

# Social Lending aus der Perspektive des Datenschutzes

Rainer Böhme und Stefanie Pöttsch

{rainer.boehme|stefanie.poetzsch}@tu-dresden.de

**Abstract:** Als *Social Lending* bezeichnet man die Kreditvergabe zwischen Privatpersonen. Beim *Online Social Lending* veröffentlichen potenzielle Kreditnehmer ihre Gesuche auf entsprechenden Webseiten im Internet, um private Investoren zu finden. Dieser Beitrag weist auf einen Konflikt zwischen wirtschaftlichen Interessen und Datenschutzzielen hin und analysiert ihn mit empirischen Daten der größten deutschen Social-Lending-Plattform *Smava.de*. Der Analyse zufolge lohnt es sich gegenwärtig nicht, mehr personenbezogene Daten zu veröffentlichen als unbedingt notwendig.

## 1 Online Social Lending: Kreditvergabe durch Crowdsourcing

Die erste kommerzielle Plattform für Online Social Lending *Zopa.com* wurde 2005 in England gegründet. Derzeit existieren mehrere Anbieter von Online-Marktplätzen für unbesicherte Konsumentenkredite, deren Businessplan am ehesten mit den aus Entwicklungsländern bekannten Mikrokrediten vergleichbar ist: Potenzielle Kreditnehmer können im Internet ihren Finanzierungswunsch vorstellen und treffen auf private Investoren, die einzeln entscheiden, in welche der angebotenen Kreditvorhaben sie investieren möchten. Bei erfolgreicher Finanzierung wird das Ausfallrisiko auf mehrere Investoren verteilt, die jeweils einen Teil der Gesamtsumme eines Projekts finanzieren und dafür später neben ihrem ursprünglich investierten Betrag auch Zinsen vom Kreditnehmer erhalten. Die Betreiber solcher Plattformen erheben vorher festgelegte Gebühren, die unabhängig vom Kreditrisiko sind. Die Plattformen unterscheiden sich in der konkreten Ausgestaltung ihrer Marktmechanismen [BG09, CGL09].

Insgesamt ist Online Social Lending eine Alternative zum klassischen Bankkredit und umgeht weitgehend das Bankensystem als Finanzintermediär bei der Kreditvergabe. Diese technologiebedingten Entwicklungen auf dem Markt für Konsumentenkredite führen zu einer Reihe von ökonomischen und sozialen Veränderungen [FS08]. Der vorliegende Beitrag beschäftigt sich nur mit einem kleinen Ausschnitt, nämlich der (Un-)Vereinbarkeit von Social Lending und Datenschutzzielen. Mit ihren Kreditgesuchen veröffentlichen Kreditnehmer im Internet eine Vielzahl personenbezogener Daten über sich selbst, aber auch über ihnen nahe stehende Personen. Im folgenden Abschnitt werden deshalb zunächst die Bedeutung von (personenbezogenen) Informationen für Kreditmärkte zusammengefasst und sich daraus ergebende Probleme aus Sicht des Datenschutzes aufgezeigt. Ergänzt wird dieser theoretische Teil im weiteren Verlauf des Beitrags mit einer empirischen Untersuchung. Mit Hilfe realer Daten der größten deutschen Social-Lending-Plattform *Smava.de* wird analysiert, inwiefern Einschnitte aus Sicht des Datenschutzes tatsächlich zu finanziellen Vorteilen, d. h. besseren Kreditkonditionen, führen. Der vorliegende Beitrag unter-

scheidet sich von veröffentlichten vorläufigen Ergebnissen durch einen erweiterten Untersuchungszeitraum und eine differenziertere Kontrolle von Drittvariablen [BP10].

## 2 Theorie

Wissenschaftler aller Disziplinen müssen einräumen, dass ein übergreifendes Verständnis der Konsequenzen von technologiebedingt immer einfacherer Erhebung, Übermittlung und Verarbeitung personenbezogener Daten fehlt. Es ist weithin unklar, wie sich die Verteilung von Informationen über Individuen in einer Gesellschaft auf Kenngrößen wie soziale Wohlfahrt oder Verteilungsgerechtigkeit auswirkt. Abgesehen von normativen Leitsätzen und der Kritik offensichtlichen Fehlverhaltens ist es deshalb schwierig, sachliche Empfehlungen zur Datenschutzregulierung zu unterbreiten. Es ist auch nicht zu erwarten, dass die Problematik hinreichend verstanden wird, bevor reale Entwicklungen Eingriffe notwendig machen. Also liegt es nahe, zunächst beherrschbarere Teilprobleme zu untersuchen.

Kreditmärkte stellen so ein Teilproblem dar, das sich aus mehreren Gründen besonders für eine Analyse eignet: Kreditmärkte nehmen eine zentrale Stellung in einem marktwirtschaftlich organisierten System ein. Weiterhin ist die Rolle von Informationen – oft personenbezogene Daten – auf Kreditmärkten in den Wirtschaftswissenschaften seit den 1970er Jahren untersucht und oft detailliert modelliert worden. Mit Online Social Lending führen technologische Innovationen außerdem zu strukturellen Veränderungen, die als exemplarisch für andere Bereiche der Gesellschaft angesehen werden können. Nicht zuletzt ist mit der Verfügbarkeit von Daten und einer klaren abhängigen Variable (Kreditkonditionen) auch eine Voraussetzung für praktische Untersuchungen erfüllt, die empirische Forschung zu Datenschutzthemen in anderen Bereichen erschwert.

### 2.1 Bedeutung personenbezogener Daten für Kreditmärkte

Zu den klassischen Aufgaben der Finanzintermediäre gehört die effiziente Transformation von Losgrößen, Risiken und Fristen im Kreditgeschäft. Nach der heute vorherrschenden Schule der Informationsökonomik ist zudem die Verfügbarkeit von Information entscheidend für Existenz und Effizienz von Kreditmärkten [Sti81]. Solche Informationen enthalten dabei fast immer auch personenbezogene Daten der Kreditnehmer. Während Datenschützer bereits den Umgang mit diesen Daten bei gemeinhin als vertrauenswürdig angesehenen Institutionen wie Banken oder Kreditauskunfteien kritisieren [Jen07], dürften sich diese Bedenken verschärfen, wenn beim Online Social Lending kreditrelevante personenbezogene Daten im Internet für alle potenziellen Investoren einsehbar sind.

Der Einfluss personenbezogener Daten von Kreditnehmern auf Kreditentscheidungen lässt sich theoretisch mit einer Reihe von Mechanismen erklären: *Asymmetrische Information* zum Zeitpunkt der Kreditvergabe verhindert, dass Investoren zwischen „guten“ und „schlechten“ Risiken unterscheiden können [Ake70]. Bessere Information über Kreditnehmer und deren geplante Vorhaben erlauben eine genauere Schätzung der Ausfallwahrscheinlichkeit und beeinflussen somit die Kreditkonditionen. Personenbezogene Daten der Kreditnehmer erleichtern weiterhin die Überwachung laufender Kreditverträge durch Investoren sowie den Einsatz von Rechtsmitteln im Falle eines Zahlungsausfalls. Allein die-

se Möglichkeit kann bereits Fehlverhalten (engl. *Moral Hazard*) der Kreditnehmer unterbinden. Eine ähnliche Präventivwirkung wird durch die (implizite) Mithaftung von dem Kreditnehmer nahe stehenden Personen und damit verbunden möglichen *sozialen Sanktionen* erreicht [BC95]. Das soziale Umfeld eines Kreditnehmers, beispielsweise Familienmitglieder, kann einen gewissen Druck auf den Kreditnehmer ausüben, die Raten pünktlich zurückzuzahlen, da die Reputation der gesamten Gruppe auf dem Spiel steht. Solche komplexen sozialen Mechanismen setzen voraus, dass der Kreditnehmer zuvor nicht nur Informationen über sich selbst, sondern auch personenbezogene Daten über Personen aus seinem Umfeld offenbart hat. Außerdem gilt es, indirekte Wirkungen der Kommunikation sensibler Informationen zu berücksichtigen. Der Austausch personenbezogener Daten kann stets auch als sozialer Hinweis verstanden werden und fördert den *Vertrauensaufbau* zwischen Akteuren [Kol00, PB10]. Schließlich nutzen Kreditnehmer verschiedene *Argumentationstechniken*, um potenzielle Investoren von ihrem Projekt zu überzeugen. Aussagen, die beispielsweise die Plausibilität des Vorhabens unterstreichen sollen oder an das Hilfeverhalten der Investoren appellieren, enthalten nebenbei oft personenbezogene Daten.

Obwohl diese Aufzählung unvollständig sein mag, verdeutlicht sie bereits das Spannungsfeld zwischen wirtschaftlichen Interessen und Datenschutz beim Online Social Lending. Nahe liegende Forschungsfragen betreffen daher die Analyse des Verhaltens von Akteuren, die diesem Zielkonflikt ausgesetzt sind, sowie den Entwurf von technischen und organisatorischen Maßnahmen, um ihn zu entschärfen ohne dabei den Kreditmarkt zu zerstören.

## 2.2 Stand der Forschung

Seit seiner Entstehung wird Online Social Lending vor allem von wirtschaftswissenschaftlicher Forschung begleitet. Datenquelle ist dabei meistens die US-amerikanische Plattform *Prosper.com*. Allerdings sind den Verfassern weder im Bezug auf *Prosper.com* noch generell andere Forschungsvorhaben bekannt, die Datenschutzaspekte bei Online Social Lending untersuchen.

In [Rav07] und [PS08] wurde die Diskriminierung bestimmter Gruppen bei der Kreditvergabe untersucht. Die Autoren finden einen Einfluss von Hautfarbe, Alter, Geschlecht und Gewicht auf Kreditkonditionen, wenn auch nicht immer statistisch signifikant. In einer weiteren Untersuchung, die der hier betrachteten Datenschutzproblematik einen Schritt näher kommt, haben Wissenschaftler den Detailgrad verschiedener Angaben in Projektbeschreibungen anhand einer dreistufigen Skala gemessen [HADL08]. Sie berichten, dass neben grundsätzlichen Parametern wie Kreditbetrag, Zinssatz und Dauer bis zur Finanzierung auch der Detailgrad zusätzlich angegebener Informationen den Finanzierungserfolg eines Projekts beeinflusst. Die Autoren unterschieden zwischen demographischen Fakten, Finanzinformationen (z. B. Kreditscore) und allgemeinen Aufwandsmaßen (z. B. Länge der Projektbeschreibung). Detaillierte Informationen der beiden letzten Kategorien haben einen nachweisbaren Einfluss auf den Erfolg eines Kreditgesuchs.

Theoretische und empirische Aspekte des Datenschutzes bei der herkömmlichen Kreditvergabe durch Geschäftsbanken werden aus Platzgründen nicht vertieft [KW06, KW08]. Für eine Zusammenfassung weiterer Arbeiten zu Online Social Lending unter Berücksichtigung von sozialen Netzwerken zwischen Akteuren sei auf [BP10] und [PB10] verwiesen.

### 2.3 Hypothese

Aus Perspektive des Datenschutzes ist eine der grundlegenden Forschungsfragen zu Online Social Lending die Rolle von personenbezogenen Daten in Projektbeschreibungen und deren Auswirkungen auf Kreditkonditionen. Vor dem Hintergrund der unter 2.1 genannten Mechanismen lautet die in diesem Beitrag zu untersuchende Hypothese wie folgt:

Kreditnehmer, die mehr personenbezogene Daten veröffentlichen, erhalten ihren Kredit zu einem günstigeren Zinssatz.

Die Hypothese basiert auf der Annahme, dass freiwillig überwiegend personenbezogene Daten angegeben werden, die den Kreditnehmer in einem positiven Licht erscheinen lassen [Sti81]. Die Überprüfung der Hypothese ist mit einer Reihe von Schwierigkeiten verbunden. Eine davon ist – sowohl konzeptionell als auch praktisch – die Quantifizierung von personenbezogenen Daten in Freitexten und Bildern. Außerdem stellen personenbezogene Daten in Projektbeschreibungen „weiche“ Informationen dar, und es ist zu erwarten, dass diese im Vergleich zu „harten“ Fakten wie Creditscore und Laufzeit einen schwächeren Einfluss haben [Pet04]. Letztere müssen also ebenfalls erfasst und beim Hypothesentest kontrolliert werden. Um den Effekt personenbezogener Daten in seiner Größenordnung einschätzen zu können, ist zusätzlich ein Vergleich mit der Wirkung anderer „weicher“ Informationen erforderlich. Dazu werden in der vorliegenden Untersuchung auch die in den Projektbeschreibungen verwendeten Argumentationstechniken berücksichtigt.

## 3 Datenerhebung und Auswertung

Der für die Analyse zur Verfügung stehende Datensatz besteht aus 1530 Kreditprojekten, die zwischen 01. November 2008 und 31. Oktober 2009 auf *Smava.de*, der größten deutschen Plattform für Online Social Lending, veröffentlicht wurden. Die Projekte repräsentieren ein Kreditvolumen von 13 Millionen Euro. Das entspricht etwa 50 % des Gesamtkreditvolumens, welches über die Plattform seit ihrem Start im März 2007 bis März 2010 vermittelt wurde. Der Zeitraum wurde gewählt, um einen möglichst dichten und homogenen Datensatz zu erhalten und gleichzeitig Anomalien innerhalb der relativ langen Startphase vom *Smava.de* zu vermeiden. Außerdem wurde so der kritische Bereich um den Zusammenbruch der US-Investmentbank Lehman Brothers im September 2008 von der Analyse ausgeschlossen, der sich auch auf deutsche Kreditmärkte auswirkte.

*Smava.de* bietet Kreditsuchenden die Möglichkeit, ihr Kreditprojekt vorzustellen und selbst einen Vorschlag zu unterbreiten, zu welchen Konditionen (Zinssatz, Laufzeit, Betrag) sie einen Kredit aufnehmen möchten. Die Betreiber der Plattform prüfen die Identität des potenziellen Kreditnehmers und ergänzen anschließend verifizierte demographische Angaben (Alter, Geschlecht, beruflicher Status) sowie den Creditscore der SCHUFA und eine Abschätzung der Kapitaldienstfähigkeit (KDF), die sich aus dem erwarteten Verhältnis der Tilgungsrate zum Nettoeinkommen errechnet. Diese Fakten werden zusammen mit der Projektbeschreibung des Kreditnehmers und optional bereitgestellten Bildern auf der Internet-Plattform veröffentlicht. Potenzielle Investoren können alle auf der Plattform vorhandenen Kreditprojekte inklusive darin genannter Detailinformationen einsehen und zu

deren Finanzierung durch Gebote in Schritten von 250 Euro beitragen. Sobald ein Projekt vollständig finanziert ist oder nach Ablauf einer Frist von 14 Tagen wird der (Teil-)Kredit durch eine Bank vergeben, die mit *Smava.de* kooperiert. Das Einbeziehen der Bank an dieser Stelle ist aufgrund der in Deutschland geltenden gesetzlichen Bestimmungen der Bankenaufsicht notwendig. Nach Bewilligung durch ausreichend viele Investoren hat die Bank angeblich keinen Einfluss auf die Kreditentscheidung oder Konditionen. Als weiterer Partner von *Smava.de* stellt die SCHUFA für jeden potenziellen Kreditnehmer den entsprechenden Kreditscore bereit. Zur Eintreibung säumiger Tilgungsraten kooperiert *Smava.de* im Interesse der Investoren außerdem mit einem Inkasso-Unternehmen.

Sollte ein eingestelltes Kreditprojekt zunächst nur auf geringeres Interesse bei Investoren stoßen, haben Kreditnehmer die Möglichkeit, den angebotenen Zinssatz nachträglich zu erhöhen, um das Angebot attraktiver zu gestalten. Sowohl Kreditnehmer als auch Investoren sind auf der Plattform unter selbst gewählten Benutzernamen aktiv, ihre tatsächliche Identität ist den Betreibern von *Smava.de* jedoch bekannt und durch ein Postident-Verfahren im Vorfeld verifiziert.

Von den ursprünglich 1530 Kreditprojekten im Untersuchungszeitraum wurden 79 (5 %) nicht vollständig finanziert. Vor der weiteren Analyse wurden diese aus dem Datensatz entfernt, um eine bessere Vergleichbarkeit des Zinssatzes als abhängige Variable in der vorliegenden Untersuchung zu erreichen. Darüber hinaus wurden weitere 228 Projekte herausgefiltert, die innerhalb der ersten zwei Minuten nach Einstellen des Kreditgesuchs vollständig finanziert waren. Damit sollen alle so genannten „Sofortkredite“ von der Untersuchung ausgeschlossen werden. Diese Art von Krediten wird automatisiert bewilligt, d. h. „weiche“ Faktoren, deren Einfluss analysiert werden soll, spielen hier keine Rolle.

### 3.1 Inhaltsanalyse

Die Menge der in den einzelnen Projektbeschreibungen veröffentlichten personenbezogenen Daten wurde mit Hilfe einer Inhaltsanalyse erhoben [Hol69]. Die Kreditprojekte variierten im Detailgrad des Beschreibungstextes, den freiwilligen Angaben im Benutzerprofil sowie möglicherweise hochgeladenen Bildern. Zwei zuvor geschulte Kodierer haben unabhängig voneinander die vorhandenen Projektbeschreibungen bewertet und vorgegebenen Kategorien zugeordnet, ohne jedoch die Hypothese zu kennen. Das Codebuch dieser Inhaltsanalyse definiert zehn verschiedene Kategorien personenbezogener Daten: Name des Kreditnehmers, Kontaktinformationen (Adresse, Telefonnummer, E-Mail, etc.), finanzielle Situation, Ausbildung, Beruf, weitere besondere Fähigkeiten, Wohnsituation, Gesundheit, Hobbys und Informationen über nahe stehende Verwandte (Partner, Kinder). Jede Kategorie wurde wiederum in verschiedene Subkategorien unterteilt, um genau erfassen zu können, wie detailliert Kreditsuchende Auskunft geben. Die Kodierer bewerteten ebenso den Gesamteindruck des Kreditprojekts (positiv, neutral, negativ) auf Grundlage aller zur Verfügung stehenden Informationen. Aus forschungsökonomischen Gründen konnten 125 der verfügbaren Projekte nicht vollständig kodiert werden. Trotzdem wurde ein Teil der Kreditprojekte von beiden Kodierern bewertet, um die Inter-Koder-Reliabilität zu errechnen. Nach der bekannten Formel aus [Hol69] war diese mit durchschnittlich 90 % Übereinstimmung recht gut.

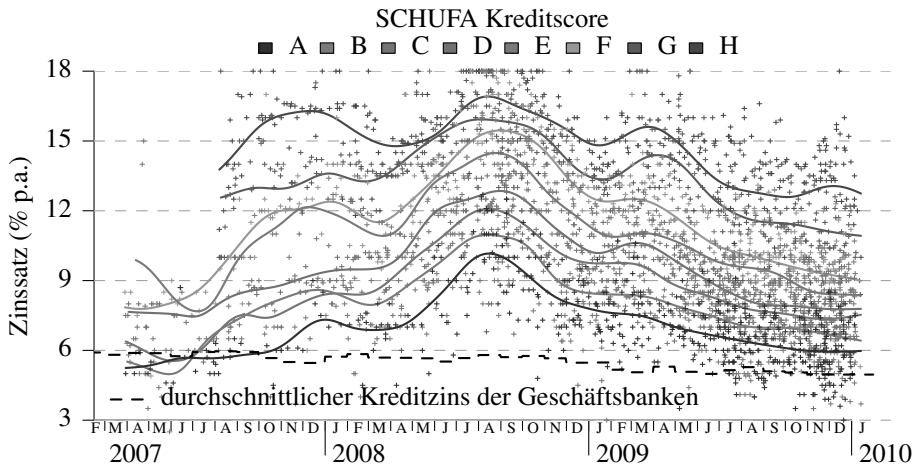


Abbildung 1: Kreditkonditionen auf *Smava.de* im Zeitverlauf, klassifiziert nach SCHUFA-Bonität

### 3.2 Weitere Datenquellen

Neben personenbezogenen Daten wurden für die Auswertung weitere Daten berücksichtigt, die die Höhe des Zinssatzes als abhängige Variable beeinflussen. Dazu zählen der von der Bundesbank ermittelte durchschnittliche Effektivzinssatz für Konsumentenkredite an private Haushalte [Bun] sowie generellen Schwankungen des Zinsniveaus auf *Smava.de*. Um die Auswirkungen dreier grundlegender Änderungen der Plattform zu kontrollieren, wurden entsprechende Dummy-Variablen generiert. Das erste Ereignis im Untersuchungszeitraum war die Erhebung von Gebühren für Kreditnehmer und Investoren im Februar 2009. Im Mai 2009 erfolgte eine Erweiterung der Funktionalitäten des Gebotsassistenten. Dieser steht Investoren zur Verfügung und ermöglicht das automatische Bieten auf Kreditprojekte mit zuvor spezifizierten Eigenschaften. Dabei ist zu beachten, dass neu eingestellte Kreditvorhaben maximal zu 50 % auf diese Weise finanziert werden können, die übrige Hälfte muss durch manuelle Gebote aufgefüllt werden. Die dritte und im relevanten Zeitraum letzte Neuerung war die Einführung von Sofortkrediten im Juli 2009. Sofortkredite werden zu 100 % automatisch mit Hilfe der Gebotsassistenten finanziert und sind deshalb für die vorliegende Untersuchung ausgeschlossen worden. Die Dummy-Variable ist dennoch notwendig, um einen möglichen Effekt dieses neuen Angebots auf das Zinsniveau der Plattform insgesamt zu erfassen.

## 4 Ergebnisse

Abbildung 1 zeigt die Zinssätze aller über *Smava.de* finanzierten Kreditprojekte im Zeitverlauf bis Januar 2010. Jedes Projekt entspricht einem Punkt in der Wolke, die farbliche Kodierung markiert die SCHUFA-Klasse von A (gute Bonität) bis H (schlecht). Die durchgezogenen Kurven sind geglättete Durchschnittszinssätze für die entsprechenden

SCHUFA-Klassen und die gestrichelte Linie zeigt den durchschnittlichen Effektivzinssatz der Geschäftsbanken für vergleichbare Kredite [Bun]. Die SCHUFA-Klassen G und H wurden erst ab Mitte 2007 zugelassen. Die Abbildung verdeutlicht die generell hohe Varianz der Kreditkonditionen sowie deren Fluktuation im Zeitverlauf. Außerdem ist ersichtlich, dass ein Großteil der Varianz durch „harte“ Information, wie die SCHUFA-Klasse, erklärt werden kann. Trotzdem verbleibt ein substanzieller Anteil an Streuung innerhalb jeder SCHUFA-Klasse, der Spielraum für Effekte der zu untersuchenden „weiche“ Informationen lässt. Weiterhin fällt auf, dass das Zinsniveau auf *Smava.de* selbst für Kredite mit guter Bonität sichtbar höher ist als das Zinsniveau der Geschäftsbanken. Im zeitlichen Verlauf lassen sich keine Korrelationen zwischen dem Zinsniveau des traditionellen und des neuen Kreditmarkts erkennen.

Die Wirkung personenbezogener Daten auf den Zinssatz aller vollständig finanzierten Kreditprojekte wurde mit Hilfe einer Regressionsanalyse geschätzt. Tabelle 1 zeigt zwei Spezifikationen mit geschätzten Koeffizienten und Tests auf statistische Signifikanz. In Modell 1 wurden nur „harte“ Informationen als Prädiktoren zugelassen und Modell 2 erweitert die Regression um „weiche“ Einflussfaktoren, wie die freiwillige Veröffentlichung personenbezogener Daten und die Verwendung bestimmter Argumentationstechniken. Mit korrigiertem  $R^2$  von über 98 % erklären die Modelle einen extrem hohen Anteil der Zinsunterschiede, allerdings ist die in Modell 2 zusätzlich erklärte Varianz sehr gering.

Es ist wenig überraschend, dass der Kreditscore am meisten Einfluss auf den Zinssatz hat. Da die Modelle ohne Konstante geschätzt wurden, können die Schätzer der SCHUFA-Klassen direkt als durchschnittliche Zinssätze für die jeweilige Bonität interpretiert werden. Der Zinssatz wird signifikant nach oben korrigiert für Kreditnehmer, deren KDF größer als 40 % ist. Zum Schutz von Kreditnehmern und Investoren erlaubt *Smava.de* keine Kreditprojekte, bei denen dieses Verhältnis über 67 % liegen würde. Die Kontrolle der Auswirkungen von Veränderungen an der Plattform zeigt, dass nach der Modifikation des Gebotsassistenten der durchschnittliche Zinssatz für alle Kreditprojekte gesunken ist. Gleiches gilt für die Einführung der Sofortkredite. Die Gebührenerhöhung hatte keinen messbaren Einfluss, was aber auch an der Konfundierung mit einem Fixed-Effect zum gleichen Zeitpunkt liegen könnte. Bezüglich der monatlichen Schwankungen des Zinssatzes im Untersuchungszeitraum ist festzustellen, dass es im März 2009 zunächst eine punktuelle Steigerung gab und ab Juli 2009 der Zinssatz für das letzte Quartal signifikant gesunken ist. Grund für Letzteres könnte eine leichte Entspannung an den Kreditmärkten sein (die sich jedoch nicht in der offiziellen Zinsstatistik niederschlägt). Eine weitere Erklärung ist die zunehmende Berichterstattung über *Smava.de* als Anlagealternative, welche zu einem Überangebot durch das Hinzukommen von mehr neuen Investoren als Kreditnehmern führte.

Bei den Merkmalen der Projektbeschreibungen zeigt sich, dass höhere Kreditsummen und längere Laufzeiten mit einem höheren Zinssatz einher gehen. Anders als in [HADL08] berichtet, wirkt sich eine (Aufwärts-)Revision des angebotenen Zinssatzes nicht negativ auf die Kreditkonditionen aus. Übereinstimmend mit der Literatur finden wir eine signifikante Verbesserung der Kreditkonditionen, wenn das Projekt mit einem individuellen Bild illustriert wurde. Bilder im nur einen Klick weiter entfernten Benutzerprofil haben dagegen keinen messbaren Einfluss. Projekte mit hohen Geboten erhalten signifikant bessere

Konditionen, allerdings ist diese Variable teilweise endogen und ein Kausalschluss damit unzulässig. Sie wurde trotzdem in die Spezifikation einbezogen, um Heterogenität aus dem Datensatz zu entfernen. Da die Variable auch als Proxy für über „weiche“ Informationen vermittelte Qualitäten des Projekts interpretiert werden kann, wurde sie in einer separaten Analyse ausgeschlossen. Die Ergebnisse für die zusätzlichen Prädiktoren in Modell 2 blieben davon unberührt. Im Abschnitt zur Demographie zeigt sich erwartungskonform eine Benachteiligung von besonders alten oder besonders jungen Kreditnehmern. Der Alterskoeffizient im Modell wurde als absolute Abweichung in Jahren vom Median aller Kreditnehmer berechnet. Dieser lag bei 43 Jahren. Es zeigt sich keine geschlechterspezifische Benachteiligung, allerdings müssen Selbstständige tendenziell mehr Zinsen bezahlen.

Prädiktor	Modell 1	Modell 2
<b>Kreditscore</b>		
SCHUFA Klasse A (gut)	5.90 *** (0.813)	6.16 *** (0.821)
SCHUFA Klasse B	6.68 *** (0.813)	6.96 *** (0.821)
SCHUFA Klasse C	7.71 *** (0.811)	7.99 *** (0.819)
SCHUFA Klasse D	8.22 *** (0.810)	8.47 *** (0.817)
SCHUFA Klasse E	9.36 *** (0.819)	9.59 *** (0.827)
SCHUFA Klasse F	10.09 *** (0.815)	10.38 *** (0.824)
SCHUFA Klasse G	11.93 *** (0.811)	12.18 *** (0.819)
SCHUFA Klasse H (schlecht)	13.33 *** (0.815)	13.62 *** (0.824)
<b>Kapitaldienstfähigkeit (KDF)</b>		
20–40 %	0.21 (0.138)	0.21 (0.138)
40–60 %	0.30 ** (0.132)	0.27 ** (0.132)
60–67 %	0.63 *** (0.133)	0.61 *** (0.133)
<b>Zeitabhängige Faktoren</b>		
Zinssatz der Geschäftsbanken	0.55 (1.130)	0.40 (1.142)
Dummy für Gebühren (1. Feb. 2009)	0.76 (0.704)	0.57 (0.707)
Dummy für Gebotsassistent (6. Mai 2009)	−0.65 ** (0.299)	−0.70 ** (0.300)
Dummy für Sofortkredite (16. Juli 2009)	−0.36 * (0.187)	−0.32 * (0.187)
Fixed-Effect: Nov. 2008	1.19 (1.101)	1.18 (1.113)
Dez. 2008	1.04 (0.893)	0.98 (0.903)
Jan. 2009	0.22 (0.868)	0.11 (0.876)
Feb. 2009	−0.24 (0.288)	−0.20 (0.289)
März 2009	0.72 ** (0.364)	0.81 ** (0.367)
Mai 2009	−0.01 (0.302)	0.04 (0.301)
Juni 2009	−0.46 (0.460)	−0.43 (0.466)
Juli 2009	−1.55 *** (0.475)	−1.45 *** (0.476)
Aug. 2009	−2.18 ** (0.917)	−2.06 ** (0.922)
Sep. 2009	−1.42 *** (0.408)	−1.39 *** (0.408)
Okt. 2009	−1.70 *** (0.395)	−1.66 *** (0.395)

... Tabelle wird auf der nächsten Seite fortgesetzt ...

Tabelle 1: Regressionsanalyse (OLS), abhängige Variable: Zinssatz p. a.



Prädiktor	Modell 1	Modell 2
<b>Merkmale der Projektbeschreibung</b>		
Laufzeit (Dummy für 60 statt 36 Monate)	0.49 *** (0.082)	0.47 *** (0.082)
Betrag (Vielfaches von 250 Euro)	0.01 *** (0.002)	0.01 *** (0.002)
durchschnittliches Gebot (Schrittweise 250 Euro)	-0.14 *** (0.049)	-0.14 *** (0.049)
gewerbliches Kreditprojekt	-0.00 (0.103)	0.06 (0.108)
Revision des angebotenen Zinssatzes	-0.09 (0.090)	-0.06 (0.090)
Individuelles Projektbild vorhanden	-0.18 ** (0.090)	-0.15 * (0.091)
Bild im Benutzerprofil vorhanden	0.11 (0.124)	0.10 (0.132)
<b>Demographische Angaben</b>		
Selbstständige Tätigkeit	0.13 (0.097)	0.18 * (0.097)
Alter (absolute Abweichung vom Median)	0.02 *** (0.004)	0.02 *** (0.005)
Geschlecht (weiblich)	-0.00 (0.078)	-0.02 (0.080)
<b>Angabe weiterer personenbezogener Daten</b>		
Benutzerprofil ausgefüllt		0.12 (0.112)
Name und Kontaktinformationen [0..20]		0.00 (0.038)
Finanzielle Situation [0..13]		0.00 (0.029)
Beruf und Qualifikationen [0..23]		-0.05 *** (0.016)
Wohnsituation [0..4]		-0.02 (0.032)
Gesundheit [0..5]		-0.00 (0.067)
Hobbys [0..4]		0.02 (0.043)
Partner, Familienmitglieder [0..19]		-0.00 (0.018)
<b>Stilmittel und Argumentation</b>		
Schilderungen, die Mitleid hervorrufen		0.42 ** (0.188)
Direkter Aufruf zu helfen		-0.22 * (0.122)
Eigene Hilfsbereitschaft (Reziprozität)		-0.09 (0.177)
Aussicht auf Bankkredit (Konkurrenz)		-0.22 * (0.128)
Schulden bereits angefallen (Dringlichkeit)		0.16 * (0.082)
Modellgüte: korrigiertes $R^2$ (Fälle)	98.72 (1098)	98.74 (1098)

Standardfehler in Klammern; Signifikanzniveaus: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Tabelle 1: (fortgesetzt) Regressionsanalyse, abhängige Variable: Zinssatz

#### 4.1 Wirkung personenbezogener Daten

Die freiwillige Angabe personenbezogener Daten in den Projektbeschreibungen wurde nicht nur dichotom kodiert, sondern wie in Abschnitt 3.1 beschrieben, in verschiedenen Kategorien und Teilkategorien erfasst. Die Anzahl der Teilkategorien ist in Tabelle 1 in eckigen Klammern angegeben. Im Widerspruch zur postulierten Hypothese zeigt sich, dass die Veröffentlichung personenbezogener Daten kaum Einfluss auf die Höhe des Zinssatzes hat. Der durchschnittliche Zinssatz ist etwas niedriger, wenn Kreditnehmer detaillierte An-

gaben zu ihrer Ausbildung, ihrem Beruf sowie sonstigen Qualifikationen und Fähigkeiten machen. Bezüglich der Verfügbarkeit von personenbezogenen Daten aus anderen Kategorien konnte kein Effekt gefunden werden. Damit ist die Hypothese mit den Daten dieser Untersuchung kaum vereinbar.

## 4.2 Argumentationstechnik und Plausibilitätsprüfung

Außer der generellen Ungültigkeit der Hypothese könnte dieser schwache Befund auch durch Schwierigkeiten bei der Messung von „weichen“ Informationen oder durch unzureichende Kontrolle von Drittvariablen entstehen. Um Ersteres zu prüfen, wurden neben personenbezogenen Daten auch Stilmittel und Argumentationstechniken in den Projektbeschreibungen in das Regressionsmodell einbezogen.<sup>1</sup> Hier konnten deutlichere Effekte von „weichen“ Informationen gemessen werden: Kreditnehmer, die direkt an das Hilfeverhalten der Investoren appellieren und solche, die auf die Möglichkeit hinweisen, alternativ einen Bankkredit zu bekommen, erhalten bessere Kreditkonditionen. Die Konditionen verschlechtern sich dagegen, wenn Kreditnehmer mit emotionalen Argumenten übertreiben und offensichtlich Mitleid bei den Kreditnehmern erregen wollen. Geht aus der Projektbeschreibung hingegen hervor, dass die Schulden bereits angefallen sind und der Kredit deshalb unbedingt notwendig ist, resultiert das in einem höheren Zinssatz. Interessant ist, dass Effekte in beide Richtungen sowohl bei eher rationalen (Konkurrenz, Dringlichkeit) als auch bei eher emotionalen Argumentationstechniken (Hilfsbereitschaft, Mitleid) beobachtbar sind. Lediglich das Betonen der eigenen Hilfsbereitschaft der Kreditnehmer und damit verbunden die Hoffnung auf reziprokes Verhalten [Kol00] hat keinen Effekt.

Eine Grundannahme der vorliegenden Untersuchung ist, dass Kreditnehmer grundsätzlich Daten veröffentlichen, die über die absolut notwendigen Angaben hinausgehen. Eine Überprüfung dieser Annahme bestätigt, dass praktisch alle Kreditprojekte eine spezifische Projektbeschreibung haben und dass 93 % dieser Projektbeschreibungen mindestens 50 Zeichen enthalten. 83 % der Projektbeschreibungen enthalten freiwillig angegebene personenbezogene Daten. Zur Kontrolle der Qualität der angegebenen personenbezogenen Daten haben die Kodierer den Gesamteindruck eines Projekts, basierend auf allen verfügbaren Informationen des Kreditnehmers und der Projektbeschreibung, bewertet. Wie die Ergebnisse in Tabelle 2 belegen, vermittelt die Mehrheit aller Projekte mit zusätzlichen personenbezogenen Daten einen positiven Eindruck. Wurden dagegen freiwillig keine personenbezogenen Daten in der Beschreibung genannt, wird der Eindruck im Durchschnitt nur als neutral bewertet. Dieser Unterschied ist statistisch signifikant. Das bedeutet, Kreditnehmer veröffentlichen vor allem solche personenbezogenen Daten, die sie selbst in einem positiven Licht erscheinen lassen und die zu einem positiven Gesamteindruck des Projekts bei Investoren beitragen. Damit kann ausgeschlossen werden, dass Kreditsuchende einfach die „falschen“ personenbezogenen Daten auswählen und deshalb die Veröffentlichung von Details nicht zu besseren Kreditkonditionen führt.

Trotz der nicht zu unterschätzenden Schwierigkeiten bei der Quantifizierung von personenbezogenen Daten aus unstrukturiertem Textmaterial ist es also unwahrscheinlich, dass der Befund allein auf Messfehler zurückzuführen ist. Damit bleibt die Verbesserung der

<sup>1</sup>Ein schrittweiser Einschluss der Prädiktorblöcke ändert an den Ergebnissen nichts.

Gesamteindruck	<i>Beschreibung enthält zusätzliche personenbezogene Daten?</i>	
	nein ( <i>n</i> = 163)	ja ( <i>n</i> = 627)
positiv	19 %	66 %
neutral	52 %	28 %
negativ	28 %	5 %

$\chi^2 = 141, df = 2, p < 0.001$

Tabelle 2: Gesamteindruck der Kreditprojekte (Kontrollvariable kodiert für 790 Fälle)

Drittvariablenkontrolle als Aufgabe für weitere Forschung. Explorative Analysen deuten darauf hin, dass sich in den Daten der dynamisch wachsenden Plattform *Smava.de* noch einige unerklärte Heterogenität befindet, die subtile Wirkungen personenbezogener Daten überdecken könnte.

## 5 Diskussion und Zusammenfassung

Dieser Beitrag dokumentiert einen ersten Versuch, Konflikte zwischen Datenschutzziele und wirtschaftlichen Interessen am Beispiel von alternativen Kreditmärkten systematisch zu untersuchen. Der Nutzen von freiwilliger Veröffentlichung personenbezogener Daten zur Verbesserung der Kreditkonditionen wurde mit empirischen Daten der größten deutschen Plattform für Online Social Lending analysiert. Die Ergebnisse zeigen, dass Kreditnehmer, die Details bezüglich ihrer Ausbildung, Berufs sowie weiterer Qualifikationen angeben, tatsächlich etwas weniger Zinsen bezahlen und so im Mittel einen kleinen finanziellen Vorteil erzielen. Weitere personenbezogene Daten verbessern die Kreditkonditionen dagegen nicht messbar. Einer Offenlegung personenbezogener Daten gegenüber allen potenziellen Investoren, im Fall von Online Social Lending sind das derzeit *alle Internetnutzer*, stehen offenbar nur sehr geringe wirtschaftliche Vorteile gegenüber. Gleichzeitig birgt dieser freigiebige Umgang mit sensiblen Daten jedoch eine Reihe von Datenschutzrisiken (Datenmissbrauch, Stalking, etc.). Kreditnehmer sollten also genau prüfen, welche personenbezogenen Daten sie im Internet einer sehr breiten Öffentlichkeit zugänglich machen. Mittelfristig sind Wissenschaftler sowie Anbieter von Online Social Lending aufgefordert, spezifische datenschutzfreundliche Technologien zu entwickeln und einzusetzen, die die Veröffentlichung von personenbezogenen Daten minimieren, ohne den Marktmechanismus irreparabel zu beschädigen.

Nicht zuletzt verspricht Online Social Lending, Finanzintermediation zu revolutionieren und Zugang zu Kredit zu „demokratisieren“. Es sei dahingestellt, ob diese Vorstellungen realistisch sind. Es wird jedoch deutlich, dass die Reduzierung von Ungleichheiten beim Zugang zu Kredit neue Ungleichheiten bei der Realisierung von Datenschutzziele zur Folge hat: Wirtschaftlich Schwächere werden häufiger als Kreditnehmer auftreten und im Gegensatz zu Investoren mehr Auskunft über sich und ihr Leben erteilen müssen. Privatsphäre bleibt also vermutlich auch durch diesen Mechanismus ein „Luxusgut“ [BK07].

**Danksagung** Wir danken Alessandro Acquisti, Michael Ehrmann und Dagmar Schönfeld für ihre wertvollen Kommentare zu früheren Fassungen dieses Manuskripts sowie Ulrike Jani, Kristina Schäfer, Michael Sauer und Hagen Wahrung für ihre Hilfe bei der Datenerhebung. Der Erstautor wurde durch ein Postdoc-Stipendium des Deutschen Akademischen Austauschdienstes (DAAD) gefördert. Die vorliegende Arbeit wurde zum Teil durch die Europäische Kommission innerhalb des siebten Rahmenprogrammes (FP7/2007-2013, Vertragsnr.: 216483) finanziell unterstützt.

## Literatur

- [Ake70] G. A. Akerlof. The Market for “Lemons”: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 84(3):488–500, 1970.
- [BC95] T. Besley und S. Coate. Group lending, repayment incentives and social collateral. *Journal of Development Economics*, 46:1–18, 1995.
- [BG09] S. Berger und F. Gleisner. Emergence of Financial Intermediaries in Electronic Markets: The Case of Online P2P Lending. *BuR - Business Research*, 2(1):39–65, 2009.
- [BK07] R. Böhme und S. Koble. 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, 2007.
- [BP10] R. Böhme und S. Pötzsch. Privacy in Online Social Lending. In *AAAI Spring Symposium on Intelligent Privacy Management*, Seiten 23–28, Stanford University, 2010.
- [Bun] Deutsche Bundesbank. Zeitreihe SUD114.
- [CGL09] N. Chen, A. Ghosh und N. Lambert. Social Lending. In *ACM Conference on Electronic Commerce*, Seiten 335–344, Stanford University, 2009. ACM Press.
- [FS08] A. Frerichs und M. Schuhmann. *Peer to Peer Banking – State of the Art*. Arbeitsbericht Nr. 02/2008, Institut für Wirtschaftsinformatik, Universität Göttingen, 2008.
- [HADL08] M. Herzenstein, R. Andrews, U. Dholakia und E. Lyandres. The Democratization of Personal Consumer Loans? Determinants of Success in Online Peer-to-Peer Lending Communities, June 2008. <http://ssrn.com/abstract=1147856>.
- [Hol69] O. R. Holsti. *Content analysis for the social sciences and humanities*. Addison-Wesley, 1969.
- [Jen07] N. Jentzsch. *Financial Privacy*. Springer-Verlag, Berlin Heidelberg, 2007.
- [Kol00] S.-C. Kolm. The Logic of Good Social Relations. *Annals of Public and Cooperative Economics*, 71(2):171–189, 2000.
- [KW06] M. Kamp und T. Weichert. *Scoringssysteme zur Beurteilung der Kreditwürdigkeit. Chancen und Risiken für Verbraucher*. Unabhängigen Landeszentrum für Datenschutz Schleswig-Holstein (ULD), Kiel, 2006.
- [KW08] D. Korczak und M. Wilken. *Scoring in Praxistest: Aussagekraft und Anwendung von Scoringverfahren in der Kreditvergabe und Schlussfolgerungen*. GP Forschungsgruppe, München, 2008.
- [PB10] S. Pötzsch und R. Böhme. The Role of Soft Information in Trust Building: Evidence from Online Social Lending. In A. Acquisti, S.W. Smith und A.-R. Sadeghi, Hrsg., *TRUST 2010*, LNCS 6101, Seiten 381–395, Berlin Heidelberg, 2010. Springer-Verlag.
- [Pet04] M. A. Petersen. Information: Hard and Soft. Mimeo, 2004.
- [PS08] D. Pope und J. Sydnor. What’s in a Picture? Evidence of Discrimination from Prosper.com, 2008. <http://wsomfaculty.case.edu/sydnor/prosperpaper.pdf>.
- [Rav07] E. Ravina. Beauty, Personal Characteristics and Trust in Credit Markets, Dezember 2007. <http://ssrn.com/abstract=972801>.
- [Sti81] J. E. Stiglitz. Credit Rationing in Markets with Imperfect Information. *American Economic Review*, 71(3):393–410, 1981.