	Contains "currency"	8.9%
	Contains "buying"	12.1%
	Contains negative words	18.4%
	Contains positive words	36.4%
	Contains date	4.4%
	Contains "budget"	65.7%
Loan Purpose	Car	3.5%
	Credit card	11.1%
	Debt consolidation	38.2%
	Educational	3.1%
	Home Improvement	7.2%
	House	1.5%
	Major Purchase	6.5%
	Medical	2.0%
	Moving	1.6%
	Other	14.0%
	Renewable Energy	0.2%
	Small Business	7.7%
	Vacation	0.8%
	Wedding	2.5%

Table 4: Model for Description Decision															
		Grade A		**Grade B**		**Grade C**		**Grade D**		**Grade E**		**Grade F**		**Grade G**	
		Estimate	**Std. Dev**	**Estimate**	**Std. Dev**	**Estimate**	**Std. Dev**	**Estimate**	**Std. Dev**	**Estimate**	**Std. Dev**	**Estimate**	**Std. Dev**	**Estimate**	**Std. Dev**
Poisson count model	Intercept	**5.409**	0.032	**4.746**	0.075	**4.395**	0.032	**3.124**	0.032	**2.567**	0.042	**1.835**	0.052	**1.707**	0.092
	FICO	**0.002**	0.001	**0.001**	0.001	0.000	0.001	**0.002**	0.001	**0.003**	0.001	**0.004**	0.001	**0.004**	0.001
	Loan length	**0.147**	0.017	**0.201**	0.006	**0.205**	0.008	**0.164**	0.006	**0.302**	0.008	**0.027**	0.010	**0.142**	0.019
	Loan amount	**-0.018**	0.001	**-0.013**	0.004	**-0.010**	0.000	**-0.006**	0.001	**-0.004**	0.001	**-0.003**	0.001	**-0.009**	0.001
Zero-inflation model (binomial with logit). 0 = no description	Intercept	**-2.503**	0.924	**-1.622**	0.743	**1.902**	0.992	-1.431	1.153	**-3.223**	1.859	**-2.248**	3.095	-5.194	0.527
	FICO	**-0.002**	0.000	0.000	0.001	**-0.006**	0.001	**-0.005**	0.002	**-0.004**	0.002	**-0.005**	0.002	-0.003	0.001
	Loan length	**0.483**	0.194	**0.663**	0.085	**0.989**	0.099	**0.734**	0.098	**1.050**	0.111	**1.115**	0.184	**1.416**	0.303
	Loan amount	-0.003	0.009	**-0.017**	0.005	**-0.011**	0.005	-0.011	0.006	-0.001	0.008	-0.004	0.011	0.017	0.017

***Bold estimates indicate signifance at the 0.05 level.**

Table 5: Model of Lending Decisions

Credit Grade		A		B		C		D		E		F		G	
		Estimate	**SD**	**Estimate**	**SD**	**Estimate**	**SD**	**Estimate**	**SD**	**Estimate**	**SD**	**Estimate**	**SD**	**Estimate**	**SD**
	Intercept	**2.856**	0.759	**2.194**	0.568	**2.265**	0.647	**2.374**	0.727	1.374	0.827	-0.285	0.312	-0.575	0.311
	# of open credit lines	0.007	0.019	0.006	0.014	-0.013	0.014	0.019	0.016	**0.056**	0.022	-0.025	0.035	-0.034	0.034
	# of delinquencies in past 2 years	**-0.199**	0.234	-0.018	0.104	**-0.180**	0.031	0.053	0.082	0.099	0.113	**-0.408**	0.176	**-0.286**	0.060
	Revolving balance utilization	-0.337	0.328	**-0.908**	0.183	**-1.185**	0.178	**-0.726**	0.201	**-1.152**	0.299	0.205	0.465	0.472	0.596
	Monthly income	**0.000**	0.000	**0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.000**	0.000
	Debt-to-income ratio	-0.380	0.964	**-1.657**	0.715	-1.129	0.756	**-1.964**	0.885	-1.239	1.250	**-4.736**	1.831	**-5.013**	2.547
	Total credit lines	0.010	0.008	0.006	0.005	0.009	0.006	0.012	0.007	-0.013	0.010	0.017	0.015	0.004	0.019
	Mortage	0.680	0.357	**0.820**	0.249	**1.141**	0.298	0.399	0.304	0.440	0.462	**1.850**	0.585	0.211	0.776
Home Onwership (compared to none)	Rent	-0.275	0.364	0.271	0.258	**0.715**	0.307	0.027	0.318	-0.354	0.484	0.929	0.836	0.676	0.859
	Own	**0.538**	0.156	**0.738**	0.248	**1.069**	0.296	0.460	0.300	0.083	0.457	0.987	0.783	-0.238	0.771
	Inquiries in last 6 months	**-0.138**	0.045	-0.022	0.028	**-0.043**	0.022	-0.021	0.024	**-0.059**	0.021	-0.068	0.045	-0.059	0.050
	Loan amount	**-0.128**	0.010	**-0.042**	0.004	**-0.154**	0.005	**-0.046**	0.006	**-0.056**	0.008	**-0.059**	0.013	-0.037	0.017
	Term (60 months = 1)	**-0.665**	0.208	**-0.558**	0.092	**-1.302**	0.110	**-0.800**	0.110	-0.233	0.147	**-0.704**	0.251	0.417	0.460
Time trend	Y2008	**-2.965**	0.603	**-3.198**	0.469	**-3.757**	0.522	**-3.400**	0.596	**-4.033**	0.324	**-3.285**	0.421	-1.769	0.485
	Y2009	**-2.603**	0.599	**-3.039**	0.466	**-3.740**	0.519	**-3.544**	0.592	**-4.120**	0.324	**-3.332**	0.485	-1.797	0.422
	Y2010	**-2.898**	0.599	**-3.214**	0.466	**-3.765**	0.520	**-3.685**	0.593	**-4.550**	0.355	**-4.112**	0.522	-1.837	0.442
	Credit card	**0.648**	0.181	**0.848**	0.163	**0.781**	0.190	**0.574**	0.242	0.113	0.372	-0.505	0.719	2.036	1.283
Loan Purpose	Debt consolidation	**0.740**	0.158	**0.848**	0.148	**0.720**	0.172	**0.943**	0.223	0.132	0.337	0.056	0.665	2.014	1.236
	Educational	-0.282	0.228	0.014	0.203	-0.111	0.224	-0.155	0.300	-0.413	0.429	-1.019	0.870	1.387	1.344
	Home Improvement	**0.748**	0.185	**0.824**	0.166	**0.528**	0.197	**0.707**	0.262	-0.345	0.377	-0.364	0.729	0.934	1.297
	House	0.104	0.345	**0.580**	0.253	-0.207	0.283	-0.320	0.352	-0.425	0.520	0.085	0.874	2.578	1.547
	Major Purchase	0.110	0.167	0.284	0.169	0.270	0.199	0.415	0.255	-0.365	0.391	-0.743	0.766	0.478	1.401
	Medical	-0.459	0.242	-0.010	0.227	-0.016	0.259	0.159	0.325	-0.286	0.475	-0.931	0.966	-0.205	1.512
	Moving	0.152	0.272	0.232	0.246	-0.126	0.276	0.316	0.367	-0.303	0.498	-0.237	0.884	0.982	1.635
	Other	**0.392**	0.163	**0.445**	0.155	**0.337**	0.178	0.408	0.234	-0.225	0.352	-0.184	0.694	1.578	1.265
	Renewable Energy	**1.306**	0.323	0.337	0.500	-0.742	0.576	**0.588**	0.425	-1.786	2.226	-1.572	3.956	-0.163	0.205
	Small Business	-0.133	0.223	-0.176	0.177	-0.337	0.198	-0.051	0.241	**-0.846**	0.358	-0.960	0.679	1.255	1.245
	Vacation	-0.170	0.307	0.221	0.333	0.164	0.381	0.027	0.469	-0.778	0.645	-1.988	1.175	1.041	1.645
	Wedding	**0.462**	0.269	0.287	0.207	**0.589**	0.246	**0.825**	0.302	-0.013	0.460	-0.655	0.861	-0.332	0.398
	Contains "currency"	-0.105	0.145	-0.047	0.106	-0.071	0.108	0.204	0.127	-0.539	0.485	-0.359	0.307	-0.134	0.361
	Contains "buying"	**0.240**	0.119	-0.066	0.094	0.039	0.102	0.176	0.121	-0.016	0.167	-0.249	0.273	0.068	0.318
Description Content	Contains negative words	-0.015	0.115	-0.075	0.078	-0.089	0.082	0.017	0.095	-0.075	0.135	-0.288	0.229	-0.072	0.266
	Contains positive words	0.048	0.091	-0.013	0.065	0.068	0.069	-0.025	0.081	-0.108	0.116	0.335	0.190	0.342	0.536
	Contains date	-0.144	0.205	-0.053	0.154	-0.170	0.160	0.319	0.181	-0.209	0.255	0.012	0.402	-0.274	0.305
	Contains "budget"	**0.325**	0.104	**0.192**	0.078	**0.242**	0.085	**0.240**	0.101	**0.363**	0.143	0.377	0.236	0.003	0.008
Description Length	Description length	**0.010**	0.003	**0.011**	0.002	**0.009**	0.002	**0.008**	0.002	**0.018**	0.003	**0.009**	0.005	**0.006**	1.44E-04
	Description length^2	**-1.34E-05**	0.000	**-1.41E-05**	0.000	**-1.48E-05**	0.000	**-1.17E-05**	0.000	**-2.07E-05**	0.000	**-9.30E-05**	0.000	**-9.14E-04**	4.35E-01
Counter signaling	No description	**1.378**	0.290	**1.415**	0.213	**1.275**	0.254	**0.832**	0.300	**0.651**	0.214	**0.426**	0.120	**0.522**	3.82E-03
	# of open credit lines	**3.00E-04**	2.15E-05	**2.18E-04**	3.40E-05	**1.75E-04**	3.44E-05	**1.77E-04**	4.50E-05	**7.62E-04**	2.27E-04	**4.84E-04**	2.60E-04	1.82E-04	3.82E-04
	# of delinquencies in past 2 years	-1.22E-03	3.54E-03	-1.35E-03	8.32E-04	2.80E-05	8.88E-04	**-1.08E-03**	1.01E-04	**-2.01E-03**	3.69E-04	**-2.22E-03**	5.41E-04	**-2.52E-03**	2.05E-04
Interaction terms. Descrition Length x	Revolving balance utilization	-1.62E-03	3.90E-03	7.99E-05	1.74E-03	**-3.79E-03**	1.73E-03	-1.83E-03	1.76E-03	**-5.75E-03**	2.65E-03	-1.50E-03	3.40E-03	1.96E-03	5.72E-03
	Monthly income	5.68E-08	1.31E-07	4.23E-08	5.20E-08	-1.10E-07	8.33E-08	3.60E-08	7.56E-08	-6.03E-08	1.63E-07	-8.98E-08	1.90E-07	**5.42E-06**	3.14E-07
	Debt-to-income ratio	-5.47E-03	1.18E-02	4.41E-04	7.25E-03	-5.92E-03	7.63E-03	-8.30E-03	8.02E-03	-7.30E-03	1.16E-02	**-2.19E-02**	3.69E-03	-1.42E-02	2.87E-02
	Total credit lines	1.34E-04	9.69E-05	4.82E-05	5.38E-05	8.61E-05	5.71E-05	4.01E-05	6.40E-05	2.84E-04	2.04E-04	-2.24E-04	1.22E-04	4.93E-05	1.89E-04
	Inquiries in last 6 months	-5.92E-04	5.38E-04	-2.48E-05	2.73E-04	8.53E-05	2.47E-04	1.72E-04	2.54E-04	**-6.46E-04**	3.06E-04	5.69E-04	3.88E-04	-3.55E-05	5.76E-04
	Residual from first stage	**0.066**	0.016	**0.090**	0.030	0.022	0.025	**0.071**	0.018	**0.065**	0.022	0.045	0.027	**0.075**	0.022

***Bold estimates indicate signifance at the 0.05 level.**

Table 6: Delinquency By Credit Grades

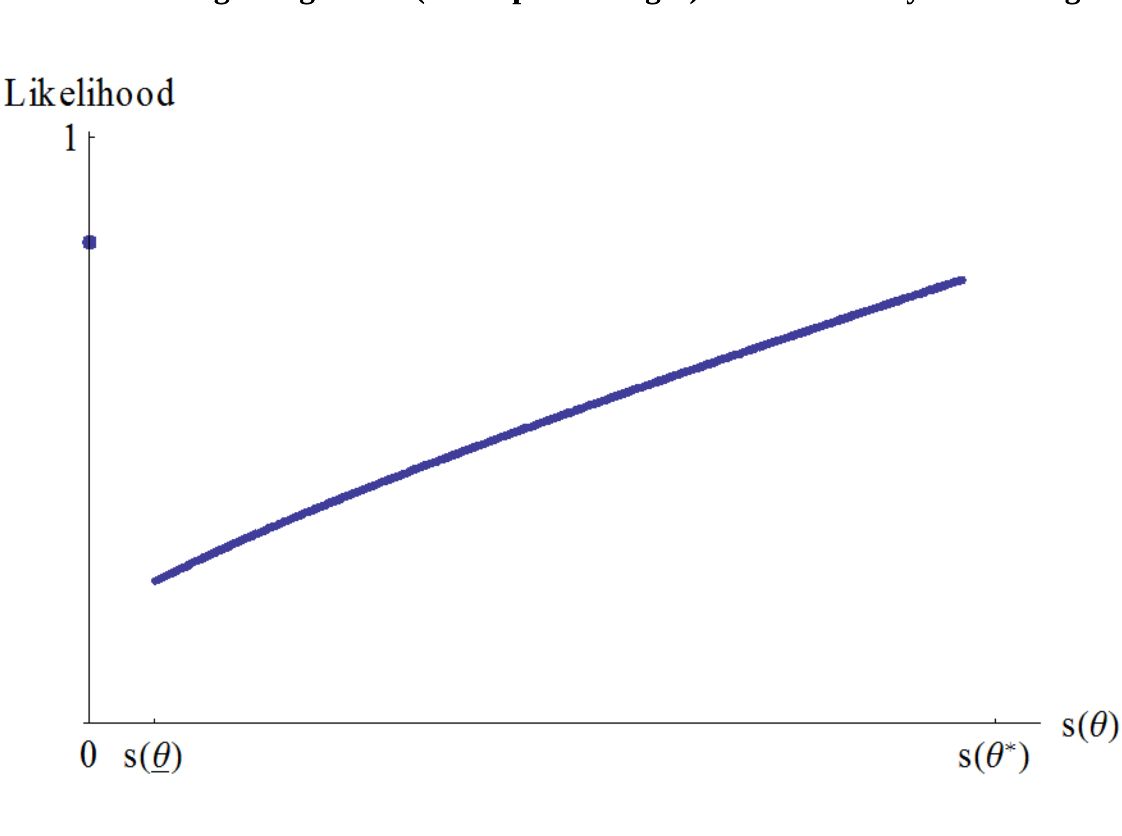
Credit Grade	number of deliquency	% deliquency	% deliquency with 0 word	% deliquency >0 and < median words	% deliquency with > median words
A	187	6.0%	2.9%	11.2%	4.7%
B	275	6.4%	2.6%	13.9%	5.5%
C	229	6.5%	3.6%	12.2%	4.8%
D	174	6.9%	3.9%	15.7%	5.4%
E	87	7.8%	3.9%	14.0%	5.3%
F	35	8.3%	5.8%	14.1%	6.3%
G	29	12.0%	7.5%	20.8%	8.8%

Table 7: Model of Delinquency

Credit Grade		A		B		C		D		E		F		G	
		Estimate	**SD**	**Estimate**	**SD**	**Estimate**	**SD**	**Estimate**	**SD**	**Estimate**	**SD**	**Estimate**	**SD**	**Estimate**	**SD**
	Intercept	-0.264	1.615	0.270	0.920	0.269	1.172	1.396	1.378	-0.168	0.102	-0.149	0.192	-0.160	0.220
	# of open credit lines	**0.100**	0.037	**0.078**	0.030	0.007	0.029	-0.044	0.038	**0.075**	0.023	**0.126**	0.009	**0.073**	0.010
	# of delinquencies in past 2 years	0.434	0.431	-0.329	0.282	-0.119	0.203	-0.242	0.171	0.061	0.129	0.143	0.537	0.501	0.513
	Revolving balance utilization	-0.719	0.603	-0.052	0.387	0.321	0.397	**0.448**	0.048	-1.003	0.724	-0.141	1.321	2.458	2.025
	Monthly income	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Debt-to-income ratio	-0.739	2.051	1.097	1.578	-1.935	1.694	-1.565	2.081	-2.175	2.741	-2.931	4.688	-5.837	7.318
	Total credit lines	0.027	0.014	-0.001	0.012	0.016	0.011	0.021	0.015	0.031	0.025	0.028	0.042	-0.012	0.053
Home Onwership (compared to none)	Mortage	0.549	1.045	-0.554	0.523	-0.274	0.778	0.194	0.779	0.152	0.122	-0.158	0.592	-0.190	0.420
	Rent	0.877	1.062	-0.110	0.550	-0.045	0.804	-0.406	0.847	0.163	0.113	0.140	0.392	-0.180	0.412
	Own	0.599	1.046	-0.385	0.517	-0.163	0.775	0.110	0.774	0.153	0.132	0.154	0.142	0.189	0.420
	Inquiries in last 6 months	0.000	0.066	0.014	0.038	0.024	0.030	-0.030	0.031	-0.007	0.035	-0.058	0.078	0.052	0.062
	Loan amount	0.024	0.023	-0.009	0.012	**0.036**	0.014	-0.027	0.017	0.004	0.025	-0.047	0.039	**0.120**	0.058
	Term (60 months = 1)	-0.619	1.072	-0.231	0.444	-0.866	0.612	-0.585	0.640	-0.224	0.848	-0.522	1.176	-1.582	2.125
Time trend	Y2008	**-1.020**	0.315	**-0.975**	0.245	**-0.499**	0.242	**-0.851**	0.265	-0.416	0.323	**-0.878**	0.579	0.968	0.739
	Y2009	**-1.716**	0.300	**-1.653**	0.241	**-0.597**	0.232	**-1.625**	0.269	**-1.415**	0.382	**-2.091**	0.685	**-2.365**	0.586
	Y2010	**-2.570**	0.328	**-2.673**	0.268	**-0.844**	0.275	**-1.875**	0.349	**-1.972**	0.667	**-2.674**	0.944	-1.213	1.799
Loan Purpose	Credit card	0.439	0.428	0.701	0.519	-0.414	0.507	0.643	0.800	0.295	1.195	0.201	1.458	-0.883	1.895
	Debt consolidation	0.074	0.407	0.513	0.501	-0.315	0.473	0.430	0.774	1.023	1.120	0.388	1.353	-2.733	1.795
	Educational	0.164	0.617	0.901	0.579	-0.502	0.615	0.744	0.898	0.286	1.281	1.540	1.614	0.981	2.246
	Home Improvement	0.034	0.461	0.265	0.555	-0.111	0.524	0.109	0.865	0.430	1.279	1.067	1.666	-1.443	1.989
	House	0.377	0.852	1.026	0.660	0.250	0.729	-0.402	1.315	2.031	1.379	3.532	2.081	-0.204	0.202
	Major Purchase	0.483	0.451	0.774	0.537	-1.039	0.643	0.097	0.869	1.340	1.280	2.344	1.902	-0.182	0.359
	Medical	0.105	0.672	0.829	0.652	-0.138	0.687	0.569	0.945	-0.047	1.534	0.131	0.185	-0.184	0.300
	Moving	-0.085	0.725	1.207	0.665	-0.774	0.872	0.078	1.094	-0.138	0.807	-0.135	0.152	-0.187	0.449
	Other	0.255	0.414	0.536	0.511	-0.117	0.484	-0.175	0.808	0.737	1.146	-0.971	1.679	**4.542**	1.850
	Renewable Energy	2.382	1.244	1.221	1.182	-0.135	0.103	-0.124	0.213	1.307	1.182	2.206	1.445	1.221	1.182
	Small Business	-0.143	0.659	-0.204	0.631	-0.748	0.592	0.153	0.842	-2.540	1.605	-2.955	2.620	**3.912**	1.186
	Vacation	1.170	0.650	0.035	1.153	-0.141	0.468	-0.142	1.100	1.454	1.373	1.990	1.865	0.035	1.153
	Wedding	-0.045	0.654	0.380	0.643	0.317	0.591	-0.142	0.432	0.380	0.643	0.555	0.461	-0.197	0.432
Description Content	Contains "currency"	0.282	0.279	-0.133	0.270	0.136	0.277	0.040	0.308	0.194	0.473	-0.587	1.116	2.401	1.518
	Contains "buying"	0.363	0.275	0.355	0.218	-0.179	0.278	0.044	0.311	-0.561	0.504	-0.029	0.824	1.996	1.343
	Contains negative words	0.123	0.248	0.051	0.189	-0.126	0.207	0.264	0.224	0.396	0.329	-1.026	0.702	0.239	0.938
	Contains positive words	-0.182	0.192	-0.118	0.156	0.068	0.170	-0.141	0.203	-0.139	0.309	-0.255	0.525	-0.332	0.752
	Contains date	0.300	0.437	-0.532	0.426	0.444	0.383	-1.048	0.753	-0.093	0.790	0.138	1.057	1.695	1.772
	Contains "budget"	-0.212	0.213	**-0.386**	0.175	-0.223	0.202	-0.359	0.238	0.078	0.357	-0.306	0.690	-0.561	0.707
Description Length	Description length	**-1.00E-02**	5.06E-03	**-1.34E-02**	4.29E-03	**-7.66E-03**	2.58E-03	**-1.98E-02**	6.75E-03	**-9.84E-03**	8.77E-03	**-1.17E-02**	1.56E-03	**-1.23E-02**	4.70E-03
	Description length^2	**2.50E-05**	9.47E-06	**3.31E-05**	6.46E-06	**2.93E-05**	7.38E-06	**3.48E-05**	2.16E-06	**2.15E-05**	3.78E-06	**2.50E-05**	5.49E-06	**2.89E-05**	6.66E-06
Counter signaling	No description	**-1.96E+00**	6.99E-01	**-1.51E+00**	6.10E-01	**-7.35E-01**	2.81E-01	**-1.98E+00**	4.44E-01	**-8.86E-01**	3.56E-01	**-8.26E-01**	2.48E-01	**-1.99E+00**	4.87E-01
	Residual from first stage	**0.022**	0.003	**0.019**	0.004	**-0.030**	0.020	0.032	0.005	0.004	0.002	**0.072**	0.009	0.035	0.019

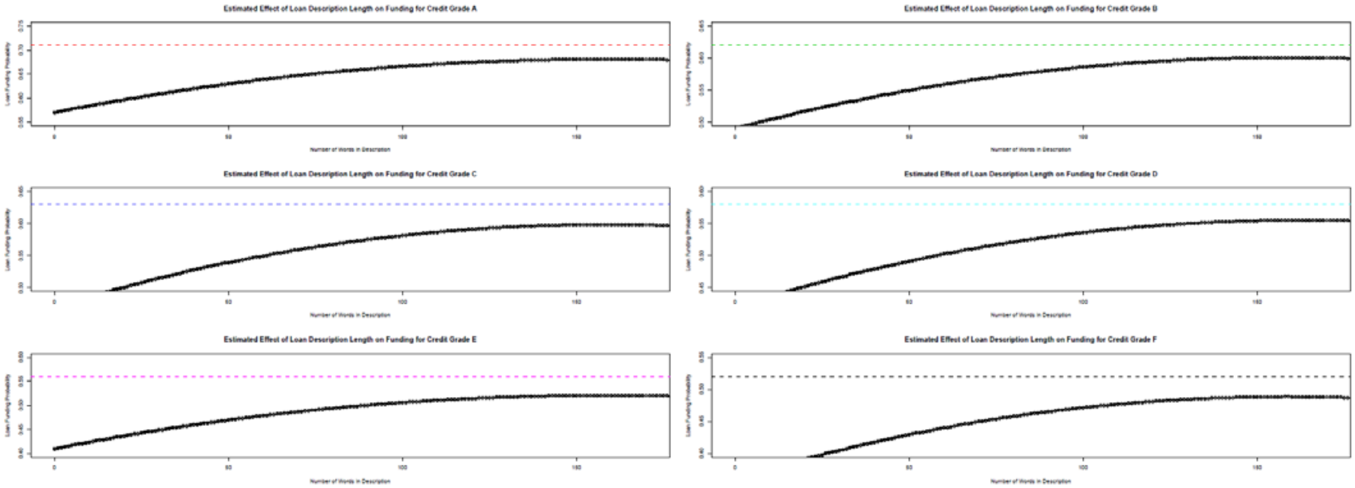
***Bold estimates indicate signifance at the 0.05 level.**

Figure 1: Counter-signaling Theoretical Prediction for the Relationship between Signaling Effort (Description Length) and Probability of Funding



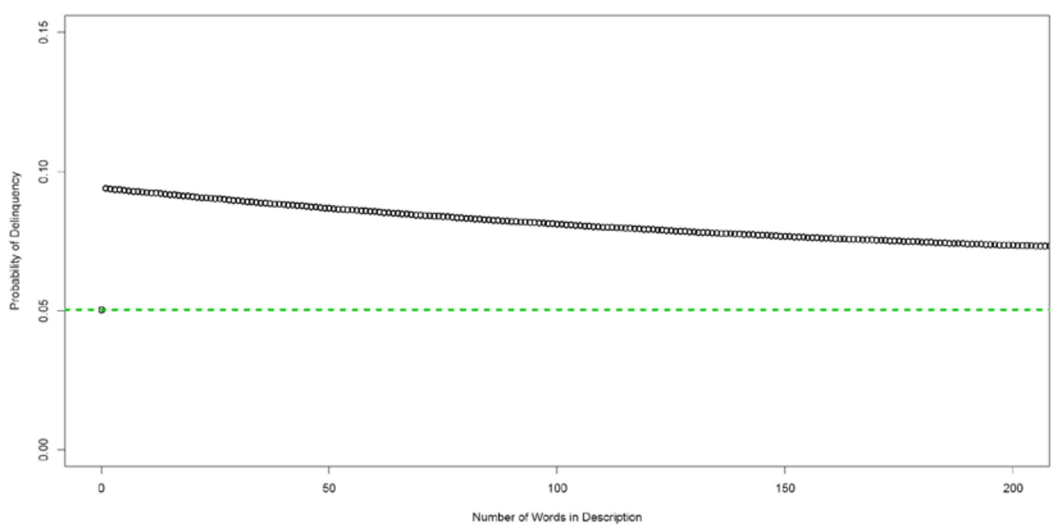
Note: The parameter values used for this graph are λ=.7, k=1/2000, V=200, γ=1/2, $\underline{\theta}$, =0.1, $\overline{\theta}$ =0.9, a=1/20. The appearance of the graph is robust to an array of parameter values.

Figure 2: Effect of Countersignaling on Loan Funding Probability



Notes: To create the graphs, we run separate models for each credit grade and use the credit-grade specific parameters. The graphs show description lengths ranging from 0 to 180 words, which account for 95% of the cases in our data. Dashed lines denote the level of funding predicted with no description.

Figure 3: Effect of Countersignaling on Delinquency Probability



Note: To create this graph, we varied the number of words in the description from 0 to 200, and the remaining independent variables are fixed at their mean levels. Credit grade has no significant effect on delinquency probability, and thus the graph is pooled across all credit grades. Dashed lines denote the level of delinquency predicted with no description.