

Working Papers

RESEARCH DEPARTMENT

WP 18-15

April 2018

<https://doi.org/10.21799/frbp.wp.2018.15>

The Roles of Alternative Data and Machine Learning in Fintech Lending: Evidence from the LendingClub Consumer Platform

Julapa Jagtiani

Federal Reserve Bank of Philadelphia Supervision, Regulation, and Credit

Catharine Lemieux

Federal Reserve Bank of Chicago



ISSN: 1962-5361

Disclaimer: This Philadelphia Fed working paper represents preliminary research that is being circulated for discussion purposes. The views expressed in these papers are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Any errors or omissions are the responsibility of the authors. Philadelphia Fed working papers are free to download at: <https://philadelphiafed.org/research-and-data/publications/working-papers>.

The Roles of Alternative Data and Machine Learning in Fintech Lending: Evidence from the LendingClub Consumer Platform

Julapa Jagtiani*
Federal Reserve Bank of Philadelphia

Catharine Lemieux
Federal Reserve Bank of Chicago

April 2018
Supersedes Working Paper 17-17

Abstract

Fintech has been playing an increasing role in shaping financial and banking landscapes. There have been concerns about the use of alternative data sources by fintech lenders and the impact on financial inclusion. We compare loans made by a large fintech lender and similar loans that were originated through traditional banking channels. Specifically, we use account-level data from LendingClub and Y-14M data reported by bank holding companies with total assets of \$50 billion or more. We find a high correlation with interest rate spreads, LendingClub rating grades, and loan performance. Interestingly, the correlations between the rating grades and FICO scores have declined from about 80 percent (for loans that were originated in 2007) to only about 35 percent for recent vintages (originated in 2014–2015), indicating that nontraditional alternative data have been increasingly used by fintech lenders. Furthermore, we find that the rating grades (assigned based on alternative data) perform well in predicting loan performance over the two years after origination. The use of alternative data has allowed some borrowers who would have been classified as subprime by traditional criteria to be slotted into “better” loan grades, which allowed them to get lower-priced credit. In addition, for the same risk of default, consumers pay smaller spreads on loans from LendingClub than from credit card borrowing.

Keywords: fintech; LendingClub; marketplace lending; alternative data; shadow banking; P2P lending; peer-to-peer lending

JEL classification: G21, G28, G18, L21

* julapa.jagtiani@phil.frb.org or 215-574-7284. The authors thank Ana Oppenheimer, Leigh-Ann Wilkins, Raman Quinn Maingi, and John Nguyen for their dedicated research assistance. They also thank Onesime Epouhe for his assistance with the stress test data. Helpful comments and suggestions from Tracy Basinger, Robin Prager, Joe Hughes, Bob Hunt, Robert Wardrop, Raghu Rau, Paul Calem, Chris Cumming, Kathleen Hanley, and participants at the FDIC conference and the AEA conference are appreciated.

Note: Working Paper 17-17 is also superseded by Working Paper 18-13.

Disclaimer: This working paper represents preliminary research that is being circulated for discussion purposes. The opinions expressed in this paper are the authors’ own views and do not necessarily represent the views of the Federal Reserve Bank of Philadelphia, Federal Reserve Bank of Chicago, or the Federal Reserve System. Any errors or omissions are the responsibility of the authors. Philadelphia Fed working papers are free to download at <https://philadelphiafed.org/research-and-data/publications/working-papers>.

I. Introduction

Consumer credit not secured by real estate has increased steadily over the last five years and stood at over \$3.8 trillion as of November 2017. Of that amount, 26.6 percent was credit card debt, and 68 percent was student loan and auto-related debt; the remaining 5.4 percent was unsecured personal loans (Federal Reserve, 2017). Despite increases in consumer credit over the last several years, Bricker et al. (2017) found that, based on the 2016 Survey of Consumer Finance, 20.8 percent of families felt credit constrained, and this result has been fairly consistent over recent years. One explanation could be information asymmetries between borrowers and lenders. Oliver Wyman (Carroll and Rehmani, 2017) estimates that 45 million to 60 million people do not have sufficient credit information in their credit file to have a credit score, essential information for any consumer lending decision. Fintech lending platforms have entered the unsecured personal loan space and have the potential to fill this unmet demand for credit.

Over the past decade, online alternative lenders have evolved from platforms connecting individual borrowers with individual lenders¹ to sophisticated networks featuring institutional investors, direct lending (on their balance sheet), and securitization transactions. While the alternative data sources and algorithms used by online alternative lenders have allowed for faster and lower-cost credit assessments, these innovations could potentially carry a risk of disparate treatment and fair lending violations. We explore some of the potential consumer benefits that could come from these new algorithms.

Regulators and policymakers have raised several questions around these issues. Can the use of alternative data (e.g., to build internal credit rating systems such as the one designed by LendingClub) increase access to credit for consumers by allowing lenders to better assess their creditworthiness? Do these data allow fintech firms to better risk-price credit so that some borrowers can get loans from fintech firms at a lower cost than from traditional banks? The use of alternative data sources, big data and machine learning technology, and other new artificial intelligence models could reduce the cost of making credit decisions and/or credit monitoring and lower operating costs for lenders. Fintech lenders could pass on the benefits of lower lending costs to their borrowers. We demonstrate in this paper that, over the years, alternative sources of information used by fintech firms to evaluate credit applications have increasingly contained additional information not embedded in the traditional credit approval criteria.

Several alternative data sources have been used by fintech lenders. While it is not known exactly what alternative data is being used by a specific fintech firm, some examples that have been

¹ Frequently referred to in prior research as peer-to-peer (P2P).

mentioned include information drawn from utility payments, electronic records of deposit and withdrawal transactions, insurance claims, bank account transfers, use of mobile phones or the Internet, and other personal data such as consumer's occupation or details about their education. Crosman reports in *American Banker* (June 14, 2016) that SoFi no longer uses FICO scores when determining loan qualifications. In addition, Kabbage claims that FICO scores are not part of its creditworthiness determination (although FICO scores are used for benchmarking and investor reporting). A quote in this *American Banker* article by Ron Suber, president of Prosper Marketplace, states that "Prosper gets 500 pieces of data on each borrower; the FICO score is just one data point." The company uses FICO scores to screen borrower candidates; a score of at least 640 is needed to be considered for a loan. Prosper analyzes additional data to determine its ultimate credit decision. These data sources were not normally used by traditional lenders. There have been policy questions around the use of big data and the appropriate policies that would protect consumers without harming the innovation process.

The Consumer Financial Protection Bureau (2017) released a request for information to explore the impact of alternative data sources, including data from mobile phones; rent payment histories; and electronic transactions such as deposits, withdrawals, and transfers on building credit histories, and increasing credit access. Concerns about the potential risks posed by these data sources have arisen because they may be biased and could potentially have an adverse impact on credit access to low-income and underserved communities.² In March 2017, Richard Cordray, director of the Consumer Financial Protection Bureau, pointed out some potential benefits to consumers through the use of these alternative data sources (Consumer Financial Protection Bureau, 2017):

By filling in more details of people's financial lives, this information may paint a fuller and more accurate picture of their creditworthiness. So adding alternative data into the mix may make it possible to open up more affordable credit for millions of additional consumers. ...

Online fintech lenders often rely on their own algorithms for credit underwriting. Our work shows that some of the information used in their algorithms may include nontraditional information (not used by traditional banks in their lending decisions). Some fintech lenders have developed their own online lending platforms that use big data in their own proprietary algorithms that they developed to evaluate borrowers' credit risk. Through this new approach to credit risk evaluation, some consumers could potentially enhance their credit access. For example, consumers

² There may also be a risk that online fintech lenders could use these new data sources and data mining techniques to identify consumers who are less sophisticated and vulnerable to exploitation.

with a short credit history may not satisfy a bank's traditional lending requirements, but these same consumers could potentially get a loan from an online alternative lender that uses alternative data sources.³ Concerns emerged that consumer privacy may be compromised in the process if information such as insurance claims, utility bills, bank account transactions, and social network details are used by lenders without the borrower's consent.

In this paper, we shed more light on the role of alternative information sources and their relationship with traditional credit scores. Many believe that the role of big data and alternative information will increase exponentially in the future. Issues around consumer privacy and the disparate treatment of protected classes still need to be explored. The rest of the paper is organized as follows. In Section II, we present the literature review. Section III describes our data from various sources. Section IV explores the pricing (credit spreads) of loans originated by a fintech platform versus traditional origination. Section V investigates the relationship between pricing and loan performance, focusing on the roles of alternative data sources. Section VI presents a regression analysis. Section VII concludes and discusses policy implications.

II. The Literature

Information asymmetries have been an important issue in the banking literature. Jaffee and Russell (1976) and Stiglitz and Weiss (1981) explained how information asymmetries between borrowers and lenders can lead to a market equilibrium in which credit is rationed. Frame, Srinivasan, and Woosley (2001) and Einav, Jenkins, and Levin (2013) find that older technologies such as credit scoring reduce information asymmetries between borrowers and lenders and expanded credit availability in the small business and auto loan markets, respectively. Morse (2015) reviewed the existing literature developing around fintech lending with a focus on whether the type of technologies employed by fintech firms can mitigate information frictions in lending. She posits that better capturing soft information contained in proximity information and better profiling of loan applicants could improve the access to or price of credit.

Many papers have found that relationships and soft information can provide advantages in borrower screening and can reduce information asymmetries in banking — see Petersen and Rajan (1994); Boot and Thakor (2000); Berger and Udell (2002); Petersen (2004); Berger, Miller, Petersen, Rajan and Stein (2005); Stein (2002), Karlan(2007); Iyer and Puri (2012); and Schoar (2012). Researchers are beginning to look at this issue for fintech lending. Freedman and Jin (2017) demonstrate the value of friends of the applicant committing to invest in the loan and show that

³ See more discussion in Demyanyk and Kolliner (2014).

this signal is more pronounced in lower credit grades. Everett (2010) finds that loans funded by investor groups perform better if someone in the group is personally connect to borrowers. Lin, Prabhala, and Viswanathan (2013) find that the credit quality of one's friends is related to improved success in fundraising, lower interest rates, and a lower default rate. Lu, Gu, Ye, and Sheng (2012) find that the reverse is also true; there is a positive relationship between a friend's default and a borrower's default. However, inferring credit risk from one's social network does present issues related to the consumer regulations around fair and equal access to credit that need to be addressed in the use of such data.

Another issue with the use of this type of information is that, if funding is limited to connections with friends, there is a limited ability to improve credit conditions in the aggregate. Researchers have investigated identifying other soft information that could be leveraged in an online loan application. Michels (2012) finds that voluntary disclosure of hard information such as income, income source, education, and other debt yields lower interest rates. Herzenstein, Sonenshein, and Dholakia (2011), through text analysis of borrower narratives, find limited usefulness. Gao and Lin (2012) use text mining and find that more complex narratives correlate with higher default rates. Ravina (2012); Pope and Sydnor (2011); and Duarte, Siegel, and Young (2012) analyzed photo-based discrimination. The results are mixed; some findings of bias lean toward attractive or trustworthy faces and against racial minorities. A central issue to the value of this line of research is that, once borrowers understand lenders are using this information, they can choose to alter the way they submit text or photo information.

Another way to leverage proximity is to use local economic information as a proxy for personal knowledge. Crowe and Ramcharan (2013); Bertsch, Hull, and Zhang (2016); Buchak, Matvos, Piskorski, and Seru (2017); Havrylchyk, Mariotto, Rahim, and Verdier (2018); Chen, Hanson, and Stein (2017), and Jagtiani and Lemieux (2018) are a few of the studies that have found local economic information as a possible relevant source of nontraditional information by fintech lenders.

Any reduction in information asymmetries will benefit lenders. It is important to investigate whether fintech lenders pass on the savings to consumers with lower credit costs and whether the pricing is appropriate for the risk taken.⁴ A few studies have attempted to compare lending rates

⁴ Morse (2015) explores a number of issues related to fintech disruption and financial disintermediation. The paper concludes that at least some cost savings seem to accrue to investors (since 80 percent of P2P funds come from institutional investors) and that the borrowers' social circles and local economic indicators are useful in predicting credit risk.

from online alternative platforms with traditional sources, but those studies have been subject to significant data limitation, and the results have been mixed.

Mach, Carter, and Slattery (2014) report that P2P small business borrowers paid higher rates for fintech loans compared with loans obtained from traditional sources. However, they used data from LendingClub's consumer platform that were identified as small business purposes and were less likely to be comparable with small business loans made by traditional banks. Demyanyk and Kolliner (2014) find that more creditworthy consumers receive preferred rates using a P2P lender over borrowing with a credit card. However, they used aggregate market rates as the comparison.

In Germany, De Roure, Pelizzon, and Tasca (2016), using data from Auxmoney, a German P2P lending site, find that interest rates are comparable with loans made by P2P alternative lenders and those made by traditional banks; but again, the interest rates used as a comparison were market rates. Buchak, Matvos, Piskorski, and Seru (2017), using mortgage data, find evidence that fintech customers are among the borrowers who value fast and convenient services and that fintech lenders command an interest rate premium for their services. Another interesting study that looked at risk pricing by LendingClub did not compare rates charged by the firm to those charged by traditional lenders; they find that the rates charged by higher-risk borrowers were not large enough to compensate for a higher probability of default — see Emekter, Tu, Jirasakuldech, and Lu (2014). Using loan-level data from both LendingClub and traditional banks, this paper is able to overcome many of the data limitations of these studies and compare how credit is priced by fintech lenders and traditional banks.

We use a unique data set that allows us to compare online alternative lending rates with traditional credit card loans. We compare account-level credit card data that large banks submitted to the Federal Reserve for stress testing with online consumer loans that were made for credit card (and debt consolidation) purposes. These data will allow us to investigate the determinants for risk pricing used by LendingClub, and the performance of these loans over time and to compare these loans with similar loans made by traditional banks.

III. The Data

We use four main sources of data in this paper: data on loans that were originated through online alternative channels (loan-level data from the LendingClub consumer platform); data on loans that were originated from traditional banking channels (loan-level data from the Y-14M reports submitted by bank holding companies with over \$50 billion in total assets); deposit market

concentration data and bank branch information, based on the FDIC Summary of Deposits database; and economic factors from the U.S. Census Bureau and Haver Analytics database.

Online Alternative Lending Channel

Our research on fintech consumer lending focuses on the LendingClub for two reasons. First, the company is one of the few lenders that has made its data publicly available. Second, LendingClub is one of the larger, more established alternative lenders in this space, and therefore the results here are likely to apply more broadly. We use loan-level data (with detailed information about the loan and the borrower) that were originated in 2007–2017 from LendingClub’s consumer platform.⁵ The loan-level database contains loan-specific information (i.e., loan rate, maturity, origination date), risk characteristics of the borrowers (i.e., FICO scores, employment, debt-to-income (DTI) ratio, homeownership), other risk characteristics, and monthly payment and performance of the loans.

Our analysis is based on data from the LendingClub consumer loan platform. We focus on loans that were specified for two purposes: credit cards and debt consolidation. As of 2015, about 90 percent of LendingClub consumer loans portfolio are in these categories, either for paying off credit card balances or for debt consolidation purposes, as shown in Jagtiani and Lemieux (2018).

To evaluate the differences in credit access and pricing between traditional versus alternative lending channels, we compare these loans (for credit cards and debt consolidation) with account-level credit card loan data from banks. We observe the differences between these two lending channels in terms of credit risk rating, price of credit, and loan performance.

Traditional Lending Channels

To explore comparable loans made by traditional banks, we use loan-level (account-level) credit card loan data from the Federal Reserve’s Y-14M reports, which are reported monthly bank holding companies with least \$50 billion in assets. From this data set, we focus on the reporting period 2014–2017 and include only those accounts that were originated in 2014–2015 (allowing for up to a two-year performance period until 2017).⁶ We do not include accounts that were originated prior to 2014 to avoid sample selection bias in our analysis. Accounts that were originated earlier and

⁵ For part of the analysis, we use data from 2010 to 2015 origination vintages, with two years’ performance window up to 2017. Data from 2007 to 2009 origination vintages are less reliable, and the volume was small during the initial period.

⁶ We note that these data are constrained by the limited number of reporters and thus may not represent the entire population of firms that issue credit cards. However, Y-14M reporters do represent over 80 percent of all credit cards issued by commercial banks.

were closed (owing to default or other reasons) would have been dropped from the Y-14M reports in 2014–2017.

We do not include charge cards in the analysis because there is no associated credit limit for these cards. In addition, for credit cards, we only include consumer cards that were issued for general purposes and private-label cards (business cards and corporate cards are not included). Only credit card accounts that are revolvers (in which consumers are actually taking a loan from the issuing bank) are included in our analysis. Since consumers report that they borrow from LendingClub to pay off their credit cards, we compare the average price and performance of LendingClub loans with consumer card loans made by traditional banks reported on the Y-14M, using card credit loan balance and LendingClub origination amounts as control factors (along with other relevant risk factors).

It is important to note that reported credit card balances are balances as of a specific reporting date, rather than balances as of the end of a statement (which varies across card accounts). The reported card balances mostly reflect spending rather than extensions of credit. To correctly compare fintech platform loans versus traditional credit card loans, we identify whether each card account is a revolver or a transactor. We observe customers' payments and fees monthly. Cardholders are considered revolvers for the month if they did not pay off the entire balance as of the end of the statement period and were subject to finance charges at some point in the last 12 months. The revolver flag for the month is removed when the entire balance is paid off and the customer was no longer subject to finance charges. Most cardholders are transactors, and they do not actually borrow from the bank; thus, we only include revolvers in our analysis of risk pricing and loan performance.

For the most part, the data from Y-14M reports contain similar information on the borrowers and other risk characteristics as those reported in the LendingClub database (i.e., origination date, origination amount, location of the borrowers, borrowers' credit scores). A few key variables are reported for LendingClub loans that are not reported by banks in Y-14M database, such as homeownership and DTI ratio at origination. It is important to note that the credit card loans from Y-14M reports and LendingClub consumer loans that are used to pay off credit card loans (or for debt consolidation) are the most comparable products. However, some credit cards have rewards (cash back or points) and/or some period of low-rate promotion period (e.g., in the first six months) to encourage balance transfers from other cards. We control for the promotion period and the rewards in our analysis.

We calculate the level of market concentration for consumer loans at the 3-digit zip code level based on account-level Y-14M data. The share of outstanding credit card loans (revolvers' balance) by each banking firm in each zip code is used to calculate the Herfindahl-Hirschman Index (HHI), a commonly accepted measure of market concentration. The calculated HHI approximates the degree of market concentration (or degree of competition) in the credit card loan market.⁷ We also estimate a similar market concentration based on deposit data, but the credit card loan market HHI is more appropriate for this study.⁸

Economic Factors

We collect various economic factors from the U.S. Census Bureau database and the Haver Analytics database. For example, we use data on economic factors including local unemployment, local average household income, local home price index, and local population. We use the most appropriate and most granular level (3-digit zip code, 5-digit zip code, or county) of economic factors in the analysis.⁹

IV. The Roles of Alternative (Nontraditional) Data Used by Fintech Lenders

One of the attractive features of getting credit from alternative lenders is how quickly lending decisions are made. An important advantage for fintech lenders is that they have access to nontraditional data sources that are not used (or not available) to traditional bank lenders. The additional sources of information are consumers' payment history (utility, phone, PayPal, Amazon), their medical and insurance claims, their social network, and so forth. These are not factors that are reflected fully in the traditional credit scores.

In the case of LendingClub, consumers are assigned a rating grade from A to G based on the full set of information (after the loan has been approved). The loan application process is as follows: (1) the application is submitted online, (2) LendingClub's credit model immediately grades and prices the loans at application, and (3) the applicant receives immediate feedback about the loan

⁷ The U.S. Department of Justice defines a concentrated market as one that has an HHI above 2,500. An HHI of less than 1,500 indicates an unconcentrated (or competitive) banking market, an HHI between 1,500 and 2,500 indicates moderate concentration, and an HHI above 2,500 indicates a highly concentrated banking market.

⁸ We do not report results based on deposit HHI in this paper because we find little or no relationship there. We also note that some banks have been booking their deposits at certain branches even though the deposits may actually be coming from many different locations.

⁹ Note that LendingClub loan-level data are reported at the three-digit zip code level; thus, three-digit zip code level of economic factors is used in these cases.

terms for which they are qualified. Additionally, the verification process takes place before funding. For example, if the credit model data sources indicate the application is fraudulent, the application may be declined. If not, after an offer is presented, further income or employment verification may be requested. LendingClub has its own proprietary models that identify whether each of the loan applications should be verified. As of 2015, about 70 percent of all loans made through LendingClub platform were verified.

We explore the correlation between LendingClub rating grades and FICO scores as of loan origination. We convert LendingClub's rating grades to numerical values, where A is 7, B is 6,... and G is 1. It is interesting to note that while the rating grades and FICO scores were highly correlated with about an 80 percent correlation as of origination date for loans originated in 2007, the correlation has weakened over the years. The plot in Figure 1 shows the correlation between FICO scores and loan grades at the time of loan origination for loans that were originated in 2007 to 2015. While we do not know how LendingClub defines their credit grades, it is obvious that these credit grades are increasingly defined using additional metrics beyond FICO scores. The correlation has declined from over 80 percent for loans that were originated in 2007 to only approximately 35 percent for loans that were originated in 2014–2015.¹⁰ This evidence indicates that LendingClub is relying more and more on additional information over the years.

The rating grades are assigned based on LendingClub's credit model, which looks beyond FICO scores to estimate the likelihood of default. The model attempts to identify applicants with FICO scores that do not reflect their true credit quality, and thus the risk could have been mispriced based on FICO scores alone.¹¹ What are the implications for consumers? Some consumers with low FICO scores (below 680) could end up being rated A by the LendingClub's credit model, especially in later years (2014–2015 origination years). Figures 2A, 2B, and 2C present the composition of loans for each rating grade and how the composition has evolved over the years for loans originated in 2007, 2011, and 2015, respectively. Some consumers who would be considered subprime are slotted into the "better" loan grades. For loans originated in 2015 (see Figure 2C), over 25 percent of the B-rated borrowers have FICO scores in the subprime range. About 8 percent

¹⁰ We also tried calculating the correlation when both the rating grades and the FICO scores are grouped into segments. The FICO score is 1 if the FICO score is lower than 680; the FICO score is 2, 3, and 4 if it is between 680 and 700, 700 and 750, and above 750, respectively. We find the same correlation between rating grades and the FICO scores that fell from 81 percent for loans that were originated in 2007 to 36 percent for loans that were originated in 2015.

¹¹ LendingClub has documented that its credit models have the Kolmogorov–Smirnov scores that outperform generic scores by identifying strong borrowers with lower FICO scores and vice versa. See the link from the LendingClub site for more details at <https://www.lendingclub.com/public/income-verification.action>.

of loans that were assigned an A rating had FICO scores below 680. This provides evidence that the use of additional information sources could allow some borrowers with low FICO scores to get access to credit and potentially get a lower price than if FICO scores were the only criteria.

We further explore those LendingClub borrowers with FICO scores below 680 (so-called subprime) who were slotted into different rating grades (from A to G) in Figure 2C. We observe their credit performance during the 12 months and 24 months after loan origination date. Figure 3A shows the probability of loans becoming delinquent (at least 60 days past due (DPD)) within 12 months after origination for these subprime borrowers. Interestingly, their default probabilities vary significantly, even though they were all rated below 680 based on FICO scores. LendingClub's use of alternative data seems to enhance its ability to identify those subprime borrowers who are actually not risky. Those borrowers who were rated A and B by LendingClub have the probability of default (PD) below 0.03 compared with the average PD of about 0.19 for those who were rated G by LendingClub.

Beyond subprime borrowers, we see that Figure 3B presents the average 12-month PD for borrowers with the various FICO brackets that were assigned ratings ranging from A to G by LendingClub in 2014–2015. The borrowers' performance reflects the rating grade assigned by LendingClub, in which A-rated borrowers have a very small average PD, and F- and G-rated borrowers have a high average PD of about 20 percent or more, regardless of their FICO scores.

To ensure robust testing, we perform the same analysis, looking beyond 12 months (up to 24 months) after origination, and observe the average PD of borrowers who were assigned ratings A to G by LendingClub, compared with their FICO scores. Figure 4A shows that subprime borrowers who were assigned an A rating or B rating by LendingClub in 2014–2015 have a very small average PD over 24 months following loan origination date. It appears that alternative data have allowed these “subprime” borrowers who are not risky to be separated out and to receive a loan at a better price. Similarly, Figure 4B illustrates all FICO segments and confirms that average PDs over the 24-month window after origination are consistent with LendingClub rating grades regardless of FICO scores. Superprime borrowers with FICO scores above 750 who were slotted into the F- and G-rated segments by LendingClub perform poorly with an average PD of about 40 percent.¹²

Figures 5A and 5B demonstrate that average PDs over the periods 12 and 24 months after origination, respectively, are consistent across all loans with the same rating grades A or B, regardless of their FICO scores. The average PD for all borrowers with FICO scores below 680, and FICO scores

¹² A small number of superprime borrowers (with FICO scores above 750) were identified by LendingClub as being high risk; they were rated F (64 observations) and G (eight observations).

680–699 are significantly higher (much more likely to default) than those exceptions identified by LendingClub (with the same low FICO scores, but rated A or B by LendingClub) in 2014–2015.

V. Fintech Lending and Pricing of Credit

In this section, we explore the pricing of LendingClub loans versus similar loans from traditional lenders. Pricing is measured in terms of the credit spread between the reported interest rate and the matching Treasury rates for the same time to maturity. LendingClub’s own rating grades (from A to G), based on the internal proprietary rating system (which is used to price loans) seem to demonstrate the risk-price rank ordering consistently throughout the sample period, in which better-rated borrowers receive lower prices (smaller credit spreads), as shown in Figure 6. LendingClub uses loan grades to differentiate interest rates offered to borrowers. We observe a tight relationship between the loan grades and the interest rate spreads on the loans in the regression analysis, even after controlling for other relevant risk and economic factors.

We observe in Figure 6 that while the rating grade and spreads are consistently in rank order over the years, the spread differential between the A- and G-rated borrowers widened significantly to approximately 20 percent for loans originated in 2015, when more alternative data were being used in credit decisions (compared with earlier vintages). In 2015, the subprime borrower who was slotted into a B-rated loan grade (from Figure 2C) would have had to pay approximately 25 percent over Treasuries instead of 9 percent over Treasuries (a meaningful difference), if he had been slotted into the G-rated loan grade.¹³ (For earlier years, the difference still exists but would be smaller.) The use of additional information allows some borrowers who would be classified as subprime by traditional criteria to be slotted into “better” loan grades and therefore obtain lower-priced credit. It does not appear that this credit is “mispriced” in terms of default risk.

To summarize, we have so far observed a tight relationship between the LendingClub’s own credit spreads and the proprietary rating grades assigned by LendingClub. We have also observed that the relationship between the rating grades and FICO scores has declined dramatically over the years, from about 80 percent initially to only about 35 percent for loans that were originated in 2014–2015, indicating an increasing role of alternative nontraditional information sources used by LendingClub. As shown in Figure 2C, for loans that were originated in 2015, some of the A-rated borrowers actually had FICO scores below 680 and were able to access credit at a lower rate. As

¹³ In the next section, we see that the higher probability of default is observed for loans that were appropriately subject to larger credit spreads (higher price). The interest rate spreads appear to have a strong relationship with the likelihood of becoming delinquent.

shown earlier (Figures 3A, 3B, 4A, and 4B) rating grades assigned by LendingClub have been effective in predicting delinquency (within 24 months after origination).¹⁴

Table 1 shows the comparison of interest rate spreads that borrowers are charged on LendingClub loans (to pay off credit card balances) versus spreads that borrowers are charged on traditional credit card loans for the same FICO scores. We find that for loans originated in 2014–2015, credit spreads on credit card loans are significantly higher than LendingClub loans (for either maturity in three or five years). The spread differentials (the savings to consumers) range from about 8 percent for those with FICO scores lower than 680 to more than 10 percent for those superprime borrowers with FICO scores of 800 or above.¹⁵

Figure 7 demonstrates average PD within 12 months after origination for loans that were originated in 2007–2015 by rating grades. In the initial period, when the rating grades were highly correlated with FICO scores, the average PD across rating grades were not in rank order as strictly as they were in later years when LendingClub incorporated more of its own alternative data. The consistently declining average PDs over the years are also illustrated in Figure 7 across all rating grades, probably reflecting the improving economic environment overall.

To further understand the roles of alternative data, we focus on loans that were originated in 2014–2015, with up to a two-year observation period for credit performance up to 2017. Recall that these origination dates are during the period when the rating grades (assigned by the LendingClub) and FICO scores are less correlated; that is, the rating grades contain different information than that which is contained in the FICO scores. Figures 8A and 8B compare delinquency rates across the credit spread brackets for LendingClub consumer loans (loans for credit cards and debt consolidation purposes) versus traditional credit cards loans (issued by large U.S. banks). Only loans originated between January 2014 and December 2015 are included in the analysis for both LendingClub and Y-14M data.¹⁶ In Figure 8A, we measure delinquency during the initial 12 months after loan origination. Figure 8B expands the performance window from 12 months to 24 months after origination. For credit card delinquency (from Y-14M data), we include

¹⁴ LendingClub consumer loans only come in two different maturities: either three or five years.

¹⁵ LendingClub interest rates (as reported on the LendingClub website) do not include origination fees, which range from 1 percent to 5 percent of the origination amount, depending on the rating grades of the borrowers. The origination fee is usually deducted from the total loan amount. The interest rate from Y-14M data is an annual percentage rate.

¹⁶ We do not include credit card accounts from the Y-14M database that were originated prior to 2014 — to avoid the sample survival bias — because cards that defaulted and were closed before 2014 would not be included in the Y-14M reports (as of 2014).

only cards that carry a balance (revolvers). Cards that involved the initial promotion low interest rates are excluded from this analysis.

Delinquency rates and credit spreads line up very well for LendingClub loans, in which higher credit spreads correspond to higher delinquency rates — for both measures of delinquency in Figures 8A and 8B. The plots show that the average delinquency rates are higher for LendingClub loans than for average PD for credit card loans, controlling for the same credit spreads. These results indicate that given the same credit risk (i.e., for borrowers with the same expected delinquency rate), consumers would be able to obtain credit at a lower rate through LendingClub than through traditional credit card loans offered by banks.

We find that, given the same rating grades, homeownership may also play a role in determining default risk. Figure 9A (with a 12-month performance window) and Figure 9B (with a 24-month performance window) show that, among all LendingClub borrowers, homeowners are less likely to become delinquent than nonhomeowners on average, holding the rating grade constant.

Figures 10A and 10B show that for loans that were originated in the same period (2014–2015) and in the same FICO score brackets, the delinquency rate is slightly higher for LendingClub loans than for Y-14M credit card loans.¹⁷ In Figure 10A, we measure delinquency during the initial 12 months after loan origination. To ensure robust results, Figure 10B expands the performance window from 12 months to 24 months after origination. For credit card delinquency (revolvers from Y-14M data), we only include accounts with a balance. All cards that involved the initial promotion low interest rates are excluded from this analysis. These results imply that for consumers with the same FICO scores, those who borrow from LendingClub have a higher risk of becoming delinquent on average. In other words, borrowers in the same FICO score brackets at LendingClub tend to be more risky on average than those who stick with credit card loans through traditional lending channels.

VI. Regression Analysis

Our analysis so far indicates that LendingClub's rating grades A to G, which are assigned to each loan (that were originated in more recent years starting in 2014–2015) based on information that is not highly correlated with the borrowers' FICO scores, seem to do a good job of identifying riskier borrowers. The rating grades are highly related to the borrowers' probability of becoming delinquent on their loans within two years of loan origination.

¹⁷ Note that a small number of the credit card loans reported on Y-14M have missing FICO scores at origination and are noted in the missing FICO category in Figures 10A and 10B.

Based on our logistic regression analysis (coefficients are not reported here) of default probability (being at least 60 DPD within two years after origination), we present the Receiver Operating Characteristic (ROC) curves in Figure 11. We plot the ROC curves for four different default probability model specifications based on the following sets of explanatory variables: 1) FICO scores only, 2) rating grades only, 3) FICO scores and other control factors, and 4) rating grades and the same set of other control factors. The results are consistent with earlier findings that the rating grades assigned by LendingClub are more powerful in predicting the borrower's default probability than a set of FICO scores, other traditional risk variables, and economic factors combined.

Furthermore, the regression results presented in Table 2 demonstrate that the rating grades are also highly correlated with interest rates that the borrowers are charged. Those subprime borrowers who are identified by LendingClub as being less risky would not only be assigned a better rating (such as A or B), but they would also be given access to credit from LendingClub at a much lower cost than what they would otherwise have had to pay.

The dependent variable is the interest rate spread (on LendingClub consumer loans that are specified for credit cards and debt consolidation purposes), which is calculated as the difference between the interest rate charged on the loans and the equivalent risk-free loans (Treasury rate of securities with the same time to maturity). The key independent variables are the various rating grades in column 1 and FICO score segments in column 3. The results indicate that there is a strong relationship between rating grades and credit spreads, with adjusted R-square of almost 90 percent, as shown in column 1. The coefficients for rating grades are all statistically significantly positive and in rank order, in which the coefficients are positive for B-rated and positive, and largest for G-rated loans. Unlike the rating grades assigned by LendingClub, the relationship between credit spreads and FICO scores at origination is not as tight, with an adjusted R-square of only about 18 percent. The coefficients for FICO scores are, as expected, positive, statistically significant, and in rank order. These results confirm that FICO scores have been used by fintech alternative lenders as an initial broad measure of credit risk, but FICO alone is not granular enough to sufficiently predict each consumer's default probability.

We include additional control factors that are intended to capture risk characteristics of the borrowers and the local economic environments in columns 2 and 4, such as DTI ratio at origination, whether the consumer owns a home, consumer's length of employment, income at origination, loan amount, and the number of consumer's credit inquiries during the period prior to loan origination. Economic factors included in the analysis are local unemployment rate, local home price index, year dummies, and the HHI index measure of credit market concentration in the

borrower's zip code. Most importantly, in columns 2 and 4, we also include a dummy indicating whether the loan actually defaulted (being at least 60 DPD within 24 months after loan origination) and another dummy indicating whether the loan was (1) originated in 2014 or 2015 and (2) defaulted within 24 months after origination date.

The coefficient of the default indicator, $D(\text{Default within 24-Mo After Origination})$, is positive and significant across all columns, indicating the positive relation between credit spreads and default probability. However, the coefficient is much larger in column 4 than in column 2 (1.3499 in column 4 and 0.2197 in column 2), implying that the default dummy picks up some of the risk factors specific to the loan and the borrower that are not captured by the FICO scores in column 4, even after controlling for a set of factors. Also, the adjusted R-square is much smaller in column 4 than in column 2 (34 percent versus 93 percent).

The second dummy indicator that identifies loans made in 2014–2015 that defaulted, $D(2014-15)*D(\text{Default within 24-Mo After Origination})$, is significantly negative in column 2 but significantly positive in column 4. In column 2, when the rating grades and other control factors are included in the analysis, a combination of these two coefficients adds to a very small number ($0.2197 - 0.0979$) compared with the equivalent number in column 4 ($1.3499 + 0.3387$) in which FICO scores are included in the analysis (instead of the rating grades). For loans that were originated in later years (2014–2015), when more alternative data were used in assigning rating grade and credit pricing, the risk factors in column 2 capture much of the risk for specific loans and borrowers. In contrast, when FICO scores are included in the analysis in column 4, more of the specific risk was not fully captured by FICO, resulting in significantly larger positive coefficients of the default dummies. Again, these results indicate that the relationship is much tighter with rating grades (A to G) than with the FICO scores, even after controlling for all the other risk characteristics of the borrowers and economic conditions.

Our regression results so far confirm that the pricing of credit risk seems to be more accurate with the use of nontraditional alternative data. A combination of the accurate risk pricing and the effectiveness of rating grade in predicting defaults suggest that alternative data could benefit consumers by providing increased access to credit at lower cost to those creditworthy individuals who have thin credit history or have poor FICO scores. We caution, however, that fintech lenders should be cautious about which alternative data to use and to keep in mind that

some set of alternative data that may work well for some groups of consumers may not be representative and stable to be used for others, depending on how the data were collected.¹⁸

Our control variables are mostly significant with the expected signs across all columns in Table 2. For example, we observe a significantly positive relationship between interest rate spreads that LendingClub charges and loan amount and the number of credit inquiries by the borrowers within six months prior to loan origination (measuring how desperately the borrowers need additional credit). In addition, we observe that LendingClub charges smaller credit spreads to borrowers who own a home, have been employed for more than 10 years, and have higher income. The market concentration variable, *D(Y-14M Card Loans HHI>2500)*, is either negative or insignificant, implying that LendingClub is likely to offer loans at the lower rate to consumers who live in the zip codes that have high consumer loan market concentration (areas that would benefit from more lenders including fintech alternative lenders).

VII. Conclusions

Fintech has been playing an increasing role in shaping the financial and banking landscapes. Technology has allowed both banks and fintech lenders to serve small businesses and consumers without brick-and-mortar investments. The FDIC and the Consumer Financial Protection Bureau have expressed concerns about impacts on consumer credit access and privacy around credit provided by fintech lenders.

In this paper, we explored the impact of fintech lending on consumers' ability to access credit and the price of credit. In addition, we explored the role of alternative information sources potentially used by these nonbank alternative lenders. Since our results are derived based on loans originated on the LendingClub platform (the largest personal unsecured installment lenders), one should be cautious in extrapolating the interpretation of our findings to all loans originated through other online alternative platforms.¹⁹ We would note that the Y-14M data are constrained by the limited number of reporters and do not include credit card lending by bank holding companies under \$50 billion in total assets or credit unions.²⁰

¹⁸ See Jagtiani, Vermilyea, and Wall (2018) for further discussion on the use of alternative data, big data, and machine learning in credit decisions.

¹⁹ Different fintech lending platforms tend to have access to different proprietary alternative data sources. Our example from LendingClub platform may be viewed as a case study that may or may not be applicable to other lending platforms.

²⁰ Based on Y-9C reports, as of 2015, the combined credit card balances across all large bank holding companies account for approximately 85 percent of all credit card balances across all U.S. bank holding companies that file Y-9C reports.

To investigate the impact on the price of credit, we explored interest rate spreads for similar loans — loans made through the LendingClub consumer platform (with specific purposes to pay off credit card balances or for debt consolidation) versus traditional bank card loans (revolvers only). We found evidence that credit spreads are priced accurately based on the expected delinquency of the loans. In addition, given that for the same risk of default, consumers pay smaller interest rate spreads on loans from LendingClub than from traditional lending channels, indicating that fintech lending can provide credit to consumers at a lower cost.

We also found that the use of nontraditional information from alternative data sources has allowed consumers with fewer or inaccurate credit records (based on FICO scores) to have access to credit. Some creditworthy consumers (but with poor FICO scores) have been identified using additional information and have been rated as low-risk borrowers by LendingClub. The correlation between rating grades and FICO scores declined steadily from over 80 percent (for loans that were originated in 2007) to about 35 percent for loans originated in 2015. Interestingly, these rating grades (with only 35 percent correlation with FICO) continued to serve as a good predictor for future loan delinquency over the next two years. There is additional (soft) information in LendingClub’s own internal rating grades that is not already incorporated in the obvious traditional risk factors. This has enabled some borrowers to be assigned better loan ratings and receive lower-priced credit.²¹

Our previous research in Jagtiani and Lemieux (2018) presented evidence that fintech lenders fill credit gaps in areas where bank offices may be less available and provide credit to creditworthy borrowers that banks may not be serving. Our further research in this paper finds that loans from fintech lenders seem to be “appropriately” risk priced. Banks are responding to these innovations by partnering with fintech firms. This relationship is evolving quickly.

Our results provide policy implications related to consumer protection. While consumers’ information and privacy should be protected by laws and regulations, certain alternative information could play a key role in allowing lenders to fully understand credit quality of the potential borrowers and allowing certain consumers access to credit that would not have been granted otherwise. Banks could potentially benefit from the alternative data sources and big data through partnership with online fintech lenders. Further research remains to be done to fully answer the question about other aspects of risks to borrowers presented by these new innovations and whether these fintech lending innovations have allowed consumers to become excessively leveraged.

²¹ Similarly, some borrowers who would have been overrated in the A grade by FICO scores are rated in lower grades by LendingClub (because their credit risk was not embedded in traditional credit scores) and they were charged a higher spread.

References

- Berger, A. N., N. H. Miller, M. A. Petersen, R. G. Rajan, and J. C. Stein (2005) "Does Function Follow Organizational Form? Evidence from the Lending Practices of Large and Small Banks," *Journal of Financial Economics*, 76(2):237–269.
- Berger, A. N., and G. F. Udell (2002) "Small Business Credit Availability and Relationship Lending: The Importance of Bank Organizational Structure," *Economic Journal*, 112(477):F32–F53.
- Bertsch, C., I. Hull, and X. Zhang (2016) "Monetary Normalizations and Consumer Credit: Evidence from Fed Liftoff and Online Lending," Sveriges Riksbank Working Paper 319.
- Boot, A. W. A., and A. V. Thakor (2000) "Can Relationship Banking Survive Competition?" *Journal of Finance*, 55(2):679–713.
- Bricker, J., L. J. Dettling, A. Henriques, J. W. Hsu, L. Jacobs, K. B. Moore, S. Pack, J. Sabelhaus, J. Thompson, and R. A. Windle (2017) "Changes in U.S. Family Finances from 2013 to 2016: Evidence from the Survey of Consumer Finance," *Federal Reserve Bulletin*, 103(3):1–42.
- Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2017) "Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks," NBER Working Paper 23288.
- Carroll, P., and S. Rehmani (2017) "Alternative Data and the Unbanked," Oliver Wyman Insights. <http://www.oliverwyman.com/our-expertise/insights/2017/may/alternative-data-and-the-unbanked.html>.
- Chen B. S., S. G. Hanson, and J. C. Stein (2017) "The Decline of Big Bank Lending to Small Banks: Dynamic Impacts on Local Credit and Labor Markets," NBER Working Paper 23843.
- Consumer Financial Protection Bureau (2017) "CFPB Explores Impact of Alternative Data on Credit Access for Consumers Who Are Credit Invisible," Press Release, available at <https://www.consumerfinance.gov/about-us/newsroom/cfpb-explores-impact-alternative-data-credit-access-consumers-who-are-credit-invisible/>.
- Crosman, P. (2016) "Will Fintechs Kill the FICO Score?" *American Banker*, June 14, 2016, available at <https://www.consumerfinance.gov/about-us/newsroom/cfpb-explores-impact-alternative-data-credit-access-consumers-who-are-credit-invisible/>.
- Crowe, C., and R. Ramcharan (2013). "House Prices and Household Credit Access: Evidence From Prosper.Com*," *Journal of Money, Credit and Banking*, 45(6), September 2013, 1085–1115.
- Demyanyk, Y., and D. Kolliner (2014) "Peer-to-Peer Lending Is Poised to Grow," Federal Reserve Bank of Cleveland, available at <https://www.clevelandfed.org/newsroom-and-events/publications/economic-trends/2014-economic-trends/et-20140814-peer-to-peer-lending-is-poised-to-grow.aspx>.
- De Roure, C., L. Pelizzon, and P. Tasca (2016) "How Does P2P Lending Fit into the Consumer Credit Market," Deutsche Bundesbank Discussion Paper 30, available at

https://www.bundesbank.de/Navigation/EN/Bundesbank/Research/Research_centre/Discussion_papers/discussion_papers.html.

Duarte, J., S. Siegel, and L. Young (2012) "Trust and Credit: The Role of Appearance in Peer-to-Peer Lending," *The Review of Financial Studies*, 25(8):2455–2484.

Einav, L., M. Jenkins, and J. Levin (2013) "The Impact of Credit Scoring on Consumer Lending," *RAND Journal of Economics*, 44(2):249–274.

Emekter, R., Y. Tu, B. Jirasakuldech, and M. Lu (2015) "Evaluating Credit Risk and Loan Performance in Online Peer-to-Peer (P2P) Lending," *Applied Economics*, 47(1):54–70.

Everett, C. R. (2010) "Group Membership, Relationship Banking and Loan Default Risk: The Case of Online Social Lending," *Banking and Finance Review*, 7(2).

Federal Reserve (2017) Consumer Credit, G.19, available at <https://www.federalreserve.gov/releases/g19/current/>.

Frame, W. S., A. Srinivasan, and L. Woosley (2001) "The Effect of Credit Scoring on Small-Business Lending," *Journal of Money, Credit and Banking*, 33(3):813–825.

Freedman, S., and G. Jin (2017) "The Information Value of Online Social Networks: Lessons from Peer-to-Peer Lending," *International Journal of Industrial Organization*, 51, 185–222.

Gao, Q., and M. Lin (2012) "Linguistic Features and Peer-to-Peer Loan Quality: A Machine Learning Approach," Working Paper, available at https://www.researchgate.net/publication/314805006_Linguistic_Features_and_Peer-to-Peer_Loan_Quality_A_Machine_Learning_Approach.

Gonzalez, L., and Y.K. Loureiro (2014) "When Can a Photo Increase Credit?: The Impact of Lender and Borrower Profiles on Online P2P Loans," *Journal of Experimental and Behavioral Finance*, (forthcoming), Fordham University Schools of Business Research Paper 2442416.

Havrylchyk, O., C. Mariotto, T. Rahim, and M. Verdier (2018) "What Drives the Expansion of the Peer-to-Peer Lending?" Available at SSRN at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2841316.

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

Iyer, R., and M. Puri (2012) "Understanding Bank Runs: The Importance of Depositor-Bank Relationships and Networks," *American Economic Review*, 102(4):1414–1445.

Jaffee, D. M., and T. Russell (1976) "Imperfect Information, Uncertainty, and Credit Rationing," *Quarterly Journal of Economics*, 90(4):651–666.

Jagtiani, J., and C. Lemieux (2018) "Do Fintech Lenders Penetrate Areas That Are Underserved by Traditional Banks?" Federal Reserve Bank of Philadelphia Working Paper 18-13, March 2018.

Jagtiani, J., T. Vermilyea, and L. Wall (2018) "The Roles of Big Data and Machine Learning in Bank Supervision," *Banking Perspectives*, The Clearing House publication, Q1:2018.

Karlan, D. S. (2007) "Social Connections and Group Banking," *Economic Journal*, 117(517):F52–F84.

Lin, M., N. R. Prabhala, and S. Viswanathan (2013) "Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending," *Management Science*, 59(1):17–35.

Lu, Y., B. Gu, Q. Ye, and Z. Sheng. (2012) "Social Influence and Defaults in Peer-to-Peer Lending Networks," Conference on Information Systems and Technology (CIST 2012) White Paper.

Mach, T., C. Carter, and C. Slattery (2014) "Peer-to-Peer Lending to Small Businesses," Federal Reserve Board of Governors, Finance and Economics Discussion Series 2014-10.

Michels, J. (2012) "Do Unverifiable Disclosures Matter? Evidence from Peer-to-Peer Lending," *Accounting Review*, 87(4):1385–1413.

Morse, A. (2015) "Peer-to-Peer Crowdfunding: Information and the Potential for Disruption in Consumer Lending," NBER Working Paper, available at www.nber.org/papers/w20899.pdf.

Liberti, J. M., and M. A. Petersen (2017) "Information: Hard and Soft," Working Paper, Northwestern University, March 2017. Available at <http://www.kellogg.northwestern.edu/faculty/petersen/htm/papers/hard%20and%20soft%20information.pdf>

Petersen, M. A., and R. G. Rajan (1994) "The Benefits of Lending Relationships: Evidence from Small Business Data," *Journal of Finance*, 49(1):3–37.

Pope, D., and J. Sydnor. (2011) "What's in a Picture? Evidence of Discrimination from Prosper.com," *Journal of Human Resources*, 46(1):53–92.

Ravina, E. (2012) "Love & Loans: The Effect of Beauty and Personal Characteristics in Credit Markets," Working Paper, Northwestern University, November 2012. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1107307.

Schoar, A. (2012) "The Personal Side of Relationship Banking," Working Paper, MIT Sloan School of Management, January 2012. Available at http://www.haas.berkeley.edu/groups/finance/RelationshipBanking_v7.pdf.

Stein, J. C. (2002) "Information Production and Capital Allocation: Decentralized vs. Hierarchical Firms," *Journal of Finance*, 57(5):1891–1921.

Stiglitz, J. E., and A. Weiss (1981) "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 71(3):393–410.

Table 1

**Comparing Price of Credit:
LendingClub Loans versus Y-14M Credit Card Loans (Revolvers Only)**

Sample Period: Loans Originated in 2014–2015 Only

FICO Segment at Origination	% Average Spread LendingClub		% Average Spread	Significant Difference at the 1% Level?	
	3-Year Maturity	5-Year Maturity	Bank Y-14M (Rollovers Only)		
				3-Year	5-Year
660–679	12.0646 N=139,337	15.7089 N=64,359	20.1923 N=6,812	Yes	Yes
680–699	10.7630 N=100,033	14.3937 N=54,030	19.8465 N=7,067	Yes	Yes
700–719	9.3477 N=64,271	13.0239 N=36,313	19.1418 N=6,637	Yes	Yes
720–739	8.12608 N=32,512	11.7484 N=17,071	18.4180 N=5,930	Yes	Yes
740–759	7.16102 N=15,403	10.5891 N=6,823	17.6569 N=5,383	Yes	Yes
760–779	6.5303 N=8,081	9.7955 N=3,015	16.8312 N=4,701	Yes	Yes
780–799	6.0904 N=4,458	9.2009 N=1,436	16.1820 N=4,586	Yes	Yes
800+	5.6408 N=2,509	8.6312 N=837	16.1668 N=12,070	Yes	Yes

Note: Credit spreads on credit card loans are significantly higher than consumer loans from LendingClub (regardless of the loan maturity), even after controlling for the borrower's FICO score.

Table 2

**Regression Results — LendingClub Consumer Loans:
Important Factors That Determine Credit Spreads**

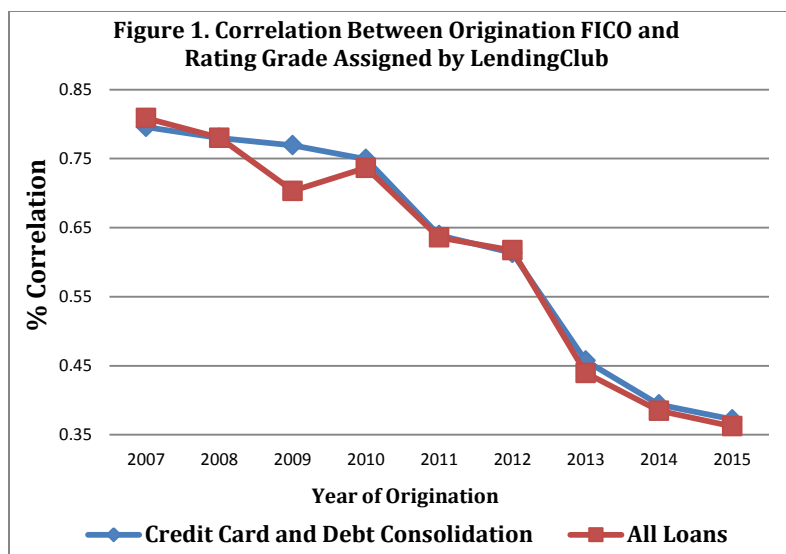
Sample Period: 2010–2017

Data are at loan level from LendingClub’s consumer platform (for credit cards or debt consolidation only). All loans were originated in 2007–2015, with a two-year performance period ending in 2017 or earlier. Dependent variables are interest rate spreads, which are calculated as the difference between the interest rates charged on the loans and the equivalent risk-free loans (U.S. Treasury rate of securities with the same time to maturity). The variables *Rating Grade A* and *FICO at Origination Greater Than 800* serve as the base case. The ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

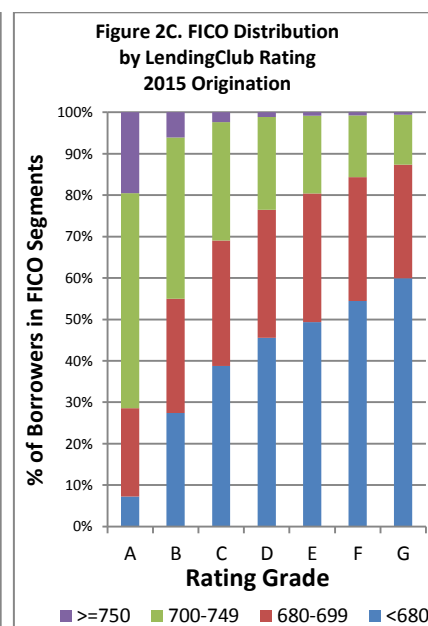
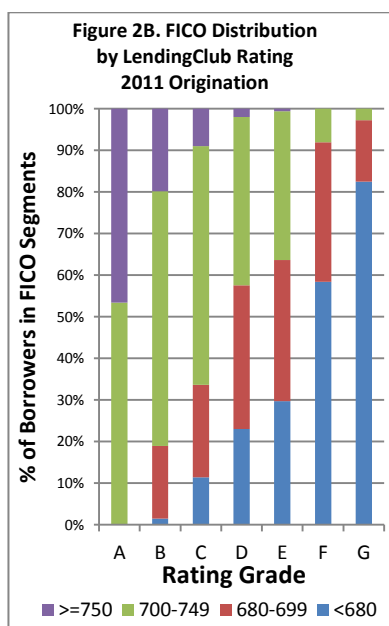
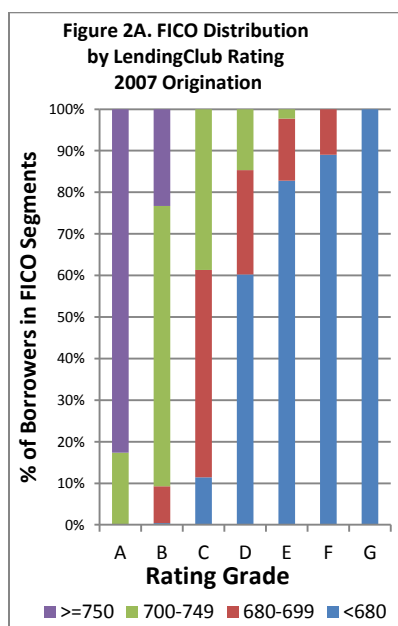
	LendingClub Rating Grades A to G		Origination FICO Scores	
Independent Variable	(1)	(2)	(3)	(4)
<i>Intercept</i>	6.3339*** (0.0001)	9.7106*** (0.0001)	6.4811*** (0.0001)	8.1175*** (0.0001)
<i>D(Default within 24-Mo After Origination)</i>	—	0.2197*** (0.0001)	—	1.3499*** (0.0001)
<i>D(2014-15)*D(Default within 24-Mo After Origination)</i>	—	−0.0979*** (0.0001)	—	0.3387*** (0.0001)
<i>D(Rating Grade B)</i>	3.5325*** (0.0001)	3.2572*** (0.0001)	—	—
<i>D(Rating Grade C)</i>	6.5685*** (0.0001)	6.3532*** (0.0001)	—	—
<i>D(Rating Grade D)</i>	9.6647*** (0.0001)	9.4415*** (0.0001)	—	—
<i>D(Rating Grade E)</i>	12.2472*** (0.0001)	12.1488*** (0.0001)	—	—
<i>D(Rating Grade F)</i>	15.9334*** (0.0001)	15.8618*** (0.0001)	—	—
<i>D(Rating Grade G)</i>	18.0956*** (0.0001)	18.1991*** (0.0001)	—	—
<i>D(650<FICO at Origination<680)</i>	—	—	7.2882*** (0.0001)	6.0992*** (0.0001)
<i>D(680<FICO at Origination<700)</i>	—	—	6.0733*** (0.0001)	4.8362*** (0.0001)

<i>D(700<FICO at Origination<750)</i>	—	—	3.9252*** (0.0001)	2.9388*** (0.0001)
<i>D(750<FICO at Origination<800)</i>	—	—	1.0541*** (0.0001)	0.6505*** (0.0001)
<i>D(Homeownership)</i>	—	-0.0819*** (0.0001)	—	-0.2051*** (0.0001)
<i>D(Employment>10 Yrs)</i>	—	0.0082*** (0.0044)	—	0.1540*** (0.0001)
<i>Debt-to-Income Ratio at Origination</i>	—	0.0037*** (0.0001)	—	0.0647*** (0.0001)
<i>Log(Borrower's Income)</i>	—	-0.1292*** (0.0001)	—	-1.5971*** (0.0001)
<i>Log(Origination Loan Amount)</i>	—	-0.0493*** (0.0001)	—	1.7489*** (0.0001)
<i>Number of Credit Inquiries 6-Mo Before</i>	—	0.1076*** (0.0001)	—	0.8983*** (0.0001)
<i>Home Price Index (3-digit Zip)?</i>	—	0.0001*** (0.0001)	—	-0.0000 (0.8678)
<i>Unemployment Rate (3-digit Zip)?</i>	—	0.0064*** (0.0001)	—	0.0280*** (0.0001)
<i>D(Origination Year 2014)</i>	—	-1.3258*** (0.0001)	— —	-1.2583*** (0.0001)
<i>D(Origination Year 2015)</i>	—	-2.1937*** (0.0001)	—	-2.2765*** (0.0001)
<i>D(Y-14M Card Loans HHI>2500)</i>	—	0.0029 (0.5825)		-0.0749*** (0.0001)
Adjusted R ² Observation Number(N)	88.63% 725,800	93.42% 663,576	17.62% 725,800	34.25% 663,576

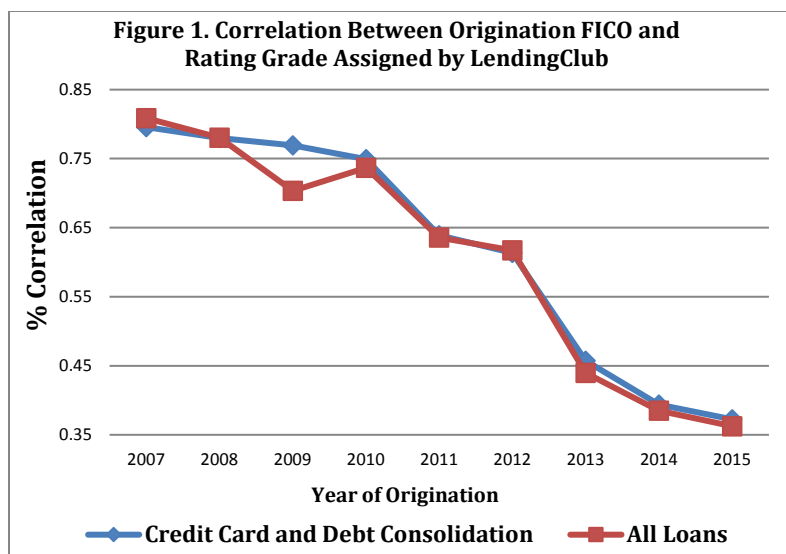
Note: The sample period starts in 2013 in columns 2 and 4 owing to unavailability of reliable Y-14M data prior to 2013. The data are used to calculate the HHI market concentration measure. Note also that all loans were originated up to 2015 to allow 24 months of performance period to observe the loans' default behavior.



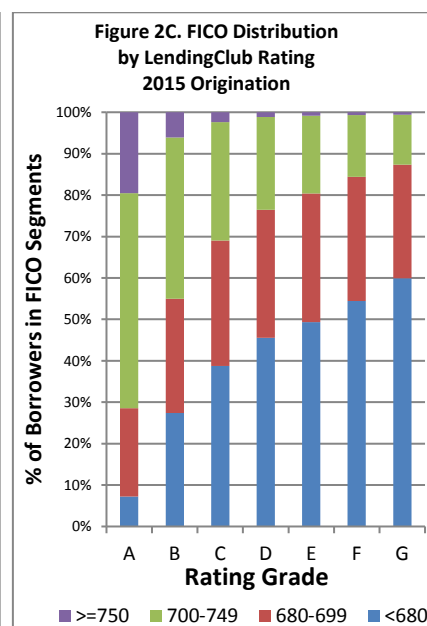
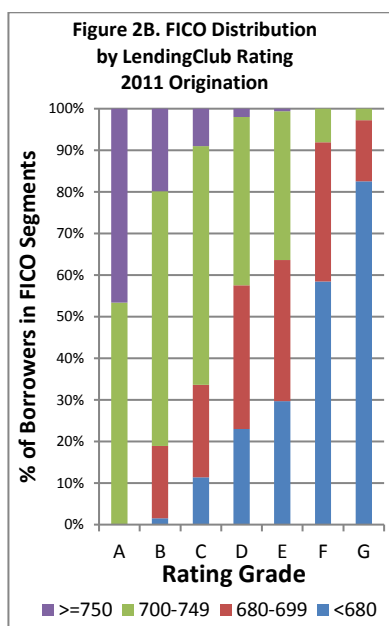
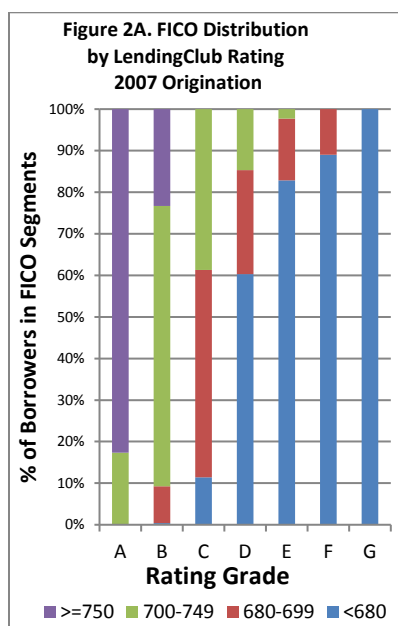
Source: LendingClub data



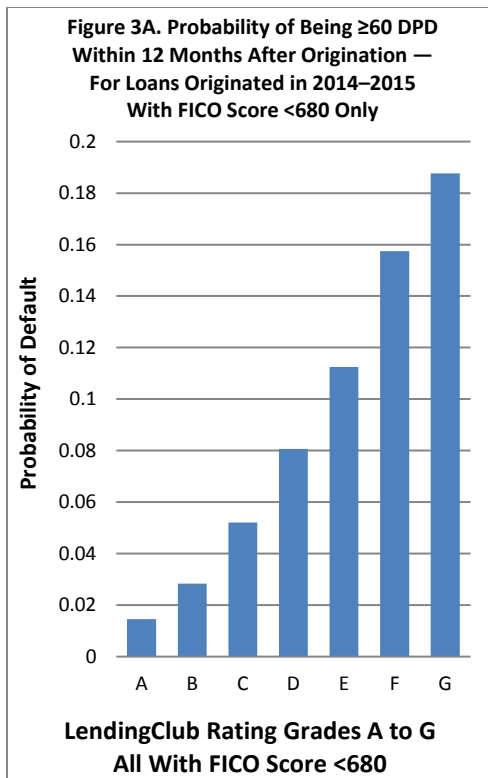
Source: LendingClub data



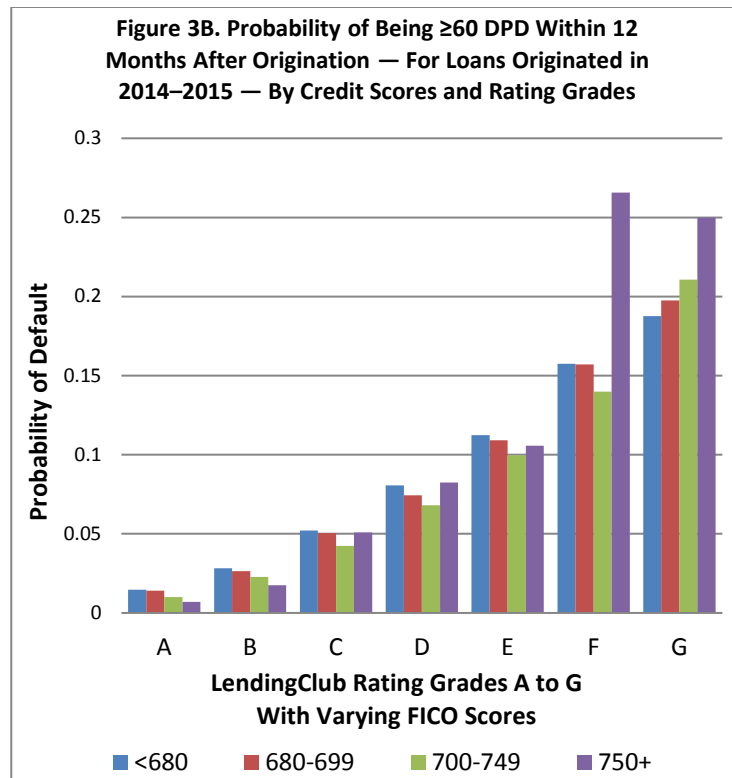
Source: LendingClub data



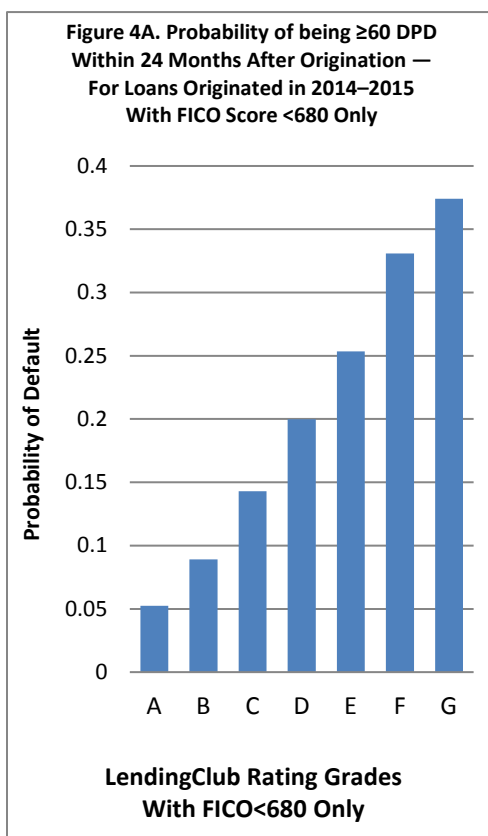
Source: LendingClub data



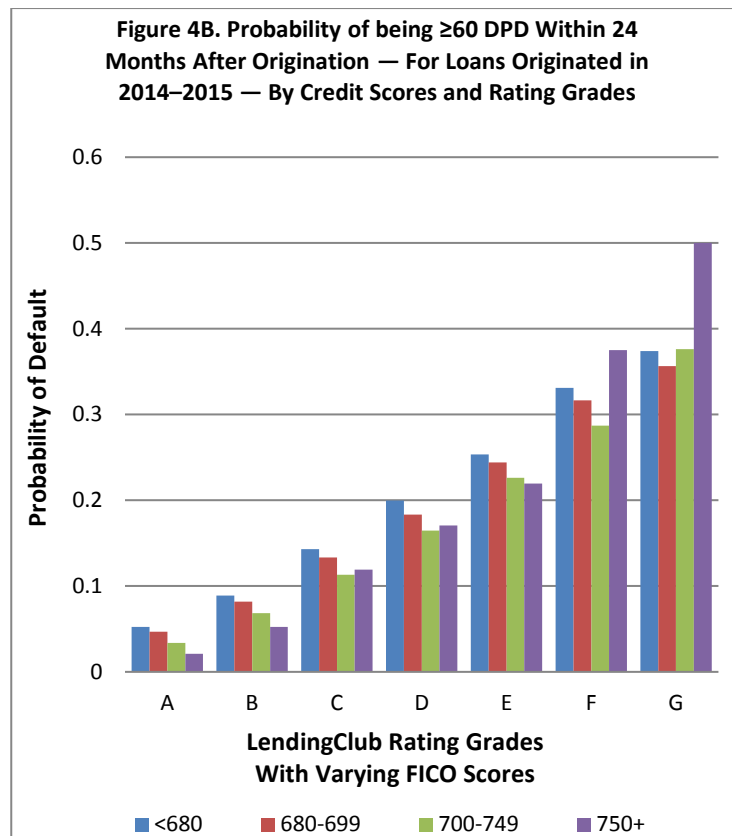
Source: LendingClub



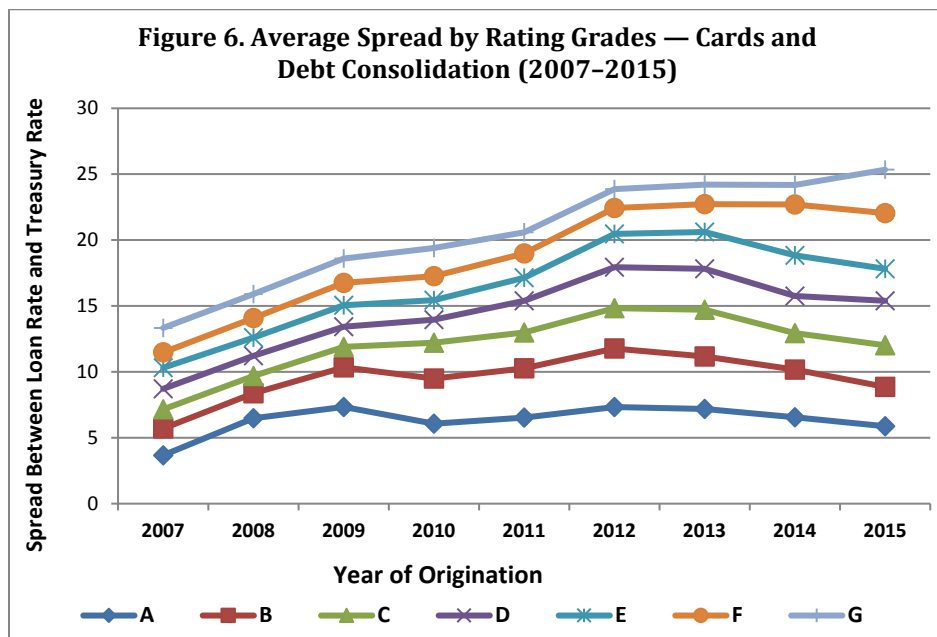
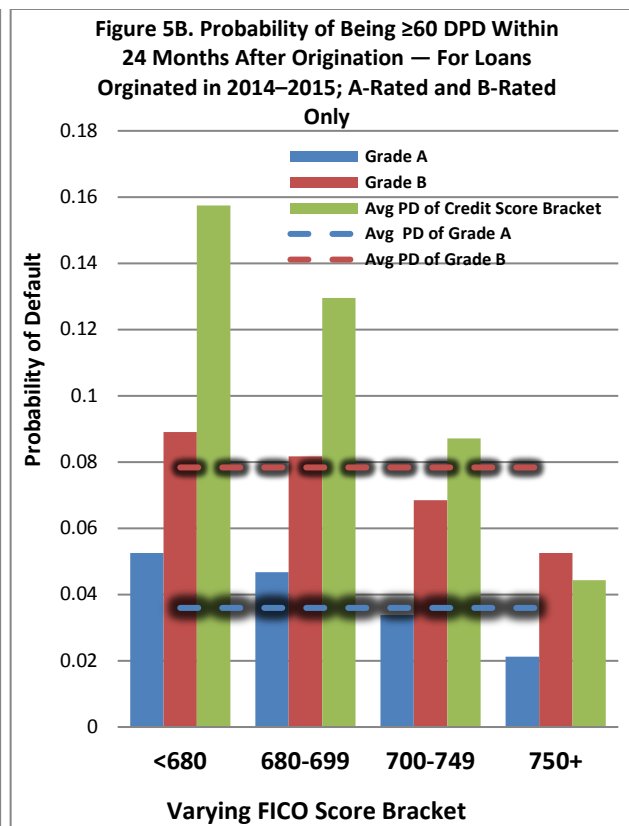
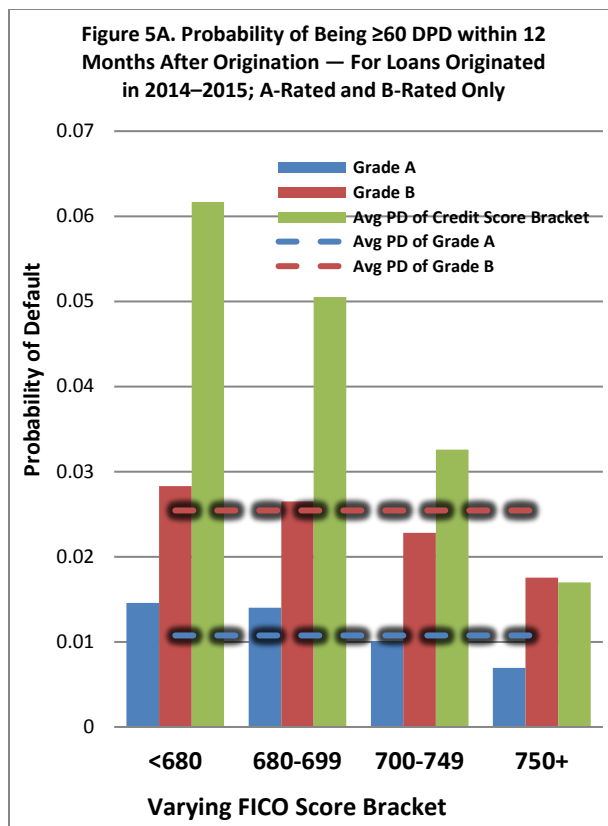
Source: LendingClub

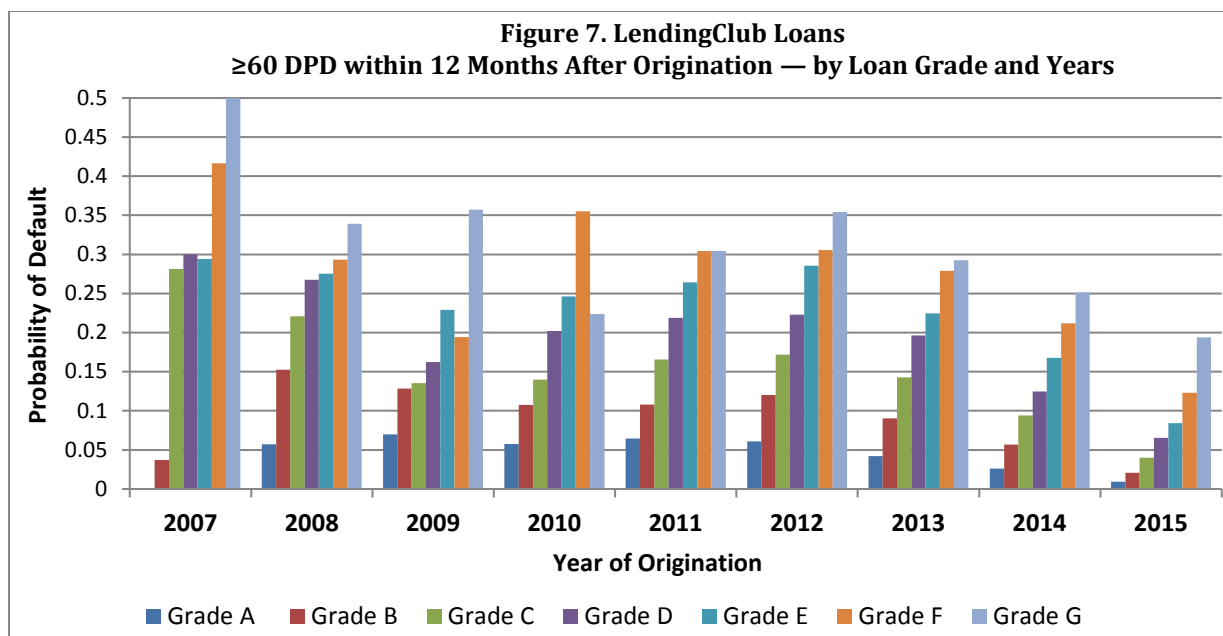


Source: LendingClub

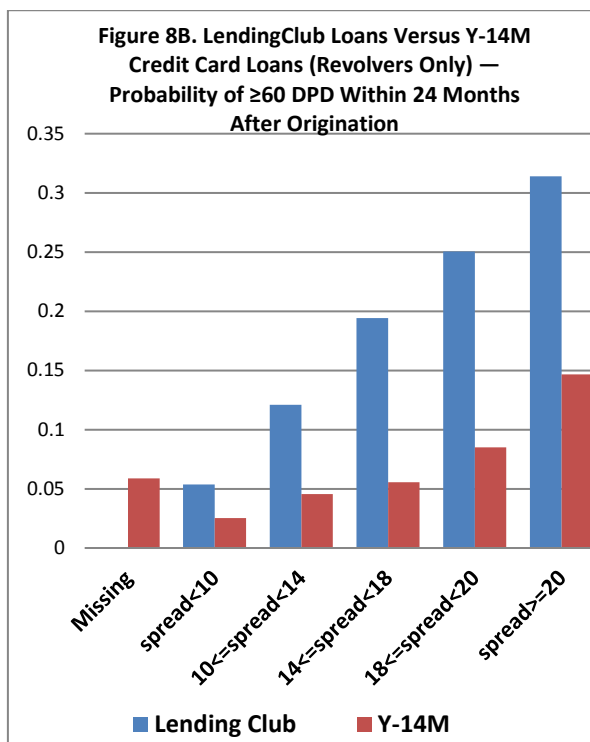
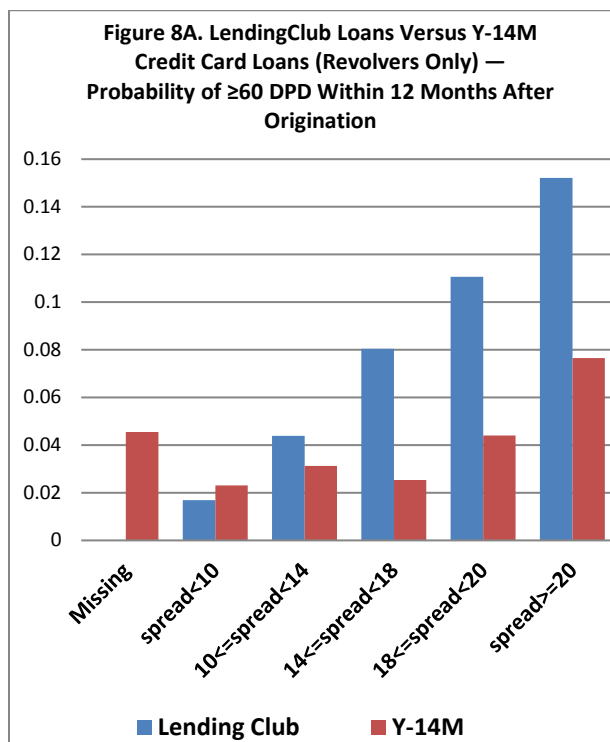


Source: LendingClub

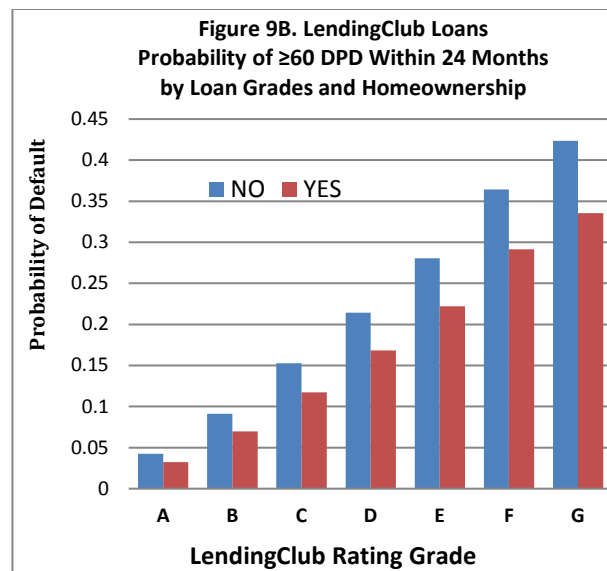
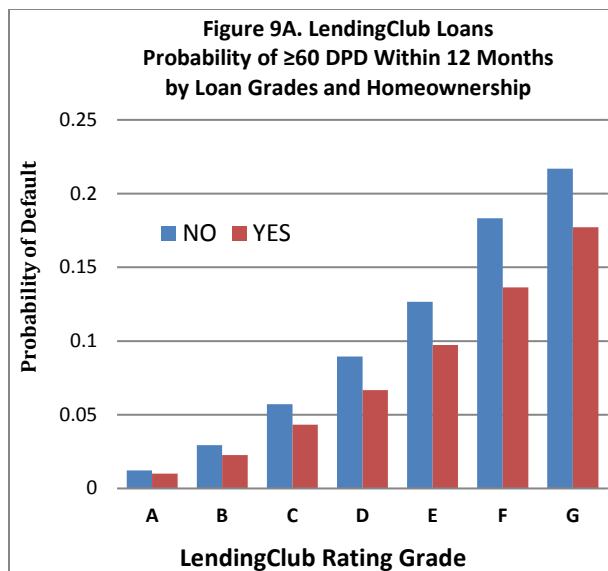




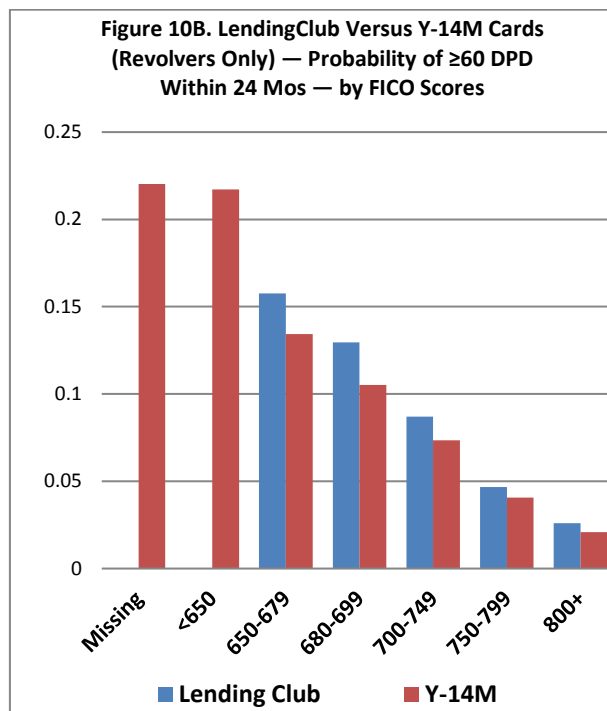
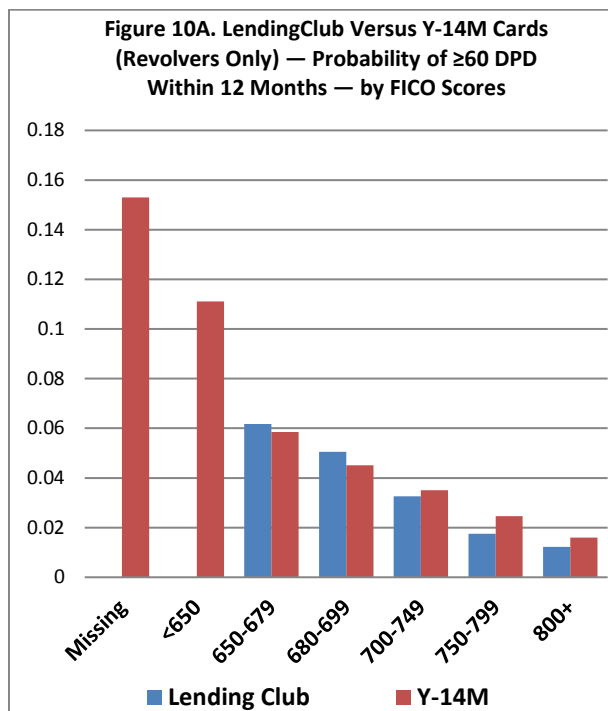
Source: LendingClub loans (cards and debt consolidation purposes only)



Sources: LendingClub loans (cards and debt consolidation purposes only) and Y-14M data on credit cards.
Note: All loans were originated during the period from January 2014 to December 2015. Delinquency status (became ≥ 60 DPD) is observed for the period within 12 months after loan origination.



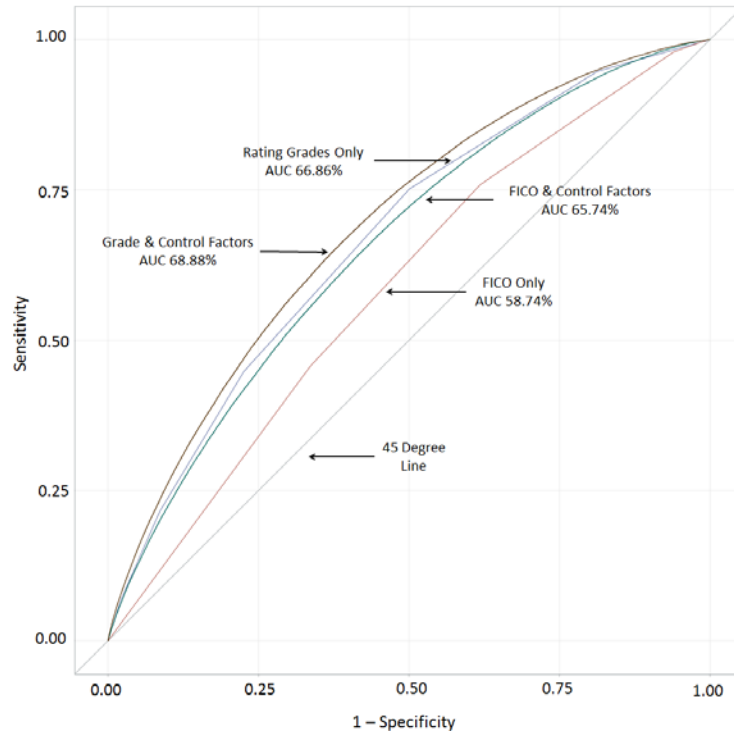
Sources: LendingClub loans (cards and debt consolidation purposes only) that were originated in 2014 and 2015 only; Y-14M data on credit card accounts were issued to consumers during 2014–2015



Sources: LendingClub loans (cards and debt consolidation purposes only) that were originated in 2014 and 2015 only; Y-14M data on credit card accounts were issued to consumers during 2014–2015.

Figure 11

This figure illustrates the discriminatory power of four different models of default probability specifications by providing the Receiver Operating Characteristics curve (ROC-curve) and the Area Under Curve (AUC). The ROC-curves are estimated using a logit regression of the default dummy (being at least 60 DPD within two years after origination) on 1) FICO scores only, 2) Rating Grades only, 3) FICO scores and other control factors, 4) Rating Grades and the same set of other control factors.



Source: LendingClub data and economic factors from U.S. Census Bureau and Haver Analytics database