Title	Information sharing and user behavior in Internet-enabled Peer-to-Peer Lending Systems: an empirical study
Author(s)	Gleasure, Robert; Treacy, Stephen; Feller, Joseph
Publication date	2017-05-30
Original citation	Feller, J., Gleasure, R. and Treacy, S. (2017) 'Information sharing and user behavior in Internet-enabled Peer-to-Peer Lending Systems: an empirical study', Journal of Information Technology, 32(2), pp. 127-146. doi: 10.1057/jit.2016.1
Type of publication	Article (peer-reviewed)
Link to publisher's version	http://dx.doi.org/10.1057/jit.2016.1 Access to the full text of the published version may require a subscription.
Rights	**© 2016, Palgrave Macmillan on behalf of the Association for Information Technology Trust. All rights reserved. This is a post-peer-review, pre-copyedit version of an article published in Journal of Information Technology. The definitive publisher-authenticated version [Feller, J., Gleasure, R. and Treacy, S. (2017) 'Information sharing and user behavior in Internet-enabled Peer-to-Peer Lending Systems: an empirical study', Journal of Information Technology, 32(2), pp. 127-146. doi: 10.1057/jit.2016.1] is available online at: https://doi.org/10.1057/jit.2016.1**
Item downloaded from	http://hdl.handle.net/10468/6264

Information Sharing and User Behavior in Internet-enabled Peer-to-Peer Lending Systems: An Empirical Study

Joseph Feller, Rob Gleasure, and Stephen Treacy

Cork University Business School, University College Cork, Ireland.

Published in the Journal of Information Technology

Abstract

Internet-based information systems (IS) have enabled a variety of forms of collective action, such as aggregating globally distributed development effort (e.g. open source software), problem solving (e.g. innovation marketplaces), or resources (e.g. crowdfunding). Within the domain of crowdfunding, *Internet-enabled Peer-to-Peer Lending Systems* (IP2PLS) have emerged as a disruptive technological genre, with implications for the financial services sector, business capitalization strategies, and personal and community development. IP2PLS have captured the attention of the Information Systems research community, and studies of user behavior in IP2PLS have revealed the saliency of social identity, personal transparency, and information sharing in such systems. We argue that the current state of knowledge is limited by methodological bias towards the study of particular IP2PLS providers, and report the findings of a study of a very large but under-researched platform. Through an analysis of 116,669 loan records, and a subsequent analysis of 1000 records manually coded through a content analysis process, we investigate the impact of information sharing on user behavior in an IP2PLS, revealing relationships that frequently contradict the findings from prior research. The paper discusses the contributions of our work towards a more complete and heterogeneous picture of user behavior in IP2PLS, implications for researchers of other crowdfunding systems and/or other emerging forms of IS-enabled collective action, as well as for system designers, users and providers.

Keywords: Internet-enabled Peer-to-Peer Lending Systems, Crowdfunding, Collective Action, Information Sharing, Personal Transparency, Social Identity

INTRODUCTION

This paper presents the findings of an empirical study theorizing the relationships between social identity, personal transparency, information sharing, and consequent user behavior in an emerging form of networked information system (IS) designed to support collective action, namely *Internet-enabled Peer-to-Peer Lending Systems (IP2PLS).*

We categorize IP2PLS as a member of a growing portfolio of Internet-based IS that enable collective action on a global scale. Such systems include the communication, coordination and collaboration tools used to aggregate distributed skill and knowledge in the development of open source software (e.g. Linux) and open content (e.g. Wikipedia), the contest platforms used to aggregate creativity (e.g. Threadless) and innovation (e.g. InnoCentive), the infrastructures that enable collective micro-work (e.g. Mechanical Turk) and collective information processing (e.g. tagging systems), and the systems to enable the global pooling of computing resources (e.g. BOINC). Specifically, IP2PLS are platforms that enable the aggregation and exchange of financial resources and are thus related to other electronic marketplace platforms (e.g. Kickstarter) designed to support the process of *crowdfunding*.

All of the phenomena above have attracted enormous scrutiny from the IS community due to their disruptive and transformative nature. These Internet-native systems (among others) have enabled the emergence of new forms of creativity, productivity, and problem solving that challenge our thinking about the nature of innovation and work. As generative and evolving phenomena, they enable IS researchers to explore some of the fundamental questions of the IS discipline, specifically (as Benbasat and Zmud (2003) phrase it) the "human behaviors reflected within, and induced through both the (1) planning, designing, constructing, and implementing, and (2) direct and indirect usage of [IS] artifacts … [and the] … the impacts (direct and indirect, intended and unintended) of these artifacts on the humans who directly (and indirectly) interact with them."

Within the domain of crowdfunding, IP2PLS have likewise emerged as a disruptive technological genre, with implications for the financial services sector, business capitalization strategies, and personal and community development. IP2PLS support Peer-to-Peer Lending (P2PL), a form of crowdfunding in which lending occurs "between private individuals on online platforms where financial institutions operate only as intermediates required by law" (Bachman *et al.,* 2011, p.2). The origins of P2PL are associated with Zopa.com in the UK, the first large platform dedicated to P2PL (Iacobuzio 2006, Kupp and Anderson 2007, Briceno Ortega and Bell 2008) and since then, various other platforms have emerged in countries around the world. Such platforms include Prosper and Lending Club in the US (Wang and Greiner 2011), PPDai and My089 in China (Chen and Han 2012), Popfunding in South Korea (Jeong *et al.*, 2012), and Smava in Germany (Pötzsch and Böhme 2010).

Within IP2PLS, the interplay of the technology with the subtle human biases and social cues that influence lending behavior (how personal and humanizing data are shared and interpreted) is particularly salient, and has attracted researcher attention from both IS (e.g. Greiner and Wang, 2009) and cognate disciplines including Economics (e.g. Dezső and Loewenstein, 2012), Human Resources (Pope and Sydnor, 2011), Business Education (Livingston and Glassman, 2009), and Marketing (e.g. Herzenstein *et al.*,, 2011b).

The current work is driven by this interplay between the human and the technological. While we acknowledge the economic importance of IP2PLS as an electronic marketplace (c.f. Chen et al., 2009), in this study we posit that it is the ability of users to humanize other users (specifically of lenders to humanize borrowers, and the ability to infer the borrower's trustworthiness and identity from their social affiliations), that is fundamental to understanding the dynamics of these systems. Problematically, while there is consensus around the importance of the impact of personal information in P2PL, it is often discussed in terms that assume homogeneity across different IP2PLS, despite the evidence that the functionalities of these systems (and the motivation and behavior of their users) can vary widely (Wang *et al.*, 2009, Bachmann *et al.*, 2011). This is further complicated by the bias in the empirical research literature towards a small number of P2PL platforms, primarily Prosper, thus leading to an incomplete picture of the phenomenon.

In response to this situation, our study has been specifically designed to target an under-investigated P2PL platform, Lending Club, which would appear to differ in several significant ways from the dominant subjects of prior research, e.g. by offering less scope for unstructured personal information, photographs, and formal within-platform social ties. Consequently, it less clear how borrowers' social identity can be inferred in such IP2PLS, and several important and unanswered questions emerge. Thus, through an analysis of 116,669 loan records, and a subsequent analysis of 1000 records manually coded through a content analysis process, we investigate the impact of information sharing on user behavior in the Lending Club IP2PLS, revealing relationships that frequently contradict the findings from prior research.

Our paper is structured as follows. First, we conceptualize the phenomenon of interest through a synthesis of prior research, and develop the theoretical model used in the study. Specifically, we utilize Social Identity Theory to theorize information sharing and information sharing expectancy within and between IP2PLS. Second, we present our study design. We describe our sampling strategy and the selection of our field site, discuss the nature of our transaction data set and describe our data coding process (with coding examples also given in Appendix A). Third, we present our findings, in two parts. Iteration 1 describes the regression analysis methods employed and presents our findings on how user sharing of "hard" data predicts user behavior, based on an analysis of 116,669 loan records. Iteration 2 describes the regression analysis methods employed and presents our findings on how user sharing of "soft" data predicts user behavior, based on an analysis of 1000 randomly selected and manually coded loan records. Finally, we conclude the paper with a discussion of the work's contribution towards a more complete and heterogeneous picture of user behavior in IP2PLS (an important emerging type of IS), and implications for researchers of other crowdfunding information systems and/or other emerging forms of IS-enabled collective action, as well as for system designers, users and providers.

SOCIAL IDENTITY THEORY, IP2PLS AND INFORMATION SHARING

A number of researchers have investigated the factors impacting on the likelihood of borrowers attracting investment (e.g. Iyer *et al.,* 2009, Herzenstein *et al.,* 2011a) and the likelihood of loans being repaid (e.g. Klafft 2008, Luo *et al.,* 2011) in IP2PLS. In line with such studies, we argue that social identity and personal cues (via information sharing) are key factors effecting the behavior of users of such systems.

Social Identity Theory and IP2PLS

The construct of social identity (aka 'collective identity' (Ashmore et al. 2008)) emerges in its contemporary form from studies of inter-group conflict that revealed an individual's appetite for social mobility vs. social change/competition was influenced by the extent to which they identified with their dominant in-groups (e.g. Tajfel 1979, Tajfel and Turner 1980). This idea (sometimes referred to synonymously) was subsequently explored further and formalized as 'social identity theory' (SIT) positing that individuals comprise multiple identities, each relating to different social networks in which that person interacts (Stryker 1980, Hoelter 1983). These within-network identities are role-based, the details of which are negotiated over the course of ongoing interaction with other network members (c.f. Stets and Burke 2000). While these roles vary in detail, recurring patterns have been identified based upon the self-image of network members, e.g. self-doubters vs. soldiers, strugglers vs. stencils (Alvesson 2010). As an individual's roles become more embedded within the overall network dynamic, so their identification with that network and personal attachment to it increases (Stets and burke 1999).

To date, the concept of social identity has featured in IP2PLS research in two ways. The first of these regards the inference of the reliability of borrowers, based on social networks *within the IP2PLS* in question (e.g. Berger and Gleisner 2009, Greiner and Wang 2009, Bachman *et al.,* 2011). These studies assume that the social networks associated with borrowers, as well as the identity of borrowers within those networks, play an important role in explaining both lenders' investment decisions and borrowers' likelihood of repaying that investment. This assumption is supported by observations of lending and repayment behavior within the explicit borrower groupings facilitated on Prosper (Freedman & Jin 2008, Herrero-Lopez 2009, Lin 2009). In these contexts, greater investment occurred because lenders made inferences about the identity of the borrower based upon known characteristics of the group, while greater repayment occurred as a result of the borrowers' commitment to that group and consequently, their higher susceptibility to the normative pressures to repay. However, a more developed social identity is not always a positive predictor of behavior. Within the Korean lending platform Popfunding.com, it was shown that failed loan postings damage borrowers' social identity in terms of their social and financial credibility, and so inhibit future investment in them (Jeong *et al.,* 2012). Similarly, within PPDai, a large IP2PLS platform based in China, it has been shown that borrowers who observe their online friends default are twice as likely to default in the future (Lu *et al.*, 2012).

Second, social identity has been used to understand the inference of the reliability of borrowers based on *offline, external* social identities. The number of such identity claims is higher among borrowers with poor credit scores, suggesting the intention of borrowers may be to present a less finance-specific identity to lenders (Herzenstein *et al.,* 2011b). However, the impact of these identity claims on lending is not always positive, even those apparently relevant to borrowers' ability and/or intention to repay (c.f. Pötzsch and Böhme 2010, Herzenstein *et al.,* 2011b, Larrimore *et al.,* 2011).

One key aspect of SIT that has not been addressed in existing IP2PLS research surrounds the role of language and information expectations in how lenders and borrowers interact. The greater an individual's commitment to a network the more *salient* that role-identity becomes for the individual's self-definition, compared to other roles they perform (Stryker and Serpe 1994). Higher role-identity *salience* allows an individual to more effortlessly gauge role expectations from other members of the network and adjust their behavior accordingly (Callero 1985, Haslan *et al.*, 1999, Shih *et al.*, 1999). This adjustment is a mutual and reciprocal process that gradually accumulates within the interaction within a network, producing a culture of implicit and explicit communication signals (c.f. Eliasoph and Lichterman 2003). Thus, the behavior of an individual who is less embedded within some social network may be perceived as less competent that that of another individual, not because of rational judgments but because their method of interacting is more consistent with expectations (Stryker 1968, Haslan *et al.*, 1999, Stets and Burke 2000). Put differently, one network member's trust in another's ability to perform a role may be influenced by their fluency with 'structured symbolic interactions' (Stryker 2008). This is illustrated in Figure 1.

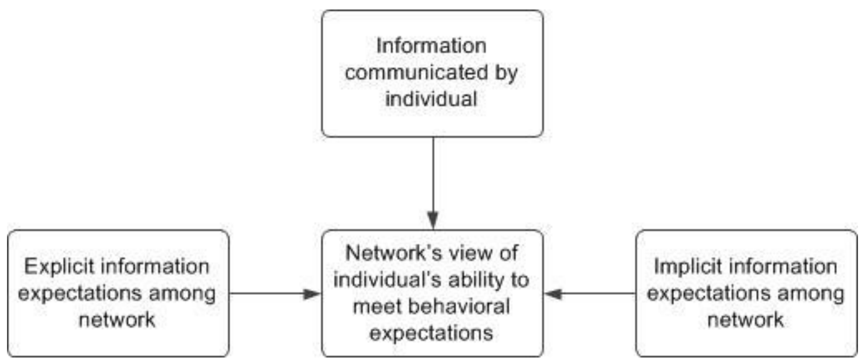


Figure 1. SIT and Information Sharing Expectations

As network boundaries deepen and the movements of individuals from one network to another decrease, so between-network competition and discrimination are likely to occur

(Tajfel and Turner 1980). In online communities, a more extreme version of this tendency has been observed in the formation of what is termed ‘xenonetworks’ by Reay Atkinson et al. (2013, 2014), whereby diversity-intolerant bubble communities form around some core indicators of *likeness*. Those authors note that such *likeness* among members is typically communicated within less structured information in the form of ‘weak’ signals (c.f. Granovetter 1973). These weak signals are often subsequently encoded in formal informational procedures for that network and hence act as useful predictors of future information structuring and governance developments (Ansoff 1975, Hiltunen 2008).

Such a perspective explains accounts where functionally interchangeable behaviors are acceptable in some networks but not others. This has been observed in a variety of situations, for example within intra-organizational social relationships (Ashforth and Mael 1989, Dutton *et al.,* 1994, Haslam *et al.,* 2000), and in relationship marketing (Arnett *et al.,* 2003, Bhattacharya and Sen 2003, Michalski and Helmig 2008). It has also been used to account for user participation in online communities in settings outside of IP2PLS, including general knowledge contribution (Shen *et al.,* 2010), as well as consumer product discussion and reviews (Forman *et al.,* 2008).

From an IP2PLS perspective, this raises an important question as to whether the various platforms such as Zopa, Prosper, Lending Club, PPDai, Popfunding, and Smava share a common community/network of borrowers and lenders, or whether each possesses a distinct group of participants. As each of these platforms represents a competing marketplace, the latter appears more likely. Under such conditions where users of a network benefit from the number of other users present, network externalities act to create oligarchies from individual competitors, and eventually a monopoly (c.f. Katz and Shapiro 1985). This occurs because while the requirements for new markets are more open, standards emerge over time to increase market efficiency, resulting in social, economic, and technological convergence around market leaders (Katz 1996, Uzzi 1996). This has important implications for individual IP2PLS providers, as it suggests that as the market matures, IT capabilities will converge around a subset of platforms and between-network competition for mainstream survival and dominance will increase, to the detriment of social mobility among their users. Thus, IP2PLS providers are not only competing for market share, they are also competing to establish the future paradigm of IT capabilities dictating the types of explicit and implicit information expected by lenders (see Figures 2 and 3).

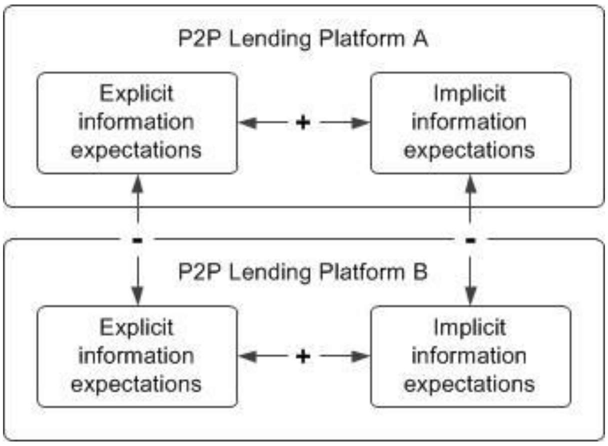


Figure 2. Information Sharing Expectations within and between Systems

This observation means that it is important to understand both the types of information explicitly and implicitly expected from/by users in of an IP2PLS, as well as the impact that information sharing through the IP2PLS has on user behavior.

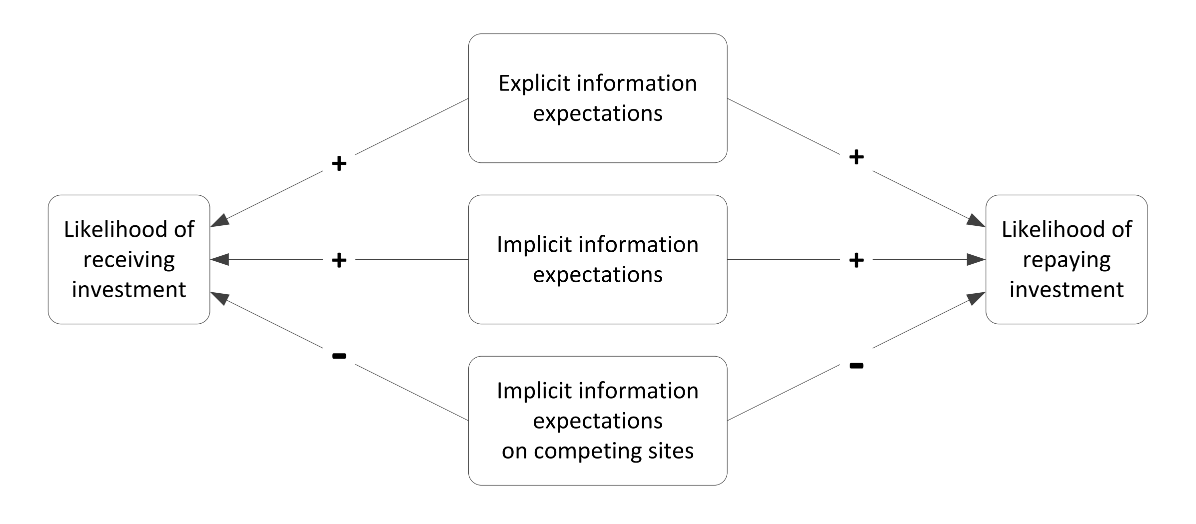


Figure 3. Information Sharing Expectations for Some Specific System

The types of explicitly expected 'hard' data that predict lending behavior

The perceived risk associated with particular individuals is often a factor of the accumulation of past experience between transacting parties (Pavlou and Gefen 2004). However, the nature of P2PL means that such an accumulation of interactions may not be possible, as "most loans are 'one-shot' events, that is, borrowers usually apply for (and receive) a loan just once. Therefore, the main avenue available in relational settings to engender confidence in another's behavior, i.e., through exemplary behavior over time, is not available" (Herzenstein *et al.*, 2013, p.12). Hence, the standard required credit information used to portray a borrower's financial identity in IP2PLS reflects accumulated financial interactions across a range of domains. One of the most prominent pieces of information in this regard is borrowers' credit scores, which are considered a useful means of predicting loan applicants' likelihood of repayment in traditional lending decisions (c.f. Mester 1997). Such credit scores have been shown to impact positively upon both investment (Klafft 2008b, Iyer *et al.*, 2009, Puro *et al.*, 2011) and repayment (Freedman and Jin 2008, Livingston and Glassman 2009) in studies of behavior on Prosper. IP2PLS platforms may include both internally and externally generated credit scores. For example, Lending Club provides one credit score that it calculates based on the information provided by borrowers, but also a FICO credit score, which is generated by a US-based third party (see http://www.myfico.com/myfico/Credit Central/ScoringWorks.asp). It is anticipated that better credit scores, both internally and externally generated, will predict higher levels of both investment and repayment.

IP2PLS also provide lenders with other historic data concerning applicants' previous financial behavior. This may come in the form of the number of public records held by the applicant, as well as a record of any account delinquencies with which they are associated. As each of these reflects negative previous experiences that other lenders have experienced with the applicant, it is predicted that their frequency should be detrimental to both their likelihood of receiving and repaying investment. Such data are further contextualized with data describing the number of months since an applicant's most recent record or delinquency. Given that lenders are more likely to be concerned with recent transgressions that those taking place several years ago (Guiral-Contreras *et al.*, 2007), it is also predicted that a higher number of months since an applicant's most recent record or delinquency will increase both their likelihood of receiving and repaying investment.

Lenders in IP2PLS are also influenced by explicitly requested personal information about the borrower's ongoing financial situation. This includes the borrower's income (Puro *et al.*, 2010, Lu *et al.*, 2012), their homeownership status (Herrero-Lopez 2009, Larrimore *et al.*, 2011), and their number of years in employment (Livingston and Glassman 2009, Larrimore *et al.*, 2011). These factors allow borrowers to demonstrate the strength of their financial identity, and so should support both investment and repayment.

Borrowers' existing usage of credit from other parties have further been shown to impact upon decision making in IP2PLS, including data relating to the credit lines possessed by an applicant and an applicant's current utilization of their revolving credit (Herrero-Lopez 2009, Iyer *et al.,* 2009, Puro *et al.,* 2010). Lenders may also infer borrowers' difficulties in managing their debt using information such as those borrowers' debt-to-credit ratios and the number of credit inquiries for the borrower made by potential lenders (Klafft 2008a, Lin 2009, Larrimore *et al.,* 2011). These records afford an opportunity to validate a borrower's financial identity by showing that they can manage debt appropriately.

Finally, 'hard' data relating to the terms and conditions of the loan are typically presented in IP2PLS contexts. This includes the loan amount, the loan duration, the interest rate, and the monthly repayments. A higher loan amount and duration demand more investment to meet the loan requirements, meaning there is intuitively less chance of meeting investment goals and more difficulty in repayment. Given that lenders are at least partly motivated by financial returns, higher interest rates and monthly repayments are likely to attract greater investment. However, as these factors place additional financial responsibility on borrowers, they appear less likely to encourage repayment. Interestingly, while the latter three pieces of data have been shown to impact upon lending behavior (Puro *et al.,* 2010, Ceyhan *et al.,* 2011, Herzenstein *et al.,* 2011a), data on Prosper show no impact on either investment or repayment for the stated purpose of the loan (Pope and Syndor 2011).

The types of implicitly anticipated 'soft' data that predict lending behavior

In addition to the explicitly required information for borrowers on P2PL platforms, individuals have the opportunity to make other information transparent in the context of their loan descriptions. Viewed in trust terms, providing additional information may reduce information asymmetry between borrowers and lenders, and so can be interpreted as a gesture of benevolence intended to convey a borrower's strong intentions of repayment (Pötzsch and Böhme 2010). This is supported by existing P2PL research, which shows that longer loan descriptions are generally associated with higher levels of lending (Larrimore *et al.*, 2011). Yet, from a SIT perspective such disclosures may not necessarily be advantageous. This is because the provision of data that is not anticipated by lenders damage the sense of shared meaning (Burke and Reitzes 1991) and distance the applicant from the 'identity standard' (Burke 1991) for loan applicants possessed by lenders. Thus, deviating from the norm, even with the intention of demonstrating benevolence, may undermine lenders' confidence in the applicants' commitment to their role. This means that while the quality of required hard data is likely to determine its impact on lending behavior, it is the normalcy of optional soft data that may moderate its impact. This makes it challenging to determine in advance whether various forms of information will have positive or negative impacts on investment and repayment behavior. However, this study will begin with the assumption that the competitive nature of P2PL platforms means that implicitly expected information for some platforms will have a negative impact on investment and repayment behavior elsewhere. In the case of Lending Club, this implies all of the soft data observed to positively influence behavior on peer-to-peer lending in existing research should theoretically have a negative impact (because existing research to date targeted only competitors).

The forms of optional soft data described in existing literature in P2PL range across three dimensions, namely additional credit information, personal or humanizing information, and direct appeals made to lenders. In the context of appeals made to lenders, these appear to have a negative effect, e.g. there is evidence from Prosper to suggest that justifications for a borrower's current financial situation discouraged lenders (Larrimore *et al.*, 2011). This may be because Prosper, unlike microfinance sites such as Kiva, presents itself as a financially motivated platform; hence, such claims may appear out of place. This is also supported by findings on Smava (Pötzsch and Böhme 2010), which show a negative response by lenders to statements arousing pity about the borrower's situation.

Conversely, positive additional credit information about one’s financial identity should have the opposite effect, as illustrated in the context of user narratives on Prosper (c.f. Herzenstein *et al.*, 2011b) where positive claims concerning how an individual conducted themselves financially in the past, as well as how they intend to behave in the future, encouraged investment (though interestingly made repayment less likely). These positive claims have also been observed on Prosper in the form of explanations/acknowledgements of previous financial transgressions, or alternatively explanations/denials of previous financial transgressions – both of which positively influence lenders by reframing or qualifying negative information presented in borrowers’ credit histories and financial records (Sonenshein *et al.*, 2011).

Other soft data that can influence lending behavior relates to descriptions of borrowers’ employment status and occupation. This is interesting because, while Prosper began presenting a borrower’s occupation with that borrower’s listing in 2007 (Freedman and Jin 2008), this information is not presented on other sites such as Lending Club. Yet, this information plays an important role in Prosper, particularly in the formation of advantageous borrower groups (Everett 2010). Such information also plays a role in demonstrating a Prosper borrower’s stability, which is considered a crucial predictor of repayment (e.g. Livingston and Glassman 2009). For similar reasons, borrowers also frequently list their educational background, although only this manifests only a marginal impact on lenders’ preferences on both Smava (Pötzsch and Böhme 2010) and on PPDai.org (Lu *et al.*, 2012).

The types of humanizing information that borrowers utilize to attract investment also varies significantly. As noted already, Prosper members typically use affiliations between members to build trust by association (Greiner and Wang 2009, Herrero-Lopez 2009, Bachman *et al.*, 2011). However, platform-external borrower data are also observed to have positive impacts on several platforms. For example, studies of photographs on *Prosper* show that lenders are influenced by borrowers' age (Duarte *et al.*, 2012, Ravina 2012), gender, ethnicity (Pope and Syndor 2011, Duarte *et al.*, 2012), obesity, perceived happiness (Pope and Syndor 2011), marriage status, children/dependents (Duarte *et al.*, 2012), and physical attractiveness (Ravina 2012). There is also some evidence to suggest that that claims of morality, religiousness, and political-mindedness can also benefit borrowers on Prosper (Herzenstein *et al.*, 2011b), as are claims by borrowers on Smava of past or future kindness that were enabled by the loan and/or debt being consolidated (Pötzsch and Böhme 2010). While some of these claims appear to demonstrate lender prejudice, e.g. gender, ethnicity, others appear to be interpreted as accepted normative indicators of borrowers' personal reliability and responsibility.

STUDY DESIGN

The previous sections have identified the relationships between different forms of information sharing and user behavior in IP2PLS (summarized in Figure 4).

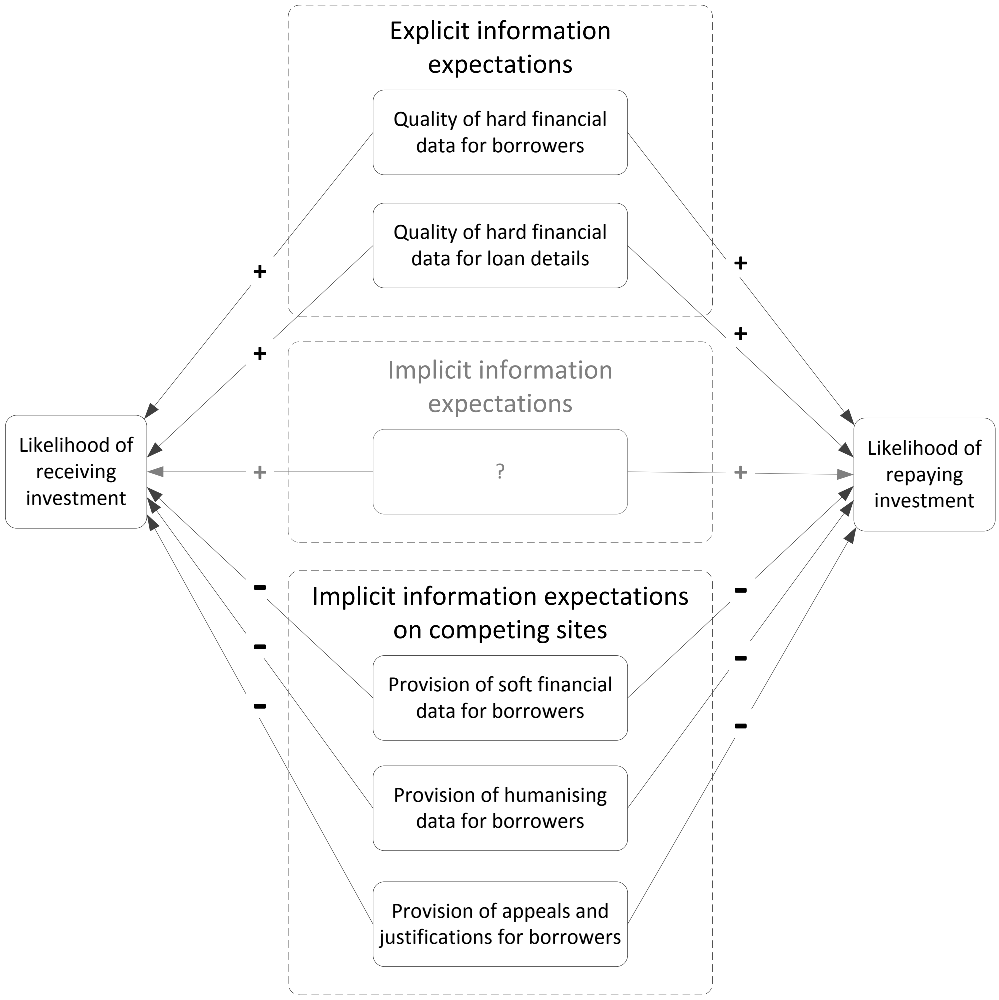


Figure 4. Preliminary Model of Relationship between Information Sharing and User Behavior

These relationships emerged from observations of user behavior made in existing research across a range of IP2PLS, including Prosper, Smava, Popfunding, and PPDai.org. It must now be investigated whether these relationships are consistent across IP2PLS (meaning the social networks of lenders and borrowers do not differ significantly across the range of platforms) or whether users of different IP2PLS interpret information sharing in fundamentally different ways. To address this question, this study first looks at the different IP2PLS, in order to find a suitable environment to compare against existing research situations. An approach to data gathering is then described through which to test these relationships.

IP2PLS Site Selection

As noted, a variety of IP2PLS have been studied in existing research. This is reflected in Table 3, which lists the platforms studied in all existing academic research known to the authors, and in which the likelihood of lender investment and/or borrower repayment in P2PL is investigated.

Table 3. The online platforms used for data gathering by existing research that investigates the likelihood of lender investment and/or borrower repayment	
Prosper	Freedman and Jin (2008), Klafft (2008a), Klafft (2008b), Berger and Gleisner (2009), Duarte *et al.*, (2009), Greiner and Wang (2009), Herrero-Lopez (2009), Iyer *et al.*, (2009), Lin (2009), Livingston and Glassman (2009), Collier and Hampshire (2010), Greiner and Wang (2010), Puro *et al.*, (2010), Wang and Greiner (2010), Bergovich (2011), Ceyhan *et al.*,(2011), Herzenstein *et al.*, (2011a), Herzenstein *et al.*, (2011b), Larrimore *et al.*, (2011), Luo *et al.*, (2011), Pope and Syndor (2011), Puro *et al.*, (2011), Sonenshein *et al.*, (2011), Chaffee and Rapp (2012), Chen and Han (2012), Duarte *et al.*, (2012), Ravina (2012), Zhang and Liu (2012)
Popfunding	Do *et al.*, (2012), Jeong *et al.*, (2012), Lee and Lee (2012), Yum *et al.*, (2012)
Kiva	Sinanan (2009), Barry (2012), Moodie (2013)
PPDai	Chen and Han (2012), Lu *et al.*, (2012)
Zopa	Bachmann *et al.*, (2011)
My089	Chen and Han (2012)
Smava	Pötzsch and Böhme (2010)

By far the most popular IP2PLS for researchers appears to be Prosper. Interestingly, Lending Club does not feature in any of these studies, although it is mentioned in some broader discussions of IP2PLS, e.g. in the context of legal regulations (Chaffee and Rapp 2012), or as a contrast to Prosper (Wang and Greiner 2011). This is surprising, given the scale and maturity of Lending Club as a platform, which, like Prosper, was launched in 2006. Although Zopa.com is often thought of as the first legitimate IP2PLS (Hulme and Wright 2006), Prosper is credited with encouraging most scientific interest in the subject by making the platform's data public in 2007 (Bachmann *et al.*, 2011). Yet, in the past number of years, the number of loans issued by Lending Club has significantly surpassed that of Prosper (Cunningham 2013), as illustrated in Figure 5.

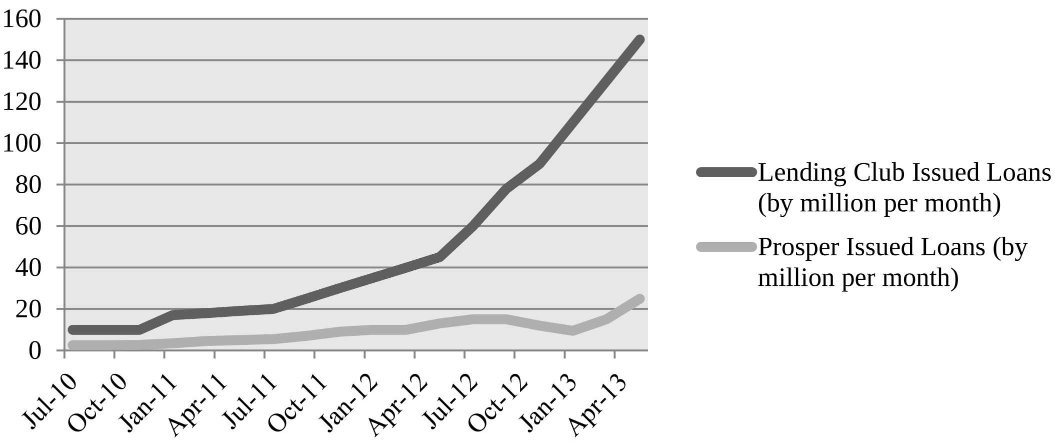


Figure 5. The value of loans issued on the Prosper and Lending Club platforms

In terms of loan value, both Prosper and Lending Club offer the ability to apply for a loan between $2,000 and $35,000. The minimum credit scores are also comparable, with Lending Club and Prosper requiring minimum scores of 660 and 640, respectively. Both Lending Club and Prosper charge borrowers a fee to take out a loan in the range of 1-5% depending on the borrower (Prosper calls this a "Closing Fee", whereas Lending Club calls it an "Origination Fee") and as of 2012, the returns for investors were similar (9.79% for Lending Club and 9.26% for Prosper (LendStats.com, retrieved Dec 2014)).

Yet, three significant differences exist in the lending mechanics employed by Prosper and Lending Club. Firstly, while interest rates in Lending Club are determined by the platform itself according to the borrowers' credentials, Prosper adopts an auction system that allows lenders to bid amounts at varying interest rates. Secondly, the listing mechanism for Prosper allows borrowers to post additional personal information, such as pictures of borrowers and their families. This has been demonstrated to introduce a number of significant social and cultural antecedents impacting upon lender and borrower behavior, such as a borrower's gender, age, ethnicity, whether they have children, whether they are part of a couple, a borrower's obesity, and a borrower's physical attractiveness (Klafft 2008b, Pope and Syndor 2011, Duarte *et al.*, 2012). Thirdly, Prosper allows members to form groups, which means that established members can endorse borrowers, hence lenders can infer trustworthiness from borrowers' social networks (Berger and Gleisner 2009, Greiner and Wang 2009, Herrero-Lopez 2009).

While the auction mechanism is important for a number of economic reasons (c.f. Chen *et al.*, 2009), the latter two differences are especially important in the context of social identity. The ability to humanize the borrower and the ability to infer the borrower's trustworthiness and identity from their social affiliations, inform the majority of IP2PLS studies concerned with social identity. Thus, several important and unanswered questions emerge surrounding a IP2PLS like Lending Club, which provides fewer system features for sharing personally and socially rich information; consequently it is not clear how a borrowers' social identity can be inferred by lenders. This makes Lending Club a valuable source of comparison with the more commonly studied P2PL platforms, and in particular with Prosper.

Data Gathering

Given that Lending Club presents significantly less system support for personal information sharing as regards soft data, two possibilities must be considered. The first is that such soft data does not feature in any meaningful way on Lending Club, meaning that borrowers' relevant social identity (and to an extent the behavior of Lender-users) is determined from only ‘hard’ data relating to the loan being requested. The second possibility is that ‘soft’ personal or social data is incorporated in some other format (unanticipated behavior by Borrower-users). To address these two possibilities, two iterations of data gathering and analysis were employed, each of which made use of publically available transactions records from Lending Club (downloaded from www.lendingclub.com on March 20th 2013). The purpose of the first iteration was to determine the explanatory power of the hard data and whether it alone is sufficient to explain user behavior on Lending Club. Conversely, the purpose of the second iteration was to determine the frequency with which different form of soft personal data was shared, as well as the relationship between this form of information sharing and user behavior.

The first iteration used the complete set of available records to analyze the predictive power of borrowers’ hard data. These records included 119,419 loan applications from June 7, 2007, to June 19, 2013 (retrieved July 23rd 2013 via a publically available file download from the Lending Club website). Data cleaning required that 2,749 of these be removed to ensure the integrity of the overall data set, as these related to loans that did not meet current policy requirements. A further three listings were removed that did not contain standard required information, such as loan ID or applicant usernames.

The second iteration analyzed the predictive power of borrower’s soft data sharing based on information voluntarily provided in the “Loan Description” field attached to loan requests. As described already, the Lending Club IP2PLS does not present users with the same opportunity as sites such as Prosper to divulge additional personal and social data. However, this Loan Description field offers borrowers some opportunity to divulge such information, if they so choose. Hence, these fields were manually coded by gathering a random sample of 1,000 records. Records were coded independently by two research assistants in blocks of 10, 40, 50, 100, 200, 200, 200, 250, and 150. When coding was completed for each block, the research assistants compared their results to identify incongruities. Resolution of these incongruities was straightforward in some cases, while others were resolved in consultation with the authors of this study. The authors also independently evaluated the results of coding by independently coding a random sub-set of each block to verify their accuracy and identify any additional issues. This random validation sampling demonstrated a coding consistency of >95% by the end of the process.

Appendix A provides a detailed table of the codes used along with example text from loan descriptions. Eight codes reflected additional credit information. These were coded at three ordinal levels, namely “no new information,” “some new information,” and “substantial new information.” The codes were

1. Additional loan purpose information
2. Justification of financial situation
3. Claims of future financial responsibility
4. Claims of past financial responsibility
5. Acknowledgement/Denial of financial transgressions
6. Education details
7. Employment status
8. Occupation details

Fourteen more codes reflected key types of personal or humanizing information that might be provided by borrowers. These were coded in binary terms, i.e. they were either present or they were not. These included

1. Links to other members
2. Links to other online presences
3. Links to photographs

4. Age
5. Gender
6. Ethnicity
7. Marriage details
8. Children and dependents
9. Hobbies
10. Health and obesity
11. Religious or political views
12. Claims of kindness
13. Please or thanks
14. Existence of follow-up comments

These additional measures allowed us to explore relationships observed on IP2PLS such as Prosper, PPDai and Smava, as well as observations made by Lee and Lee (2012) on the Popfunding platform (namely that the number of postings made by borrowers on the Q&A board positively influenced lender-users' behaviors).

DATA ANALYSIS AND FINDINGS

Iteration 1: How hard data predicts lending behavior on Lending Club

Loan records (N =116,669) included an internal Lending Club credit grade ranging from a high of A1 to a low of G5, with a five-number summary of (A1, B1, B5, C5, G5). These numbers are represented numerically for the purposes of testing from 1-35, where A1 = 1 and G5 = 35. All but 2 of the listings (N =116,667) also include an external FICO range, varying from 660-664 to 846-850, with a five-number summary of [660-664, 680-684, 695-699, 725-

729, 846-850]. These FICO ranges were represented by the middle number in each range for the purposes of testing, e.g. such that 660-664 was replaced with 662.

111,046 (95.2%) of applicants possessed no records on file, while 5,230 (4.5%) possessed only one record on file, and only 343 (0.2%) possessed more than one record. A five-number summary of the 5,623 listings involving applicants possessing one or more public records is [1, 81, 95, 107, 129]. Only 54 applications (0.05%) were made by applicants who listed previous account delinquencies, with a five-number summary of the delinquent amount among these 54 applications of [17.00, 53.25, 862, 4526.75, 65000] and a five-number summary of the months since the delinquencies occurred of [0, 1, 1, 2, 11]. Full homeowners comprised 9,026 (7.7%) of the applications (not including those currently paying off mortgages), while a five-number summary of the current number of years in employment and monthly income listed among applications was [0, 1, 2, 5, 9] and [333.33, 3583.33, 5000.00, 7083.33, 595148.17], respectively.

To test the impact of the hard data available for lenders in the Lending Club platform, a multiple logistic regression was used to determine the impact of these variables on whether or not the full proportion of funding was received. Another multiple logistic regression was used to determine the impact of these variables on the likelihood of a loan being charged off. The results of these tests are illustrated in Table 4.

Table 4. Regression Output for Hard Data

Predictor Variables	Explanatory Power for Proportion Funded (Nagerkerle R^2 10.5%) **B**	Explanatory Power for Loan Defaults (Nagerkerle R^2 12.1%) **B**
Credit Grade	**-.106*** **	**.136*** **
FICO Grade	**-.025** **	**.026*** **
Public Records	**.327***	-1.255 NS
Months Since Last Record	**.005** **	.007 NS
Delinquent Amounts	<.001 NS	-.002 NS

Months Since Last Delinquency	.002 NS	.005 NS
Home Ownership	-.016 NS	**.310***
Monthly Income	<.001 NS	**<.001****
Employment Length	.028	-.052 NS
Debt to Income Ratio	**2.588****	-1.556 NS
Open Credit Lines	.016 NS	-.040 NS
Total Credit Lines	-.003 NS	.023 NS
Credit Inquiries Past 6 Months	.004 NS	.112 NS
Revolving Credit Balance	**<.001***	<.001 NS
Revolving Line Utilization	-.405 NS	.646 NS
Loan Amount	**<.001*****	<.001 NS
Loan Length	**.046*****	-.017 NS
Interest Rate	**2.558 ***	**-9.576*****
Monthly Payment	**.005*****	<.001 NS
* = p<0.05, ** = p<0.01, *** = p<0.001, NS = not significant		

The hard data explains 10.5% of the variance in funding. The data relating to borrowers show that an applicant's likelihood of receiving all of the requested investment is positively correlated with a lower (better) credit grade, a lower (better) FICO range, fewer public records on file, more months since the most recent public record, a lower debt-to-income ratio, and a higher revolving credit balance. The data relating to the terms and conditions of the loan show that the loan amount, the loan duration, the interest rate, and the monthly repayments also impact an applicant's likelihood of receiving investment. No other variables demonstrate any significant impact.

The hard data on loan repayment explains 12.2% of the variance borrowers for whom loans have been charged off. The data relating to borrowers show that an applicant's likelihood of defaulting on investment is negatively correlated with a lower (better) credit grade, a lower (better) FICO range, and a higher monthly income. The data relating to the terms and conditions of the loan show that an applicant's likelihood of defaulting on investment is negatively correlated with higher interest rates. The data also show that the home ownership status of borrowers is a significant predictor of defaulting on loans. This was due to a slightly lower proportion of borrowers with mortgages were among those defaulting (39.6%) as opposed to those not defaulting (47.5%), and a slightly higher proportion of borrowers renting their homes defaulting (51.6%) as opposed to those not defaulting (44.7%). No other variables demonstrate any significant impact.

The findings surrounding borrowers’ for whom loans were charged off show that little of the standard credit information for a borrower reliably predicts their likelihood of defaulting on investment. Similarly, from the quantitative terms and conditions of the loan, neither the loan amount, nor the loan duration, nor the monthly repayments predict borrowers' likelihood of defaulting.

These findings are interesting for several reasons. Firstly, credit scores arguably represent a summary of all other data relating to the financial history and identity of a borrower, thus it is not surprising they present significant predictive power for lending. Yet, the general lack of support for data relating to specifics of borrowers’ financial histories suggests that borrowers are not concerned with the details of a borrower's past and present financial data. Secondly, the importance of the four factors relating to the terms and conditions of the loan illustrate that this may be of higher interest to lenders than the details of the borrowers themselves. Thirdly, many of the variables predicting investment do not actually predict repayment, with the exception of borrowers’ credit scores and the interest rate of the loan. Fifthly, and perhaps most interestingly, the overall explanatory power from all of these variables is surprisingly low, suggesting that lenders do not make decisions based on this hard data in isolation (assuming selections are not random and that lenders share some decision-making criteria). Thus, lending behavior on the Lending Club platform may only be moderately informed by hard data relating to the borrower in question and the loan being requested. All of this reinforces the possibility that soft data on Lending Club plays an important part in predicting lending behavior.

Iteration 2: How soft data predicts lending behavior on Lending Club

The data show that 44,982 (38.6%) of applicants chose to leave the loan description field blank, with the remainder of borrowers making some additional effort to contextualize their request. The provision of soft data by borrowers on Prosper has been argued to represent a desire to reduce information asymmetry and increase the perceived trustworthiness of those borrowers (Pötzsch and Böhme 2010, Herzenstein *et al.,* 2011b). Similarly, longer loan descriptions may exaggerate this effect by representing even greater effort on the part of borrowers (Larrimore *et al.,* 2011). Of those that provide loan descriptions on Lending Club, the lengths of these descriptions (measured in number of characters) vary notably, with a five figure summary of [1, 110, 207, 333, 3966]. However, contrary to findings on Prosper, a hierarchical regression shows that the provision (vs non-provision) of a loan description has no significant impact on a borrower's likelihood of receiving investment for borrowers using Lending Club, and that the relative length of loan descriptions if provided has negligible explanatory power (see Table 5). This suggests that - if anything - it is the content of these descriptions that is important for lenders' investment decisions; not their existence or length.

Table 5. Summary of Hierarchical Regression Analysis for Variables Predicting The Proportion of a Loan Request Received

Variable	Model 1	Model 2	Model 3
	B	*B*	*B*
Credit Grade	**-.106*****	**-.106*****	**-.104*****
FICO Grade	**-.025****	**-.025****	**-.024****
Public Records	**.327***	**.322***	**.316***
Months Since Last Record	**.005****	**.005****	**.005****
Delinquent Amounts	<.001 NS	<.001 NS	<.001 NS

Months Since Last Delinquency	.002 NS	.002 NS	.002 NS
Home Ownership	-.016 NS	-.017 NS	-.024 NS
Monthly Income	<.001 NS	<.001 NS	<.001 NS
Employment Length	.028	.028	.028
Debt to Income Ratio	**2.588****	**2.580****	**2.471****
Open Credit Lines	.016 NS	.016 NS	.016 NS
Total Credit Lines	-.003 NS	-.003 NS	-.003 NS
Credit Inquiries Past 6 Months	.004 NS	.005 NS	.004 NS
Revolving Credit Balance	**<.001***	**<.001***	**<.001***
Revolving Line Utilization	-.405 NS	-.398 NS	-.372 NS
Loan Amount	**<.001*****	**<.001*****	**<.001*****
Loan Length	**.046*****	**.046*****	**.045*****
Interest Rate	**2.558 ***	**2.528 ***	**2.427 ***
Monthly Payment	**.005*****	**.005*****	**.005*****
Loan Description Provided		-.068 NS	.055 NS
Length of Loan Description			**<.001***
Nagerkerle R^2	0.105	0.105	0.109
* p < 0.05, ** p < 0.01, *** p < 0.001, NS = Not Significant			

Research on Prosper also suggests borrowers' desire to reduce information asymmetry may have positive indications of their likelihood to repay investment (Herzenstein *et al.*, 2011b, Sonenshein *et al.*, 2011). Unlike the mixed impact of loan descriptions on investment within Lending Club, the impact of the provision of loan description on repayment is significant. However, unlike Prosper, the effect on Lending Club is negative, i.e. those providing loan descriptions were significantly more likely to default on loans. A hierarchical logistic regression demonstrates that both the provision of a loan description, as well as the length of that loan description, predict more loans being charged off (see Table 6).

Table 6. Summary of Hierarchical Logistic Regression for Variables Predicting the Likelihood of a Loan Being Charged Off

Variable	Model 1	Model 2	Model 3

	B	*B*	*B*
Credit Grade	**.136*** **	**.142*** **	**.138*** **
FICO Grade	**.026*** **	**.027*** **	**.025*** **
Public Records	-1.255 NS	-1.148 NS	-1.121 NS
Months Since Last Record	.007 NS	.007 NS	.007 NS
Delinquent Amounts	-.002 NS	-.002 NS	-.002 NS
Months Since Last Delinquency	.005 NS	.005 NS	.005 NS
Home Ownership	**.320***	**.329***	**.345***
Monthly Income	**<.001***	**<.001***	**<.001***
Employment Length	-.052 NS	-.051 NS	-.053 NS
Debt to Income Ratio	-1.556 NS	-1.442 NS	-1.166 NS
Open Credit Lines	-.040 NS	-.043 NS	-.043 NS
Total Credit Lines	.023 NS	**.025***	**.024***
Credit Inquiries Past 6 Months	.112 NS	.106 NS	.105 NS
Revolving Credit Balance	<.001 NS	<.001 NS	<.001 NS
Revolving Line Utilization	.646 NS	.538 NS	.486 NS
Loan Amount	<.001 NS	<.001 NS	<.001 NS
Loan Length	-.017 NS	-.018 NS	-.017 NS
Interest Rate	**-9.576*** **	**-9.419*** **	**-9.312*** **
Monthly Payment	<.001 NS	-.001 NS	-.001 NS
Loan Description Provided		**.872*** **	**.668** **
Length of Loan Description			**.001** **
Nagelkerle R^2	0.121	0.141	0.149
* p < 0.05, ** p < 0.01, *** p < 0.001, NS = Not Significant			

However, it is difficult to draw conclusions based upon this finding, as this may be due to the increasing pressure for borrowers to address concerns when they are struggling to meet their repayments (and thus communicating with creditors through the loan description field).

The mixed impact of loan descriptions on lenders' investment decisions that is observed on Lending Club resonates with existing research of the impact of pictures in borrower profiles on Prosper. Studies have found that it is not necessarily the presence of pictures that influences lenders but rather the content of those pictures (Pope and Sydnor 2011, Duarte *et al.*, 2012). This suggests the types of data shared in the Loan Description field are what determine the impact of personal transparency on lenders' decision making. The frequency of each type of data shared in the loan descriptions in our data set is illustrated in Table 7.

Table 7. Frequencies of codes in 1000 sample records

Ordinal variables	No new information	Some new information	Substantial new information
Additional loan purpose information	351	551	98
Justification of financial situation	925	56	19
Claims of future financial responsibility	882	107	11
Claims of past financial responsibility	769	200	31
Acknowledgement of financial transgressions	977	23	-
Denial of financial transgressions	994	5	1
Education details	944	46	10
Employment status	733	140	127
Occupation	902	18	80
Binary variables	**No mention**		**Mentioned**
Links to other members	1000		-
Links to other online presences	1000		-
Links to photographs	1000		-
Age	997		3
Gender	969		31
Ethnicity	1000		-
Marriage details	932		68
Children and dependents	950		50
Hobbies	995		5
Health and obesity	955		45
Religious or political views	998		2
Claims of kindness	969		31
Please or thanks	818		182
Existence of follow-up comments	826		174

As for the previous investigation of hard data, the impact of the coded soft data was tested using a multiple logistic regression to determine the impact of these variables on whether or not the full proportion of funding was received. Another multiple logistic regression was used to determine the impact of these variables on the likelihood of a loan being charged off. The results of these tests are illustrated in Table 8.

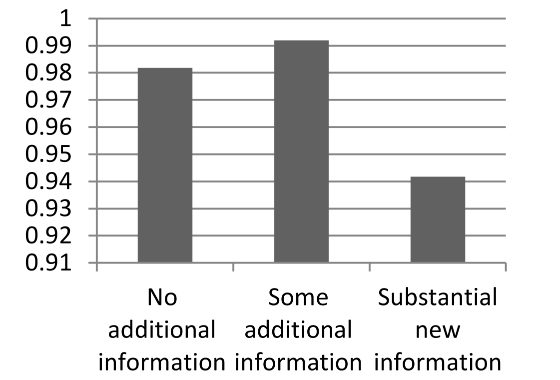
Table 8. Regression Output for Soft Data

Predictor Variables	Explanatory Power for Proportion Funded (Nagerkerle R^2 12.5%) **B**	Explanatory Power for Loan Defaults (Nagerkerle R^2 12.3%)

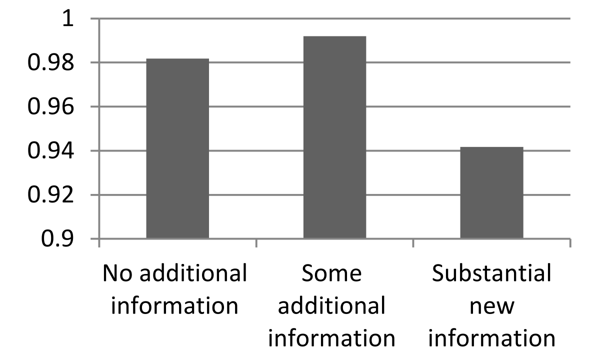
		B
Additional loan purpose information	**-.371*** **	.389 NS
Justification of financial situation	-.085 NS	**.725***
Claims of future financial responsibility	-.190 NS	**1.114*** **
Claims of past financial responsibility	.098 NS	-.272 NS
Acknowledgement of previous transgressions	-.032 NS	-1.191 NS
Denial of previous transgressions	-3.339 NS	1.560 NS
Age	-20.855 NS	-19.271 NS
Gender	-.329 NS	.153 NS
Education details	.164 NS	-1.532 NS
Employment status	**-.363*** *	-.068 NS
Occupation details	**-.401*** *	.252 NS
Hobbies	-.424 NS	-18.830 NS
Marriage details	**-.287 ***	.111 NS
Children and dependents	.182 NS	.081 NS
Health and obesity	.379 NS	-.707 NS
Religious or political views	.466 NS	-21.124 NS
Claims of kindness	.627 NS	-.314 NS
Please or thanks	-.117 NS	.608 NS
Existence of follow-up comments	.172 NS	.128 NS
* = p<0.05, ** = p<0.01, *** = p<0.001, NS = not significant		

This regression predicts 12.5% of the variance in funding. The data show that an applicant's likelihood of receiving investment is negatively correlated with the mention of their marriage status, additional information concerning the purpose of the loan, the borrower's occupation, or their current employment status. No other soft data demonstrate any significant impact.

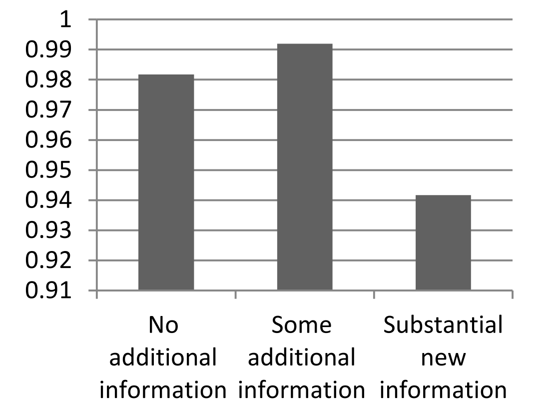
The four significant binary variables contain qualitatively different information. Breakdowns of the types of information coded are presented in Figure 6. These breakdowns illustrate two trends. Firstly, lenders not only responded more negatively to borrowers providing more information about the purpose of the loan, their occupation, and their employment status, this negative response increased as borrowers' level of detail increased. Secondly, borrowers mentioning their marriage or intention to get married were less likely to receive investment.



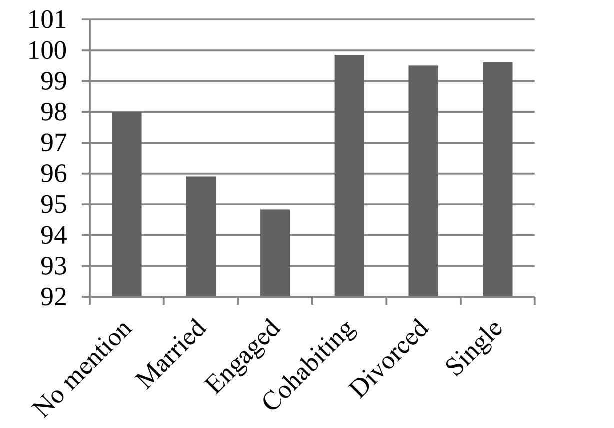
Information given regarding the purpose of the loan.



Information given regarding borrower's employment status.



Information given regarding the borrower's occupation.



Information given regarding borrower's marital status.

Figure 6. Breakdown of binary variables in terms of proportion of funding received.

The logistic regression for the soft data on loan repayment predicts 12.6% of the variance in the number of borrowers for whom loans are charged off. The data show that an applicant's likelihood of repaying investment is negatively correlated with justifications of their current financial situation, claims of future responsibility, and statements of gratitude. No other variables demonstrate any significant impact. Where these statements of gratitude are broken down, the data suggest that borrowers who thank lenders in advance of their investment are those most likely to default on their loans, $\chi 2$ (5, N = 1,000) = 13.93, p = 0.001 (see Table 9).

Table 9. Contingency table of borrowers' statement of gratitude and loan default

	No mention	Thanks	Please	Total
Loans not charged off	797 (82.7%)	153 (15.9%)	14 (1.5%)	964
Loans charged off	21 (58.3%)	14 (38.9%)	1 (2.8%)	36
Total	818	167	15	

CONCLUSIONS

In this paper we have explored the relationships between user information sharing, information sharing expectations, and user behavior, through a two-iteration analysis of historical transactional data from Lending Club, a large *Internet-enabled Peer-to-Peer Lending System* (IP2PLS). Such systems are of key interest to the Information Systems research community, as they have emerged as a disruptive technological genre within the financial services area and are analogous to other Internet-based information systems supporting collective intelligence and action (such as open source software and innovation marketplaces) that have far reaching organizational and social implications.

In line with extant research we conceptualize such systems in terms of their ability to support humanizing behaviors and exchanges, not simply financial transactions. Thus, our study was driven by the saliency of social identity, personal transparency, and information sharing in such systems (and the role played by the system in supporting such sharing. Arguing that current knowledge is limited by bias towards particular IP2PLS providers, we chose to analyze a very large but under-researched platform. Our choice of site was driven specifically by the evidence that the functionalities and participants of different P2PL platforms vary widely (Wang *et al.*, 2009, Bachmann *et al.*, 2011).

Our study makes several key contributions towards a more complete and heterogeneous picture of user behavior in IP2PLS, and has implications for researchers of other crowdfunding systems and/or other emerging forms of IS-enabled collective action, as well as for system designers, users and providers.

Our final model, in which we theorize the impact of social identity and information provision on the behavior of users in the IP2PLS, is illustrated in Figure 7. It is noteworthy that our findings vary considerably from the observations of other IP2PLS previously investigated in the literature.

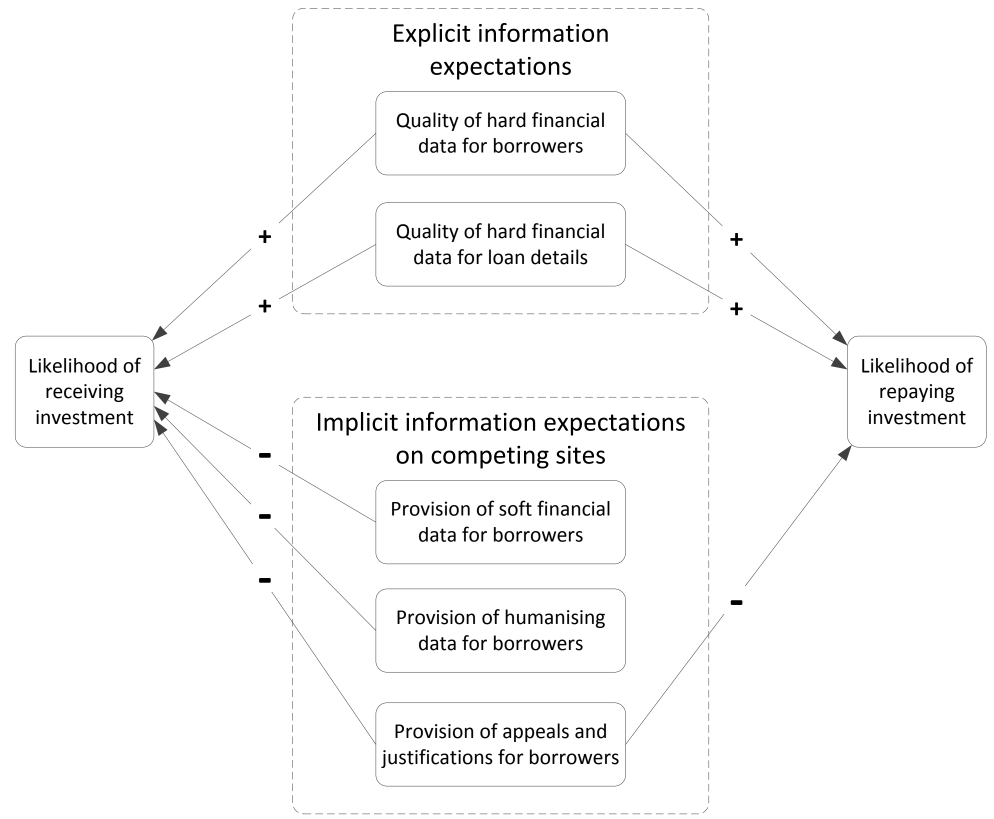


Figure 7. Revised Model of Relationship between Information Sharing and User Behavior in Lending Club

First, we see that the kinds of hard financial data that have been identified as important on other platforms, have limited predictive power for Lending Club, both in terms of borrower repayment but especially in terms of lender-users behavior and investment decision-making. Second, we find that the hard data that best predicts borrower repayment behavior in Lending Club is not necessarily the same data that predicts the likelihood of borrowers receiving investment. For example the home ownership status of borrowers is a useful predictor of repayment neglected in borrower decision making, whereas a borrower’s revolving credit balance appears to have little or no predictive power over repayment, yet demonstrates a significant influence on lenders. These first two sets of findings suggest that much of the hard financial data known to impact upon lending behavior in platforms such as Prosper, does not impact lending behavior on Lending Club.

Third, we found that the ‘soft’ financial and personal data revealed by borrowers within descriptions of the proposed loans, it is highly noteworthy that this soft data explained more of the variance in investment behavior than lenders’ hard financial data, which demonstrated no significant predictive power with the smaller sample size ($p = .696$). However, and equally notable, the majority of the predictive factors identified on other lending platforms had no significant impact on Lending Club. Indeed while some soft data did demonstrate significant predictive power, much of the data that benefitted borrowers on other platforms was in fact harmful to both investment and repayment on Lending Club.

Fourth, none of the factors that were predictive of lending were predictive for repayment.

Fifth, soft data that humanize borrowers and/or characterize them as family-oriented, while beneficial on platforms such as Prosper, had the opposite impact on Lending Club. Rather than demonstrating the reliability and/or trustworthiness of a borrower, such data appeared to dilute that borrower’s financial identity on Lending Club, to the detriment of their likelihood of receiving investment.

These findings, taken together, expand our understanding of the wide variety of IP2PLS user behaviors between systems. These observed variations are our primary data contribution and significantly extend our empirical knowledge of IP2PLS user behavior.

More critically, they expand our conceptual understanding of IP2PLS, and demonstrate the usefulness of viewing IP2PLS, through the lens of Social Identity Theory, as systems that *support user decision making by enabling humanizing/social information exchanges*. This conceptualization forms our primary theoretical contribution. Specifically, and consistent with existing research on identity theory, our findings suggest that the impact of information sharing on user behavior is related to *platform-specific sharing expectations*. Indeed, we see that unlike on platforms such as Prosper, *revealing detailed personal information appears to be outside of the norm on Lending Club (i.e. occurs in a minority of requests and is not explicitly supported by the platform)*.

Thus, lender-users appear troubled by unexpectedly transparent borrowers who disclose such information, rather than simply presenting themselves in purely financial terms. Interestingly, we note again that this unexpected humanizing soft data that appears to deter investment behavior has no actual predictive power over borrowers' likelihood of repayment. In other words, the poor reception received by borrower-users that share "too much" data cannot simply be understood as a financial risk response, but is instead to be understood as a violation of community expectations.

Thus our study (1) challenges the assumption that IP2PLS are homogenous by revealing contrary findings, (2) transforms our understanding of IP2PLS from purely transactional market platforms to social identity and information exchange platforms, and (3) consequently extends our understanding of IP2PLS user behavior. For future researchers, our work implies that a more heterogeneous view of IP2PLS is required, as well as a richer appreciation of the interplay between user identity behaviors, community expectations and platform functionalities.

Our work thus also has practical implications for the designers, developers and managers of IP2PLS. First, the heterogeneous nature of IP2PLS communities revealed by the comparison of our work to other IP2PLS studies suggests that there is not a one-size-fits-all ideal design for IP2PLS. Rather IP2PLS design and functionality must be driven by user information requirements and user information sharing expectations (that factors that are significant in one IP2PLS may not be significant in another (or indeed, may have an inverse relationship, where data are *standard* in one domain, but are *unexpected* in another). Second, the perspective shift from pure financial transaction platform to social identity platform suggests that the design of IP2PLS may better serve its users by facilitating more structured and thus "normal" presentation of social identity information within the platform. This would allow borrower-users to express identity in ways that may allow lender-users to evaluate them in a manner that is more reliable and less prone to cognitive bias.

REFERENCES

Ansoff, H. I. (1975) Managing Strategic Surprise by Response to Weak Signals. California Management Review 18(2): 21-33.

Ashforth, B. E., & Mael, F. (1989) Social identity theory and the organization. *Academy of management review*, 14(1): 20-39.

Bachmann, A., Becker, A., Buerckner, D., Hilker, M., Kock, F., Lehmann, M., Tiburtius, P. & Funk, B. (2011) Online Peer-to-Peer Lending -- A Literature Review. *Journal of Internet Banking & Commerce* 16(2): 1-18.

Barry, J. J. (2012) Microfinance, the Market and Political Development in the Internet Age. *Third World Quarterly* 33(1): 125-141.

Berger, S. C. & Gleisner, F. (2009) Emergence of Financial Intermediaries in Electronic Markets: The Case of Online P2P Lending. *Business Research* 2(1): 39-65.

Benbasat, I., & Zmud, R. W. (2003). The Identity Crisis within the IS Discipline: Defining and Communicating the Discipline's Core Properties. MIS Quarterly, 27(2), 183–194. P.186

Bradford, C. S. (2012) Crowdfunding and the Federal Securities Laws. *Columbia Business Law Review* 2012(1): 1-150.

Briceno Ortega, A. C. & Bell, F. (2008) Online social lending: Borrower-generated content. In: In: Proceedings of The Americas Conference For Information Systems, Toronto, Canada.

Burke, P. J. (1991) Identity processes and social stress. *American Sociological Review* 56(6): 836-849.

Burke, P. J., & Reitzes, D. C. (1991) An identity theory approach to commitment. *Social Psychology Quarterly* 54(3): 239-251.

Bushman, R., & Smith, A. (2003). Transparency, financial accounting information, and corporate governance. Financial Accounting Information, and Corporate Governance. *Economic Policy Review* 9(1): 65-87.

Ceyhan, S., Shi, X. & Leskovec, J. (2011) Dynamics of bidding in a P2P lending service: effects of herding and predicting loan success. Proceedings of the 20th international conference on World wide web. Hyderabad, India.

Chaffee, E. C. & Rapp, G. C. (2012) Regulating Online Peer-to-Peer Lending in the Aftermath of Dodd-Frank: In Search of an Evolving Regulatory Regime for an Evolving Industry. *Washington and Lee Law Review* 69(2): 485-533.

Chen, D. & Han, C. (2012) A Comparative Study of online P2P Lending in the USA and China. *Journal of Internet Banking & Commerce* 17(2): 1-15.

Chen, N., Ghosh, A. & Lambert, N. (2009) Social Lending. Proceedings of the 10th ACM conference on Electronic commerce. Stanford, California, USA.

Collier, B. C. & Hampshire, R. (2010) Sending mixed signals: multilevel reputation effects in peer-to-peer lending markets. Proceedings of the ACM conference on Computer supported cooperative work. Savannah, Georgia, USA.

Cunningham, S. (2013) Lending Club vs Prosper: The Ultimate Showdown. *Lending Memo* online publication 25 March, http://www.lendingmemo.com/lending-club-vs-prosper/

Dezső, L. & Loewenstein, G. (2012) Lenders' blind trust and borrowers' blind spots: A descriptive investigation of personal loans. *Journal of Economic Psychology* 33(5): 996-1011.

Do, H., Jeon, S., Banker, R., Lee, B. & Yoo, B. (2012) Is the Leaderboard Information Useful to Investors?: The Leaderboard Effect in P2P Lending. In: Proceedings of the International Conference for Information Systems, Orlando, USA.

Duarte, J., Siegel, S. & Young, L. (2012) Trust and Credit: The Role of Appearance in Peer-to-peer Lending. *Review of Financial Studies* 25(8): 2455-2483.

Freedman, S., & Jin, G. Z. (2008). Do social networks solve information problems for peer-to-peer lending? Evidence from prosper.com. . *NET Institute Working Paper* (No. 08-43).

Granados, N., Gupta, A., & Kauffman, R. J. (2010). Research Commentary—Information Transparency in Business-to-Consumer Markets: Concepts, Framework, and Research Agenda. *Information Systems Research* 21(2): 207-226.

Granovetter, M. S. (1973) The strength of weak ties. American journal of sociology 78(6): 1360-1380.

Greiner, M. E. & Wang, H. (2009) The Role of Social Capital in People-to-People Lending Marketplaces. Proceedings of the International Conference for Information Systems. Phoenix, USA.

Greiner, M. E. & Wang, H. (2010) Building Consumer-to-Consumer Trust in E-Finance Marketplaces: An Empirical Analysis. *International Journal of Electronic Commerce* 15(2): 105-136.

Herrero-Lopez, S. (2009) Social interactions in P2P lending. Proceedings of the Workshop on Social Network Mining and Analysis. Paris, France.

Herzenstein, M., Dholakia, U. M. & Andrews, R. L. (2011a). Strategic Herding Behavior in Peer-to-Peer Loan Auctions. *Journal of Interactive Marketing* 25(1): 27-36.

Herzenstein, M., Sonenshein, S. & Dholakia, U. M. (2011b) Tell Me a Good Story and I May Lend You Money: The Role of Narratives in Peer-to-Peer Lending Decisions. *Journal of Marketing Research* 48: S138-S149.

Hiltunen, E. (2008) The future sign and its three dimensions. Futures 40(3): 247-260

Iacobuzio, T. (2006) Can’t get a bank loan? Try asking your online peers. *Bank Technology News* 19(7): 25-27.

Iyer, R., Khwaja, A. I., Luttmer, E. F., & Shue, K. (2009). Screening in new credit markets: Can individual lenders infer borrower creditworthiness in peer-to-peer lending? *John F. Kennedy School of Government Working Paper Series* (No. w15242).

Jeong, G., Lee, E. & Lee, B. (2012) Does borrowers' information renewal change lenders' decision in P2P lending?: an empirical investigation. Proceedings of the International Conference on Electronic Commerce. Singapore, Singapore.

Kappel, T. (2008) Ex ante crowdfunding and the recording industry: A model for the us. *Loyola of Los Angeles Entertainment Law Review* 29(3): 375-385.

Katz, M. L. (1996) Remarks on the economic implications of convergence. Industrial and Corporate Change (5:4): 1079-1095

Katz, M. L., & Shapiro, C. (1985) Network externalities, competition, and compatibility. The American economic review 75(3): 424-440.

Klafft, M. (2008a). Online peer-to-peer lending: a lenders' perspective. In Proceedings of the International Conference on E-Learning, E-Business, Enterprise Information Systems, and E-Government, Las Vegas, USA.

Klafft, M. (2008b). Peer to peer lending: Auctioning microcredits over the internet. In Proceedings of the International Conference on Information Systems, Technology and Management, Dubai, UAE.

Kupp, M. & Anderson, J. (2007) Zopa: Web 2.0 meets retail banking. *Business Strategy Review* 18(3): 11-17.

Larrimore, L., Jiang, L., Larrimore, J., Markowitz, D. & Gorski, S. (2011) Peer to Peer Lending: The Relationship Between Language Features, Trustworthiness, and Persuasion Success. *Journal of Applied Communication Research* 39(1): 19-37.

Lee, E. & Lee, B. (2012) Herding behavior in online P2P lending: An empirical investigation. *Electronic Commerce Research and Applications* 11(5): 495-503.

Lin, M. (2009) Peer-to-peer lending: An empirical study. In: Proceedings of The Americas Conference For Information Systems, San Francisco, USA.

Lin, M., Prabhala, N. R. & Viswanathan, S. (2009) Can social networks help mitigate information asymmetry in online markets. In: Proceedings of the International Conference for Information Systems, Phoenix, USA.

Livingston, L. S. & Glassman, T. (2009) Creating A New Type Of Student Managed Fund Using Peer-To-Peer Loans. *Business Education & Accreditation* 1(1): 1-14.

Lu, Y., Gu, B., Ye, Q. & Sheng, Z. (2012) Social Influence and Defaults in Peer-to-Peer Lending Networks. In: Proceedings of the International Conference for Information Systems, Orlando, Florida.

Luo, C., Xiong, H., Zhou, W., Guo, Y. & Deng, G. (2011) Enhancing investment decisions in P2P lending: an investor composition perspective. In: Proceedings of The 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Diego, USA.

Moodie, M. (2013) Microfinance and the Gender of Risk: The Case of Kiva.org. *Signs* 38(2): 279-302.

Pope, D. G. & Sydnor, J. R. (2011) What's in a Picture? Evidence of Discrimination from Prosper.com. *Journal of Human Resources* 46: 53-92.

Pötzsch, S. & Böhme, R. (2010) The role of soft information in trust building: evidence from online social lending. The 3rd International Conference on Trust and Trustworthy Computing, Berlin, Germany.

Puro, L., Teich, J. E., Wallenius, H. & Wallenius, J. (2010) Borrower Decision Aid for people-to-people lending. *Decision Support Systems* 49(1): 52-60.

Puro, L., Teich, J. E., Wallenius, H. & Wallenius, J. (2011) Bidding strategies for real-life small loan auctions. *Decision Support Systems* 51(1): 31-41.

Reay Atkinson, S., Tavakoli Taba, S., Goodger, A., Caldwell, N. H., & Hossain, L. (2014) Synthetic Standards in Managing Health Lifecycles and Cyber Relationships. International Journal on Advances in Life Sciences 6(1 and 2): 41-51

Reay Atkinson, S., Tavakoli Taba, S., & Hossain, L. (2014) Seeding Trusts and [In.] Tolerance in Cyber Communities. In: Proceedings of the Second ASE International Conference on Big Data Science and Computing, Stanford, USA.

Sinanan, J. (2009) Lenders, borrowers and fellows: personal narrative and social entrepreneurship in online microfinance. In: Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7, Melbourne, Australia.

Sonenshein, S., Herzenstein, M. & Dholakia, U. M. (2011) How accounts shape lending decisions through fostering perceived trustworthiness. *Organizational Behavior and Human Decision Processes* 1

Tajfel, H. (1979). Individuals and groups in social psychology*. British Journal of Social and Clinical Psychology, 18(2), 183-19015(1): 69-84.

Tajfel, H., & Turner, J. C. (1979) An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), The social psychology of intergroup relations (pp. 33–47). Brooks/Cole, USA

Uzzi, B. (1996) The sources and consequences of embeddedness for the economic performance of organizations: The network effect. American sociological review 61(4): 674-698

Wang, H., Greiner, M. & Aronson, J. E. (2009) People-to-people lending: The emerging e-commerce transformation of a financial market. In: Proceedings of The Americas Conference For Information Systems, San Francisco, USA.

Wang, H. & Greiner, M. E. (2010) Herding in Multi-winner Auctions. In Proceedings of the International Conference on Information Systems, St. Louis, USA.

Wang, H. & Greiner, M. E. (2011) Prosper—The eBay for Money in Lending 2.0. *Communications of the Association for Information Systems* 29(1): 243-258.

Yum, H., Lee, B. & Chae, M. (2012) From the wisdom of crowds to my own judgment in microfinance through online peer-to-peer lending platforms. *Electronic Commerce Research and Applications* 11(5): 469-483.

Zhang, J. J. & LIU, P. (2012) Rational Herding in Microloan Markets. *Management Science* 58(5): 892-912.

APPENDICES

Appendix 1. List of hypotheses predicting how explicitly expected hard data may impact investment and repayment
H#1: A higher Lending Club credit score will have a positive impact on both a borrower's likelihood of investment and likelihood of repayment
H#2: A higher FICO credit score will have a positive impact on both a borrower's likelihood of investment and likelihood of repayment
H#3: A higher number of public records will have a negative impact on both a borrower's likelihood of investment and likelihood of repayment
H#4: A higher number of account delinquencies will have a negative impact on both a borrower's likelihood of investment and likelihood of repayment
H#5: A lower number of months since the last public record will have a negative impact on both a borrower's likelihood of investment and likelihood of repayment
H#6: A lower number of months since the last account delinquency will have a negative impact on both a borrower's likelihood of investment and likelihood of repayment
H#7: A higher monthly income will have a positive impact on both a borrower's likelihood of investment and likelihood of repayment

H#8: Homeownership will have a positive impact on both a borrower’s likelihood of investment and likelihood of repayment
H#9: A higher number of years in employment will have a positive impact on both a borrower’s likelihood of investment and likelihood of repayment
H#10: A higher number of total credit lines will have a positive impact on both a borrower’s likelihood of investment and likelihood of repayment
H#11: A higher debt-to-income ratio will have a negative impact on both a borrower’s likelihood of investment and likelihood of repayment
H#12: A higher number of open credit lines will have a negative impact on both a borrower’s likelihood of investment and likelihood of repayment
H#13: A higher number of credit inquiries in the past 6 months will have a negative impact on both a borrower’s likelihood of investment and likelihood of repayment
H#14: A higher revolving credit balance will have a negative impact on both a borrower’s likelihood of investment and likelihood of repayment.
H#15: Higher revolving line utilization will have a negative impact on both a borrower’s likelihood of investment and likelihood of repayment
H#16: A higher loan amount will have a negative impact on both a borrower’s likelihood of investment and likelihood of repayment
H#17: A higher length of loan duration will have a negative impact on both a borrower’s likelihood of investment and likelihood of repayment
H#18: A higher interest rate will have a positive impact on a borrower’s likelihood of investment but a negative impact on their likelihood of repayment
H#19: A higher monthly repayment will have a positive impact on a borrower’s likelihood of investment but a negative impact on their likelihood of repayment

Appendix 2. List of hypotheses predicting how possible implicitly expected soft data may impact investment and repayment
H#20: Additional information about a loan’s purpose will have a negative impact on a borrower’s likelihood of investment and likelihood of repayment
H#21: Justifications for a borrower’s current financial situation will have a negative impact on a borrower’s likelihood of investment and likelihood of repayment
H#22: Claims of future financial responsibility will have a negative impact on a borrower’s likelihood of investment and likelihood of repayment
H#23: Claims of past financial responsibility will have a negative impact on a borrower’s likelihood of investment and likelihood of repayment
H#24: Acknowledgement of financial transgressions will have a negative impact on a borrower’s likelihood of investment and likelihood of repayment
H#25: Denial of financial transgressions will have a negative impact on a borrower’s likelihood of investment and likelihood of repayment
H#26: Denial of a borrower’s education will have a negative impact on a borrower’s likelihood of investment and likelihood of repayment
H#27: Denial of a borrower’s occupation will have a negative impact on a borrower’s likelihood of investment and likelihood of repayment
H#28: Denial of a borrower’s employment status will have a negative impact on a borrower’s likelihood of investment and likelihood of repayment
H#29: Links to other members will have a negative impact on a borrower’s likelihood of investment and likelihood of repayment
H#30: Links to other online presences possessed by the borrower will have a negative impact on a borrower’s likelihood of investment and likelihood of repayment

H#31: Links to photographs will have a negative impact on a borrower's likelihood of investment and likelihood of repayment
H#32: Details of the borrower's age will have a negative impact on a borrower's likelihood of investment and likelihood of repayment
H#33: Details of the borrower's gender will have a negative impact on a borrower's likelihood of investment and likelihood of repayment
H#34: Details of the borrower's ethnicity will have a negative impact on a borrower's likelihood of investment and likelihood of repayment
H#35: Details of the borrower's marriage status will have a negative impact on a borrower's likelihood of investment and likelihood of repayment
H#36: Details of the borrower's children and dependents will have a negative impact on a borrower's likelihood of investment and likelihood of repayment
H#37: Details of the borrower's hobbies will have a negative impact on a borrower's likelihood of investment and likelihood of repayment
H#38: Details of the borrower's health or obesity will have a negative impact on a borrower's likelihood of investment and likelihood of repayment
H#39: Details of the borrower's religious or political views will have a negative impact on a borrower's likelihood of investment and likelihood of repayment
H#40: Claims of borrower kindness associated with the loan will have a negative impact on a borrower's likelihood of investment and likelihood of repayment
H#41: Expressions of gratitude will have a negative impact on a borrower's likelihood of investment and likelihood of repayment

FIGURE AND TABLE LEGENDS

(FIGURES AND TABLES WERE EMBEDDED IN DOCUMENT ABOVE FOR EASE OF REFERENCE DURING REVIEW PROCESS)

Figure Legends

Table Legends

Table 1. List of hypotheses predicting how explicitly expected hard data may impact investment and repayment

Table 2. List of hypotheses predicting how possible implicitly expected soft data may impact investment and repayment

Table 3. The online platforms used for data gathering by existing research that investigates the likelihood of lender investment and/or borrower repayment

Table 4. Regression Output for Hard Data

Table 5. Summary of Hierarchical Regression Analysis for Variables Predicting The Proportion of a Loan Request Received

Table 6. Summary of Hierarchical Logistic Regression for Variables Predicting The Likelihood of a Loan Being Charged Off

Table 7. Frequencies of codes in 1000 sample records

Table 8. Regression Output for Soft Data

Table 9. Contingency table of borrowers' statement of gratitude and loan default

APPENDIX A. MANUAL CODING EXAMPLES

Code Category	**No New Information**	**Some New Information**	**Substantial New Information**
	Nothing, or simply repeats structured data found elsewhere in request.	*Terse or vague information not found elsewhere in request.*	*Longer or more specific information not found elsewhere in request.*
Additional loan purpose information	No mention or restates loan category.	"Funds to be used for Development and Pre-Production costs on a feature film"	"Seeking a consolidation loan for 3 debts totalling $1141.83 and current monthly minimum payments of $564.85"
Justification of financial situation	No mention or restates loan category.	"Just finished year one of my career and it involved a lot of moving and relocating and large miscellaneous purchases"	"I have been helping my oldest daughter pay down her student loan from college- she works full time, but low-paying job. My son (graduating college this May) also works, but getting fewer and fewer hours, so my mom and I share helping him with his living expenses. He was able to get scholarships and grants for the first 3 years of college, but I am paying the loan on his senior year and 1/2 his rent/room/board. My middle

			daughter has also been relying on my help. Her husband has abandoned her and their 3-year old daughter. She is working full time, but cannot afford her apartment but also was not allowed out of her lease”
Claims of future financial responsibility	No mention.	"This (loan) will enable me to spread the repayment of these loans over a longer period so that I can free up cash"	"I make $2350 every 2 weeks and only have car and rent payment”
Claims of past financial responsibility	No mention or restates credit rating.	"All of the cards...have never missed a payment or been late"	"I have knocked down $15,000 in debt in less than 4 years"
Acknowledgement/Denial of financial transgressions	No Mention	"Large part of the debt has come from poor decisions regarding a house we purchased in 2005"	"The loan terms were not favorable looking out a few years, but I expected to be able to refinance my way out before it got bad. I was working as an IT manager and my credit record was steadily improving. Unfortunately, my company's parent company sold us, and I was laid off in the process. I took a chance and tried to go into private consulting with a partner.

			Without going into detail, the partnership didn't work - we hardly made money. We broke up, and I was left with his share of six months' expenses, which I had covered. I also had to deal with the eviction of a bad tenant and the associated financial pain. Meanwhile, my $300k 1st mortgage rate climbed to 8%, nearly as high as my equity loan ($33k at 9%). I focused on keeping the mortgage payments current, which meant I made some hard (bad?) decisions to get by, including running up credit card debt, delaying payments, and becoming late on my property taxes (which were not escrowed.)"
Education details	No Mention	"I hold a finance degree"	"I recently received my CPC and graduated class valedictorian... In the professional school I had attended"
Employment status	No Mention	"have been steadily employed for the past 8 years"	"Employed as a Director at COMPANY X for six years"

Occupation	No Mention	***"I am in the medical field and have my own practice"***	*"I am a GS-11 Step 5 Information Technology specialist"*
Code Category	**Not Present**	**Present**	
Links to other members	Absent	Listed account names, real names, URLs.	
Links to other online presences	Absent	Listed account names and URLs	
Links to photographs	Absent	*Not found in sample*	
Age	Absent	Explicitly Stated Age	
Gender	Absent	Explicitly Stated Gender	
Ethnicity	Absent	*Not found in sample*	
Marriage details	Absent	Explicitly stated or implied, e.g. "My wife and I keep seperate accounts..."	
Children and dependents	Absent	Explicitly stated or implied, e.g. "Ever since she had our first child..."	
Hobbies	Absent	Explicitly stated or implied, e.g. "I am a huge Suze Orman fan"	
Health and obesity	Absent	Explicitly stated or implied, e.g. "Over 12 years, I have averaged taking less than 2 sick days per year"	
Religious or political views	Absent	Explicitly stated or implied, e.g. "We are old school Americans who believe in paying off their debt vs. bankruptcy"	
Claims of kindness	Absent	Explicitly stated or implied, e.g. "One of the reasons I am moving is so I can...start helping the poor in my community"	
Please or thanks	Absent	Explicitly stated, e.g. "Please assist me"	
Existence of follow-up comments	Absent	Follow up comments existed	