Privacy in Online Social Lending

Rainer Böhme
International Computer Science Institute
1947 Center Street, Suite 600
Berkeley, CA 94704, USA
rainer.boehme@icsi.berkeley.edu

Stefanie Pötzsch
Technische Universität Dresden
Institute of Systems Architecture
01062 Dresden, Germany
stefanie.poetzsch@tu-dresden.de

Abstract

Online social lending is the Web 2.0's response to classical bank loans. Borrowers publish credit applications on websites which match them with private investors. We point to a conflict between economic interests and privacy goals in online social lending, empirically analyze the effect of data disclosure on credit conditions, and outline directions towards efficient yet privacy-friendly alternative credit markets.

"There is no practice more dangerous than that of borrowing money," said George Washington in 1797. For sure he has not accounted for the privacy risks of 21st century online social lending, the topic of this paper.

Online social lending, also known as *peer-to-peer lending*, has grown rapidly after the launch of the first commercial platform, UK-based *Zopa.com*, in 2005. Drawing on concepts of (offline) micro-finance, the idea of social lending is to provide a marketplace for unsecured personal loans: an online platform lets borrowers advertise credit projects to individual lenders, who decide in which project they invest. Credit risk is shared in project-specific pools of lenders; each member funds a small share of the financed amount. As compensation for taking risk, interest is paid to the lenders, whereas platform operators typically charge fixed (i. e., risk-free) fees. The exact market mechanism differs between platforms and has very recently been subject to research in mechanism design (Chen, Ghosh, and Lambert 2009). Independent of the specific mechanism, matching borrowers' demand with lenders' supply online sidesteps the traditional role of banks as intermediaries in credit markets.

Obviously, this technology-driven paradigm shift in organizing credit markets has a string of economic and social consequences. This paper covers only a small part, namely the role of personal data and the impact of social lending on borrowers' *informational privacy*. Borrowers' privacy is affected since credit applications entail personal data being irrevocably disclosed on the Internet. We will briefly revisit the role of information in credit markets from a privacy protection point of view and contrast it with empirical results from data collected on *Smava.de*, the largest German social lending platform. To the best of our knowledge, this is the first attempt to study social lending from a privacy angle.

AAAI Spring Symposium on Intelligent Information Privacy Management, Stanford University, Palo Alto, CA, March 2010

Background and Research Question

Adequate and coherent privacy regulation requires deep understanding of how the distribution of personal data (i. e., knowledge about individuals' attributes) in a society affects social welfare. It is unlikely that in the near future, a single unified model will be ripe enough to guide policy makers. Therefore it is advisable to break the large problem into smaller, more tractable sector-specific ones. Credit markets form a particularly relevant sector for at least three reasons: first, they are core to modern economies' allocation of capital; second, credit markets are driven by information and thus exhibit a clear link to questions of privacy regulation; and third, advances in technology are about to change the shape of credit markets substantially, as witnessed by the uptake of online social lending over the past couple of years. A welcome side-aspect is the availability of empirical data.

Personal Information in Credit Markets

Aside from traditional functions of financial intermediaries in credit markets, such as size, risk, and maturity transformation, modern economic theory recognizes that information is crucial to prevent market failure (Stiglitz 1981). At the same time, useful information almost always consists of personal data of borrowers. While privacy activists were already concerned about the use of such information by hopefully trustworthy institutions, such as banks and credit bureaus (Jentsch 2007), the privacy problem exacerbates when credit-relevant personal data is disclosed to *every potential lender*. For current platforms, this means all Internet users.

Borrowers' personal data may influence lenders' credit decisions by several mechanisms:

- Most importantly, *information asymmetries* (Akerlof 1970) preclude lenders from distinguishing between good and bad risks. Detailed information on borrowers and their envisaged projects helps to assess the likelihood of timely debt service and to adjust credit conditions accordingly. This improves the overall allocation of capital.
- Knowing personal details of creditors, including their identity, facilitates *contract monitoring* and legal recourse in case of default. The mere possibility to do so can prevent *moral hazard* over the payback period.
- This feedback channel exists also indirectly through *joint liability* and *social sanctions* (Besley and Coate 1995).

Debt service is encouraged by pressure from the borrower's close social network (e. g., family, community) whose reputation is at risk in the case of default. These mechanisms require the disclosure of personal data of the borrower's peers. A borrower's decision to seek credit thus affects the privacy of others, too.

- One has to consider *behavioral effects* as well: trust between market participants can be established by communicating personal attributes as social cues (Kolm 2000).
- Borrowers may use *persuasion* techniques to convince potential lenders of their credit projects with rational (consistency, plausibility) or emotional arguments (helping tendency, altruism, goodness). Such arguments often tell a lot about the borrower and thus contain personal data.

This list is most likely incomplete, but good enough to point out that online social lending exhibits an outright conflict between economic goals and privacy protection goals. This calls for research in at least two directions: first, to study how people deal with the conflict, and second, can we find technical, organizational, or legal provisions to alleviate it?

Related Work on Online Social Lending

Online social lending has recently attracted interest of scholars in economics and social sciences. Unless otherwise stated, all prior art is based on data of *Prosper.com*, the major online social lending platform serving US residents.

Ravina (2007) as well as Pope and Syndor (2008) look at discrimination in credits decisions made on social lending platforms. They report effects of race, age, gender, and weight on credit conditions, though not always statistically significant. Credit conditions were operationalized by (inverse) interest rate, or the probability of (full) funding success. The predictors were either extracted from categorial data of the project description, or manually assigned by evaluating text or pictures. A study by Herzenstein et al. (2008) comes one step closer to privacy aspects. The authors measured the level of detail of the project description on a 3-step ordinal scale and found it to be a major influencing factor for funding success after controlling for fundamental financial parameters, such as amount, starting interest rate in an auction market, and the (endogenous) duration of the listing. In terms of predictive power, the researchers distinguished between demographic factors (e. g., gender), financial factors (e. g., credit score), and effort measures (e. g., length of project description). The first category was found to have only very little effect on funding success, whereas the latter two categories were found to be decisive.

Another string of research has focused on social network theory and tries to identify decision pattern and peer-influence within the social network of registered users of a social lending platform. An early study found that endorsement by group leaders—typically lenders—has the most positive effect on both funding success and total number of bids (Ryan, Reuk, and Wang 2007). This indicates that in social lending, a personal recommendation can be more important than 'hard' facts, such as credit scores. This finding was confirmed by Berger and Gleisner, who analyzed the role of group leaders as new "financial intermediaries" (Berger and Gleisner 2009). Freedman and Jin (2008) as well as Herrero-Lopez (2009) studied whether the social network *between borrowers* can contribute to reducing information asymmetries and thus is helpful to make good investments. They report empirical evidence that a borrower's affiliation with a reputed group increases the chance of full funding and results in lower interest rates. When looking at the realized default rates, however, the picture becomes more complicated: according to a study by Everett (2008), mere group membership tends to *increase* default rates. Only when distinguishing between groups with supposedly strong and weak interpersonal relations (alumni of one school vs mere profession), a positive outcome (i. e., lower default rate) is observable for groups with strong interpersonal relations. The author interprets this as evidence for social sanctions.

Hypothesis

A very basic empirical research question on privacy in online social lending is to measure the amount of personal data in credit project descriptions and analyze its effect on credit conditions. In line with the above arguments, we expect:

> *Borrowers who disclose more personal data pay lower interest rates.*

This hypothesis is based on the hidden presumption that due to the voluntary disclosure, borrowers publish positive data and withhold less favorable details (Stiglitz 1981).

If our hypothesis can be retained, it can be interpreted as evidence for the conflict between economic and privacy goals. Testing this hypothesis is not very easy, though. One challenge is the conceptual and practical difficulty to quantify personal data disclosure in textual descriptions and pictures. Moreover, personal data is 'soft' information and its effect may be subtle compared to the influence of 'hard' information, like credit score and maturity (Petersen 2004). So some effort is required to control for as much as possible hard information. Lastly, to gauge the effect size of personal data disclosure, a comparison base is needed. We decided to compare with the effect of argument styles in project descriptions because they constitute soft information which is not necessarily confounded with personal data disclosure.

Data and Method

Our data consists of 672 credit projects advertised on the largest German social lending site *Smava.de* between November 01, 2008 and June 12, 2009, representing a total credit amount of 5.5 million euro (US$ 8.25 million). This is about 30 % of all loans on the platform from its start in March 2007 until mid-October 2009.

We have limited the time range to obtain a densely populated sample and avoid singularities in the data during the long launch phase. It also helps to avoid heterogeneity before and after the collapse of Lehman Brothers in September 2008, the climax and turning point of the financial crisis. Aside from language issues, homogeneity over time was reason for preferring *Smava.de* data over the much larger US-based platform *Prosper.com* (loans of US$ 180 million since February 2006). Data of the latter exhibit breaks and instability over time due to several business interruptions, rule changes, and lawsuits with financial supervisors between late 2008 and mid-2009 (Eisenbeis 2009).

Smava.de lets potential borrowers propose credit conditions (amount, interest rate, and maturity of 36 or 60 months), checks their identity and publishes on its website verified demographic information (age, gender, state) along with a credit score and a rough debt service-to-income ratio (so-called *KDF* indicator), as well as a user-provided project description and optional pictures. Lenders can review this information and contribute to its funding in step sizes of 250 euros. When the project is fully funded or after two weeks, whatever is earlier, the (partial) loan is granted via a commercial bank, who partners with *Smava.de* to comply with the local financial supervision regulations. Apparently the bank has no say in the credit decision and immediately securitizes the loan, thereby transferring credit risk to the pool of borrowers. The platform also partners with SCHUFA, the leading private German credit bureau, which provides the credit scores, and with a debt collection agency to handle distressed debt. Borrowers can revise the interest rate upwards if their initial offer receives little response. Borrowers and lenders can appear on the platform under self-chosen nick names, however their full identity is known to and verified by *Smava.de*.

Of the 672 credit projects, 81 (12 %) were not fully funded and excluded from the analysis to make interest rates—the dependent variable in this paper—more comparable between projects. Note that this ratio is as high as 78 % for *Prosper.com* (Ryan, Reuk, and Wang 2007), so studies based on this data source could not exclude partially funded projects without introducing a substantial bias. In our study, the remaining cases were divided by their category label into 439 private and 152 commercial credit projects, as different economic and social mechanisms might apply when making credit decisions of either type. One may argue that privacy is not an issue for commercial credits and rather look at it from a trade secret point of view. However, due to *Smava.de*'s cap at 25,000 euro, all credit projects classified as 'commercial' belong to small non-incorporated businesses, whose owners have privacy interests as well.

Content Analysis

We conducted a content analysis (Holsti 1969) to measure the amount of personal data in credit applications. Variation in personal data disclosure can be found in the textual project descriptions, the voluntary categories of the borrower profile page, and possibly associated pictures. Two trained coders independently rated the textual descriptions and assigned it to categories without knowing our hypothesis. The underlying code book distinguishes between ten categories of personal data, namely the borrower's name, financial situation, education, profession, further special skills or knowledge, housing situation, health, hobbies, contact information (address, phone, e-mail, etc.), and information about close relatives (children or partner). Each category has several subcategories that encode in which detail borrowers disclose personal data of the respective category. For a consistency check, we also asked the coders to rate on a 7-point scale the likelihood that a borrower can be identified based on the complete set of disclosed personal data. Individual ratings were collected for several levels of prior knowledge, i. e., identifiability by relatives, neighbors, colleagues or arbitrary persons with access to a search engine.

Apart from personal data, other soft information was measured by flagging the occurrence of precisely defined arguments in the textual description. The categories include:

- helping tendency, i. e., direct appeals for help;
- emotional statements, such as potentially making the reader feel pity about the borrower's situation;
- reference to alternative funding by commercial banks.

Due to resource constraints, only about 50 % of the available credit projects were coded in full detail. This explains the differences in the number of cases as indicated in the result tables. Nevertheless, a subset of projects has been coded by both coders to calculate the inter-coder reliability. According to Holsti's (1969) popular formula, our overall reliability of 87 % is reasonably good.

Auxiliary Data

To control for fluctuations in the economic environment, we have added monthly data on the average effective interest rates charged by German commercial banks for consumer credit of comparable maturity (Bundesbank time series code: SUD114). We further created two dummy variables to capture fundamental changes on *Smava.de*. First, the platform increased fees for both borrowers and lenders in February 2009. The second invention was the introduction of a *bidding assistant* in May 2009. The new function places bids on behalf of a lender and distributes a given amount of investment capital on several credit projects.

Preliminary Results

Effects of personal data disclosure on credit conditions are expected to be rather subtle, so we have to take a two-stage approach. In the first step, a set of predictors from hard information is regressed on the final interest rate of all fully funded credit projects. The best-fitting specification for this baseline model—after elimination of some weak factors unnecessarily consuming degrees of freedom—is listed in Table 1, along with estimates and statistical significance tests. Private and commercial credit projects appear in separate columns. Space constraints do not allow us to report descriptive statistics, nor elaborate on each of the 2×26 predictors. Most importantly, the overall goodness-of-fit, measured by the adjusted R^2 statistic, signals that the baseline model removes the lion's share of heterogeneity. This is essential to make effects of personal data disclosure measurable.

As to the predictors, not surprisingly the credit score is most influential. Since the model is fitted without intercept, the estimates for SCHUFA classes can be directly interpreted as average rates for the respective class. This rate is further corrected upwards for debt service-to-income ratios above 20 %, relatively more so for commercial projects. (To protect borrowers and lenders alike, *Smava.de* does not allow credit projects with KDF above 80 %.)

Table 1: Baseline regressions (fitted with OLS)

Predictors	Estimated coefficients (%-pts. of interest rate) private	commercial
Credit score fixed effects		
SCHUFA class A (best)	6.59*** (1.031)	4.23 (2.988)
SCHUFA class B	7.34*** (1.029)	5.00 (3.013)
SCHUFA class C	8.44*** (1.033)	6.30* (3.015)
SCHUFA class D	9.55*** (1.038)	6.82* (3.013)
SCHUFA class E	10.42*** (1.029)	7.89** (2.981)
SCHUFA class F	11.30*** (1.023)	9.11** (2.976)
SCHUFA class G	12.89*** (1.027)	11.08*** (3.025)
SCHUFA class H (worst)	14.20*** (1.027)	12.18*** (2.998)
Debt service-to-income ratio		
KDF bracket 20–40 %	0.31 (0.186)	1.21* (0.520)
KDF bracket 40–60 %	0.41* (0.182)	1.37** (0.506)
KDF bracket 60–80 %	0.65*** (0.183)	1.98*** (0.491)
Demographics		
Age (absolute deviation from median)	0.03*** (0.005)	0.02 (0.018)
Gender (1=female)	0.01 (0.094)	0.22 (0.244)
Self-employed (1=yes)	0.25* (0.123)	−0.30 (0.378)
Economic environment		
Time trend (month in sample)	−0.28*** (0.054)	−0.58*** (0.139)
Commercial bank rate	0.53 (0.461)	0.49 (0.948)
Fee raise (step dummy)	0.86*** (0.257)	2.14*** (0.592)
Bidding assistant (step dummy)	−0.24 (0.173)	0.39 (0.411)
Properties of credit project		
Total amount ($\log_2$)	0.01 (0.094)	0.22 (0.244)
Maturity (dummy, 1=60 months)	0.65*** (0.102)	0.62** (0.227)
Length of description ($\log_2$)	−0.07 (0.035)	−0.12 (0.101)
Has own project picture (1=yes)	0.01 (0.110)	−0.31 (0.236)
Has borrower profile (1=yes)	0.06 (0.095)	0.14 (0.226)
Has borrower profile picture	−0.00 (0.146)	0.16 (0.380)
Number of bids	0.03** (0.008)	0.02 (0.014)
Revision of initial rate (1=yes)	0.45*** (0.105)	0.40 (0.240)
Model summary (adj. R^2)	99.3 ($n = 439$)	99.0 ($n = 152$)

$^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, standard errors in brackets

In the demographics section, we do find evidence for discrimination by age in the expected direction: very young and very old private borrowers have to pay higher rates. The trend is also visible for commercial projects, though not significant due to lack of variation and fewer cases. (The age coefficient is estimated for the absolute deviation in years from the sample median of 43.) No evidence can be found for discrimination by gender. Data on race was not available as it is of minor relevance in Germany. Among several (loosely verified) categories of professions, only self-employed private borrowers were penalized with significantly higher interest rates. (All other categories could be removed from the model specification without compromising goodness-of-fit.) As self-employment should be default for commercial credits, the coefficient is very noisy.

Regarding the environment variables, interest rates declined on average over the sample. This can be explained by the monetary easing of the period, although it is somewhat surprising that this development is not captured by the official statistic. The fee rise caused a hike in interest rates (stronger for commercial credit), so lenders seem to be in a better position to pass the cost on to borrowers.

Properties of credit projects appear to matter barely. Higher rates are charged for longer maturities. This is in line with the normal shape of the yield curve over the sample. Conditions for private credit projects tighten significantly when interest rates were revised (possibly endogenous), and with the number of bids (but not the order of magnitude of the total amount). No significant relationship can be found between crude effort measures (length of description, pictures) and the interest rate (unlike Herzenstein et al. 2008).

After controlling for hard information in the baseline model, we extracted its residuals and use them as dependent variable in a sequence of individual models (M1–M7), each of which includes a single predictor. We chose this two-stage approach over a direct inclusion of the privacy variables in the large model for the higher number of cases in the baseline model, as some privacy variables were only collected for a sub-sample. Moreover, this approach considerably shortens the result tables and allows for a better comparability between different indicators on the second stage.

Main Results for Personal Data Disclosure

Models M1 to M4 test our hypothesis with different definitions for the amount of personal data. M1 uses the number of broad categories in which at least one entry has been coded, and M2 repeats this logic for the sub-categories. M3 includes only personal information about other people than the borrower (i. e., family members). Finally, M4 uses a more comprehensive indicator combining the predictors of M2 and M3. The results are reported in Table 2, again broken down by private and commercial credit projects. Overall we find that the amount of personal data, regardless of the indicator, has no measurable influence on the interest rate of private credit projects. Conversely, we do find significant negative effects on the interest rate for commercial credit projects, again regardless of the indicator, and despite fewer cases. For example, data disclosed on all eight categories on average reduces the interest rate by two percentage points

Table 2: Indicators of personal data disclosure regressed on interest rate residuals (independent models in rows)

Indicator (one per model)	private credit projects estimate	adj. R^2	n	commercial credit projects estimate	adj. R^2	n
M1: Number of disclosed data categories [0...9]	0.01 (0.036)	0.0	213	-0.25^{**} (0.099)	6.7	78
M2: Number of disclosed personal details [0...66]	0.00 (0.014)	0.0	213	-0.05^{*} (0.028)	2.4	78
M3: Number of disclosed details about family members [0...20]	0.01 (0.025)	0.0	213	-0.15^{**} (0.063)	5.4	78
M4: Additive index of indicators in M2 and M3 [0 ...86]	0.00 (0.011)	0.0	213	-0.05^{**} (0.023)	5.4	78

$^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$, std. errors in round brackets, ranges in square brackets, coefficient estimates in %-pts. of interest rate

Table 3: Occurrence of arguments in project descriptions regressed on interest rate residuals (independent models in rows)

Indicator (one per model)	private credit projects estimate	adj. R^2	n	commercial credit projects estimate	adj. R^2	n
M5: Helping tendency – direct appeal for help	-0.13 (0.125)	0.0	439	-0.04 (0.294)	0.0	152
M6: Emotion – statements that make readers feel pity	0.30^{**} (0.150)	0.7	439	0.08 (0.548)	0.0	152
M7: Competition – borrower claims to be eligible for bank loan	-0.72^{***} (0.180)	3.3	439	0.32 (0.449)	0.0	152

$^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$, standard errors in brackets, coefficient estimates in %-pts. of interest rate

compared to a similar project without personal data in its description. For a typical commercial loan, this can make a difference of several hundred euros.

Interestingly, neither effect size nor direction differ between personal data of the borrower and data of family members. On the one hand, this means that different mechanisms to combat moral hazard (direct vs social sanctions) cannot be told apart in our data (unlike Everett 2008). On the other hand, this finding raises a flag for privacy regulation. The way incentives are set makes it very difficult to ensure that family members are asked for consent before their personal data is exposed on the Internet. This is so because borrowers seem to improve *their own* credit condition at the expense of *others'* privacy. It will be tricky to deal with this externality.

So our hypothesis is only supported by data for commercial credit projects. It must be rejected for private credit projects. We can only speculate about the reasons, which might include conceptual difficulties in the quantification of personal data with content analysis, or uncontrolled heterogeneity thwarting our expected relationship for private borrowers. The latter is more likely because our subsequent analyses show that credit conditions for private borrowers are in fact sensitive to soft information.

Cross-check with Other 'Soft' Information

Results in Table 3 reveal that private credit decisions can be influenced by soft information, such as argument style. While direct appeals for help (M5) are neither rewarded nor punished, lenders seem to dislike if borrowers overdo it and arouse pity (M6). The average 0.3 %-pts. surcharge can also be explained with the rational expectation that borrowers who happen to maneuver themselves into pitiful situations might not show exceptional effort when it comes to repaying their debt. The strongest effect can be found for rational arguments that refer to alternative funding from the banking sector (often with a quoted interest rate; M7). If such claims are credible (though not verified), lenders reward it with significantly better conditions. Observe that effects of arguments are not significant for commercial credit projects, most likely due to absence of variance in the predictors.

Of course, our results suffer from the usual limitations. Explorative data analysis suggests that there is still some unexplained heterogeneity, and we might have overlooked a third variable that changes the picture. Moreover, our measures of data disclosure are noisy and—despite internally reliable—not necessarily valid for the concept we are trying to quantify. There is a general need for advances in the quantification and valuation of privacy, and online social lending seems to be a good domain to refine and validate methods. So far, our empirical results should be considered as preliminary and not be over-interpreted in a policy context. We are currently extending the data and plan a more comprehensive presentation of a more robust analysis in the near future.

Discussion and Conclusion

This paper is a first approach to understand privacy challenges in emerging online social lending platforms. Empirical data from the largest German social lending platform only partially supports our hypothesis that disclosure of more personal data is rewarded on the market place with lower interest rates (after controlling for hard economic determinants). In particular, no effect is measurable for private borrowers. This might be attributed to measurement error, or deviation from economic theory. Either lenders act irrationally and do not use all available information, or the presumption that borrowers select only positive information for voluntary disclosure does not hold. Resolving this puzzle is up to further research, possibly extended to other platforms or by additional forms of data collection, such as surveys of borrowers and lenders, or ideally, summary data on direct inquiries of borrowers via personal communication channels.

Temporarily, our results suggest a privacy-friendly recommendation for private borrowers on *Smava.de*: don't disclose personal details – it's simply not worth it.

In the medium term, online social lending might be the domain where interesting questions in the intersection of economic policy and privacy regulation emerge first. While social lending seems to facilitate access to credit for borrowers who would not be served (at reasonable conditions) by the banking system (Freedman and Jin 2008), thereby fostering equality and democratizing credit markets, it comes at the cost of borrower privacy. So one form of inequality is replaced by another, potentially more subtle one: socially disadvantaged members of society are more likely to act as borrowers and thus are in a worse position to protect their informational privacy. This corroborates the notion of "privacy as a luxury good", a finding that emerges from economic analyses of other markets, too (Varian, Wallenberg, and Woroch 2004; Böhme and Koble 2007).

Thinking of the long term, one might ask for technological and organizational measures to mitigate the privacy problems of current online social lending, while at the same time not compromising its benefits. Computer scientists are invited to conceive privacy-enhancing technologies and protocols that solve the match-making problem on credit markets without compromising (sincere) borrowers' privacy. A first line of thinking on the organizational side could be a model borrowed from Lloyd's of London, namely *lead underwriting* of insurance policies. The idea is that not every lender is an expert for each type of credit. Oftentimes investment decisions can be made by just following the example of other, better informed lenders. If this becomes standard, it will be sufficient to disclose credit-relevant personal data only to the first lender for investigation, rather than to the disperse community of every potential lender. If the incentives of such a system are properly balanced, it can be more efficient (better credit decisions at lower cost) and more privacy-friendly at the same time. Most platforms already provide basic recommender systems drawing on the social network of lenders, so a logical next step is to leverage these features to enable more selective data disclosure.

We conclude with a quote of Benjamin Franklin: "If you would know the value of money, go and try to borrow some," and remark that people borrowing money now offer a great research opportunity to learn to know the value of privacy.

Acknowledgments

Thanks are due to Hagen Wahrig and our coders Ulrike Jani and Kristina Schäfer for their excellent research assistance.

Part of the research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 216483. The first author was supported by a post-doctoral fellowship of the German Academic Exchange Service (DAAD).

References

Akerlof, G. A. 1970. The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics* 84(3):488–500.

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

Besley, T., and Coate, S. 1995. Group lending, repayment incentives and social collateral. *Journal of Development Economics* 46:1–18.

Böhme, R., and Koble, S. 2007. On the viability of privacy-enhancing technologies in a self-regulated business-to-consumer market: Will privacy remain a luxury good? *Workshop on the Economics of Information Security (WEIS)*, Carnegie Mellon University, Pittsburgh, PA. http://www.inf.tu-dresden.de/~rb21/publications/BK2007_PET_Viability_WEIS.pdf.

Chen, N.; Ghosh, A.; and Lambert, N. 2009. Social lending. In *Proc. of ACM Conference on Electronic Commerce*, 335–344. Stanford, CA: ACM Press.

Eisenbeis, H. 2009. The government crackdown on peer-to-peer lending. *The Big Money*. http://www.thebigmoney.com/print/1674.

Everett, C. R. 2008. Group Membership, Relationship Banking and Default Risk: The Case of Online Social Lending. http://ssrn.com/abstract=1114428.

Freedman, S., and Jin, G. 2008. Do Social Networks Solve Information Problems for Peer-to-Peer Lending? Evidence from Prosper.com. http://ssrn.com/abstract=1304138.

Herrero-Lopez, S. 2009. Social Interactions in P2P Lending. Workshop on Social Network Mining and Analysis. Paris, France.

Herzenstein, M.; Andrews, R.; Dholakia, U.; and Lyandres, E. 2008. The Democratization of Personal Consumer Loans? Determinants of Success in Online Peer-to-Peer Lending Communities. http://ssrn.com/abstract=1147856.

Holsti, O. R. 1969. *Content analysis for the social sciences and humanities*. Addison-Wesley.

Jentsch, N. 2007. *Financial Privacy*. Berlin Heidelberg: Springer-Verlag.

Kolm, S.-C. 2000. The Logic of Good Social Relations. *Annals of Public and Cooperative Economics* 71(2):171–189.

Petersen, M. A. 2004. Information: Hard and soft. Mimeo.

Pope, D., and Syndor, J. 2008. What's in a Picture? Evidence of Discrimination from Prosper.com. http://wsomfaculty.case.edu/sydnor/prosperpaper.pdf.

Ravina, E. 2007. Beauty, Personal Characteristics and Trust in Credit Markets. http://ssrn.com/abstract=972801.

Ryan, J.; Reuk, K.; and Wang, C. 2007. To Fund Or Not To Fund: Determinants of Loan Fundability in the Prosper.com Marketplace. http://www.prosper.com/Downloads/Research/Prosper_Regression_Project-Fundability_Study.pdf.

Stiglitz, J. E. 1981. Credit rationing in markets with imperfect information. *American Economic Review* 71(3):393–410.

Varian, H. R.; Wallenberg, F.; and Woroch, G. 2004. Who signed up for the do-not-call list? *Workshop on Economics and Information Security (WEIS)*, University of Minnesota. http://www.dtc.umn.edu/weis2004/varian.pdf.