Adverse Selection on Maturity: Evidence from Online Consumer Credit*

Andrew Hertzberg   Andres Liberman   Daniel Paravisini*

December 2016†

Abstract

Longer loan maturity provides borrowers with insurance against future changes in the price of credit. The present paper examines whether, consistent with theories of insurance markets with private information, maturity choice leads to adverse selection. Our estimation compares two groups of observationally equivalent borrowers that took identical unsecured 36-month loans, only one of which had also a 60-month maturity choice available. We find that when long maturity is available, fewer borrowers take the short-term loan, and those that do, default less. Additional findings suggest borrowers self-select on private information about their future ability to repay. The findings imply that maturity can be used to screen borrowers on this private information.

Keywords: Adverse Selection, Loan Maturity, Consumer Credit.
JEL codes: D82, D14.

---

*Hertzberg is at Columbia University, email: ah2692@gsb.columbia.edu. Liberman is at New York University, email: aliberma@stern.nyu.edu. Paravisini is at London School of Economics, email: D.Paravisini@lse.ac.uk. We thank Sumit Agarwal, Asaf Bernstein, Emily Breza, Tony Cookson, Anthony DeFusco, Theresa Kuchler, Adair Morse, Holger Mueller, Christopher Palmer, Mitchell Petersen, Philipp Schnabl, Antoinette Schoar, Amit Seru, Felipe Severino, Johannes Stroebel, and participants at AFA (San Francisco), Bocconi University, Columbia University, Credit and Payments Markets Conference (Federal Reserve Bank of Philadelphia), Crowdfunding Symposium (Berkeley), CUHK, Dartmouth University (Tuck), EFA (Oslo), Financial Intermediation Research Society Conference (Lisbon), HKU, HKUST, LSE (Economics and Finance departments), Melbourne Business School, Monash Business School, NBER Corporate Finance 2015 Fall meeting (Stanford), NBER Household Finance Summer Institute 2015, NYU (Stern), NYU-Columbia Junior Faculty Seminar, UBC (Sauder), UNC (Kenan-Flagler), and University of New South Wales. We thank Siddharth Vij for outstanding research assistance. All errors and omissions are ours only.

†First version: July 2015

Loan maturity provides borrowers with insurance against future changes in the price of credit that may arise, for example, if the borrower's observed credit quality deteriorates or credit supply dries up. A short maturity borrower who wishes to delay payment must return to the credit markets to borrow at an uncertain prevailing rate in the future, while a long maturity borrower can delay payment at a rate predetermined at issuance.[1] When borrowers have private information about the value they place on this insurance (e.g., about their future observable ability to repay, their risk aversion, or the timing of their cash flows), the market for loan maturity may not be characterized by a single price at which borrowers can buy all the insurance—maturity—they require (Rothschild and Stiglitz (1976)). In particular, when borrowers are privately informed about their future probability of repayment, low-risk borrowers may choose contracts that forgo insurance (maturity) in order to avoid the cost of pooling with higher-risk types. Thus, assessing whether and how borrowers' private information affects maturity choice is crucial for understanding the functioning of the market for loan maturity.

While the theory of maturity as a screening device dates back to Flannery (1986), there is, to date, no evidence of this role. The present paper fills this gap. Doing so requires addressing two empirical challenges. The first is demonstrating that borrowers have *private* information about their own future probability of default. By definition, this can only be addressed by measuring outcomes that are not observed at the time of origination, e.g., by looking at the ex-post default performance of observationally equivalent borrowers. This leads to the second challenge: demonstrating that, given a choice between contracts, borrowers self-select among maturity options using this private information. Demonstrating the screening role of any contract dimension *cannot* be achieved by comparing across borrowers that chose different contracts, because the different contracts characteristics (e.g., maturity, price, installment amount) may affect the ex post behavior of ex ante identical borrowers.[2] For this reason, empirically identifying the consequences of maturity selection on repayment requires comparing how selected and non-selected borrower samples behave when facing the *same* contract. This is the basis of our empirical strategy and our main departure from the existing empirical literature examining the link between asymmetric information and loan maturity, which has focused on showing how maturity choice varies with borrowers' observable creditworthiness or ex-ante proxies for the degree of private information.[3]

[1]The insurance role of long term contracts is not exclusive to credit markets, as it is also played, for example, by long-term employment contracts (Holmstrom (1983)) and long-term health and care insurance (Cochrane (1995), Finkelstein, McGarry, and Sufi (2005)).

[2]Papers that make this comparison, such as Goyal and Wang (2013) and Gopalan, Song, Yerramilli, et al. (2014), are therefore unable to isolate the role of either selection or maturity on default.

[3]For examples of the first see Barclay and Smith (1995), Guedes and Opler (1996), Johnson (2003) and for the second see Berger, Espinosa-Vega, Frame, and Miller (2005). Much of this empirical work is motivated by the theory of Diamond (1991) who uses a framework with asymmetric information to predict a link between observable creditworthiness and the type of maturity that all borrowers will pool on in equilibrium. As such, these papers cannot rule out the possibility that all observably equivalent borrowers select the same loan regardless of their private information. By isolating selection on private information, our paper is also distinct to theories of maturity choice that are unrelated

To illustrate the above point and provide a motivation for our empirical strategy, consider the idealized setting for identifying selection on maturity depicted in Figure 1. Suppose we observe two groups of prospective borrowers, A and B, before they take a loan. Group A is offered only a short-maturity loan at an interest rate of r_{ST}. The default rate of these borrowers is γ_A^{ST}. Group B is offered two options: the same short-maturity loan as group A (at rate r_{ST}), and a long-maturity loan for the same amount at a rate of r_{LT}. Group B borrowers that choose the short-term (long-term) loan default at a rate γ_B^{ST} (γ_B^{LT}). Borrowers from group B who take the short-term loan are *selected on maturity*: they could have taken a long-term loan, but chose not to. Group A borrowers, in contrast, are an *unselected* group. Further, both group A and group B short-term borrowers face identical loan terms (interest rate, amount, and maturity). Thus, any difference in the repayment of the short-term loans between group B and group A borrowers, $\gamma_B^{ST} - \gamma_A^{ST}$, must be driven by the selection induced by the long-maturity loan. Under the null hypothesis γ_B^{ST} is equal to γ_A^{ST}, which would suggest either that borrowers do not have any private information about their future default rate or, if they do, that their choice of maturity does not depend on this private information. A rejection of this null hypothesis indicates instead that individuals are privately informed about their probability of default and that their choice of maturity is related to this private information. In particular, $\gamma_B^{ST} - \gamma_A^{ST} < 0$ would indicate that borrowers with a higher privately observed default risk select into the long-maturity loan.

We exploit the staggered roll-out of long-maturity loans by an online lending platform, Lending Club (hereafter, LC), as an empirical setting that closely resembles this idealized one. When a borrower applies for a loan at LC she is assigned to a narrow risk category based on FICO score and other observable characteristics. All the borrowers in a risk category are offered the same menu of loan choices, e.g. the same interest rate for every amount and maturity combination. Loans of amounts between \$1,000 and \$35,000 are available in either short—36 months—or long maturities—60 months. Before 2013 the long-maturity loan was available only for amounts above \$16,000. During 2013, the available menu of long-term loan options expanded twice: 1) to loans amounts between \$12,000 and \$16,000 in March 2013, and 2) to loan amounts between \$10,000 and \$12,000 in July 2013. Crucially for our analysis, during our analysis sample period LC did not change of any of the loan terms that were available in the menu of borrowing options before the addition of the new long-term loans, nor the screening criteria to qualify for a loan.

Our empirical strategy compares the default rate of short-term loans between \$10,000 and \$16,000 issued before and after the availability of the long-maturity option at the corresponding amount, within borrowers assigned by LC to the same risk category, which approximate groups A and B of

to ex-ante asymmetric information such as: asset maturity matching (e.g., Myers (1977), Hart and Moore (1994)), agency problems (e.g., Hart and Moore (1995)), market conditions (e.g., Barry Bosworth (1971), Taggart (1977)), minimize rollover risk (e.g., Graham and Harvey (2001)), predictable violations of the expectations hypothesis (e.g., Baker, Greenwood, and Wurgler (2003)), and government behavior (e.g., Greenwood, Hanson, and Stein (2010)).

the idealized setting of Figure 1.[4] Before-after comparisons within risk categories are potentially confounded by changes over time in the composition of borrowers on the LC platform. To account for these changes, we estimate a difference-in-differences specification that exploits the staggered roll-out of the long-term loans, and that uses short-term loans of amounts just above and just below the \$10,000 to \$16,000 interval to construct counterfactuals. Intuitively, our main test compares, amongst borrowers that look ex ante identical in all observables, the default rate of loans between \$10,000 and \$16,000 that were issued before and after the long-maturity loan became available at these amounts, relative to the same change in the default rate of loans between \$5,000 and \$10,000 or between \$16,000 and \$20,000 issued during the same period. The identification assumption is that any change in the composition of borrowers within a risk category that occurs for reasons other than the menu expansion, for example due to changes in the credit supply by other lenders, did not affect diferentially loans between \$12,000 and \$16,000 in March 2013 and between \$10,000 and \$12,000 in July 2013, relative to other amounts in the analysis sample at those dates. To further insure that all comparisons are done across observationally equivalent borrowers, we include in our specifications month-of-origination, 4-point FICO range, state fixed-effects, and controls for all the borrower characteristics recorded by LC at origination.[5]

We begin by documenting that the bulk of self-selection into long-maturity loans occurs among borrowers who would have borrowed between \$10,000 and \$16,000. We find that the number of short-maturity loans between \$10,000 and \$16,000 drops by 14.5% after the long-maturity loans become available, relative to loans issued at amounts just above and below this interval. Further, the decline was permanent and occurred on the same month the 60-month loan appeared in the menu for the corresponding amount.

Then we explore how selection on maturity relates to ex-post performance. We find that the average default rate of short-maturity loans decreases by 0.8 percentage points when a long-maturity loan is available at origination, relative to when it is not. This implies that borrowers that look identical ex ante from the investors' perspective but that have a higher default risk *self-select* out of short-term loans and into long-term ones. Assuming that the difference in short-term loan performance is due to the 14.5% of borrowers who self-select into long maturity, these self-selected borrowers would have had a default rate 5.5 percentage points higher (0.8/14.5) than the average 36-month

[4]For example, borrowers choosing a 36-month \$10,000 loan *before* July 2013 resemble those in group A of Figure 1: these borrowers did not have a long term option in the menu at the time of making the choice. Borrowers choosing a 36-month \$10,000 loan *after* July 2013 resemble borrowers in group B: they chose the 36-month loan when a longer maturity loan was available, and are thus a sample selected on maturity.

[5]The LC setting has several additional advantages that underline the robustness of our estimates. First, loans offered on the LC platform are funded by investors at the terms set by LC's pricing algorithm. These terms compare favorably to other investments of similar risk, thereby ensuring that all loans are funded. This rules out that selection is occurring based on supply side screening decisions. Second, LC charges an upfront origination fee between 1.1 and 5% of a borrower's loan amount (subtracted from the amount borrowed). Thus, borrowers who took a short-maturity loan prior to the expansion could not costlessly swap them for long maturity ones after the expansion. This ensures that the pool of borrowers who select the short-maturity loan prior to the expansion is not impacted by the expansion itself.

borrower in our sample (9.2%). The findings are consistent with the joint hypotheses that LC borrowers have private information related to their future repayment probability, and that this private information affects loan maturity choice. The large economic magnitude suggests that selection on maturity provides a powerful device for identifying, among a pool of observationally identical borrowers, those with the poorest repayment prospects.[6]

Having established that borrowers select maturity based on private information that correlates with their repayment prospects, we turn to understanding the economic nature of this private information. In theory, borrowers who are privately informed about their own high risk aversion will select the higher insurance provided by longer maturity loans (De Meza and Webb (2001)). However, since risk averse borrowers are expected to default less, self-selection on risk aversion is inconsistent with the higher default rate exhibited by long maturity borrowers. In addition, it is unlikely that borrowers are privately informed about interest rate risk, the probability of credit supply shocks, or other macro determinants of the future cost of borrowing. It follows that borrowers who select long-maturity loans privately place higher value on the insurance it provides either because: 1) they are more exposed to future shocks to their observable creditworthiness (e.g., the probability of job loss or illness) or because 2) they are more exposed to rollover risk due to privately observed differences in the timing of their income.

The two explanations have different predictions regarding the timing and level of default by borrowers who self-select into long maturity. Regarding the timing of default, borrowers that self-select into long maturity because their income arrives later will tend to default less over time, as their income realizes. In contrast, borrowers who self-select into long-maturity loans because they are more exposed to future shocks to their ability to repay default more over time, as the negative shocks realize. Using our empirical approach to estimate how selection affects default at different horizons, we find that selection does not significantly affect repayment during the first twelve months after origination, even though, unconditionally, more than a third of the loans that default do so during this period. In other words, we can reject the hypothesis that the propensity to default of borrowers who self-select into long maturity loans decreases over time (relative to borrowers who self-select into short maturity loans). This evidence is inconsistent with borrowers self-selecting on the basis of the timing of their income, and consistent with them self-selecting on private information about the exposure to shocks to their ability to repay.

Regarding the level of default, if borrowers prefer a long- over a short-maturity loan because their income arrives in the future, their default probability should be lower under a long-term loan that aligns payments better with the timing of income. In our setting, however, the average default probability of 60-month loans is 3 percentage points higher than that of 36-month loans

[6]Officers at LC privately expressed to us that adverse selection is one of the biggest concerns they have whenever LC modifies loan menu items, which is consistent with the large economic effects we document..

(conditioning on loan amount, month of origination, and FICO).[7] Although stylized, this evidence runs counter to the notion that the bulk of selection is driven by borrowers looking for loans that best suit the timing of their income.

We find additional evidence in support of the interpretation that borrowers select maturity based on private information about their exposure to shocks. When we estimate our main specification to identify the effect of maturity selection on the borrower's future FICO score, measured approximately two years after origination, we find that the average FICO score for the selected group (borrowers that choose short maturity when long is available) is 2.7 points higher relative to the non-selected group. Second, we show that the time-series variance of the FICO scores is higher for the sample of borrowers that choose short maturity when long was available (the selected sample).[8] Third, we find that the propensity for borrowers to prepay the short-term loan is lower in the selected group relative to the unselected group. Although this result is not statistically significant, it is inconsistent with the hypothesis that short-term loans are selected by borrowers based on private information that their income arrives sooner. These results demonstrate that borrowers are selecting maturity based on private information that is related to a higher exposure to adverse shocks to their future observable creditworthiness.

In theory, the results could also be driven by borrowers who have a preference for long term loans for behavioral reasons (e.g., borrowers may evaluate the price of a loan by the installment amount instead of by the interest rate and fees) and who, at the same time, are more likely to default. However, 87% of LC borrowers claims to use the LC loan proceeds to repay credit card debt. Since credit card debt is essentially very long-term debt, the majority of borrowers in our sample is actively choosing to lower the maturity profile of their debt and to increase the monthly installment amounts.[9] Thus, LC borrowers seem to be unconstrained enough to commit to increase their minimum monthly payments relative to those imposed by their existing credit card debt and sufficiently sophisticated to understand the difference between price and monthly payment amounts.

Moreover, it is important to note that for unconstrained sophisticated borrowers, loan maturity (a contractual feature of the loan) is distinct from the actual timing of loan repayments (a choice variable). An impatient borrower that has a short-term loan can lower the effective out-of-pocket payments by undertaking additional borrowing each period. For example, if the monthly installment amount of the short term loan is $400, the borrower could pay $300 out of pocket and borrow an addtional $100 in credit card debt to pay the balance (this is feasible for the average borrower

[7] Commensurate with this increased risk, LC charges a 3.3% higher APR for 60-month loans, holding other borrower and loan characteristics constant.

[8] Future FICO scores are measured on April 2015 for all borrowers. The time-series variance of FICO scores is measured using three observations of future FICO scores between the origination date and April 2015.

[9] For comparison, the monthly installments of a $10,000 5-year 10% APR LC loan would be $210, while the minimum repayment per month in a credit card with the same balance and APR would be $93. If the credit card APR were 20%, the minimum monthly payments would be $157, still lower than the monthly installments in the LC loan.

in our sample, who uses only 60% of her available revolving credit at origination). The *only* difference between this series of short-term loans and a long-maturity loan is that the additional borrowing must be done at market interest rates at the time of the new loan. This analysis highlights how the key difference of long- and short-maturity loans is the insurance feature we stress in our interpretation: maturity locks in, at origination, the price at which borrowers can delay monthly payments. Therefore, selection must ultimately be driven by differences in the value borrowers assign to this insurance.

We formalize this intuition in the last section of the paper. We develop a stylized model of consumer credit choice that matches our central empirical findings: borrowers with private information about their increased exposure to shocks to their observable creditworthiness select into long-maturity loans. We use this framework to discuss the conditions under which maturity is the *optimal* way to screen borrowers when screening using loan amounts is also an option. In the model, borrowers have private information about their exposure to adverse shocks in the short and long term. By lowering the minimum payment due in the interim period, long maturity debt provides borrowers with insurance against future shocks to their income and ability to repay. Lenders offer a menu of contracts so that borrowers self-select, and better borrower types can credibly separate themselves from worse types by either borrowing less or by taking shorter maturity loans. Our model demonstrates that maturity (rather than quantity) is the optimal screening device when the informativeness of borrowers' private information to predict default is increasing over time from origination. Even though the purpose of the model is not to replicate the institutional details of the empirical setting, we observe that this theoretical condition is met in the data: selection has an impact on default that is increasing in the time since origination. Intuitively, screening with maturity is optimal under this condition because it shifts payments closer to the horizon at which borrowers have less private information about their repayment capacity.[10]

Our paper relates to a literature on credit markets that follows the logic of Spence (1973), to argue that price can be used in conjunction with other contractual features to screen borrowers on their private information, partially alleviating credit rationing. Aside from maturity, screening devices that have been proposed in the theory literature include collateral (Bester (1985)), loan size (Schreft and Villamil (1992), Brueckner (2000), Adams, Einav, and Levin (2009)), inside ownership (Leland and Pyle (1977)), managerial incentives and capital structure (Ross (1977)), loan covenants (Levine and Hughes (2005)), mortgage points (Stanton and Wallace (1998)), and prepayment penalties (Bian and Yavas (2013)). Despite the wealth of theory, there is essentially no direct evidence that *any loan term* can be used to screen borrowers based on their private information. For example, Adams, Einav, and Levin (2009), and Dobbie and Skiba (2013), estimate adverse selection as a residual,

[10]This result contrasts strongly with Goswami, Noe, and Rebello (1995), the only existing paper that studies how the time structure of private information impacts loan maturity choice. The stark difference arises because that paper assumes that in equilibrium there is no screening on maturity.

given by the correlation between default and loan size that cannot be explained by the direct effect of loan size on default. De Meza and Webb (2016) highlight that such a correlation may exist under symmetric information and thus "cannot diagnose whether asymmetric information is present." [11]

This discussion highlights why our empirical setting presents a unique opportunity to analyze empirically the role of non-price loan contract terms in dealing with asymmetric information. Prior to our work, isolating adverse selection by comparing the behavior of selected and non-selected samples facing the same credit contract had only been possible through randomized controlled trials performed in developing countries (Karlan and Zinman (2009)). Our results pertain to prime borrowers in the U.S. and thus demonstrate how the functioning of consumer credit markets can be shaped by the presence of adverse selection even in a developed economy. The LC environment is particularly well suited to perform the analysis because loan contracts vary only in three dimensions: quantity, maturity and price. Our theoretical discussion provides a hint as to where maturity may serve as a screening device when contracts are more complex (e.g., mortgages): in markets where borrowers' private information is more informative about default risk in longer horizons.[12] Our paper provides the first empirical evidence of the existence of such a time structure of borrower private information, which is an essential ingredient in recent theories of debt financing under asymmetric information (see, for example, Goswami, Noe, and Rebello (1995) and Milbradt and Oehmke (2014)).

The rest of this paper proceeds as follows. Section I describes the LC platform and the data, as well as the expansion of the supply of long-maturity loans. In Section II we describe our empirical strategy and document that borrowers who self-select into long-maturity loans exhibit a higher propensity to default on the short-term loan. In Section III we evaluate what is the specific private information that is driving selection. Section IV provides a framework to develop a testable condition under which it is optimal to screen borrowers using loan maturity, and shows evidence for this condition in our data. Section V concludes.

[11] In another example, Jimenez, Salas, and Saurina (2006) show that firms who post more collateral ex ante are less likely to default ex post, conditional on observables. As with maturity, this relationship cannot isolate the screening role of collateral, because collateral is likely to impact default probabilities even in the absence of any selection. Similar to the empirical literature on maturity, the existing evidence on the role of collateral in alleviating problems stemming from asymmetric information is limited to showing how collateral correlates with ex ante measures of observable creditworthiness or proxies for asymmetric information. For examples of the first see Leeth and Scott (1989), Berger and Udell (1990) Booth (1992), Degryse and Van Cayseele (2000) and for the second see Berger and Udell (1995) and Berger, Espinosa-Vega, Frame, and Miller (2011). None of these papers establishes that for observationally equivalent borrowers, collateral choice varies depending on their own private information.

[12] For stylized evidence on maturity choice outside of unsecured consumer finance see, for example, Khandani, Lo, and Merton (2013) in mortgage markets, Gottesman and Roberts (2004) in syndicated loan markets, and Berger, Espinosa-Vega, Frame, and Miller (2005) in bank debt markets.

I. Setting

A. Lending Club

LC is the largest online lending platform in the U.S., operating in 45 states and originating $4.4B in consumer loans in 2014 alone. By comparison its nearest rival, Prosper Marketplace, originated $1.6B in the same year.[13] LC loans are unsecured amortizing loans for amounts between $1,000 and $35,000 (in $25 intervals).[14] LC loans are available in two maturities: 36 months, which are available for all amounts, and 60 months, which are available for different amounts at different points in time. Loans are funded directly by institutional and retail investors (LC holds no financial stake in the loans), and 80% of the total funds are provided by institutional investors (Morse (2015)). Since each loan is considered an individual security by the Securities and Exchange Commission, the agency that regulates online loan marketplaces in the U.S., LC is required to reveal publicly all the information used to evaluate the risk of each loan. This is an ideal institutional setting for the purposes of studying adverse selection, since we observe all the borrower information that the lenders and investors observe at the time of origination.

When a borrower applies for a loan with LC, she first enters her yearly individual income and sufficient personal information to allow LC to obtain the borrower's credit report. In most cases (e.g., 71% of all loans issued in 2013) LC verifies the yearly income that a borrower enters using pay stubs, W2 tax records, or by calling the employer. Every loan application is processed in two steps. First, LC decides whether or not a borrower is eligible for a loan on the platform. The eligibility decision is made mechanically based purely on hard borrower information observable at the time of origination. For example, during 2013 LC only issued loans to borrowers with a FICO score over 660, non-mortgage debt to income ratio below 35%, and credit history of at least 36 months. If LC determines that a borrower is eligible for a loan in the first step, she is then assigned to one of 25 risk categories (labeled by LC as risk "subgrades"). This assignment is made using a proprietary credit risk assessment algorithm that uses the hard information in a borrower's credit report (e.g., FICO score, outstanding debt, repayment status) and income. The assignment to risk category is made *prior* to the borrower selecting a loan amount or maturity and is therefore independent of both choices. The risk category determines the entire menu of interest rates faced by the borrower, for all loan amounts and for the two available maturities. That is, two borrowers assigned to the same risk category at the same time will face the same menu of interest rates for all amounts and for the two maturities. Interest rates for each subgrade are weakly increasing in amount and strictly increasing in maturity (ceteris paribus). The terms of all loans, other than interest rate, amount, and maturity, are identical. Once a borrower selects a loan from the menu, the the loan is listed on LC's

[13]Figures reported in the firms' 2014 10K reports.
[14]The maximum loan amount was increased to $40,000 after our sample period ended.

website for investors' consideration. Investors cannot affect any of the terms of the loan: they only decide whether or not to fund it. According to LC, over 99% of all listed loans are funded.[15] Thus, we ignore the supply side of funds in the analysis. As of 2013, LC charges an origination fee that varies between 1.1% and 5% of the loan amount depending on credit score, which is subtracted at origination, and a further 1% fee from all loan payments made to investors.

B. Staggered expansion of 60 month loans

Before March 2013, 60-month loans were only available for loans of $16,000 and above. A borrower could not synthetically create at 60-month loan for a smaller amount using prepayment, because prepayment reduces the number of installments without changing their amount, effectively reducing the maturity of the loan. In March 2013 LC introduced to the menu 60-month loans between $12,000 and $16,000. And in July 2013, it further expanded the available 60-month loans to include amounts between $10,000 and $12,000. The consequences of the menu expansion can be seen in Figure 2, where we plot the fraction of loans originated every month that have a 60-month maturity, by loan size groups. On December 2012, the first month of the analysis sample period, around 40% of loans between $16,000 and $20,000 are 60-month loans. This fraction remains relatively constant throughout the sample period, until October 2013. The fraction of 60-month loans is zero for loan amounts below $16,000 in December 2012, and jumps up for $12,000 to $16,000 loans in March 2013, and then for $10,000 to $12,000 loans on July 2013. By the end of the sample the fraction of 60-month loans stabilizes at around 30% for $12,000 to $16,000 loans and around 25% for $10,000 to $12,000 loans. The fraction of 60-month $5,000 to $10,000 loans remains at zero throughout the sample period. As we discuss in detail in Section II, our empirical strategy exploits the fact that loan amounts between $10,000 and $16,000 were affected by the expansion of a long maturity option, and that loan amounts outside this range were not.

C. Summary statistics

LC makes publicly available in its website all the information used to assign borrowers to risk categories, the assigned risk category, and the loan performance of all funded loans. Our main analysis is conducted using data downloaded as of April 2015. The data is a cross section of all loans originated at LC. Variables are measured either at the time of origination (e.g. date of loan, loan terms, borrower income and credit report data, state of residence) or at the time of the performance data download (e.g. loan status, time of last payment, current FICO score of borrower). We complement our main outcomes, which are measured as of April 2015, with measures of FICO

[15] See http://kb.lendingclub.com/borrower/articles/Borrower/What-if-my-loan-isn-t-fully-funded-when-my-listing-ends/?l=en_US&fs=RelatedArticle.

score obtained from two previous loan performance updates, August 2014 and December 2014.[16]

We use the origination date of each loan to restrict the sample period of the analysis to meet two criteria: 1) that it contains the dates in which the 60-month loan menu was expanded (March 2013 and June 2013) and that are the basis of our empirical analysis, and 2) that the interest rate assigned to each amount-maturity combination remained constant within each risk category (in other words, that all menu options other than the added long-term option remained constant). Thus, the beginning and ending months of our analysis sample are determined by two dates, surrounding the menu expansion events, on which we observe that LC repriced menu options (December 2012 and October 2013). We verify empirically that the interest rates of all risk category-amount pairs for 36-month loans are unchanged between these dates.[17]

We further limit the sample of loans to include those for amounts between \$5,000 (closed) and \$20,000 (open) because the interest rate schedule jumps discretely at \$5,000 and \$20,000 for all credit risk categories.[18] This interval includes all 36-month loans issued at amounts affected by the 60-month borrowing threshold reduction (\$10,000 to \$16,000), as well as amounts above and below this interval that allow us to control for any time-of-origination changes in unobserved borrower creditworthiness or credit demand. Finally, we further limit our sample to those loans where we can uniquely match the loan that a borrower chose to the menu associated with the risk category she was assigned to based on her publicly available data. We obtain this unique match for 98.6% of all loans in the sample period (we drop observations for which this matching does not yield a unique value). Our final sample has 60,514 loans.[19]

Table 1, Panel A, presents summary statistics for the subset of our sample corresponding to the 12,091 36-month loans issued by LC before the first menu expansion, that is, 36-month loans with amounts between \$5,000 and \$20,000 issued between December 2012 and February 2013. On average, loans for this sub-sample have a 16.3% APR and a monthly installment of \$380. Borrowers self report that 87% of all loans were issued to refinance existing debt (this includes "credit card"

[16] This allows us to estimate a measure of time-series volatility of FICO score for each individual.

[17] The exact dates correspond to loans listed as of December 4, 2012 and October 25, 2013. Even though we refer to months as the borders of the interval, all our analysis consider these two dates as the starting and end points of the sample period, respectively. We verify empirically that the interest rates of all risk category-loan quantities pairs are unchanged over this period. For example, Figure 12 in the Internet Appendix shows supply schedules (rate versus amount) before and after the expansion of the menu of borrowing options for borrowers assigned to risk categories B1 through B5: the graphs are identical. We establish the same point in general in a tractable way in Appendix D by regressing the interest rate of all 36 month loans in our sample on fixed effects for loan amount by risk category. The regression yields an R^2 of 99.7%, which confirms that the pricing of each menu was constant throughout the sample period for all 25 risk categories.

[18] We exclude loans whose "policy code" variable equals 2 , which have no publicly available information and according to the LC Data Dictionary are "new products not publicly available". In robustness tests, we limit the sample to loan amounts between \$6,000 and \$19,000, a \$1,000 narrower interval. Also, in some placebo tests we shift our sample to loans issued between July 2013 and May 2014.

[19] See Appendix C for details on this reverse-engineering procedure. The error in matching loans to their sub-grade does not vary systematically over the same period or by loan amount.

and "debt consolidation"). We define a loan to be in default if it is late by more than 120 days. According to this definition, 9.2% of the loans in the sub-sample are in default as of April 2015. Figure 3 shows the default hazard rate by months-since-origination for loans issued before the menu expansion.[20] The hazard rate exhibits the typical hump shape and peaks between 13 and 15 months.

Table 1, Panel B, shows borrower-level statistics of this sample. On average, LC borrowers in our sample have an annual income of $65,745 and use 17.4% of their monthly income to pay debts excluding mortgages. The average FICO score at origination is 695, and credit report pulls show that the FICO score has on average decreased to 685 approximately one year later. LC borrowers have access to credit markets: 56% report that they own a house or have an outstanding mortgage. The average borrower has $38,153 in debt excluding mortgage debt and $14,549 in revolving debt, which represents a 61% revolving line utilization rate (the average revolving credit limit is $27,464). LC borrowers have on average approximately 15 years of credit history.

To obtain a sense on how representative the LC borrowers are of the average US consumer credit user in the same FICO range, we compare our summary statistics to the credit card user statistics from Agarwal, Chomsisengphet, Mahoney, and Strobel (2015). Using the average credit card limit in the subsample of borrowers with FICO scores between 660 and 719 ($7,781) and assuming the average number of credit cards held by the average card-holder is 3.7 (according to Gallup 2014 survey) implies that the representative U.S. user of consumer credit has a revolving credit limit of $28,789, very close to the $27,464 average revolving credit limit of the LC borrowers in our sample. Thus, LC's selection criteria imply that the analysis sample is drawn exclusively from prime U.S. consumer credit users (as measured by FICO scores), but LC borrowers do not seem to be different in their revolving credit availability to the average U.S. consumer credit user in the same FICO range.

II. Measuring Selection On Maturity

We exploit the staggered menu expansion of 60-month loans during 2013 to identify adverse selection along maturity. As prescribed in the ideal experiment, LC offered new loan options at longer maturities for amounts already offered on short-term contracts prior to the expansion. Crucially, the pricing of all loan options available prior to the expansion was unchanged after the expansion for all 25 risk categories during our sample period. This ensures that the only difference in the menu of borrowing options offered to borrowers assigned to the same risk category before and after the expansion is the availability of 60-month loans in lower amounts.

As per the logic of the ideal experiment, our goal is to compare the outcomes of borrowers who took the short-term loan before the menu expanded (group A in our idealized experiment) with

[20] The date of default is determined by the last payment date, a variable that is available in the LC data.

those that were assigned to the same risk category and took it after (group B). We develop a research design that accounts for any other changes over time in the composition of borrowers within a risk category that are not driven by the menu expansion. The LC setting provides two sources of variation that allow constructing a counterfactual using a difference-in-differences approach: 1) the menu expansion was staggered over time for different loan amounts (*eventually-selected* amounts), 2) some loan amounts were never affected by the menu expansion (*never-selected* and *always-selected* amounts). The three groups of loans defined this way by the loan amount and the time of origination are represented in Figure 5.

Loans of amounts between \$10,000 and \$16,000 are eventually-selected, in the sense that they are unselected at the beginning of the sample (no long-term option available at the time of origination) and selected (long-term option available) at the end of the sample. Since the menu expansion was staggered, loan amounts between \$10,000 and \$12,000 serve as a control group for loan amounts between \$12,000 and \$16,000 that were affected by the March expansion and the reverse applies for the expansion in July. We build two additional control groups with loan amounts whose selection status was not affected by the menu expansion. The always-selected, for which the long-term loan was always available at the time of origination during the sample period (\$16,000 to \$20,000), and the never-selected, for which the long-term option never became available (\$5,000 to \$10,000). Our identification assumption is that any change in the composition of borrowers within a risk category, for example, due to changes in the economic environment, changes in the borrowing options outside of LC, or changes in how LC assigns borrowers to risk categories, does not affect differentially borrowers opting to take loans between \$12,000 and \$16,000 in March and borrowers opting to take loans between \$10,000 and \$12,000 in July relative to loans issued at control amounts. Under this assumption, comparing the change in performance of eventually-selected amounts before and after the menu expansion at those amounts with the change in performance of the control amounts in the same risk category isolates the effect of the maturity selection induced by the menu expansion. We further include a comprehensive set of granular borrower controls, which ensures that the estimations come from comparing borrowers who took loans at selected amounts to *observationally equivalent* borrowers taking loans at non-selected amounts.

Before providing evidence to support the identification assumption (see section II.C below), we discuss here its plausibility. First, even though it is unlikely that changes in economic conditions may have affected the demand for loans between \$10,000 and \$16,000 exactly at the same month of the menu expansion, to check whether there were any aggregate changes in the demand for LC loans we plot in figure 4 the total dollar amount of LC loans issued by month. There is no indication that the growth rate of LC lending changed around the dates of the two 60-month loan expansions. Second, in web searches we found no evidence of a change in the outside borrowing options that *exclusively* targeted the eventually-selected loan amounts (\$10,000 to \$16,000) in a manner that corresponds with the staggered expansion of the menu. Third, we found no evidence that LC released advertisement targeted at 60-month loans between \$10,000 to \$16,000 during the analysis sample. On the contrary, according to the information reported in the website Internet Archive, LC continued to advertise that 60-months loans were available only for amounts above \$16,000 until November 2013, after our analysis period ends. Fourth, any change in LC's screening process or assignment to risk categories cannot, by construction, affect borrower selection across different amounts within a risk category. The reason is that both eligibility for an LC loan and the assignment to risk categories are determined using borrowers' observable information *before* the borrower selects a loan amount from the menu. Nevertheless, we verify that the criteria used to determine eligibility to a LC loan (the minimum FICO score of 660, minimum credit history length of 36 months, and maximum non-mortgage debt to income threshold of 35%) remain constant over the sample period.

It is important to emphasize why our estimates rely exclusively on a comparison of 36-month loans taken before and after the expansion, and ignore any changes in the composition of borrowers that take 60-month loans. There is no appropriate counterfactual for borrower selection on the 60-month loans. The mix of borrowers taking a 60-month loan could have changed, for example, because some borrowers that take the 60-month loan would have not borrowed at all before this option became available in the menu. Since we are unable to account for such selection on the extensive margin for 60-month loans, we are limited in how much we can infer about the determinants of the performance of the 60-month loans. The focus on 36-month loans also implies that our approach for measuring the effect of selection is based on a revealed-preference argument, which relies on the axiom of independence of irrelevant alternatives. Specifically, we assume that a borrower who prefers not to borrow from LC over taking a 36-month loan when there is no 60-month option available, will not prefer to take the 36-month loan once the 60-month loan becomes available.

Finally, we note that the empirical approach is aimed at estimating the effect of selection on maturity in LC loans. If LC borrowers have access to 60-month loans between \$10,000 and \$16,000 at a similar price elsewhere during the analysis period, we should fail to reject the null hypothesis and conclude that there is no adverse selection on maturity in LC (since borrowers who wish to select long-term loans would already be taking them elsewhere). In effect, any impact of the menu expansion at LC can also be interpreted as indirect evidence that consumer credit markets are imperfectly competitive. This might be true because some intermediaries have a technology advantage over others that generates some market power or because there are search frictions in the market.[21]

A. Evidence of Selection

We start by measuring the amount of selection induced by the menu expansion: how does the number of borrowers who take the short-term loan at any given amount change after the long-term option

[21] For evidence of search frictions in consumer credit markets see Stango and Zinman (2013).

becomes available at that amount. To do so we collapse the data and count the number of loans N_{jkt} at the month of origination $(t)\times$risk category $(j)\times$\$1,000 loan amount bin (k) level for all 36-month loans issued during our sample period (amount bins measured starting from \$10,000, e.g. \$10,000 to \$11,000, \$11,000 to \$12,000, etc).[22] We define a "selected" dummy variable D_{kt} equal to one for those loan amount bin-month pairs where a 60-month option was available, and zero otherwise. That is:

$$D_{kt} = \begin{cases} 1 & if\ \$16,000 > LoanAmounts \geq \$12,000\ \&\ t \geq March\,2013 \\ 1 & if\ \$12,000 > LoanAmounts \geq \$10,000\ \&\ t \geq July\,2013 \\ 0 & otherwise \end{cases}$$

Then we estimate the following difference-in-differences regression:

$$log(N_{jkt}) = \beta'_k + \delta'_{jt} + \gamma' \times D_{kt} + \varepsilon_{jkt}. \tag{1}$$

The coefficient of interest is γ', the average percent change in the number of short-maturity loans originated for eventually-selected amounts (i.e., amounts in which a long-maturity loan was not available at the beginning of the sample and became available due to the menu expansion) relative to control amounts. We include amount bin fixed effects β'_k, which control for level differences in the number of loans in each \$1,000 bin. In turn, risk category$\times$month fixed effects δ'_{jt} control for any changes over time in the number of borrowers who are approved at each of the 25 different risk categories.

Table 2, column 1, shows the results of regression (1), estimated on the full sample of borrowers who took a 36-month loan between \$5,000 and \$20,000 during the sample period (December 2012 to October 2013). The point estimate of γ' is negative and significant, and implies that the number of borrowers who took a short-term loan is 14.5% lower once the new long-term loan option for the same amount becomes available. This estimate provides us with a magnitude for the number of borrowers who would have taken a short-term loan if the long term option had not been available.[23]

B. Selection and Repayment

Having shown that the expansion of the menu of borrowing options induced a significant amount of self-selection from short-term to long-term loans, we run our main test to uncover the unobserved quality of the borrowers who selected into the new long term contract. We estimate the following

[22] Results are insensitive to using actual loan amount instead.

[23] Standard errors for estimates of equation (1) are robust to heteroskedasticity, but other alternatives, e.g., clustering in any dimension, are irrelevant in terms of statistical significance. For example, when clustering at the risk category level (25 clusters), the standard error of the coefficient γ' in Column 1 of Table 2 is 0.028.

difference-in-differences specification on the sample of 36-month loans:

$$Default_i = \beta_i^{1000bin} + \delta_i^{jt} + \gamma \times D_i + X_i + \varepsilon_i, \tag{2}$$

where data is at the loan level i. The outcome variable, $Default_i$, is defined as a dummy that equals one if the loan is late by more than 120 days measured as of April 2015.[24] Standard errors are clustered at the state level (45 clusters).

The main explanatory variable of interest, D_i, is a dummy equal to one if the 36-month loan i is issued at a time when a 60-month loan of the same amount is also available, and zero otherwise:

$$D_i = \begin{cases} 1 & if\ \$16,000 > LoanAmount_i \geq \$12,000\ \&\ t \geq March\,2013 \\ 1 & if\ \$12,000 > LoanAmount_i \geq \$10,000\ \&\ t \geq July\,2013 \\ 0 & otherwise \end{cases}$$

The coefficient of interest, γ, measures the change in the default rate of 36-month loans for eventually-selected amounts before and after the expansion of the menu options, relative to the change of the default rate for never-selected and always-selected amounts, which were not affected by the menu expansion. We include granular month of origination $t \times$risk category j fixed effects, δ_i^{jt}, which ensure we compare borrowers who took a loan on the same month with the same contract terms and with similar observed measures of credit risk (same risk category). We also include a vector of control variables observable at origination, X_i. In our baseline specification, X_i includes 4-FICO score-at-origination bin and state fixed effects. The rich set of fixed effects ensures that we perform the difference-in-differences estimation by comparing borrowers that are observationally equivalent. We also report results including as controls the *full* set of variables that LC reports and that investors observe at origination. These variables (61 in total) include, annual income, a dummy for home ownership, stated purpose of the loan, length of employment, length of credit history, total debt balance excluding mortgage, revolving balance, and monthly debt payments to income.

Table 3, columns 1 and 2, reports results of regression (2). The negative point estimate for γ indicates that borrowers who take a 36-month loan once a 60-month option is available are significantly less likely to default than borrowers that take the same 36-month loan when the long term option is not available. The point estimate of -0.0081 means that the default rate of the borrowers that are selected on maturity is 0.8 percentage points lower than the default rate of the non-selected borrowers (column 1), and the magnitude is unchanged when we include as additional controls every single variable observable at origination in LC's dataset (column 2). The fact that our estimate is virtually unaffected by including this full suite of additional controls demonstrates that

[24] We also define a borrower to be in default if she is reported in a "payment plan". Our results are robust to not including these borrowers as in default.

the granular fixed-effect structure in our baseline regression is sufficiently comprehensive to absorb any changes in the composition of observed borrower characteristics.

This decline in the default probability is due to the borrowers that self-select into the long-term loan, which we estimated to be 14.5% of the borrowers in the not-selected sample (Table 2, column 1). Combining the two results allows us to obtain an estimate of the default probability of the borrowers that self-selected into the 60-month loan: it is $0.8\%/14.5\% = 5.5\%$ higher than those who self-select into the 36-month loan when the long-term loan is available (significant at a 10% level, based on bootstrapping with 1,000 repetitions). This is an estimate of the counterfactual probability we are after: it is the default rate that the borrowers who self-select into the 60-month loan would have had if they had taken the 36-month loan. The economic magnitude of this difference is large compared to the average default rate of 36-month loans issued before the menu expansion, between December 2012 and February 2013, is 9.2% (Table 1). The comparison implies that amongst observationally equivalent borrowers, those that self-select into a long maturity contract are 59% more likely to default than those borrowers that self-select into the short term contract, ceteris paribus (e.g., holding constant the contract characteristics).

The results suggest that maturity choice reveals unobserved heterogeneity that cannot be priced in by lenders. The lower default rate of borrowers who self-select into a short-maturity loan cannot be predicted by variables available to investors at the time of origination, as attested by the comparison between the estimates with and without controls for observables. Although we do not control for the exact FICO score but for scores *within* each 4-point FICO bin, the predictive power of FICO on default in our sample is too small for selection within 4-point FICO bins to account for our results. Indeed, a regression of $default_i$ on the high end of the FICO 4-point range at origination, including risk category by $1,000 amount bin by month fixed effects, gives a coefficient of -0.0000362. That is, a 1 point increase in FICO score at origination is correlated with a 0.004% decline in default rate, not statistically significant. Thus, variation in default rates within FICO score bins can at most account for a 0.012% difference in default rates (0.004% $\times 3$), quantitatively irrelevant next to our estimated effect of 0.8% reduction in default.

C. Identification Tests

C.1. Evidence to Support the Identifying Assumption

Our empirical strategy rests on the identifying assumption that there were no changes in unobserved borrower creditworthiness that differentially impacted borrowers taking loans between $12,000 and $16,000 in March and borrowers taking loans between $10,000 and $12,000 in July. One potential concern which could threaten this assumption is the possibility that there is a gradual shift in the composition of borrowers over 2013 that approximately matched the pattern of the staggered expansion. We test for this possibility by running an amended version of (1) using a series of dummies that become active τ months after a 60-month loan is offered at each amount. Formally, we define:

$$D(\tau)_{kt} = \begin{cases} 1 & if\ \$16,000 > Loan Amount \geq \$12,000\ \&\ t = March 2013 + \tau \\ 1 & if\ \$12,000 > Loan Amount \geq \$10,000\ \&\ t = July 2013 + \tau \\ 0 & otherwise \end{cases},$$

and we run the following regression:[25]

$$log(N_{jkt}) = \beta_k + \delta_{jt} + \sum_{\tau=-3}^{3} \gamma_\tau \times D(\tau)_{kt} + \varepsilon_{jkt}. \tag{3}$$

Figure 7 shows the results of regression 3. The results show no differential pre-trends in the three months leading up to the expansion and then show a discontinuous fall in the number of loans made in these amounts exactly at the time of the expansion. This rules out that our results are coming from pre-existing trends in borrower demand or composition unrelated to the menu expansion.

To further ensure that our results are not driven by differential trends in the demand for loans of varying amounts, we run regression (1) on a sample shifted forward to start when the 60-month loan option is available for any amount above \$10,000 (after the expansion in menus is complete). That is, we shift the definition of D_{kt} forward by 7 months and run the regression on the sample of loans originated between July 2013 and May 2014. Column 5 of Table 2 shows the results. The coefficient on D_{kt} equals -4.4% and is insignificant, and given the confidence interval we can reject the null that this coefficient equals our main estimate.

C.2. Simultaneous Choice of Maturity and Loan Amount

A second identifying assumption behind our empirical approach is that the loan amount choice is sufficiently inelastic to loan maturity in this setting. If this is not the case, the difference-in-difference estimate will be biased towards zero. This could be the case because borrowers in what we classify as eventually-selected amounts may be already selected on maturity before the menu expansion (e.g. if some borrowers who wanted to take a long-term loan at a treated amount before the expansion took a long-maturity loan at larger amount instead) or because borrowers in what we classify as never-selected amounts may be a selected group after the menu expansion (e.g., because some borrowers who wanted to take a long-term loan at a control amount after the menu expansion took a long-maturity loan at a treated amount instead).

[25] The final 60 month threshold reduction takes place in July 2013, which leaves three more months in our sample period up to October 2013. Similarly, the first 60 month threshold reduction occurs in March 2013, which leaves three months in the pre-period (from December 2012).

Let's consider first the possibility that eventually-selected amounts are selected before the menu expansion. As an example, consider borrowers who would like to take a \$10,000 60-month loan. Before the menu expansion this option is not available, and the closest alternatives are: 1) a \$10,000 36-month loan, and 2) a \$16,000 60-month loan.[26] Our empirical strategy will estimate the effect of maturity on selection if borrowers choose the first option, e.g. take a loan for the amount they prefer at a shorter maturity—36-months— when the 60-month option is not available. The reason is that these borrowers select out of the 36-month loan when the 60-month option is available, after the menu expansion.[27] If, on the contrary, borrowers choose the second option, e.g. take a 60-month maturity loan but for a larger amount, then our difference-in-differences estimate will be zero. Indeed, these borrowers will not be in the eventually-selected group of loans before or after the expansion because our estimation is based exclusively on the outcomes of 36-month loans. Thus, selection from one long-term loan to another will not affect our estimates.[28]

Now consider the second case: where the never-selected amounts are treated after the menu expansion. Take for example borrowers that would like to take a \$5,000 60-month loan, but since this option is not available before the menu expansion, they take a \$5,000 36-month loan instead. Although these borrowers are in the control group in our estimation, it is possible that they choose a \$10,000 60-month loan when this option becomes available in the menu. If this is the case, then the menu expansion will also cause self-selection into long maturity among the control group of loans, and the comparison between eventually-selected and control loans will be biased towards zero.

We investigate formally whether eventually-selected loans amounts were affected prior to the expansion or if control loan amounts were impacted after the expansion. To do this we exploit the same set-up as regression 1, which measures the change in the number of short-term loans issued at eventually-selected and control amounts, before and after the menu expansion, and compare the evolution of the number of 60-month (36-month) loans in the \$16,000 to \$20,000 (\$5,000 and \$10,000) range relative to the evolution of the number of loans in the \$20,000 to \$24,000 (\$1,000 and \$5,000) range around the menu expansion. We estimate the same difference-in-differences regressions with a modified definition of the "selected" dummy D_{kt} to equal 1 one after March 2013 or July 2013 for different loan amounts according to the timing of the menu expansion, as explained below.

[26]These borrowers may also choose not to borrow at all when their preferred option is not available in the menu, and take the 60-month loans when it becomes available. This extensive margin will not affect our estimates, since our results are based exclusively on the behavior of 36-month loans before and after the menu expansion.

[27]One way to test for whether borrowers at control amounts are selected before the menu expansion is to look for evidence of bunching at the borders of the treated interval. The top panel in Figure 6 presents the pre-period loan amount histogram at the short maturity. The histogram suggests that borrowers choose "round" numbers like \$10,000 and \$12,000 much more frequently than other intermediate amounts. In turn, this makes it very hard to find evidence of bunching at specific amounts.

[28]The bottom panel in Figure 6 presents the pre-period loan amount histogram at the long maturity. The histogram has the same pattern as the top panel. Evidence of bunching is, again, very hard to establish because of borrower's preference for round numbers.

First, using the sub-sample of 60-month loans in amounts between \$16,000 and \$24,000, we define D_{kt} to be equal to one after March 2013 for all amounts between \$16,000 and \$20,000. The coefficient on this dummy tells us whether the number of loans of amounts close to the \$16,000 expansion threshold declined relative to those farther from the threshold. If so, it would be an indication that eventually-selected loan amounts experienced selection to 60 month loans prior to the expansion. The coefficient on the interaction term is -8.25% and is not significantly different from zero (Table 2, column 2). This suggests weak evidence that our estimates may understate the degree of selection because some borrowers in eventually-selected amounts may have opted for 60-month loans above \$16,000 prior to the expansion.[29]

We repeat the exercise at the \$10,000 amount threshold using 36-month loans. We restrict the analysis to the sample of loan amounts between \$1,000 and \$10,000, and define D_{kt} equal one after July 2013 for amounts between \$5,000 and \$10,000 and zero otherwise. The coefficient on the interaction term is -3.6% and, again, not significantly different from zero (column 4). Thus, there is no evidence that borrowers who in the pre-period selected a short-maturity loan below \$10,000 would have taken a larger long-maturity loan above the \$10,000 threshold when they became available in July. In other words, we find no evidence that the control group of loans in our main empirical design were affected by the menu expansion. Taken together the results in Table 2 confirm our conjecture that the bulk of any selection to longer maturity loans induced by the expansion of the menu was in the eventually-selected amounts.

D. Robustness

We present in Table 4 several tests that demonstrate the robustness of our results. First, column 1 of Table 4 presents a counterpart to our main result in column 1 of Table 2 but limiting the sample to loan amounts between \$6,000 and \$19,000 (a \$1,000 narrower window than our main sample, which uses loans from \$5,000 to \$20,000). The results are qualitatively similar, although the estimate is noisier and significant only at a 10% level.

As we mentioned above when describing our empirical strategy, the expansion in the menu of borrowing options may have induced selection in the unaffected or control group of amounts, above and below the \$10,000 to \$16,000 interval. In Table 2 above we show that the number of loans issued at the control amounts did not change, which suggests that no such selection occurred. However, it is important to independently verify that there is no change in the credit quality of loans issued at control amounts induced by the menu expansion. Here we test for this possibility. Column 3 of

[29] In an analogous test we also check whether borrowers of 36-month loans in control amounts above the \$16,000 threshold were affected by the expansion. The coefficient on the interaction term is 5.9% and is not significantly different from zero (Table 2, column 3) indicating that the expansion of the menu did not induce selection away from short-term loans above \$16,000. Given that long-maturity loans were always available for these amounts, this is not a surprising result.

Table 4 restricts the sample to loans issued between December 2012 and October 2013, between \$16,000 and \$24,000. The independent variable of interest equals one for loans between \$16,000 and \$20,000 after March 2013. The coefficient is positive and insignificant. Column 4 of Table 4 repeats the exercise for loans between \$1,000 and \$10,000 issued between December 2012 and October 2013. Here, the independent variable of interest equals one for loans between \$5,000 and \$10,000 issued after July 2013. Here, the coefficient is negative and insignificant. In both cases, we find no significant differences in the default rate of loans issued at amounts bordering the interval of eventually-selected amounts. The results in column 2 and 3 of Table 4 also serve as placebo tests and confirm that our results are not spuriously driven by shifting creditworthiness at different loan amounts. Overall, these tests point to a robust conclusion: borrowers who self-select into long-maturity loans are unobservably more likely to default, holding the loan contract characteristics constant.

III. Interpretation: Private Information About What?

A. Time Structure of Private Information

So far, our empirical results show that borrowers who select into longer maturity loans are privately informed about their increased propensity to default on a short-term loan. We turn to understanding what is the specific private information that borrowers are selecting on. Since maturity provides insurance against future changes in the price of credit, then the private information must relate to how borrowers value this protection. It is theoretically possible that borrowers who are privately informed about their own high degree of risk aversion select into longer maturity loans (De Meza and Webb (2001)). However, selection on risk aversion is inconsistent with the higher default rate that these borrowers exhibit.

It follows that borrowers who select long-maturity loans privately place higher value on the insurance it provides either because: 1) they have a high exposure to future shocks to their observable creditworthiness (e.g., probability of a job loss or illness), or 2) they have a higher exposure to rollover risk due to privately observed differences in the timing of their income (the cash flow timing hypothesis). These two explanations differ in the horizon after origination at which borrowers become risky. Borrowers that self-select into long maturity because they are more exposed to shocks to their observable ability to repay will tend to default more the longer the horizon after origination, as the negative shocks realize. In contrast, borrowers that self-select into long maturity because their income arrives later will tend to default less with time after origination, as their income realizes. Therefore, the two selection mechanisms can be distinguished by their predictions of the time structure of default implied by the private information.

We exploit that fact that we observe when a borrower in our sample enters default to differentiate between these two accounts. To do this, we redefine our baseline measure of default and create two variables for default at different horizons: borrowers who missed their first payment within the first 12 and 24 months of loan origination (for loans that are 120 days past due in April 2015). We label these variables $default12m$ and $default24m$ respectively and use them as dependent variables in regressions that are otherwise identical to the one we estimated in column 1 of Table 3. The results are presented in columns 1 and 2 of Table 5. Column 1 shows that borrowers who self-select into long-term loans have no differential propensity to default within the first year of the loan. Since the hazard rate of default in our sample peaks at 13 months (Figure 3), this result is not mechanically driven by lack of statistical power due to a low frequency of default early in the life of the average loan (unconditionally, loans are as likely to default in the first 12 months after origination than later). Column 2 shows that the differential propensity to default is present at the 24 month horizon from origination.

We present in Figure 8 the coefficients from estimating our main specification using as the dependent variable an indicator for whether the first missed payment occurred before 1, 2, and so on, up to 24, months after origination.[30] The figure indicates that the cumulative default probability differential between the two groups of borrowers increases linearly with the months after origination. That is, borrowers who select into the 60-month loan have a propensity to default on the 36-month loan that is *increasing* in the time since origination of their loan.[31] This evidence indicates that the source of private information that is driving maturity selection is a borrower's exposure to shocks to their own future observable creditworthiness. Note that Figure 3 demonstrates that the hazard rate of default for 36 month loans peaks at 16 months.[32] This indicates that the bulk of default at either maturity occurs well before the 24 month horizon possible in our analysis, thereby ruling out the concern that our results are too near to origination to account for default behavior for either type of loan.

B. Private Information About Future Observable Creditworthiness

Our data allows us to provide additional evidence in support of our preferred interpretation. Specifically, we observe the realized observable creditworthiness measured by a credit score (FICO score) of each borrower in April 2015, roughly two years after origination. Table 5, columns 3 and 4 run our main regression model but replace the main outcome $Default_i$ with $FICO_i$, the borrower's

[30] At horizons of 19 months and further the sample used to run the regression is right censored because loans issued late in our sample do not have sufficient time to enter default at these horizons. This affects loans in the eventually-selected and control amounts in the same way and does not affect the identification strategy.

[31] The finding that information asymmetries grow with the horizon from origination is itself new and potentially important in its own right. For example this supports the assumed time structure of information asymmetry in Milbradt and Oehmke (2014).

[32] The hazard rate of default on 60-month loans issued at the same time is similarly shaped and peaks at 17 months. This indicates that repayment over the first 24 months of a loan is the crucial determinant of default at either maturity.

FICO score as reported in the latest LC data pull. In column 2 we include as controls all variables that are observable by investors at origination, as in column 2 of Table 3. The results imply a statistically significant increase of future FICO scores of approximately 2.7 points among selected short term borrowers relative to unselected ones. In economic terms this means that the average future FICO score of the 14.5% of borrowers who self-select into the long-maturity loans is $^{2.7}/_{14.5\%} = 18.6$ points lower than the average borrower that selects the 36-month loan.

To further demonstrate that borrowers are selecting maturity based on private information about their exposure to shocks to their future observable creditworthiness, we use the volatility of a borrower's future credit rating as a measure of the reclassification risk she is exposed to. If borrowers have private information about this reclassification risk, we expect borrowers that self-select into the 36-month loan to have less volatile future FICO scores. To test this hypothesis we present in column 1 of Table 5 the results of estimating our main specification using the within-individual standard deviation of the FICO score as the outcome variable, using FICO scores obtained from 4 different pulls of the LC loan performance data: at origination, as of August 2014, as of December 2014, and as of April 2015, which is the same outcome variable used in Table 2.[33] The cross sectional average and standard deviation of this measure for loans in our sample that were issued in the three preperiod months are 24.5 and 19.1, respectively.

The point estimate in column 5 of Table 5 is -0.57 and statistically significant at the 5% level. This implies that borrowers who select the 36-month loan have a future FICO score that is 2.3% (equal to 0.57/24.5) less volatile when the 60-month loan is available than when it is not. This pattern is strongly consistent with the insurance rationale behind adverse selection: borrowers who select long-maturity loans are (unobservably) more exposed to reclassification risk.

Note that $FICO_i$ measures a borrower's repayment status in *all* of their debts. In particular, it considers a borrower's performance not only on the 36-month loan with LC, but on loans of different maturities as well. Thus it is unlikely that this result is driven by the incompatibility between the short-term LC loan and the time profile of borrower's future income. Instead, this shows directly that borrowers who select long-maturity loans have private information that directly relates to shocks to their observed creditworthiness and the impact that this will have on the price at which future lending will occur.

Finally, we study how maturity choice relates to a borrower's unobserved propensity to prepay her loan prior to maturity. The LC data records loans that have been fully prepaid as of April 2015, which we code in *Prepayment*, a dummy variable. If borrowers are selecting maturity based on private information about the timing of their income, we would expect that those borrowers who select into a short-term loan would prepay at a higher rate than borrowers in an unselected group. If this were the case, the main coefficient in our regression model (2) where we replace the outcome

[33] The standard deviation is calculated as $sd(FICO_i) = \sqrt{\frac{1}{4} \times \sum_{t=1}^{4} \left(FICO_{i,t} - \overline{FICO_i}\right)^2}$.

variable $Default$ with *Prepayment* should be positive. We document the output of this regression in column 6 of Table 5. The point estimate is negative but insignificant (p-value is 0.44). Although this result is not conclusive, it does suggest that maturity choice does not seem to be driven by private information about the timing of borrowers' income shocks. It is difficult to believe that selection based on private information about the timing of income would simultaneously generate a statistically significant reduction in default but would produce a change in loan prepayment that is statistically undetectable, when the prediction about the *timing* of payment is most directly tied to the hypothesis itself.[34]

C. 60-month Loan Performance

Further evidence about the underlying private information that is driving maturity choice can be provided by looking at the default rate of borrowers who took 60-month loans. If, as we hypothesize, these borrowers are more exposed to shocks to their ability to repay then, after controlling for observables, the default rate should be higher at the longer maturity loans. In contrast, if borrowers are selecting to match the privately observed horizon of their income then the default rate should be no higher.

Before presenting this evidence, an important caveat that stems from our core empirical challenge is required. Our analysis has so far focused on the propensity to default holding the terms of the contract constant, that is, focusing exclusively on a sample of 36-month loans. Thus, our analysis tells us what the default probability of borrowers who self-select into 60-month loans *would have been* had they selected a 36-month loan. We cannot empirically identify what their default probability is for a a 60-month loan. This is because the default rate of 60-month loans is also driven by selection in the *extensive* margin: there are some borrowers who would have chosen not to take a loan at all in the absence of a 60-month option, but do so when it becomes available, and we cannot independently isolate the repayment propensity of these extensive margin borrowers.

Notwithstanding this problem, we can provide suggestive evidence by comparing the average default rate of 36-month and 60-month loans that have the same measured expected default risk (initial risk category and 4-point FICO score bin), issued the same month, and of the same size ($1,000 amount bin). The propensity to enter default by April 2015, which holds the repayment horizon equal across the two loan contracts, is 3% higher for the 60-month than for the 36-month loans. Commensurate with this increased risk, LC charges a 3.3% higher APR for 60-month loans, holding other borrower and loan characteristics constant. This provides further evidence that selection is based on private information about exposure to shocks to creditworthiness. If, alternatively, borrowers were selecting maturity based on the time horizon of their income rather than their future

[34]Our insurance based interpretation, that borrowers who are privately informed of their increased exposure to shocks to their ability to repay select into long-maturity loans, can explain the measured reduction in prepayment: positive realized shocks lead to early prepayment, negative shocks lead to default.

creditworthiness, then we should not expect to see higher default or interest rates at the longer maturity loan.

A different, but related, question is whether increased maturity impacts a borrower's propensity to repay a loan. The answer also hinges on the average ability to repay of borrowers who select to take 60-month loans on the extensive margin, which we cannot measure in our setting. If we make the stark assumption that their ability to repay is the same as borrowers who are selected away from the 36-month loan, then our results suggest that 2 more years of maturity reduces the propensity to default by 2.5% over the horizon for which we observe these loans.[35] If borrowers who take the 60-month loan on the extensive margin have a lower (higher) ability to repay then this will under (over) state the effect. This unmeasured margin could reconcile our results with Dobbie and Song (2015), who use a randomized experiment on US household credit card borrowers to show that increased maturity does not causally change a borrowers propensity to default or with Field, Pande, Papp, and Rigol (2013) who find that increased maturity induces entrepreneurs to undertake risky projects and leads to higher default.

D. Price Reaction to Selection

Our empirical analysis benefits from the natural experiment created by LC's decision to expand the availability of long-term loan contracts without changing any of the characteristics or terms of the short maturity contract. Note that this implies that, within the window of the natural experiment, the default probability of 36-month loans between \$10,000 and \$16,000 dropped while the interest rate did not change. If LC was earning a competitive return on these loans before the menu expansion, then it must have been earning rents after the expansion. In theory, competitive pressures should eventually drive the interest rate on the short-maturity loan down to reflect the lower risk of the borrowers that self-select into short maturity.

Indeed, after our analysis sample period (during which all lending terms were held constant), LC adjusted the APR of the 36-month loan in a way that is consistent with this conjecture. We show this in Figure 10 which plots the average APR charged to borrowers on 36-month loans in each month controlling for loan amount and borrower characteristics.[36] Consistent with our conjecture, we see that the APR fell by roughly 0.8% for short-term loans after long-term loans were added to the menu. This number is in the same order of magnitude to our estimate in Column 1 of Table 2 that showed the expected default rate of the 36-month loans fell by 0.8% as a result of the selection

[35] We obtain this number as 5.5%-3%=2.5%, where 5.5% is the excess default probability at the short-term loan of the 14.5% of borrowers who chose the long-term loan when it became available in the menu expansion, measured in Section II) and 3% is the average excess default rate of long-term loans over short-term loans.

[36] These characteristics are FICO score bin, annual income, and address state. Note that variation in APR before November 2013 in this graph is entirely accounted for by the fact that we do not control for the borrower initial risk category, which we cannot estimate after October 2013. This also implies that we are unable to simply compare the APR for the 36-month loan at each menu.

into long-maturity loans.

IV. When is Screening on Maturity Optimal?

Our results imply that maturity choice *can* be used to screen borrowers in consumer credit markets. The existing theoretical literature on corporate debt maturity choice under asymmetric information can rationalize this finding. In Flannery (1986), lenders can screen using maturity because bad credit risks only receive a temporary interest rate discount for pooling with good credit risks on short-term debt and are unwilling to incur the higher transaction costs needed to do so.[37] However, neither our results so far nor the extant theoretical literature can tell us *whether* maturity is the optimal screening device when there are alternative screening dimensions of the loan contract available, such as loan size. The purpose of this final section of the paper is to answer this question. We first develop a simple model of consumer finance to show that maturity can be used to screen borrowers when the value of long-term contracts is to provide insurance to risk averse households. The primary contribution of the model is that it allows lenders to screen borrowers using both loan maturity (as studied by Flannery (1986)) and loan amount (as per Schreft and Villamil (1992) and Brueckner (2000)). In unsecured consumer credit markets these are the two primary contract dimensions available to creditors. Our focus is to find when maturity will be, in equilibrium, the optimal way to screen borrowers as opposed to loan amount. We then compare this condition to the evidence provided in Section III.A on the time structure of private information, which, to the best of our knowledge, had not been measured before this paper.

A. Setup

The time-line of the model is shown in Figure 9. At $t = 1$ there is a continuum of observationally equivalent households. Borrowers wish to consume at $t = 1$ and $t = 3$ but have no income or wealth at $t = 1$.[38] Each household anticipates receiving risky income at $t = 3$ and this creates the desire to borrow in order to smooth consumption over time and between high and low income states. In the interim period $t = 2$, public information about a borrower's ability to repay is released in the form of a signal $S = \{G, M, B\}$ that indicates the probability a household will generate income at $t = 3$. A borrower for whom good news is released ($S = G$) will earn income of $I = E > 0$ with certainty. A borrower for whom intermediate news is released ($S = M$) has lower expected income–she will generate income of $I = E$ with probability $q \in (0, 1)$ and zero income otherwise. Finally, a borrower

[37] Diamond (1991) provides a theory of how maturity choice varies with a borrower's *observable* credit rating and Goswami, Noe, and Rebello (1995) shows how maturity choice will vary with the time-structure of asymmetric information. While both theories are based on borrower private information, in equilibrium all borrowers pool on identical contracts and hence no screening occurs in either paper.

[38] For simplicity we abstract from consumption at $t = 2$.

for whom bad news is released ($S = B$) will not generate any income at $t = 3$ with certainty. Income is not verifiable in court and therefore contracts cannot be made contingent on the realization of I.

Each borrower can be one of two types, high or low, indexed by $k \in \{H, L\}$. Let $\phi \in (0,1)$ be the fraction of borrowers who are the high type. A borrower's type determines the probability with which each signal is released and hence her exposure to adverse shocks to her ability to repay loans in the future. A borrower of type i will have intermediate news released at $t = 2$ with probability $p_k \in [0,1]$ and bad news released at $t = 2$ with probability $x_k \in [0,1]$ where $p_L \geq p_H$ and $x_L \geq x_H$. [39]

The supply of credit is perfectly competitive, the opportunity cost of funds is normalized to zero, and lenders are risk neutral. Lenders offer non-callable debt contracts for any amount and maturity within the three period model. Specifically, a debt contract at $t = 1$ will specify three quantities: $\left\{A_1, D_{1,2}, D_{1,3}\right\}$. $A_1 \geq 0$ is the amount received by the household at $t = 1$. $D_{1,2} \geq 0$ is the face value due at $t = 2$ and $D_{1,3} \geq 0$ is the face value due at $t = 3$. Since households do not have any income at $t = 2$ any amount due at this time must be paid out of savings or through additional borrowing. A loan made at $t = 2$ will specify two quantities: $\left\{A_2, D_{2,3}\right\}$ where $A_2 \geq 0$ is the amount received by the household at $t = 2$ and $D_{2,3} \geq 0$ is the face value due at $t = 3$ to repay this loan. The supply of loans at $t = 2$ is also perfectly competitive, and loan terms are set using the information contained in S. However we assume that, while lenders can observe S, the information it contains cannot be verified in court and hence loan contracts offered at $t = 1$ cannot be made continent on the signal.

When a loan payment is due a borrower can either make the payment or default. All loans are uncollateralized so a creditor is unable to seize any household assets upon default. We abstract from ex post moral hazard: in the event of default the household incurs a utility cost of $\Omega > 0$ that captures the inconvenience of being contacted by collection agents and the non-modeled reputation consequences of having default on the borrower's credit history. We assume that Ω is sufficiently high so as to rule out the incentive for strategic default–a borrower will repay using income or by taking on additional debt whenever possible–but small enough so as not to deter borrowing.

The objective of each household is to maximize $(1-\alpha)\,u(c_1) + \alpha u(c_3)$ where $u(c_t)$ is a strictly increasing and concave utility function and α captures the relative weight that a household places on consumption in each period. Consumption is not contractible or observable by lenders. A household will allocate the funds raised at $t = 1$ between consumption and savings. Because income is risky, households have an incentive to keep a buffer stock of savings to fund consumption when income is zero at $t = 3$. Savings are risk free and earn the opportunity cost of funds (here normalized to zero). If the household does generate income at $t = 3$ they repay all loans and consume their income and savings net of any debt payments. If the face value of debt due at $t = 3$ is greater than household income (which occurs when $I = 0$) then the household will default and consume any savings.

[39] Note that S is a sufficient statistic for estimating a borrowers expected income and probability of default. This assumption is not necessary for our results but it simplifies the analysis by eliminating the potential for additional screening at $t = 2$.

We assume that q is sufficiently high so that in equilibrium all payments due at $t = 2$ are repaid through new borrowing when $S = G, M$. Thus, $p_k(1-q)$ captures a borrower's private assessment at $t = 1$ of her probability of default at $t = 3$. Conversely, upon receiving the bad signal $S = B$, a household is unable to raise any new debt and hence must default at $t = 2$. Thus x_k captures a borrower's private assessment, at $t = 1$, of her probability of default at $t = 2$. The distinction between the two is central to the analysis which follows and its link to our empirical results.

B. Symmetric Information

Consider the benchmark case in which a household's type is known by all agents at $t = 1$. As we show in Section A of the Internet Appendix, in the unique equilibrium each household borrows the entire present value of their expected income using a long term-debt contract that requires no repayment at $t = 2$.[40] All households receive full insurance that is provided by the default feature of the debt contract.[41] Any debt contract that requires some repayment at $t = 2$ is unable to provide full insurance because the terms at which that payment is refinanced will be contingent upon the uncertain interim news about a borrowers ability to repay that is revealed at $t = 2$.

C. Asymmetric Information and Optimal Screening

Now suppose that a household's type is private information and so, from a lender's perspective, all borrowers are observationally equivalent at $t = 1$.[42] As is standard in screening models, low type households will be offered the same full insurance contract that maximizes their utility under symmetric information. The contract offered to high types will offer them the highest expected utility possible while ensuring that this loan is not chosen by low creditworthiness borrowers. The problem that characterizes the optimal lending contract offered to the high type is shown in Section B of the Internet Appendix. Characterizing the optimal contract analytically is in general infeasible due to the well known "hidden savings problem" (see Kocherlakota (2004a)). To deal with this, we first analytically solve a special case of the model in which all consumption occurs at $t = 3$ ($\alpha = 1$), and then we use numerical solutions to show that the findings are robust to a wider set of cases. Applying the logic of Rothschild and Stiglitz (1976), high types will be screened by choosing a contract that gives up some of the full insurance provided under symmetric information. Our focus is to show under what condition this is optimally achieved by rationing loan maturity as opposed to loan amount.

[40] In this equilibrium, each household of type k is offered a contract at $t = 1$ of $\left\{A_1^{**k}, D_{1,2}^{**}, D_{1,3}^{**}\right\}$ where $A_1^{**k} = E\left(1 - p_k(1-q) - x_k\right)$, $D_{1,2}^{**} = 0$, and $D_{1,3}^{**} = E$.

[41] The insurance provided by defaultable debt is stressed in theory papers such as Zame (1993) and Dubey, Geanakoplos, and Shubik (2005) and empirical papers such as Mahoney (2015) and Dobbie and Song (2014).

[42] In the Internet Appendix we show that no pooling equilibrium exists, as per Rothschild and Stiglitz (1976), hence the focus here on the optimal separating contract is without loss of generality.

C.1. Analytical Solutions with $\alpha = 1$

Suppose that households only value consumption at $t = 3$ ($\alpha = 1$), in which case the motive for borrowing is purely to build a buffer stock of savings to insure against income risk. We compare two extreme cases of the model analytically (all derivations in Internet Appendix B). First, consider the case where borrowers only differ in their propensity to default far from origination, at $t = 3$. This will be the case when $x_H = x_L$ and $p_L > p_H$. For tractability we normalize $p_H = 0$. Under these conditions there is a set of optimal contracts that can be offered to the high type borrower. These have the following characteristics: $A_1^{*H} = E(1 - x_H)$, $D_{1,2}^{*H} + D_{1,3}^{*H} = E$, $D_{1,2}^{*H} \geq \bar{D}_{1,2} \in (0, E)$ and $D_{1,3}^{*H} \geq 0$. In words, optimal screening is achieved by any contract that has a sufficiently high repayment at $t = 2$. There is no quantity rationing at all as the high type household borrows the fully present value of her expected income.

Next consider the opposite scenario where borrowers only differ in their propensity to default close to origination, at $t = 2$. This will be the case when $x_L > x_H$ and $p_L = p_H$. For tractability we normalize $x_H = 0$. Under these conditions the optimal contract that is offered to the high type borrower is a long-term loan with no repayment at $t = 2$: $D_{1,2}^{*H} = 0$, $D_{1,3}^{*H} \in (0, E)$, and $A_1^{*H} = (1 - (1 - q)\, p)\, D_{1,3}^{*H} < (1 - (1 - q)\, p)\, E$. In this case screening is achieved purely through quantity rationing--high type borrowers chose a long-maturity loan that has no repayment due at $t = 2$ for an amount that is lower than the present value of their expected income.

To interpret these analytical solutions, the difference in each borrowers propensity to default near ($t = 2$) and far ($t = 3$) from origination captures the degree of private information a borrower has at each horizon. Contrasting both cases, we see that screening is optimally achieved by targeting repayment towards the horizon from origination where the degree of asymmetric information about repayment is lowest. When the asymmetry about a borrowers ability to repay close to origination is low, then short maturity debt targets repayments to this time and optimally screens borrowers. Conversely when a borrower's private information has more power to predict the ability to repay close to origination, then short-term debt does not achieve screening. Screening must then be achieved in another way and this is when loan amount will be used.[43]

C.2. Numerical Solution

Next, we solve the model numerically and conduct a comparative static exercise that captures how the differential propensity to default between types is varying over time. We focus on this comparative static because it allows us to link the results of the model to the empirical results in

[43]It should be noted that the model abstracts from dynamic considerations of the impact of loan choice on the terms of new credit after $t = 3$. This would create an additional motivation both for high and low types to indicate they are of high credit quality. On balance this should not alter the truth-telling conditions that characterize the optimal contract and hence will not qualitatively alter the way in which screening is achieved.

Section III.A. To do this, let $p_H = \bar{p}_H + \frac{\Delta}{1-q}$ and $x_H = x_L - \Delta$, where $\bar{p}_H < p_L$ and $\Delta \in [0, (1-q)(p_L - \bar{p}_H)]$.[44] Under this parametrization, the difference in the propensity for each type to default at $t = 2$, $x_L - x_H$, is zero when $\Delta = 0$ and increases linearly to $(1-q)(p_L - \bar{p}_H) > 0$ when $\Delta = (1-q)(p_L - \bar{p}_H)$. Conversely, the difference in the propensity for each type to default at $t = 3$, $(1-q)(p_L - p_H)$, is $(1-q)(p_L - \bar{p}_H)$ when $\Delta = 0$ and decreases linearly to zero when $\Delta = (1-q)(p_L - \bar{p}_H)$. Our comparative statics exercise will show how the optimal contract offered to the high type varies with Δ. Changing Δ allows us to vary whether the difference in the propensity for each type to default is growing or falling with the horizon from origination. By construction, a change in Δ leaves the expected income of a high type household unchanged hence, any change in the amount of borrowing is not mechanically driven by changes in the level of expected income.

Figure 11 presents the comparative statics of the equilibrium lending contract varying Δ. The left axis shows the maturity of the loan (measured as the Macaulay duration of the contract) offered to high type households at $t = 1$ (the solid line).[45] The right axis measures the total amount borrowed by the high type relative to the low type at the equilibrium contract (the dotted line). These measure the degree to which screening is achieved through maturity and quantity rationing, respectively. The comparative statics show that when borrowers differ only on their propensity to default far from origination, Δ is low, then the optimal contract will screen borrowers using maturity. For example when $\Delta = 0$, high type borrowers take a loan that is 6.7% *larger* in size than low types, and hence are not quantity rationed at all. Instead they credibly distinguish themselves by accepting a loan with a shorter duration–here with a duration of 1.43, which indicates that 57% of the loan's value is repaid at $t = 2$. When borrowers differ mainly in their propensity to default near to origination, then the equilibrium contract offered to the high type household uses loan quantity rather than maturity to screen. In the numerical example, when Δ is large, the high type household accepts a long-term loan contract (duration of 2) for an amount that is 19% below the amount taken by the low type household. Intuitively, the comparative statics confirm the intuition from the special cases studied above: maturity is the optimal way to screen when the private information that borrowers possess has more power to predict default further from origination.

The comparative statics presented in Figure 11 are robust to the parameter assumptions used to derive those numerical solutions. We demonstrate this in the Internet Appendix Figure 13. We resolve the model in the case where the household values consumption at $t = 1$ by setting $\alpha < 1$ (Panel A), we vary the choice of q (Panel B), and vary the choice of utility function (Panel C). The comparative statics are qualitatively the same in each case.

[44] These inequalities ensure that $p_H \leq p_L$ and $x_H \leq x_L$. We also only consider parameters for which $x_L \geq (1-q)(p_L - \bar{p}_H)$ to ensure $x_H \geq 0$ for all Δ.

[45] Formally: $Duration = 1 \times \frac{D_{1,2}^{*H}(1-x_H)}{A_1^{*H}} + 2 \times \frac{D_{1,3}^{*H}(1-p_H(1-q)-x_H)}{A_1^{*H}}$.

D. Linking the Model to Evidence on the Time Structure of Private Information

In Section III we show that borrowers who select into short-maturity loans have a decreased propensity to default relative to borrowers who select into long-maturity loans, and that this difference is growing in the horizon from origination. Figure 8 suggests that this gap grows linearly in the time from origination. Since, in terms of our model, selection draws apart the unobserved characteristic of high and low type borrowers, this pins down how the difference in their propensity to default varies at different times from origination. We can study the prediction of our model in the special case where this gap grows linearly over time. To capture this we need to set Δ so that the difference in each types propensity to default at $t = 3$, $(1-q)(p_L - p_H)$, is twice as large as it is at $t = 2$, $x_L - x_H$. This occurs when $\Delta = \frac{1}{3}(p_L - \bar{p}_H)(1-q)$ which, in terms of the parameters used in our numerical results, occurs when $\Delta = 0.016$.

Examining the solution in Figure 11 that corresponds to this special case, we see that selection is optimally achieved through maturity, not quantity rationing. Specifically, high type borrowers take a loan that is 5% *larger* in size than low types, and hence are not quantity rationed at all. Instead they credibly distinguish themselves by accepting a loan with a shorter duration (here with a duration of 1.36), which indicates that 64% of the loan's value is repaid at $t = 2$. The direct implication is that screening using maturity is indeed optimal given the time structure of private information observed in Figure 8. This further emphasizes the usefulness of the new evidence presented in Section III to understand how asymmetric information impacts credit markets. [46]

V. Conclusion

We have documented that loan maturity may be used to screen borrowers based on unobserved creditworthiness in US consumer credit markets. Borrowers who are unobservably more exposed to shocks to their ability to repay self-select into longer maturity loans. We provide a framework that rationalizes this finding and demonstrates that screening borrowers using maturity, as opposed to loan quantity, is optimal when the power of their private information to predict default is increasing over time from origination. We confirm that this condition is indeed true in our empirical setting. Our analysis contrasts with the bulk of work since Stiglitz and Weiss (1981) that has focused on quantity rationing as the primary cost of adverse selection. Our results indicate that maturity rationing is empirically important. More broadly, our results show that information asymmetry limits the ability of financial markets to provide insurance through the provision of long term contracts that protect

[46] The only existing theory that links the time structure of private information and maturity choice is Goswami, Noe, and Rebello (1995). Our framework contrasts sharply with that paper's theoretical conclusion: short-term debt is used in equilibrium when borrowers have more private information about short-term cash flows. The fundamental difference stems from their assumption that borrowers of all types pool on the same contract in equilibrium while we study an equilibrium, consistent with our empirical results, in which screening based on private information is achieved.

borrowers from future shocks to their creditworthiness.

Our analysis also provides a framework to evaluate the implications of regulating consumer loan maturities. In particular, it highlights the the welfare cost implications of banning long-term loans in a competitive consumer credit market with asymmetric information. Our analysis implies that there will be a direct effect on consumer welfare due to the disappearance of instruments to insure against future income and consumption shocks. But there will also be an indirect effect on all borrowers through equilibrium prices and access to credit. Since high and low risk borrowers will pool into short-term loans, the ban would increase the cost of borrowing for borrowers with relatively stable future income and low probability of consumption shocks.

The institutional setting of online credit marketplaces is ideal for a first study on adverse selection on maturity, thanks to the simplicity of the contracts and the regulatory requirement that all borrower information available to the lenders to be public. It remains an open question, both from an empirical and a theoretical perspective, whether selection on maturity is also a first order determinant of equilibrium loan prices in consumer credit markets where lenders may screen on other dimensions of the contract, such as collateral in mortgage markets. Thus, understanding the scope and pervasiveness for adverse selection and screening in other consumer credit markets remains a fruitful area for future research.

References

Adams, William, Liran Einav, and Jonathan Levin, 2009, Liquidity constraints and imperfect information in subprime lending, *American Economic Review* 99, 49–84.

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Strobel, 2015, Regulating consumer financial products: Evidence from credit cards, *The Quarterly Journal of Economics* pp. 111–164.

Baker, Malcolm, Robin Greenwood, and Jeffrey Wurgler, 2003, The maturity of debt issues and predictable variation in bond returns, *Journal of Financial Economics* 70, 261–291.

Barclay, Michael J, and Clifford W Smith, 1995, The maturity structure of corporate debt, *the Journal of Finance* 50, 609–631.

Barry Bosworth, Warren Smith, Daniel Brill, 1971, Patterns of corporate external financing, *Brookings Papers on Economic Activity* 1971, 253–284.

Berger, Allen N, Marco A Espinosa-Vega, W Scott Frame, and Nathan H Miller, 2005, Debt maturity, risk, and asymmetric information, *The Journal of Finance* 60, 2895–2923.

——— , 2011, Why do borrowers pledge collateral? new empirical evidence on the role of asymmetric information, *Journal of Financial Intermediation* 20, 55–70.

Berger, Allen N, and Gregory F Udell, 1990, Collateral, loan quality and bank risk, *Journal of Monetary Economics* 25, 21–42.

——— , 1995, Relationship lending and lines of credit in small firm finance, *Journal of business* pp. 351–381.

Bester, Helmut, 1985, Screening vs. rationing in credit markets with imperfect information, *American Economic Review* 75, 850–55.

Bian, Xun, and Abdullah Yavas, 2013, Prepayment penalty as a screening mechanism for default and prepayment risks, *Real Estate Economics* 41, 193–224.

Booth, James R, 1992, Contract costs, bank loans, and the cross-monitoring hypothesis, *Journal of Financial Economics* 31, 25–41.

Brueckner, Jan K, 2000, Mortgage default with asymmetric information, *The Journal of Real Estate Finance and Economics* 20, 251–274.

Cho, In-Koo, and David M Kreps, 1987, Signaling games and stable equilibria, *The Quarterly Journal of Economics* pp. 179–221.

Cochrane, John H, 1995, Time-consistent health insurance, *Journal of Political Economy* pp. 445–473.

De Meza, David, and David Webb, 2001, Advantageous selection in insurance markets, *RAND Journal of Economics* pp. 249–262.

——— , 2016, False diagnoses: Pitfalls of testing for asymmetric information in insurance markets, *The Economic Journal* p. Forthcoming.

Degryse, Hans, and Patrick Van Cayseele, 2000, Relationship lending within a bank-based system: Evidence from european small business data, *Journal of financial Intermediation* 9, 90–109.

Diamond, Douglas W, 1991, Debt maturity structure and liquidity risk, *The Quarterly Journal of Economics* pp. 709–737.

Dobbie, Will, and Paige Marta Skiba, 2013, Information asymmetries in consumer credit markets: Evidence from payday lending, *American Economic Journal: Applied Economics* 5, 256–282.

Dobbie, Will, and Jae Song, 2014, Debt relief and debtor outcomes: Measuring the effects of consumer bankruptcy protection, Discussion paper National Bureau of Economic Research.

——— , 2015, The impact of loan modifications on repayment, bankruptcy, and labor supply: Evidence from a randomized experiment, .

Dubey, Pradeep, John Geanakoplos, and Martin Shubik, 2005, Default and punishment in general equilibrium, *econometrica* pp. 1–37.

Field, Erica, Rohini Pande, John Papp, and Natalia Rigol, 2013, Does the classic microfinance model discourage entrepreneurship among the poor? experimental evidence from india, *The American Economic Review* 103, 2196–2226.

Finkelstein, Amy, Kathleen McGarry, and Amir Sufi, 2005, Dynamic inefficiencies in insurance markets: Evidence from long-term care insurance, *American Economic Review* 95, 224–8.

Flannery, Mark J, 1986, Asymmetric information and risky debt maturity choice, *The Journal of Finance* 41, 19–37.

Gopalan, Radhakrishnan, Fenghua Song, Vijay Yerramilli, et al., 2014, Debt maturity structure and credit quality, *Journal of Financial and Quantitative Analysis* 49, 817–842.

Goswami, Gautam, Thomas Noe, and Michael Rebello, 1995, Debt financing under asymmetric information, *The Journal of Finance* 50, 633–659.

Gottesman, Aron A., and Gordon S. Roberts, 2004, Maturity and corporate loan pricing, *Financial Review* 39, 55–77.

Goyal, Vidhan K, and Wei Wang, 2013, Debt maturity and asymmetric information: Evidence from default risk changes, *Journal of Financial and Quantitative Analysis* 48, 789–817.

Graham, John R, and Campbell R Harvey, 2001, The theory and practice of corporate finance: Evidence from the field, *Journal of financial economics* 60, 187–243.

Greenwood, Robin, Samuel Hanson, and Jeremy C Stein, 2010, A gap-filling theory of corporate debt maturity choice, *The Journal of Finance* 65, 993–1028.

Guedes, Jose, and Tim Opler, 1996, The determinants of the maturity of corporate debt issues, *the Journal of Finance* 51, 1809–1833.

Hart, Oliver, and John Moore, 1994, A theory of debt based on the inalienability of human capital, *The Quarterly Journal of Economics* pp. 841–879.

——— , 1995, Debt and seniority: An analysis of the role of hard claims in constraining management, *The American Economic Review* pp. 567–585.

Holmstrom, Bengt, 1983, Equilibrium long-term labor contracts, *The Quarterly Journal of Economics* pp. 23–54.

Jimenez, Gabriel, Vicente Salas, and Jesus Saurina, 2006, Determinants of collateral, *Journal of financial economics* 81, 255–281.

Johnson, Shane A, 2003, Debt maturity and the effects of growth opportunities and liquidity risk on leverage, *Review of Financial Studies* 16, 209–236.

Karlan, D., and J. Zinman, 2009, Observing unobservables: Identifying information asymmetries with a consumer credit field experiment, *Econometrica* 77, 1993–2008.

Khandani, Amir, Andrew W. Lo, and Robert C. Merton, 2013, Systemic risk and the refinancing ratchet effect, *Journal of Financial Economics* 108, 29–45.

Kocherlakota, Narayana R, 2004a, Figuring out the impact of hidden savings on optimal unemployment insurance, *Review of Economic Dynamics* 7, 541–554.

Kocherlakota, Narayana R., 2004b, Figuring out the impact of hidden savings on optimal unemployment insurance, *Review of Economic Dynamics* 7, 541–554.

Leeth, John D, and Jonathan A Scott, 1989, The incidence of secured debt: evidence from the small business community, *Journal of Financial and Quantitative Analysis* 24, 379–394.

Leland, Hayne E, and David H Pyle, 1977, Informational asymmetries, financial structure, and financial intermediation, *The journal of Finance* 32, 371–387.

Levine, Carolyn B, and John S Hughes, 2005, Management compensation and earnings-based covenants as signaling devices in credit markets, *Journal of Corporate Finance* 11, 832–850.

Mahoney, Neale, 2015, Bankruptcy as implicit health insurance, *American Economic Review* 105, 710–46.

Milbradt, Konstantin, and Martin Oehmke, 2014, Maturity rationing and collective short-termism, *Journal of Financial Economics*.

Morse, Adair, 2015, Peer-to-peer crowdfunding: Information and the potential for disruption in consumer lending, *Annual Review of Financial Economics*.

Myers, Stewart C, 1977, Determinants of corporate borrowing, *Journal of financial economics* 5, 147–175.

Ross, Stephen A, 1977, The determination of financial structure: the incentive-signalling approach, *The Bell Journal of Economics* pp. 23–40.

Rothschild, Michael, and Joseph E Stiglitz, 1976, Equilibrium in competitive insurance markets: An essay on the economics of imperfect information, *The Quarterly Journal of Economics* 90, 630–49.

Schreft, Stacey L, and Anne P Villamil, 1992, Credit rationing by loan size in commercial loan markets, *Economic Review* 78, 3–8.

Spence, Michael, 1973, Job market signaling, *The Quarterly Journal of Economics* 87, 355–374.

Stango, Victor, and Jonathan Zinman, 2013, Borrowing high vs. borrowing higher: Sources and consequences of dispersion in individual borrowing costs, Discussion paper National Bureau of Economic Research.

Stanton, Richard, and Nancy Wallace, 1998, Mortgage choice: what's the point?, *Real Estate Economics* 26, 173–205.

Stiglitz, J.E., and A. Weiss, 1981, Credit rationing in markets with imperfect information, *The American economic review* 71, 393–410.

Taggart, Robert A, 1977, A model of corporate financing decisions, *The Journal of Finance* 32, 1467–1484.

Zame, William R, 1993, Efficiency and the role of default when security markets are incomplete, *The American Economic Review* pp. 1142–1164.

Appendix

A. Figures and Tables

Figure 1: Description of variation

This figure presents a stylized description of the identification strategy to disentangle selection from the causal effect of loans of different maturity. Rows depict two observationally equivalent groups of borrowers, A and B. Columns depict the menu options available to borrowers of Group A (only short-term loan, APR of r_{ST}) and Group B (short, APR r_{ST}, and long term, APR r_{LT}, loan options). Each cell presents the observed default rate of the set of borrowers of either group that selected into a particular loan. Our ID strategy measures $\gamma_B^{ST} - \gamma_A^{ST}$. This compares the default rate of short term borrowers of Group B, which constitute a selected sample, with the default rate of short term borrowers of Group A, which constitute a non-selected sample, holding the contract terms fixed.

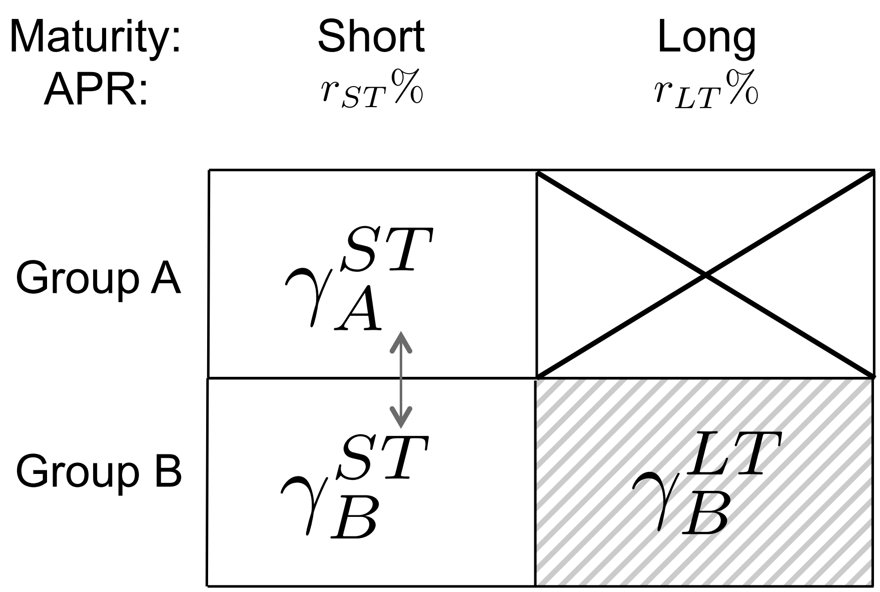


Figure 2: Staggered expansion of 60-month loans

This figure shows the time series of the number of 60-month loans by listing month for \$10,000 to \$12,000 and \$12,000 to \$16,000.

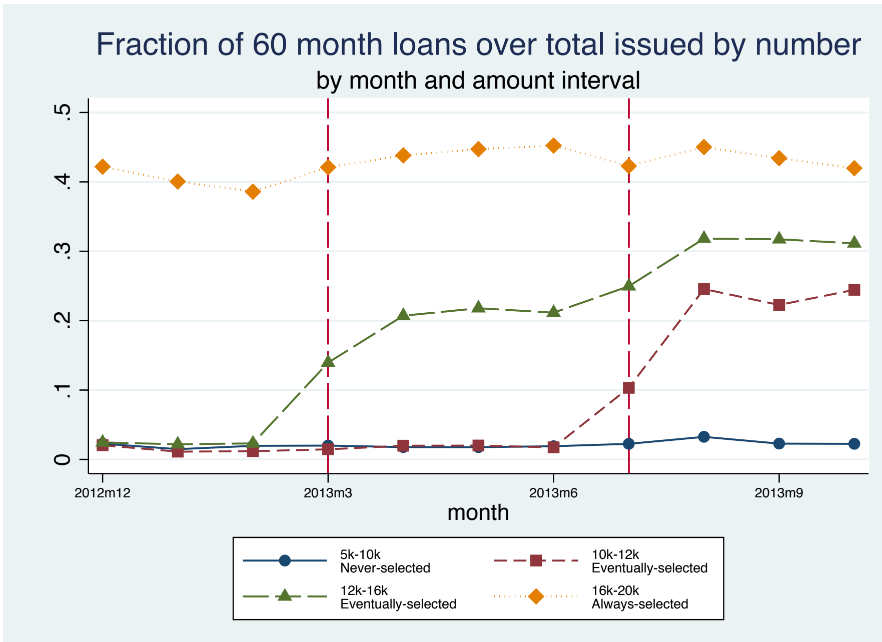


Figure 3: Hazard rate of default

This figure shows the hazard rate of default by month since origination for 36-month loans issued by LC in amounts between \$5,000 and \$20,000, between December 2012 and February 2013 (pre-period). A loan is in default if payments are 120 or more late on April 2015. The timing of default is the month, measured as time since origination in which payments were first missed. The hazard rate at horizon t is the number of loans that enter default at that horizon as a fraction of the number of loans that are in good standing at $t-1$.

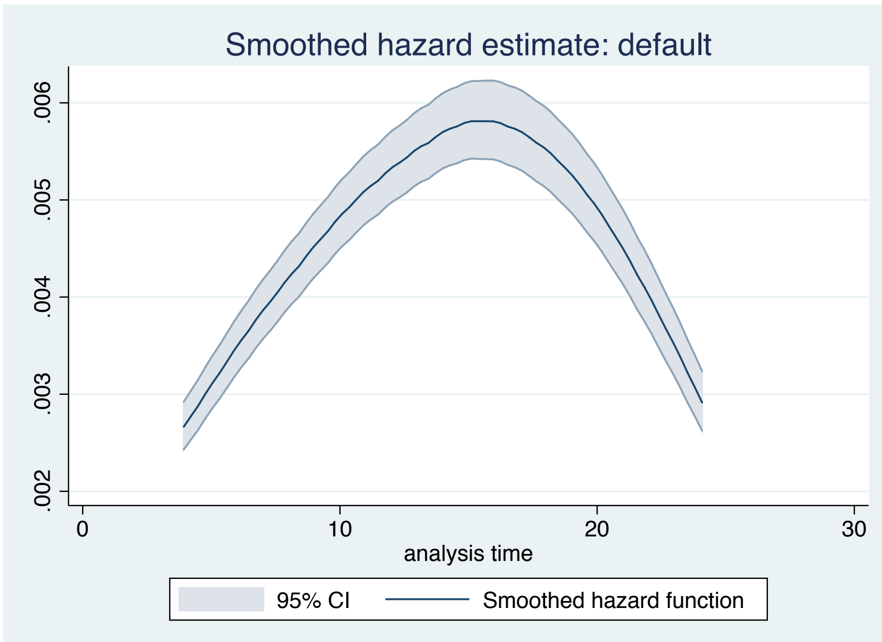


Figure 4: Total $ amount issued by LC by month of listing

This figure shows the time series of total $ amount of LC loans (of both maturities) by listing month since 2012. The vertical dashed lines show the two months in which the 60-month loan minimum amount was reduced.

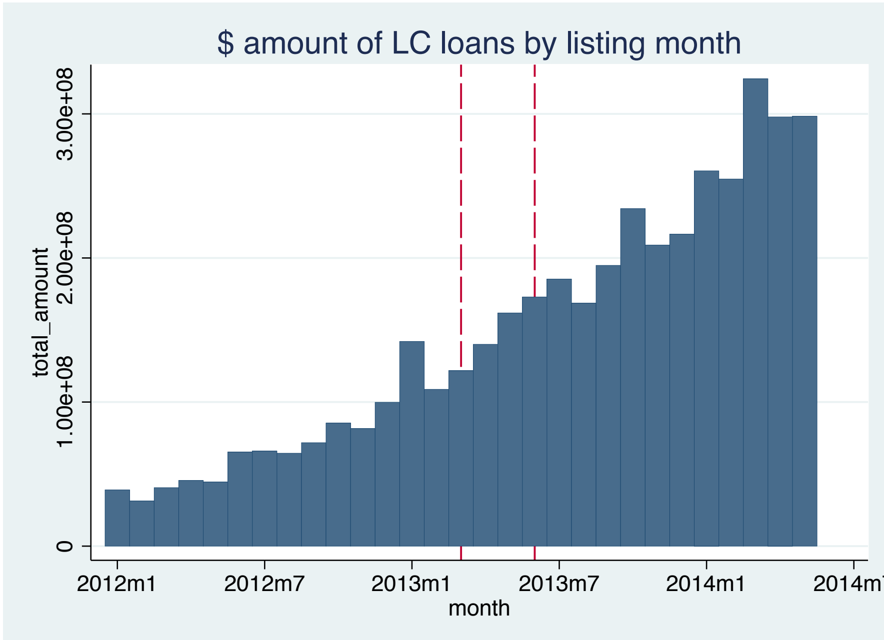


Figure 5: Stylized depiction of identification strategy

This figure shows a stylized depiction of our difference-in-differences strategy using the expansion of the menu of borrowing options.

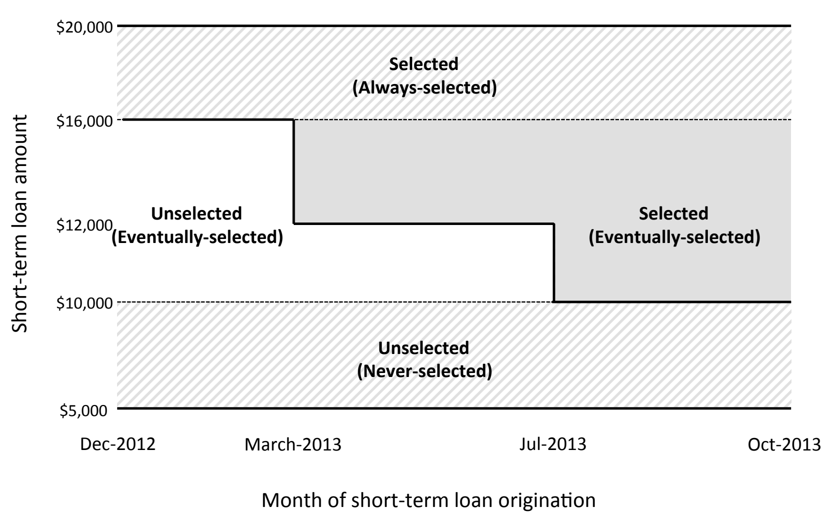
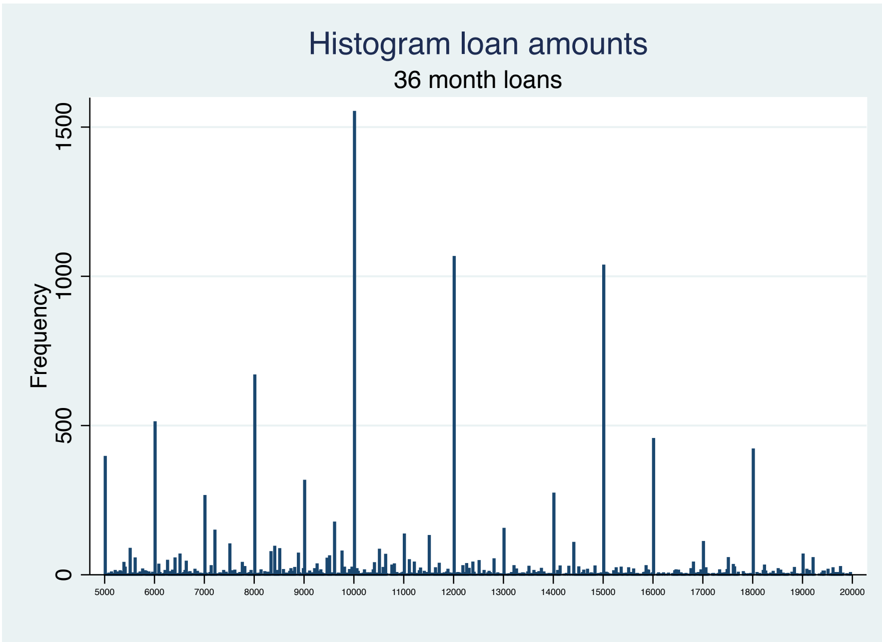


Figure 6: Pre-period loan amount histogram

This top panel shows the number of 36-month loans issued by LC by loan amount in $25 increments, between $5,000 and $25,000 between December 2012 and February 2013. The bottom panel shows the same histogram for the same period of time but for 60-month loans.



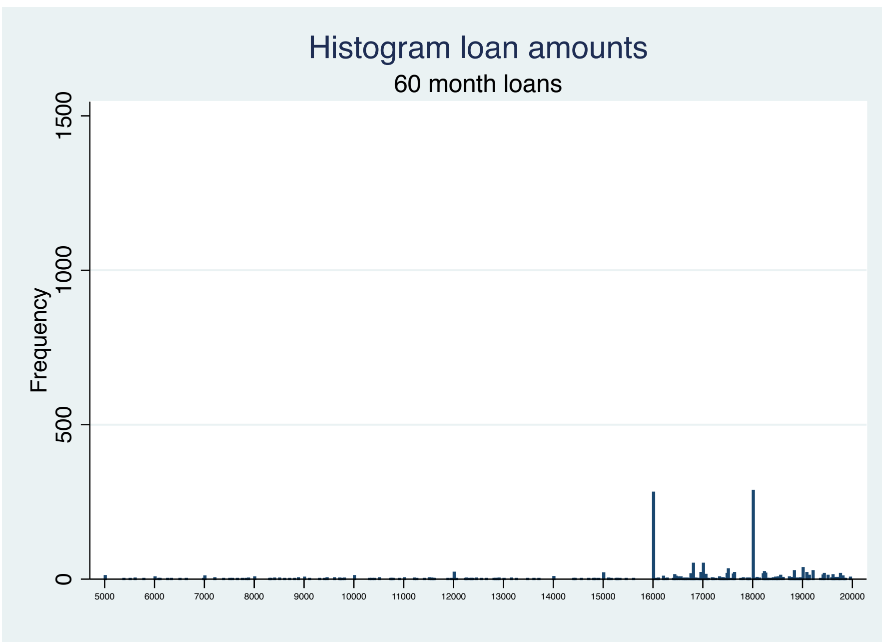


Figure 7: Pre-trends on number of loans originated

This figure shows the regression coefficients (γ_τ) and 90% confidence interval of regression:

$$log(N_{j,t,amount1000}) = \beta_{amount1000} + \delta_{j,t} + \sum_{\tau=-3}^{3} \gamma_\tau \times D(\tau)_{amount1000,t} + \varepsilon_{i,t},$$

which measures the difference in the number of loans issued between eventually-selected and control amounts τ months after the threshold expansion. Standard errors are robust to heteroskedasticity.

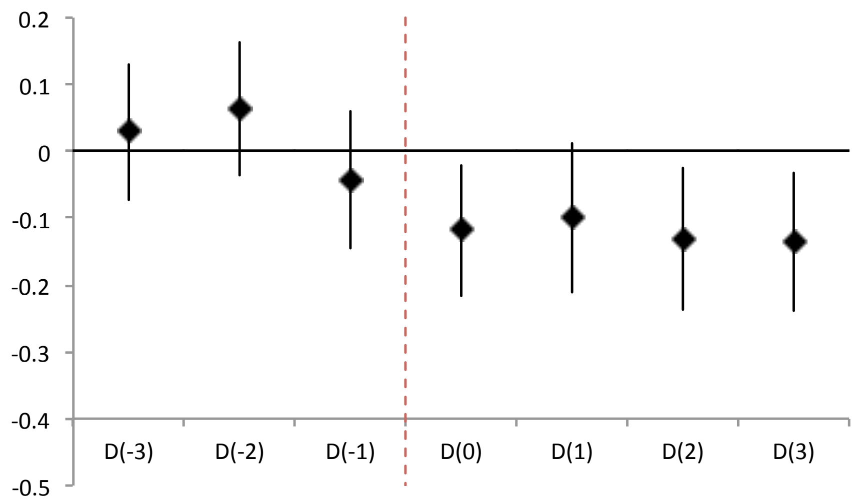


Figure 8: Default rate coefficient by number of months since origination

This figure shows the estimated coefficient and 90% confidence interval of the regression:

$$default\,(\Delta t) = \beta_{amount1000} + \delta_{j,\overline{FICO},t} + \gamma \times D_{amount100,t} + X_{i,t} + \varepsilon_i,$$

where the outcome is $default\,(\Delta t)$, a dummy that equals one if a loan is late by more than 120 days as of April 2015 and if the last payment on these loan occurred Δt months after origination, on $D_{amount1000,t}$, a dummy that captures the staggered expansion of the 60-month loans for amounts above \$12,000 and \$10,000 on March and July 2013, respectively. Standard errors are clustered at the state level. Sample includes loans issued between December 2012 and October 2013, for loan amounts between \$5,000 and \$20,000.

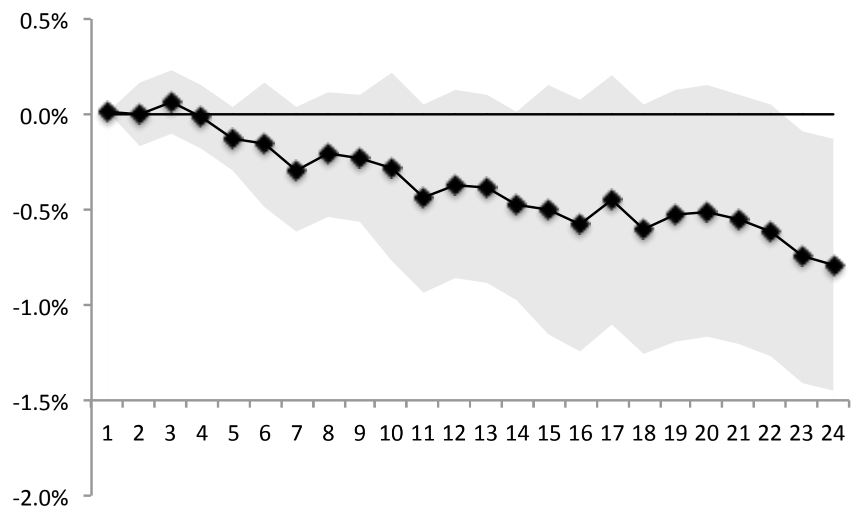


Figure 9: Model Time-line

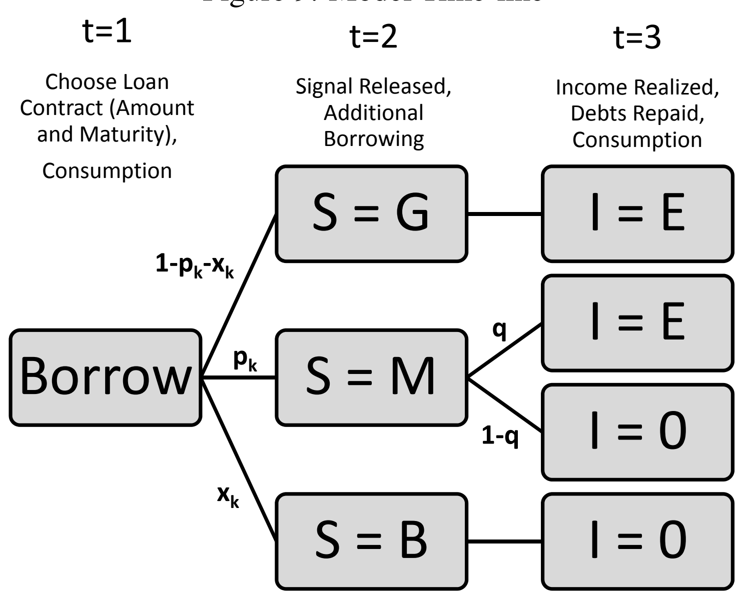


Figure 10: Reduction in APR

This figure shows the time series of the predicted residual of a regression of loan APR on $1,000 amount dummies, FICO score bin dummies, annual income, and address state dummies, by month of origination, for 36-month loans issued between $10,000 and $16,000.

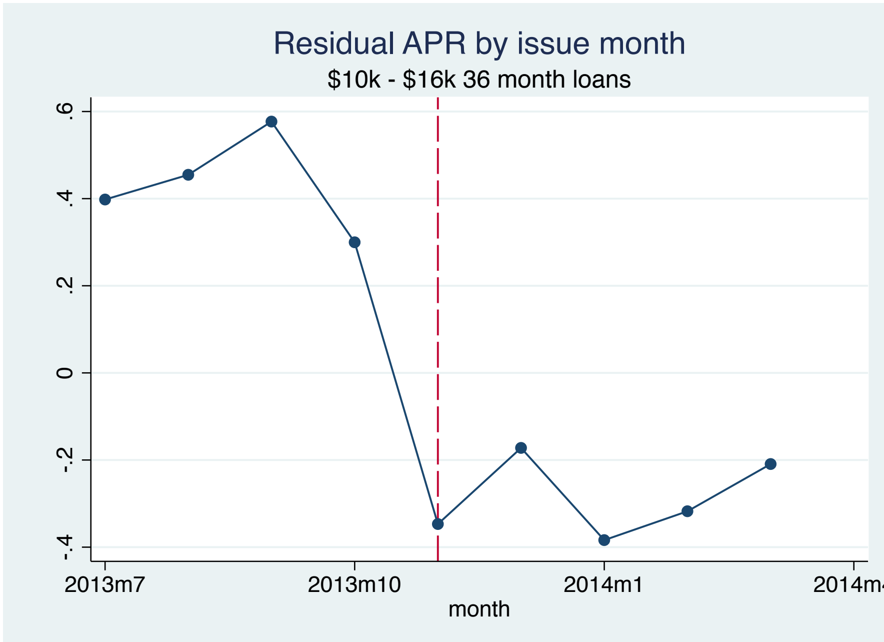


Figure 11: Model Comparative Statics

This figure shows comparative statics from numerical solutions of the theoretical framework presented in Section IV. The household utility function is assumed to be CARA: $u(c) = 1 - \frac{1}{\eta} e^{-\eta c}$. The following parameter assumptions are used: $E = 100$, $p_L = 0.3$, $x_L = 0.1$, $q = 0.75$, $\bar{p}_H = 0.1$, $\eta = 0.1$ and $\alpha = 1$. The equilibrium contract is shown for varying values of Δ. The left axis shows the degree of maturity rationing as captured by the Macaulay duration of the equilibrium loan offered to the high type: $Duration = 1 \times \frac{D_{1,2}^{*H}(1-x_H)}{A_1^{*H}} + 2 \times \frac{D_{1,3}^{*H}(1-p_H(1-q)-x_H)}{A_1^{*H}}$. The right axis shows the degree of quantity rationing as captured by the ratio of the amount lent to high and low type borrowers at $t = 1$: $\frac{A_1^{*H}}{A_1^{*L}}$.

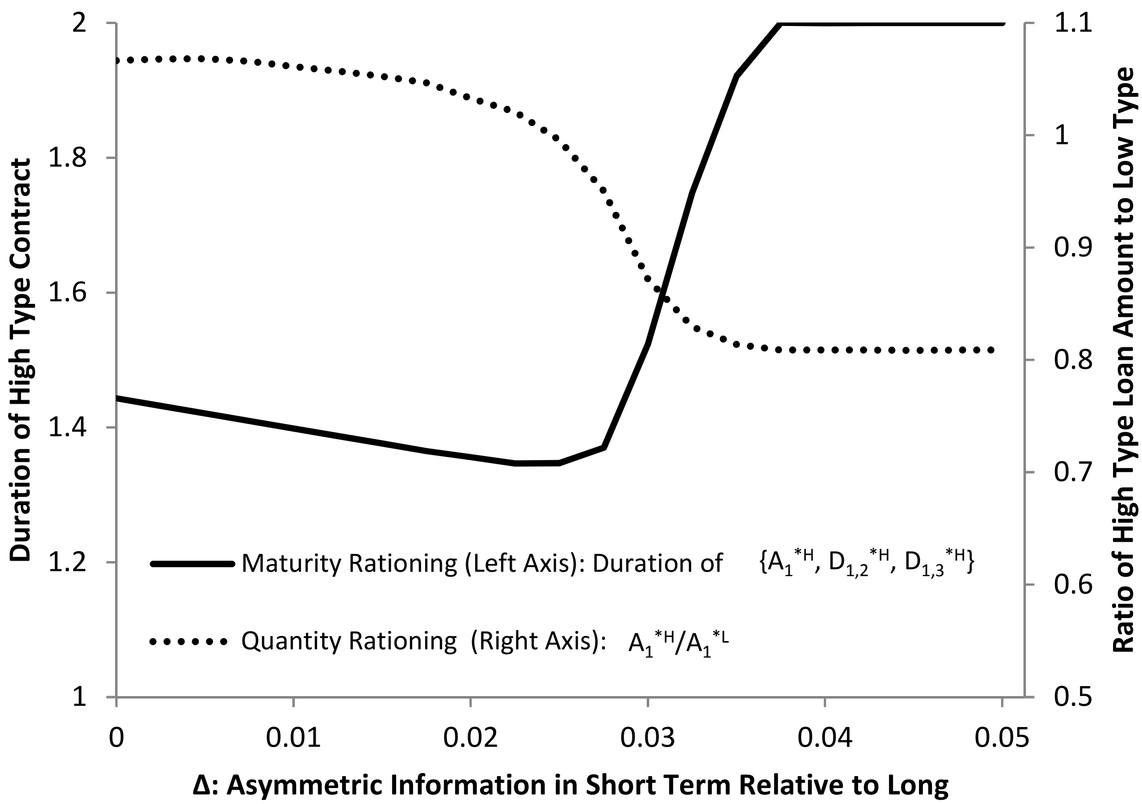


Table 1: Pre-period summary statistics

This table shows summary statistics of the main sample of Lending Club borrowers for pre-expansion months, which includes all 36-month loans whose listing date is between December 2012 and March 2013, for an amount between $5,000 and $20,000, and for which we estimate an initial risk category based on LC's publicly available information.

	mean	p50	sd
Panel A: loan characteristics			
APR (%)	16.3	16.0	4.1
Installment ($)	379.9	360.9	125.1
For refinancing (%)	87.0		
Default (%)	9.2		
Fully paid (%)	37.6		
Panel B: borrower characteristics			
Annual income ($)	65,745	57,500	74,401
Debt payments / Income (%)	17.4	16.9	7.7
FICO at origination (high range of 4 point bin)	695	689	26
FICO at latest data pull (high range of 4 point bin)	685	699	70
Home ownership (%)	55.5		
Total debt excl mortgage ($)	38,153	29,507	33,805
Revolving balance ($)	14,549	11,592	12,719
Revolving utilization (%)	60.7	62.7	21.9
Months of credit history	182	164	84
N		12,091	

Table 2: Regression results: selection into long-maturity loans

This table shows that selection into the new 60-month options was higher among borrowers who would have selected a 36-month loan of the same range of amounts as the new 60-month options. The sample corresponds to loan amounts between \$5,000 and \$20,000 whose list date is between December 2012 and October 2013. Column 1 shows the coefficient of the regression of *log(N)*, the logarithm of the number of loans at each month, credit risk risk category, and \$1,000 amount interval level, on a dummy that equals one for loan amounts at which the 60-month loan was first not available and then made available, and zero otherwise. Columns 2, 3 and 4 show the regression results on different samples where we re-define $D_{amount1000,t}$ in an ad-hoc manner for each column. Column 2 restricts the sample to 60-month loans issued in the main sample period for amounts between \$16,000 and \$24,000; $D_{amount1000,t}$ is defined as one for loan amounts between \$16,000 and \$20,000 on and after March 2013, and zero in other cases. Column 3 restricts the sample to 36-month loans issued in the main sample period for amounts between \$16,000 and \$24,000; $D_{amount1000,t}$ is defined as one for loan amounts between \$16,000 and \$20,000 on and after March 2013, and zero in other cases. Column 4 restricts the sample to 36-month loans issued in the main sample period for amounts between \$1,000 and \$10,000; $D_{amount1000,t}$ is defined as one loan amounts between \$5,000 and \$10,000 on and after July 2013 and zero in other cases. Column 5 reports the tests of a Placebo sample, which includes loan amounts between \$5,000 and \$20,000 issued between July 2013 and May 2014. Standard errors are robust to heteroskedasticity. *, ** and *** represent significance at the 10%, 5%, and 1% respectively.

	(1)	(2)	(3)	(4)	(5)
	$log(N)$	$log(N)$	$log(N)$	$log(N)$	$log(N)$
$D_{amount1000,t}$	-0.1451***	-0.0825	0.0586	-0.0355	-0.0441
	(0.033)	(0.065)	(0.064)	(0.048)	(0.028)
Sample	Main	60m, 16k - 24k	36m, 16k - 24k	36m, 1k - 10k	Placebo
Observations	3,663	1,738	1,637	2,374	3,861
R^2	0.817	0.724	0.802	0.761	0.862

Table 3: Regression results: screening with maturity

This table shows that the default rate of borrowers who selected into a short-term loan when they could take a long-term loan is higher than borrowers who could not take a long-term loan. The table shows the output of the regression of each outcome on a dummy for the staggered reduction of the minimum amount threshold for long-maturity loans on March 2013 (to $12,000) and July 2013 (to $10,000). The outcome is $Default$, a dummy that equals one if a borrower is late by more than 120 days, measured as of April 2015. The sample corresponds to loan amounts between $5,000 and $20,000 whose listing date is between December 2012 and October 2013. All regressions include risk category $\times$month, and 4-point FICO score bin, state, and $1,000 amount-bin fixed effects. Columns 2 and 4 include all borrower level variables observed by investors at the time of origination as controls. Standard errors are clustered at the state level. *, ** and *** represent significance at the 10%, 5%, and 1% respectively.

	(1)	(2)
	$Default$	$Default$
$D_{i,t}$	-0.0081**	-0.0080**
	(0.004)	(0.004)
Sample	MAIN	MAIN
Observations	60,511	57,263
R^2	0.035	0.047
# clusters	45	45

Table 4: Robustness

The table shows the output of several robustness tests. Column 1 replicates column 1 in Table 3 on a sample of loans listed between December 2012 and October 2013 and issued for amounts between \$6,000 and \$19,000 (\$1,000 narrower interval than main sample). Columns 2 and 3 report the output for regressions ran on a sample of loans listed between December 2012 and October 2013 for different loan amounts, where the independent variable is defined in an ad-hoc manner using $default$ as outcome. Column 2 restricts the sample to 36-month loans issued in the main sample period, for amounts between \$16,000 and \$24,000; $D_{i,t}$ is equal to one for loan amounts between \$16,000 and \$20,000 listed on or after March 2013, and zero otherwise. Column 3 restricts the sample to 36-month loans issued in the main sample period for amounts between \$1,000 and \$10,000; $D_{i,t}$ is equal to one for loan amounts between \$5,000 and \$10,000 listed on or after July 2013, and zero otherwise. All regressions include risk category$\times$month, and 4-point FICO score bin, state, and \$1,000 amount-bin fixed effects. Standard errors are clustered at the state level. *, ** and *** represent significance at the 10%, 5%, and 1% respectively.

	(1)	(2)	(3)
	$Default$	$Default$	$Default$
$D_{i,t}$	-0.0066*	0.0018	-0.0106
	(0.004)	(0.010)	(0.007)
Sample	6k - 19k	36m, 16k - 24k	36m, 1k - 10k
Observations	54,689	14,652	33,493
R^2	0.037	0.061	0.035
# clusters	45	45	46

Table 5: Interpretation of results

This table shows the output of the regression of each outcome on a dummy for the staggered reduction of the minimum amount threshold for long-maturity loans on March 2013 (to \$12,000) and July 2013 (to \$10,000). Outcomes include $Default12m$ and $Default24m$, dummies that equal one if a borrower is late by more than 120 days as of April 2015 and whose last payment occurred within 12 and 24 months after origination, respectively; *FICO*, the (the high end of the 4-point bin) FICO score measured as of April 2015; *sd(FICO)* the time series standard deviation of (the high end of the 4-point bin) FICO scores within an individual, using four observations per individual: at origination, as of August 2014, as of December 2014, and as of April 2015; and $Prepayment$, a dummy that equals one if the loan is fully paid as of April 2015. In column 4 we include as controls all variables that are observable by investors at origination, as in column 2 of Table 3. The sample corresponds to loan amounts between \$5,000 and \$20,000 whose listing date is between December 2012 and October 2013. All regressions include risk category×month, and 4-point FICO score bin, state, and \$1,000 amount-bin fixed effects. Standard errors are clustered at the state level. *, ** and *** represent significance at the 10%, 5%, and 1% respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	$Default12m$	$Default24m$	$FICO$	$FICO$	$sd(FICO)$	$Prepayment$
$D_{i,t}$	-0.0039	-0.0082*	2.7464**	2.6705**	-0.5764**	-0.0063
	(0.003)	(0.004)	(1.032)	(0.999)	(0.266)	(0.008)
Sample	MAIN	MAIN	MAIN	MAIN	MAIN	MAIN
Observations	60,511	60,511	60,511	57,263	60,511	60,511
R^2	0.024	0.032	0.192	0.215	0.027	0.023
# clusters	45	45	45	45	45	45

B. Mathematical Appendix for Framework

A. Symmetric Information

Here we solve the optimal lending contract when lenders and borrowers are symmetrically informed about borrower type. Consider the following optimal insurance problem

$$\max_{c_1, c_3^G, c_3^M, c_3^B} (1-\alpha)\, u(c_1) + \alpha \left[(1-p_k-x_k)\, u\left(c_3^G\right) + p_k u\left(c_3^M\right) + x_k u\left(c_3^B\right) \right] \tag{4}$$

subject to

$$c_1 + (1-p_k-x_k) \times c_3^G + p_k \times c_3^M + x_k \times c_3^B \leq E\,(1-p_k\,(1-q)-x_k) \tag{5}$$

Let λ^{Symm} be the Lagrange multiplier on the break-even constraint (5). The first order conditions for each choice variable are

$$c_1 : (1-\alpha)\, u'(c_1) - \lambda^{Symm} = 0 \tag{6}$$

$$c_3^G : \alpha\,(1-p_k-x_k)\, u'\left(c_3^G\right) - (1-p_k-x_k)\, \lambda^{Symm} = 0 \tag{7}$$

$$c_3^M : \alpha p_k u'\left(c_3^M\right) - p_k \lambda^{Symm} = 0 \tag{8}$$

$$c_3^B : \alpha x_k u'\left(c_3^M\right) - x_k \lambda^{Symm} = 0 \tag{9}$$

The first order conditions for consumption at $t=3$, (7) (8) and (9), are satisfied if and only if $u'\left(c_3^G\right) = u'\left(c_3^M\right) = u'\left(c_3^B\right)$. Given the strict concavity of $u()$ this requires $c_3^G = c_3^M = c_3^B$. Let c_3 denote this state independent level of consumption at $t=3$. Consumption at $t=3$ in each state as a function of the loan contract $\left\{A_1^{**k}, D_{1,2}^{**k}, D_{1,3}^{**k}\right\}$ is

$$c_3^G = A_1^{**k} - c_1 + E - D_{1,2}^{**k} - D_{1,3}^{**k} \tag{10}$$

$$c_3^M = A_1^{**k} - c_1 - D_{1,2}^{**k} + q\left(E - D_{1,3}^{**k}\right) \tag{11}$$

$$c_3^M = A_1^{**k} - c_1 \tag{12}$$

recalling that the household defaults when it is unable to repay the loan: $S=M, I=0$ and $S=B$. Using (10) and (11) we have that $c_3^G = c_3^M$ if and only if $E - D_{1,3}^{**k} = q\left(E - D_{1,3}^{**k}\right)$ which can only hold if $E = D_{1,3}^{**k}$. Using (11) and (12) we have that $c_3^M = c_3^B$ if and only if $D_{1,2}^{**k} = q\left(E - D_{1,3}^{**k}\right)$ and since $E = D_{1,3}^{**k}$ this implies $D_{1,2}^{**k} = 0$. Competition ensures that the the breakeven condition (5) must hold and so $A_1^{**k} = E\,(1-p_k\,(1-q)-x_k)$. Using (6) and (6) the choice of c_1 will be determined by the Euler equation:

$$(1-\alpha)\,u'(c_1) = \alpha u'\left(E\left(1-p_k(1-q)-x_k\right)-c_1\right). \tag{13}$$

B. Asymmetric Information

We start by studying the general case where households value consumption at both $t=1$ and $t=3$: $\alpha \in [0,1]$. Formally, the contract offered to high creditworthy households $\left\{A_1^{*H}, D_{1,2}^{*H}, D_{1,3}^{*H}\right\}$ will be the solution to:

$$\max_{c_1,A_1,D_{1,2},D_{1,3}} (1-\alpha)\,u(c_1)+\alpha\left[(1-p_H-x_H)\,u\left(c_3^G\right)+p_H u\left(c_3^M\right)+x_H u\left(c_3^B\right)\right] \tag{14}$$

subject to

$$c_3^G = A_1+E-D_{1,2}-D_{1,3}-c_1 \tag{15}$$

$$c_3^M = A_1+qE-D_{1,2}-qD_{1,3}-c_1 \tag{16}$$

$$c_3^B = A_1-c_1 \tag{17}$$

$$A_1 \leq (1-x_H)\,D_{1,2}+(1-(1-q)\,p_H-x_H)\,D_{1,3} \tag{18}$$

$$D_{1,2} \leq A_1-c_1+q\,(E-D_{1,3}) \tag{19}$$

$$D_{1,2} \geq 0 \tag{20}$$

$$D_{1,3} \geq 0 \tag{21}$$

$$U^{*L} \geq U^{*L'}(A_1, D_{1,2}, D_{1,3}) \tag{22}$$

where c_3^S is the consumption that will be achieved at $t=3$ for each possible realization of the interim signal.[47] The conditions 15, 16, 17 give the level of consumption that the household will have at $t=3$ in each state given the debt contract and the choice of c_1. Condition 18 ensures that a lender will break even in expectation. Condition 19 ensures that the household is able to repay the payment due at $t=2$ whenever $S=G,M$. We assume that q is sufficiently large so that this constraint does not bind. Condition 20 and 21 ensures that the contracted repayments at $t=2$ and $t=3$ are non-negative. Crucially, since the debt is defaulted on in certain states this ensures that the lender is unable to sign a contract to make payments to the borrower at $t=2$ or $t=3$ that is conditional on the signal S or the realized amount of income I. Condition 22 is the truth telling constraint that ensures low type households do not choose the loan designed for the high type. $U^{*L'}(A_1, D_{1,2}, D_{1,3})$ is the expected utility that a low type will achieve if she deviates and takes the contract designed for the high type: $\left\{A_1, D_{1,2}, D_{1,3}\right\}$. The function $U^{*L'}(A_1, D_{1,2}, D_{1,3})$ is is defined by finding the level of consumption at $t=1$, c_1', that a low type will choose if they deviate and take the contract

[47] Note that additional borrowing at $t=2$ will ensure that conditional on reaching $S=B$ all remaining income risk is insured at $t=2$ and hence independent of the realization of I.

designed for this high type household. Formally $U^{*L'}(A_1, D_{1,2}, D_{1,3})$ is the maximized objective of the following problem:

$$\max_{c_1'} (1-\alpha)\, u\left(c_1'\right) + \alpha\left[(1-p_L-x_L)\, u\left(c_3^{\prime G}\right) + p_L u\left(c_3^{\prime M}\right) + x_L u\left(c_3^{\prime B}\right)\right] \tag{23}$$

subject to

$$c_3^{\prime G} = A_1 + E - D_{1,2} - D_{1,3} - c_1' \tag{24}$$

$$c_3^{\prime M} = A_1 + qE - D_{1,2} - qD_{1,3} - c_1' \tag{25}$$

$$c_3^{\prime B} = A_1 - c_1' \tag{26}$$

where (24), (25), (26), are the counterparts to (15), (16), (17) in the problem above. Substituting (24), (25), (26)into (23) and taking the first derivative with respect to c_1' gives the following first order condition:

$$\begin{aligned}(1-\alpha)\, u'\left(c_1'\right) = {} & \alpha\,(1-p_k-x_k)\, u'\left(A_1+E-D_{1,2}-D_{1,3}-c_1'\right) \\ & + \alpha p_k u'\left(A_1+qE-D_{1,2}-qD_{1,3}-c_1'\right) \\ & + \alpha x_k u'\left(A_1-c_1'\right)\end{aligned} \tag{27}$$

As argued by Kocherlakota (2004b) this first order condition cannot in general be simply used as an additional constraint in the first problem. Also, doing so renders the problem such that analytical solutions (and often numerical solutions) are unworkable. We avoid this problem by solving the model analytically in two special cases as well as providing a range of numerical solutions below.

B.1. Consumption only at $t = 3$ ($\alpha = 1$)

We now consider the contracting problem in the case where the household only consumes at $t = 3$. This eliminates the possibility of hidden savings since all debt raised at $t = 1$ will be saved. To simplify the problem we make use of the fact that, as is standard, the zero profit condition 18 will bind at the optimal contract. Combining this with 15, 16, 17 allows us to express consumption in each state as a function of $D_{1,2}$ and $D_{1,3}$. The Lagrangian for the constrained optimization problem is:

$$
\begin{aligned}
\mathscr{L} = \max_{D_{1,2},D_{1,3}} (1-p_H-x_H)\,u\,(E-x_H D_{1,2}-(x_H+(1-q)\,p_H)\,D_{1,3}) && (28)\\
+p_H u\,(qE-x_H D_{1,2}+((1-q)\,(1-p_H)-x_H)\,D_{1,3}) && (29)\\
+x_H u\,((1-x_H)\,D_{1,2}+(1-x_H-(1-q)\,p_H)\,D_{1,3}) && (30)\\
+\lambda_1 u\,((1-p_L\,(1-q)-x_L)\,E) && (31)\\
-\lambda_1\,(1-p_L-x_L)\,u\,(E-x_H D_{1,2}-(x_H+(1-q)\,p_H)\,D_{1,3}) && (32)\\
-\lambda_1 p_L u\,(qE-x_H D_{1,2}+((1-q)\,(1-p_H)-x_H)\,D_{1,3}) && (33)\\
-\lambda_1 x_L u\,((1-x_H)\,D_{1,2}+(1-x_H-(1-q)\,p_H)\,D_{1,3}) && (34)\\
+\lambda_2 D_{1,2}+\lambda_3 D_{1,3} && (35)
\end{aligned}
$$

where (34) is the truth telling condition ensuring low types do not accept the contract designed to the high type while (35) ensures $D_{1,2}$ and $D_{1,3}$ are non-negative. The associated Lagrange multipliers, $\lambda_1, \lambda_2, \lambda_3$ are non-negative and obey the standard Kuhn-Tucker conditions. Observe that the truth-telling condition must bind for the optimal contract or else the high type agent would be given the full insurance contract but this is strictly preferred by the low type household. The first order conditions with respect to two choice variables are

$$
\begin{aligned}
D_{1,2}: \quad & -(1-p_H-x_H)\,x_H u'\left(c_3^G\right)-p_H x_H u'\left(c_3^M\right)+x_H\,(1-x_H)\,u'\left(c_3^B\right)\\
& +\lambda_1\left[(1-p_L-x_L)\,x_H u'\left(c_3^G\right)+p_L x_H u'\left(c_3^M\right)-x_L\,(1-x_H)\,u'\left(c_3^B\right)\right]+\lambda_2=0\\
D_{1,3}: \quad & -(1-p_H-x_H)\,(x_H+(1-q)\,p_H)\,u'\left(c_3^G\right)+p_H\,((1-q)\,(1-p_H)-x_H)\,u'\left(c_3^M\right)+x_H\,(1-x_H-(1-q\\
& +\lambda_1\left[(1-p_L-x_L)\,(x_H+(1-q)\,p_H)\,u'\left(c_3^G\right)-p_L\,((1-q)\,(1-p_H)-x_H)\,u'\left(c_3^M\right)-x_L\,(1-x_H-(1-q)\,p_H\right.
\end{aligned}
$$

We now use (36) and (37) to characterize the optimal contract under the two scenarios considered in the paper.

B.2. Information Asymmetry Only About Ability to Repay at $t=3$

Consider the case where $x_L = x_H = x \geq 0$ and $p_L > p_H = 0$. We conjecture that at the optimal contract $\lambda_1 = 0$. We will verify this conjecture below. Now (36) and (37) can be re-written as

$$
D_{1,2}: x\,(1-x)\left[u'\left(c_3^G\right)-u'\left(c_3^B\right)\right]=\lambda_2 \qquad (38)
$$

$$
D_{1,3}: x\,(1-x)\left[u'\left(c_3^G\right)-u'\left(c_3^B\right)\right]=\lambda_3 \qquad (39)
$$

Together (38) and (39) imply that $\lambda_2 = \lambda_3$. Suppose first that $\lambda_2 = \lambda_3 > 0$. By complementary slackness this would imply $D_{1,2} = D_{1,3} = 0$. In that case $c_3^G = E > c_3^B = 0$ which implies $u'\left(c_3^G\right) <$

$u'\left(c_3^B\right)$ and hence creates a contraction. It therefore must be that $\lambda_2 = \lambda_3 = 0$. Using (38) and (39) this implies that $c_3^G = c_3^B$ at the optimal contract. Using this together with (15) and (17) we get that the unique solution to the optimal contract problem is to set $D_{1,2}$ and $D_{1,3}$ such that

$$D_{1,2} + D_{1,3} = E \tag{40}$$

The range of allowable values of $D_{1,2}$ and $D_{1,3}$ is determined by checking that the truth-telling constraint binds. Letting $D_{1,3} = E - D_{1,2}$ the truth-telling constraint can be written as:

$$\begin{aligned} u\left(\left(1-p_L\left(1-q\right)-x\right)E\right) \geq \\ \left(1-p_L\right)u\left(E\left(1-x\right)\right) \\ +p_L u\left(E\left(1-x\right)-D_{1,2}\left(1-q\right)\right) \end{aligned} \tag{41}$$

Observe that the right hand side of (41) is strictly decreasing in $D_{1,2}$. If $D_{1,2} = 0$ then (41) is not satisfied because the right hand side becomes $u\left(E\left(1-x\right)\right)$. If $D_{1,2} = E$ then (41) must be strictly satisfied because in this case the expected consumption is the same under either contract and is risky under the high type contract. It follows immediately that there exists a $\bar{D}_{1,2} \in (0,E)$ which (41) is satisfied with equality and thus (41) is ensured to hold as long as $D_{1,2} \geq \bar{D}_{1,2}$. By complementary slackness this confirms our conjecture that $\lambda_1 = 0$. The total amount of debt raised by the high type at $t = 1$ is $A_1 = E\left(1-x\right)$ as determined by the break even condition.

B.3. Information Asymmetry Only About Ability to Repay at $t = 2$

Consider the case where $p_L = p_H = p \geq 0$ and $x_L > x_H = 0$. Now (36) and (42) can be re-written as

$$D_{1,2} : \lambda_1 x_L u'\left(c_3^B\right) = \lambda_2 \tag{42}$$

$$\begin{aligned} &D_{1,3} : p\left(1-p\right)\left(1-q\right)\left[u'\left(c_3^M\right)-u'\left(c_3^G\right)\right] \\ &+\lambda_1\left[p\left(1-p-x_L\right)\left(1-q\right)u'\left(c_3^G\right)-p\left(1-q\right)\left(1-p\right)u'\left(c_3^M\right)-x_L\left(1-\left(1-q\right)p\right)u'\left(c_3^B\right)\right] \\ &+\lambda_3 = 0 \end{aligned} \tag{43}$$

To start we prove $\lambda_2 > 0$ by contradiction. Suppose instead that $\lambda_2 = 0$. By (42) this would imply imply $\lambda_1 = 0$. This implies that (43) becomes

$$\lambda_3 = -p\left(1-p\right)\left(1-q\right)\left[u'\left(c_3^M\right)-u'\left(c_3^G\right)\right] \tag{44}$$

Note that $c_3^G \geq c_3^M \iff D_{1,3} \leq E$ and that if one of these inequalities is strict then the other must be as well. If $D_{1,3} = E$ this would recreate the symmetric information contract and the truth-telling constraint must be violated. It follows that $D_{1,3} < E$ which implies $c_3^G > c_3^M$ but this, combined with the strict concavity of $u()$, means that (44) requires $\lambda_3 < 0$ which cannot hold. Therefore by argument on contradiction it must be that $\lambda_2 > 0$ and by complementary slackness this requires $D_{1,2} = 0$ at the optimal contract. Further (42) implies that $\lambda_1 > 0$ at the optimal contract and hence the optimal choice of $D_{1,3}$ must be such that the truth-telling constraint binds with equality. With $D_{1,2} = 0$ the truth-telling constraint is

$$\begin{aligned} & u\left((1-p(1-q)-x_L)E\right) \\ \geq\ & (1-p-x_L)\,u\left(E-(1-q)\,pD_{1,3}\right) \\ & +pu\left(qE+(1-q)(1-p)D_{1,3}\right) \\ & +x_L u\left((1-(1-q)\,p)\,D_{1,3}\right) \end{aligned} \tag{45}$$

Notice that (45) is not satisfied when $D_{1,3} = E$ since this would guarantee the low type agent a higher level of consumption with certainty. Also, (45) is slack when $D_{1,3} = 0$ since this provides the same level of expected consumption and is risky. If we label the the right hand side of (45) as $\Upsilon(D_{1,3})$ then

$$\begin{aligned} \frac{\partial^2 \Upsilon(D_{1,3})}{\partial D_{1,3}^2} = & \ p^2(1-q)^2(1-p-x_L)\,u''\left(c_3^G\right) \\ & +\left[p(1-q)(1-p)\right]^2 u''\left(c_3^M\right) \\ & +x_L\left[1-(1-q)\,p\right]^2 u''\left(c_3^B\right) \quad < 0 \end{aligned} \tag{46}$$

where the strict inequality in (46) follows directly from the strict concavity of $u()$. Since $\Upsilon(D_{1,3})$ is strictly concave, and feasibility requires $D_{1,3} \leq E$, then there must be a unique value of $D_{1,3} \in (0,E)$ for which (45) is satisfied with equality.

B.4. CARA Utility and $\alpha \in [0,1]$

Assume that the household utility function exhibits constant absolute risk aversion:

$$u(c) = 1 - \frac{1}{\eta} e^{-\eta c} \tag{47}$$

where $\eta > 0$ is the coefficient of absolute risk aversion. With this assumption (27) simplifies to give the level of consumption at $t = 1$ that a household of type k will select conditional on accepting a contract of $\{A_1, D_{1,2}, D_{1,3}\}$ as

$$c_1' = \frac{-1}{2\eta} \ln \left[\frac{\alpha}{1-\alpha} \left\{ (1 - p_k - x_k) e^{-\eta A_G} + p_k e^{-\eta A_M} + x_k e^{-\eta A_B} \right\} \right] \tag{48}$$

where

$$A_G \equiv A_1 + E - D_{1,2} - D_{1,3}$$
$$A_M \equiv A_1 + qE - D_{1,2} - qD_{1,3}$$
$$A_B \equiv A_1$$

Substituting (48) into (23) and using this in (22) defines the optimal contracting problem under CARA utility. Numerical solutions to this problem are provided in Figure 11.

C. Pooling Equilibrium

The analysis in the paper has focused on characterizing the debt contracts that will arise in a separating equilibrium. The goal of this sub-section is to briefly argue that this focus has been without loss of generality because pooling equilibrium do not exist as long as out of equilibrium beliefs are reasonable in the sense of the intuitive criteria of Cho and Kreps (1987). Under this criteria if a high type household deviates from a proposed pooling equilibrium to accept a contract that a low type does not prefer to the pooling contract then they will be believed to be high type.

Competition ensures that a pooling equilibrium, if it exists, will only occur at the contract that maximized the expected utility of both types subject to the break-even constraint. Thus a pooling equilibrium, if it were to exist would have all household accepting the following contract:

$$A_1^{Pool} = [\phi (1 - p_H (1-q) - x_H) + (1-\phi)(1 - p_H (1-q) - x_H)] E$$
$$D_{1,2}^{Pool} = 0$$
$$D_{1,3}^{Pool} = E$$

This pooling equilibrium can only survive if there is no other contract $\{A_1, D_{1,2}, D_{1,3}\}$ that (i) would be preferred by high type household and not by a low type and (ii) would allow the lender offering the contract to high type households to at least break even. In fact competition would ensure that the contract which maximized the expected utility of high type households were offered if any such contract exists and thus we can characterize this deviating contract in exactly the same way as the separating contract with the only difference being that the truth telling constraint for the low type

(22) is now

$$U^{Pool} \geq U^{*L'}(A_1, D_{1,2}, D_{1,3}) \tag{49}$$

where

$$U^{Pool} \equiv u([\phi(1-p_H(1-q)-x_H)+(1-\phi)(1-p_H(1-q)-x_H)]E)$$

It must be that the truth-telling constraint (49) will bind at this contract because the optimal contract where this doesn't bind is simply the full insurance contract that arises under asymmetric information and the low type would always strictly prefer this contract. But if the low type is indifferent between both contracts then the high type must strictly prefer this new contract if it is set optimally. To avoid the hidden savings problem and thus allow an analytical characterization of the optimal contract suppose $\alpha = 1$. Take the first example we considered in the paper where $p_L > p_H$ and $x_L = x_H$. As we argued in the paper the first order conditions imply that $c_3^G > c_3^B > c_3^M$. So if (49) binds for the low type then

$$U^{Pool} = (1-p_L-x_L)\,u\left(c_3^G\right)+p_L u\left(c_3^M\right)+x_L u\left(c_3^B\right) \tag{50}$$

but then it must be that expected utility of the high type is strictly higher than this since $p_L > p_H$ and $c_3^G > c_3^M$. So this contract will break the pooling equilibrium. A similar argument applies the more general case where $p_L \geq p_H$ and $x_L \geq x_H$. For each of the numerical solutions for the general case of $\alpha \in (0,1)$ presented in 11 it is also verified that no pooling equilibrium exists by a similar argument - a deviating contract can always be found that the high type strictly prefers.

C. Inferring initial credit risk category (sub-grade) from data

LC assigns each loan's interest rate depending on the credit risk sub-grade. In the data, the variable sub-grade takes one of 35 possible values for each loan: A1, A2, ... A5, B1, ... B5, ... G5. Each grade is assigned a number: A1 = 1, A2 = 2, ... G5 = 35 ranging from least risky to most risky. Each sub-grade is then assigned an interest rate. For example, as of December 2012, A1 loans had an interest rate of 6.03%, while A2 loans had a rate of 6.62%. We take a snapshot of LC's "Interest Rates and How We Set Them" page as of December 31, 2012 from the Internet Archive.[48] According to this page, the borrower's credit risk grade is calculated in the following manner. First, "*the applicant is assessed by Lending Club's proprietary scoring models which can either decline or approve the applicant.*" If an applicant is approved by the model, she receives a Model Rank (an "initial sub-grade"), which can range from A1 (1) through E5 (25). According to the website, *"The Model Rank is based upon an internally developed algorithm which analyzes the performance of Borrower Members and takes into account the applicant's FICO score, credit attributes, and other application data.*" The initial sub-grade is then modified depending on the requested loan amount and maturity. For example, the initial sub-grade of 36-month loans was not modified, while the initial sub-grade of 60-month loans was modified by 4 grades for A borrowers (initial sub-grades 1 to 5), 5 grades for B borrowers (initial sub-grades 6 to 10) and 8 grades for all other grades. The amount modifications are publicly available for each period on LC's website, and vary over time. We choose our main sample period between December 2012 and October 2013 so that these modifications stay constant. For example, between December 2012 and October 2013, the amount modifications for each grade were as follows:

	Initial sub-grade		
	A	B	C-E
<$5,000	1	1	1
$5,000 - $15,000	0	0	0
$15,000 - $20,000	0	0	2
$20,000 - $25,000	0	1	3
$25,000 - $30,000	2	3	4
$30,000 - $35,000	4	4	5
$35,000	6	6	6

According to this table, the initial sub-grade of a borrower who requests a loan for $10,000 is the same as her final sub-grade before the modification for maturity. Instead, a borrower who was ranked initially as C1 (equivalent to 11) who requests a $16,000 loan will see her grade modified two steps to a C3 (13).

[48] See https://web.archive.org/web/20121204222850/http://www.lendingclub.com/public/how-we-set-interest-rates.action

Borrowers who share the same initial sub-grade will have very similar risk characteristics as assessed by LC's lending model, while their interest rate will only vary according to their choice of amount and maturity. Thus, our analysis above uses the initial sub-grade before amount and maturity modifications to construct fixed effects. This variable–initial sub-grade– is not observable in the data. Instead, LC only provides the credit risk sub-grade after all modifications have been made. To re-construct a borrower's initial sub-grade, we reverse engineer LC's credit risk process for every loan in our sample using their publicly available information. For example, a 36-month loan issued on January 2013 for $16,000 that appears in the data as a C4 borrower must have been assigned an initial grade of C2 (2 modifications for the loan amount, no modifications for maturity). The table below documents the fraction of loans on each final sub-grade that we cannot assign an initial sub-grade from our reverse engineering procedure for loans issued between December 2012 and October 2013, for amounts between $5,000 and $20,000:

% of loans in main sample period not assigned an initial sub-grade					
Final sub-grade	%	Total loans	Final sub-grade	%	Number of loans
A1	0.66	1,658	D1	0.73	2,465
A2	0.23	1,292	D2	0.22	2,291
A3	0.36	1,389	D3	0.83	1,814
A4	0.74	1,624	D4	0.33	1,508
A5	1.59	2,383	D5	0.32	927
B1	0.34	5,623	E1	0.30	328
B2	0.78	6,120	E2	0.00	437
B3	0.55	6,399	E3	0.00	208
B4	0.64	6,283	E4	0.60	166
B5	0.65	3,080	E5	0.00	120
C1	11.79	3,884	F1	85.29	34
C2	2.03	2,957	F2	92.31	52
C3	0.56	3,236	F3	100.00	9
C4	0.49	2,830	F4	100.00	9
C5	0.56	2,325	G1	100.00	1

First, by construction, almost all loans below an F1 rating (26) will not have an initial sub-grade because LC's model states that only 25 initial grades are issued. Second, we succeed in matching a borrower's initial sub-grade for more than 98% of the loans of each final subgrade in 24 out of the 25 top sub-grades. Grade C1 (grade 11) is slightly problematic as the success rate drops to 88.2%. The reason for this drop is that, given the algorithm presented above, we should not observe C1 loans between $15,000 and $20,000, but LC categorizes 458 of these loans during our sample period. All our results are robust to eliminating loans issued in final grade C1 and to replacing the initial sub-grade in our regression model with the observed final sub-grade.

D. Pricing of existing loans constant during expansion - evidence

A. Evidence for all 25 risk categories

We verify that interest rates by sub-grade and loan amount remain constant during our sample period. As an expample, Figure 12 shows the schedule of APRs (by loan amount) for sub-grades B1 through B5 measured before and after the menu expansion, for amounts in \$5,000 bins between \$5,000 and \$20,000 (\$15,000 includes all loans issued between \$15,000 and \$20,000), for 36 month loans. The figure shows that interest rates remain unchanged before and after the menu expansion. Formally, we regress APR on sub-grade by \$5,000 loan amount bins (rates remain fixed within \$5,000 bins) for all loans issued during our sample period,

$$APR_i = \delta_{\$5,000 amount \times subgrade} + \varepsilon_i.$$

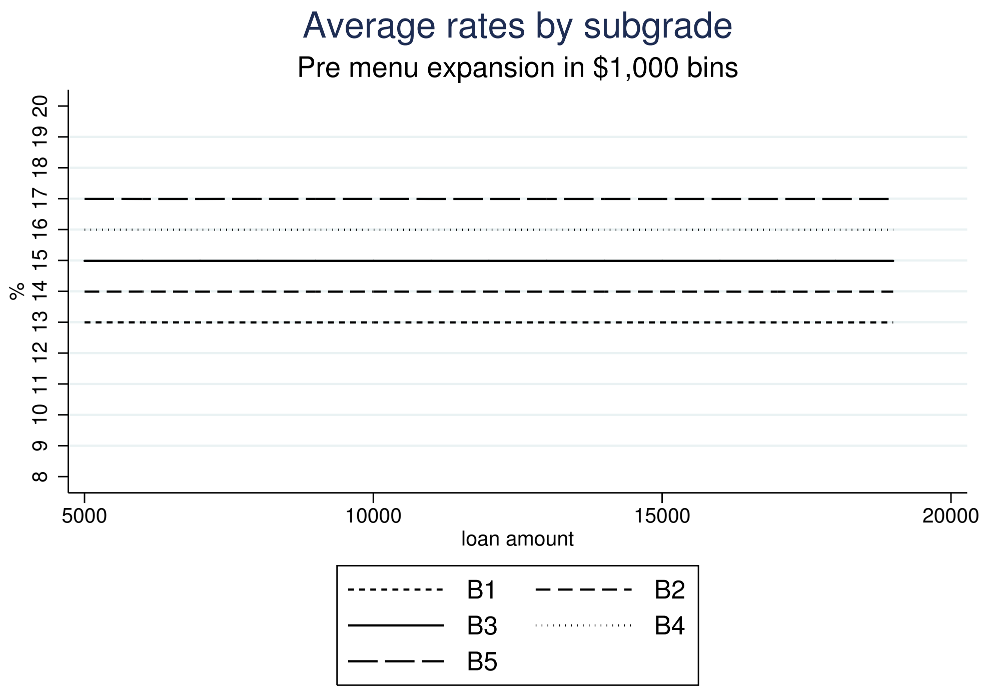
The R^2 of this regression is 99.7%, i.e., only 0.3% of the variation in APRs is explained by other variables that include month of issuance. This implies that the contract terms offered to each borrower remained constant during our sample period.

B. Example - The pricing of B grade 36-month loans

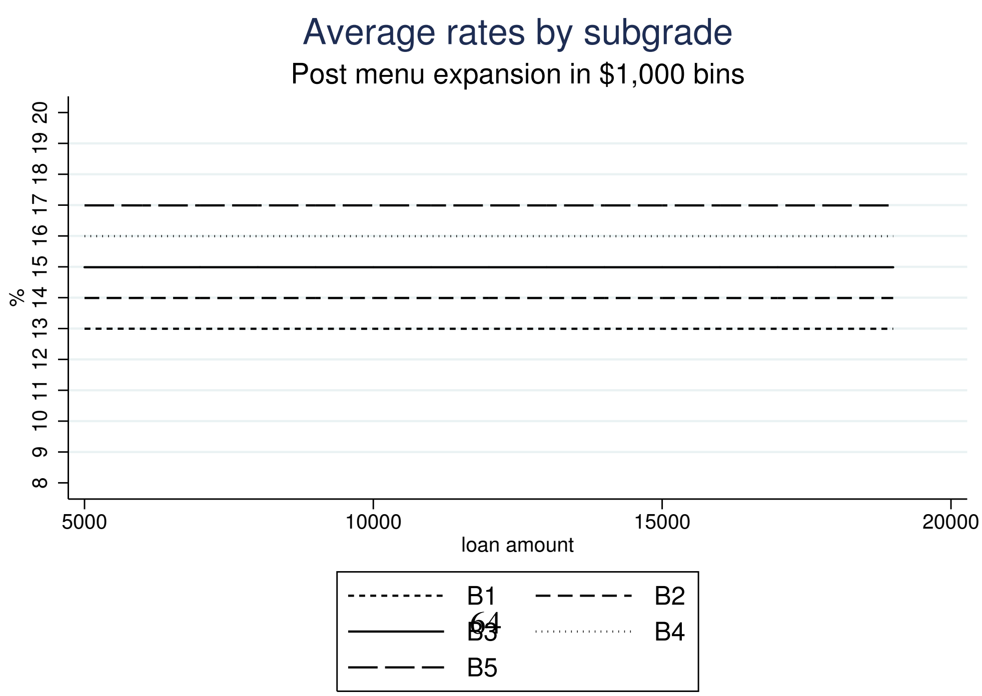
Figure 12: Interest rate schedule for B sub-grades

This figure shows that interest rates, measured as APRs, remain constant by sub-grade and loan amount among B group borrowers for 36 month loans before and after the expansion of the menu of borrowing options. The top panel shows the avergae APR by initial sub-grade for B grade borrowers (B1 through B5) for 36 month loans for loan amounts between $5,000 and $20,000 (labelled up to $15,000, as loans between $15,000 and $20,000 have the same terms) issued on May 2013. The bottom panel shows the same plot for loans issued on November 2013.

Panel A: Interest rate schedule before menu expansion



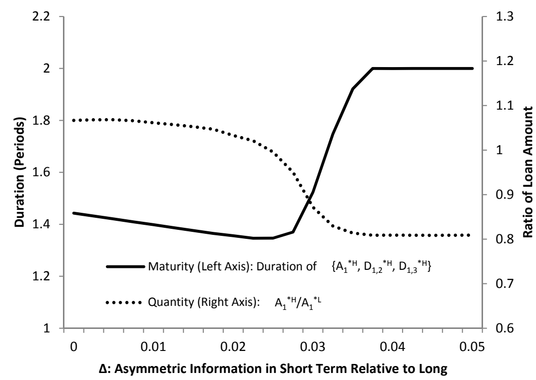
Panel B: Interest rate schedule after menu expansion



E. Comparative statics for theoretical framework

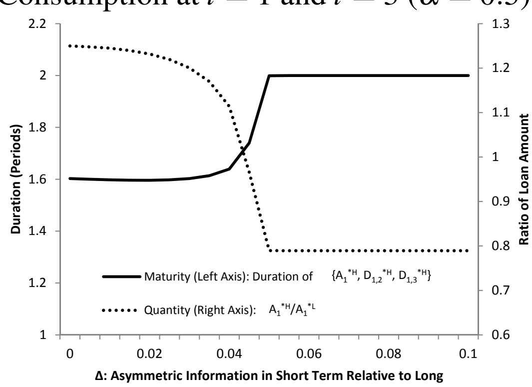
Figure 13: Model Comparative Statics - Robustness

This figure shows additional comparative statics from numerical solutions of the theoretical framework presented in SectionIV in order to demonstrate the robustness of the results in Figure 11. The following parameters are used (identical to Figure 11): $E = 100$, $p_L = 0.3$, $x_L = 0.1$, $\bar{p}_H = 0.1$. Panel A and B continue to use a CARA utility function: $u(c) = 1 - \frac{1}{\eta} e^{-\eta c}$ with $\eta = 0.1$. In Panel A and B the household values consumption at both dates equally: $\alpha = 0.5$. In Panel A $q = 0.75$ and (i.e. the same as in Figure 11) and in Panel B this is lowered to $q = 0.25$. For Panel C the CARA utility function is replaced with a CRRA utility function of $u(c) = \frac{c^{1-\eta}-1}{1-\eta}$ with $\eta = 2$. Otherwise the parameters in Figure C are identical to those in Figure 11: $q = 0.75$ $\alpha = 1$. Thus Panel A varies the concern for consumption at $t = 1$, and Panel B the probability of repayment conditional on $S = M$., and Panel C varies the utility function. The left axis in each panel shows the degree of maturity rationing as captured by the Macaulay duration of the equilibrium loan offered to the high type: $Duration = 1 \times \frac{D_{1,2}^{*H} \times (1-x_H)}{A_1^{*H}} + 2 \times \frac{D_{1,3}^{*H} \times (1-p_H(1-q)-x_H)}{A_1^{*H}}$. The right axis in each panel shows the degree of quantity rationing as captured by the ratio of the amount lent to high and low type borrowers at $t = 1$: $\frac{A_1^{*H}}{A_1^{*L}}$.



Panel A: Consumption at $t = 1$ and $t = 3$ ($\alpha = 0.5$)

Panel B: Consumption at $t = 1$ and $t = 3$ ($\alpha = 0.5$), $q = 0.25$



Panel C: Consumption only at $t = 3$ ($\alpha = 1$), CRRA Utility