ADVERSE SELECTION ON MATURITY: EVIDENCE FROM ON-LINE CONSUMER CREDIT*

ANDREW HERTZBERG†, ANDRES LIBERMAN‡, AND DANIEL PARAVISINI††

ABSTRACT. We provide evidence of adverse selection on maturity in consumer credit. Our estimation compares two groups of observationally equivalent borrowers that took identical 36-month loans, but where only one of the groups is selected on maturity: borrowers chose the 36-month loan when a 60-month maturity option was also available. Borrowers who self-select into short maturity loans default less a year after origination, and have future credit ratings that are higher and less volatile. Consistent with the insurance role of long term credit, the findings suggest borrowers self-select on their exposure to shocks to their future ability to repay.

This version: December 3, 2015

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

† Columbia University. Email: ah2692@gsb.columbia.edu.
‡ New York University. Email: aliberma@stern.nyu.edu.
†† London School of Economics. Email: D.Paravisini@lse.ac.uk.

*We thank Emily Breza, Theresa Kuchler, Adair Morse, Holger Mueller, Philipp Schnabl, Antoinette Schoar, Amit Seru, Felipe Severino, Johannes Stroebel, and participants at Bocconi University, Columbia University, Dartmouth University (Tuck), LSE (Economics and Finance departments), NBER Corporate Finance 2015 Fall meeting (Stanford), NBER Household Finance Summer Institute 2015, NYU (Stern), NYU-Columbia Junior Faculty Seminar, UBC (Sauder), and UNC (Kenan-Flagler). We thank Siddharth Vij for outstanding research assistance. All errors and omissions are ours only.

I. Introduction

Households may become a bad credit risk due to shocks to their ability to repay, such as experiencing job loss, illness, divorce, or inflexible consumption needs. If households finance themselves using a sequence of short term debt contracts, these shocks will increase the price of debt precisely in those states of the world where expected income is lower and marginal utility of consumption is higher. Thus, a fundamental function of long term debt contracts is to allow households to insure against the risk of becoming, or being re-classified as, bad risk over the life of the agreement.

The insurance role of long term contracts is not exclusive to consumer credit markets, as it is also played, for example, by long term corporate debt (Diamond (1991)), long-term employment contracts (Holmstrom (1983)), and long-term health and care insurance (Cochrane (1995), Finkelstein, McGarry, and Sufi (2005)). In theory, adverse selection, which results from households' private information about their risk exposure, may limit the ability of competitive markets to provide insurance through long term contracts.[1] Lenders facing adverse selection will offer a menu of contracts with varying maturities, such that households that are more exposed to shocks self-select into longer maturity contracts that offer more protection, and are priced accordingly. Although many common consumer loan products, such as mortgages, auto loans, and personal loans, offer borrowers a choice over loan maturity, the potential for adverse selection in this contract dimension and the role of maturity as a screening device remain, to date, largely theoretical propositions.[2]

The present paper provides the first evidence consistent with the insurance and screening roles of debt maturity in consumer credit markets. We show that borrowers self-select into long maturity loans when they are privately informed about a high exposure to shocks to their future repayment capacity. The main challenge in measuring how ex ante selection on maturity affects ex post repayment performance stems from the fact that maturity itself affects borrower behavior. Suppose we compared the performance of two observationally equivalent groups of borrowers, the first with a short term loan and the second with a long term loan. The borrowers in the second group could default at a higher rate due to adverse selection, but also due to the effect that a different maturity, installment amount, and interest rate have on the borrowers' repayment behavior. For this reason, identifying empirically the consequences of selection on repayment requires comparing how selected and unselected borrower samples behave when facing the *same* contract.

To illustrate this point and provide a motivation for our empirical strategy below, consider the idealized setting 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

[1] Rothschild and Stiglitz (1976) make this point for insurance markets in general. De Meza and Webb (2001) show there can be advantageous selection when consumers have private information about their strong taste for insurance.

[2] The role of asymmetric information in corporate loan maturity choice was first studied theoretically by Flannery (1986) and Diamond (1991). In Section IV we present a model of maturity choice in the context of consumer lending that formalizes this intuition.

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. 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 on-line lending platform, Lending Club (hereafter, LC), as an empirical setting that closely resembles this idealized setting. 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 are available in $25 increments between $1,000 and $35,000 in either short—36 months—or long maturities—60 months. In January 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, the terms of all other previously available menu items were unchanged within each risk category during this time, and the roll-out was not advertised on the LC website or accompanied by any additional marketing campaign. Borrowers would only notice the new options once they began applying for a loan.[3]

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, which approximate groups A and B of the idealized setting of Figure 1, respectively.[4] Simple before-after comparisons are potentially biased due to time-of-origination shocks. To account for these shocks 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 to construct counterfactuals. And to perform comparisons between observationally equivalent borrowers, our specifications include month-of-origination, risk category, 4-point FICO range, and state fixed-effects.[5]

[3] According to the information reported in the website Internet Archive, LC's website at the time of the expansion advertised that 60 month loans were available for amounts above $16,000 until November 2013.

[4] For example, borrowers choosing a 36-month $10,000 loan *before* July 2013 resemble those in group A of Figure: 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] Our empirical 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. Moreover all terms of the 36-month loans including interest rate remain unchanged during our sample period. This rules out that selection is

We validate the identification assumptions behind our empirical strategy by documenting that the bulk of self-selection into long maturity loans occurred among borrowers that would have borrowed between \$10,000 and \$16,000. We find that the number of short maturity loans between \$10,000 and \$16,000 dropped 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 the same month the 60-month loan appeared in the menu for the corresponding amount.

We find that the average default rate of short maturity loans decreases by 0.8 percentage points and the average future FICO score increases by 2.7 points when a long maturity loan was available at the time of origination. This implies that borrowers with unobservably higher default risk (and risk of suffering a decline in their future credit rating) *self-select* out of short-term loans and into long term ones. The increase in performance of 36-month loans is due to the 14.5% of borrowers that self-select into long maturity, which implies that these borrowers would have had a default rate that is 5.5 percentage points higher (0.8/0.145) than the default rate of the average 36-month borrower in our sample (9.2%). These results indicate 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 who select into long term loans are privately informed about their poor repayment prospects with a short term loan, we attempt to shed light on the underlying private information that drives selection along the maturity dimension. We uncover two facts. First, the future FICO scores of borrowers who self-select into the 36-month loans are less volatile. And second, the decline in the default probability of 36-month loans when the long term loan is available occurs only twelve months after origination, even though, unconditionally, close to half of the loans that default do so within twelve months of origination. This implies that borrowers who select into long term loans have a more volatile credit risk, and their incremental propensity to default (relative to those that self-selec into short-term loans) increases with time since origination. Both pieces of evidence are consistent with the insurance motivation for selection on maturity: borrowers self-select into long maturity loans when they have a higher exposure to future income or consumption shocks.

Borrowers may self-select into longer maturity loans if they face immediate liquidity constraints that they expect will be relaxed in the future. If this were the case, the larger installment amounts of short term loans will induce borrowers to default more in the short run, when the constraints are binding, and less in the long run, when income arrives and the constraints are relaxed. This is the

occurring based on supply side screening decisions or by reverse causality of loan terms impacting default. 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.

[6] Two robustness tests ensure our results are not simply capturing time-of-origination varying differences in creditworthiness for loans of different amounts. First, our results are unchanged when we reduce our sample to a narrower bandwidth around the affected amounts. Second, we test for differences in creditworthiness in the same time period but at other amounts in the borders of the interval of affected amounts. These tests find no evidence of time-varying differences in creditworthiness.

opposite of what we observe in the data: the additional propensity to default of borrowers who select into long term loans *increases* with time since origination. This highlights that not all determinants of maturity choice will lead to adverse selection on maturity, only those on which the borrower is asymmetrically informed. Our results suggest that borrowers are privately informed about their exposure to shocks to their future repayment capacity.

We formalize this intuition in the last section of the paper, where we develop a stylized model of consumer credit choice to discuss the conditions under which maturity is the *optimal* way to screen borrowers when screening using loan amounts is also an option.[7] In the model borrowers have private information about their exposure to adverse shocks in the short and long run, 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 is increasing over time from origination, a condition that we confirm empirically in the data. Intuitively, the reason is that long maturity loans shift payments to the future, when the private information is more informative about the repayment capacity of the borrower **[AH: check whether we want this last sentence or not]**.

Studying the selection response to maturity is important given the widespread choice of maturities consumers face in credit markets. Our analysis has positive implications, as it explains how asymmetric information can determine the menus of loan maturities and prices observed in consumer credit markets. Our results also have normative implications since they highlight the sources of efficiency losses that may result from regulations that impose loan maturity caps, such as Regulation Z in the U.S. mortgage market. Existing empirical work on maturity is silent on these issues because it has focused on how credit demand and repayment change with maturity after controlling for selection, or fully eliminating it through randomized controlled trials (see Karlan and Zinman (2008) and Attanasio, Koujianou Goldberg, and Kyriazidou (2008)). We share with that work the conclusion that loan maturity appears to be a first order determinant of borrower behavior in consumer credit markets.[8]

The rest of the paper proceeds as follows. Section II describes the Lending Club platform and the data, as well as the expansion of the supply of long maturity loans. In Section III we describe our empirical strategy and document that borrowers who self-select into long maturity loans exhibit

[7]This is the case in most unsecured consumer credit applications (see Stiglitz and Weiss (1981), Adams, Einav, and Levin (2009), Dobbie and Skiba (2013). In general, other contract terms that can be used to screen borrowers are collateral (Bester (1987)) and their willingness to agree to joint liability (Van Tassel (1999)).

[8]The literature that documents the importance of adverse selection in consumer credit markets has focused on other dimensions of the debt or purchase contract, such as the price, amount, and down-payment (Ausubel (1999), Adams, Einav, and Levin (2009), Agarwal, Chomsisengphet, and Liu (2010), Dobbie and Skiba (2013), Karlan and Zinman (2009), Rai and Klonner (2009) and Zinman (2014)). Evidence of adverse selection has been provided in other markets. See for example: used cars Genesove (1993), insurance Chiappori and Salanie (2000), real estate Garmaise and Moskowitz (2004), stocks Kelly and Ljungqvist (2012), and the securitized mortgage market Agarwal, Chang, and Yavas (2012).

a higher propensity to default on the short term loan. 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.

II. Setting

A. Lending Club

LC operates in 45 US states and is the largest on-line lending platform in the world. In 2014 LC originated $4.4B in consumer loans. By comparison its nearest rival, Prosper Marketplace, originated $1.6B in the same year.[9] LC loans are unsecured amortizing loans for amounts between $1,000 and $35,000 (in $25 intervals). 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 on 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 Exchange Commission, the agency that regulates on-line 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 pricing the loans.

When a borrower applies for a loan with LC she enters the following information: a non-binding estimate of the amount to be borrowed, yearly individual income, and sufficient personal information to allow LC to obtain the credit report for the borrower. 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. LC only issues loans to borrowers with a FICO score over 660 and a non-mortgage debt to income ratio below 35%. Using a proprietary credit risk assessment model that uses the information in a borrower's credit report (FICO score, outstanding debt) and income, LC assigns the borrower to one of 25 risk categories. This credit risk category determines the entire menu of interest rates faced by the borrower for all loan amounts and for the two maturities. Interest rates for each sub-grade are weakly increasing in amount and strictly increasing in maturity (ceteris paribus). Apart from the interest rate, amount, and maturity, the terms of all loans are identical. Once a borrower selects a loan and it is approved by LC, the application is listed on LC's website for investors' consideration. According to LC, over 99% of all approved loans are funded by investors and investors do not affect any of the terms of the loans.[10] Thus, we ignore the supply side of funds in the analysis. LC charges an origination fee that varies between 1.1% and 5% of the loan amount

[9] Figures reported in the firms' 2014 10K reports.

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

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. Since loan prepayment in LC reduces the number of installments and leaves installment amount unchanged, e.g. prepayment reduces the maturity of the loans, it was impossible for a borrower to synthetically create a 60-month loan of an amount smaller than \$16,000 before the menu expansion. After March 2013 the minimum threshold for a 60-month loan was lowered, first to \$12,000, and later to \$10,000. Figure 2 shows the fraction of loans originated every month that that has 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 is a 60-month loan. 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 III, our empirical strategy exploits the fact that loan amounts between \$12,000 and \$16,000 were affected by the expansion of a long maturity option, and that loan amounts outside this range were not.

We found no evidence that the expansion of 60-month loan options coincided with a marketing campaign or a surge in demand for LC loans. LC did not change its lending policy, including the menu of other loans and the rating algorithm, during the months included in our sample (described in detail in the next subsection). Further, using the Internet Archive website, we verified that LC did not advertise this change or mention it on its own primary web page, which suggests that the characteristics of the pool of applicants did not change after the 60-month menu expansion. We assess this formally by looking at LC's total issuance around the months of the expansions. Figure 3 plots the total dollar amount issued by month. There are no obvious changes in the trend of growth around the dates of the two 60-month loan expansions.

C. Summary statistics

LC's dataset is publicly available on its website. Our main analysis is conducted using data downloaded as of August 2014. We complement this data with an update on loan performance as of April 2015, which we match to our main analysis dataset with the use of the unique loan ID number. We use this update on loan performance to calculate our main outcomes, default and FICO score. The data is a cross section of loans where the 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). Finally, we complement our main outcomes, which are measured as of April 2015, with measures of FICO score obtained from two previous loan performance updates, August 2014 and December 2014.[11]

We select our main sample period such that LC's lending policies remain constant during the period. Based on an Internet search and on our analysis of the data, we found that LC changed the model it uses to assign a borrower's risk category (sub-grade) in December 2012 and in October 2013.[12] Hence, we limit our sample period to all 36-month loans with a list date (variable list_d) between and including these two months. We further limit the sample of loans to include those for amounts between $5,000 and $20,000 because the interest rate schedule jumps discretely at $5,000 and $20,000 for all credit risk categories (sub-grades).[13] This amount interval includes loan amounts affected by the 60-month threshold reduction ($10,000 to $16,000) as well as amounts above and below this interval to control for time-of-origination shocks to creditworthiness and credit demand. Finally, we limit our sample further to those loans where we can infer the borrower's risk category from the publicly available data, which account for 98.6% of all loans in the sample period (we drop the observations for which reverse engineering the risk category from the reported modified sub-grade, the loan amount and maturity does not yield a unique value). Our final sample has 60,514 loans. [14]

Table 1, Panel A, presents summary statistics for the 12,091 36-month loans issued by LC during the pre-expansion period in our sample, that is, 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" 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 these loans are in default as of April 2015. Figure 4 shows the hazard default rate by number of months since origination for loans in our sample issued in the pre-expansion period.[15] 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

---

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

[12] 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.

[13] In some placebo tests we shift our sample to loans issued between July 2013 and May 2014. 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.

[14] See Appendix C for details on this reverse-engineering procedure.

[15] The date of default is determined by the first month when a borrower failed to make a payment.

that the FICO score has on average decreased to 685 approximately one year later.[16] 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 use 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.[17] Thus, LC selection criteria imply that the analysis sample is drawn 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.

III. Measuring Selection On Loan Maturity

We exploit the staggered menu expansion of 60-month loans during 2013 to identify selection of borrowers into long maturity loans. 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. Also, crucially, LC's risk algorithm did not change over our sample and so none of the loan terms for the menu items that existed prior to the expansion, in particular interest rate, were altered. Thus, our empirical strategy compares the outcomes of borrowers who took the short term loan before and after the menu expanded, which proxy groups A and B in our idealized experiment.

Since the expansion in the menu of options occurred for all borrowers at the same time, time of origination-varying differences in creditworthiness or credit demand would make borrowers who were exposed to the menu expansion different from borrowers who were not. We address this potential problem using a difference-indifferences approach: we compare the change in the default rate of borrowers in the affected ("treated") amounts, $10,000 to $16,000, before and and after the expansion, to the same difference for borrowers at "control" loan amounts, which are amounts just above and just below the affected range: $5,000 to $10,000 and $16,000 to $20,000. Since the menu expansion was staggered, we can also use only eventually-treated amounts to build a counterfactual, e.g. use loan amounts between $10,000 and $12,000 before July 2013 as a control

[16] This "last FICO score" variable is updated every time LC discloses new information except for borrowers who have fully paid their loans or who have been charged off.

[17] We calculate this figure as the average of the pre-period revolving balance divided by the pre-period revolving utilization, for borrowers in our main sample.

for loan amounts between \$12,000 and \$16,000. Our identifying assumption is that time varying shocks to observationally equivalent borrowers are the same for control and treated amounts.

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

$$outcome_i = \beta_i^{1000bin} + \delta_i^{jt} + \gamma \times D_i + X_i + \varepsilon_i, \tag{1}$$

where data is at the loan level i. The outcome measures include $default_i$, a dummy that equals one if the loan is late by more than 120 days, and $FICO_i$, the high end of borrower's FICO score 4 point bin, both variables measured as of April 2015.[18] We also report regression results for $fullypaid_i$, a dummy that equals one if the loan has been prepaid as of April 2015. In all our regression, standard errors are clustered at the sub-grade level (25 clusters).

The main explanatory variable of interest, D_{it}, 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:

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

The coefficient of interest, γ, measures the change in the outcome of 36-month loans for affected amounts before and after the expansion of the menu options, relative to the change of the outcome for control amounts that were not affected by the menu expansion. We include granular month of origination $(t) \times$sub-grade (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 sub-grade). We also include a vector of control variables observable at origination, X_{it}. 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 include, for example, 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, among others (more than 58 variables).

While our analysis focuses on the loan menu expansion that occurred at LC, it is important to note that borrowers potentially had access to consumer credit loans with other intermediaries. As one example, Prosper Marketplace, LC's largest rival, offered three- and five-year maturity loans during our sample period. To the extent that credit markets are perfectly competitive, the availability of other five year loans will bias the estimates towards finding no effect of the expansion, since borrowers who wish to select long term loans would already be taking them elsewhere. However, we

[18] Our main dataset corresponds to the LC update as of September 2014. We merge these data with the latest default information using the unique ID variable. We also define a borrower to be in default if she is reported as in a "payment plan". Our results are robust to not including these borrowers as in default.

do not need to control explicitly for the entire set of loan options available to households elsewhere. As long as the outside borrowing options did not change at the same time as the LC menu expansion and did not target the treated loan amounts ($10,000 to $16,000), then our empirical strategy will fully account for the influence of the competitive environment. We are not aware of any such change elsewhere in the consumer credit market during our sample period. In effect, any impact of the menu expansion at LC can also be interpreted as evidence that consumer credit markets are imperfectly competitive. This might be true because some intermediaries have a technology advantage over others which generates some market power or because there are search frictions in the market.[19]

A. Identification Tests

A second identifying assumption behind our empirical approach is that the loan amount choice is inelastic to loan maturity in this setting. If this is not the case, the difference-in-difference estimate will be biased towards zero since some of the "control" borrowers are in fact "treated." Control borrowers may be treated both before and after the menu expansion. Let's consider first the possibility that control amounts are treated before the menu expansion. As an example, consider borrowers that 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.[20]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 —36-month—maturity 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.[21] If, on the contrary, borrowers choose the second option, e.g. take a loan on the 60-month maturity but for a larger amount, then our difference-in-differences estimate will find a zero. The reason is that these borrowers will not be either in the treatment or in the control group of loans in our estimation (since our estimation is based exclusively on the outcomes of 36-month loans, selection from one long term loans to another will not affect our estimates).[22]

Now consider the second case: where the control 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

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

[20]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 they are based exclusively on the behavior of 36-month loans.

[21]One way to test for whether control amounts are treated by selecting a $10,000 short term loan before the menu expansion is to look for evidence of bunching at the borders of the treated interval. The top panel in Figure 5 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.

[22]The bottom panel in Figure 5 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.

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 treatment and control loans will by biased towards zero.

We explore formally whether the expansion induced selection on maturity at the treated amounts. To do so we collapse the data and count the number of loans N_{jkt} at the month of origination $(t) \times$sub-grade $(j) \times$\$1,000 loan amount bin (k) level for all 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)..[23] We define a "treatment" 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:

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

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

$$log(N_{jkt}) = \beta'_k + \delta'_{jt} + \gamma' \times D_{kt} + \varepsilon_{jkt}. \tag{2}$$

The coefficient of interest is γ', the average percent change in the number of short maturity loans originated for affected amounts (i.e., amounts in which a long maturity loan was introduced as a option) 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, sub-grade$\times$month fixed effects, δ'_{jt}, control for the terms of the contract offers.[24]

Table 2, column 1, shows the results of regression (2), 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 the magnitude 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.[25]

For the reasons explained earlier in this subsection, this estimate represents a lower bound on the selection effect on the number of 36-month loans in the \$10,000 to \$16,000 range induced by the introduction of the 60-month loan option. We can investigate further whether the expansion induced

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

[24] Results are qualitatively and quantitatively unchanged by collapsing the data at the month of origination $t \times k$ level instead and not including δ'_{jt} fixed effects.

[25] Standard errors for estimates of equation (2) 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 sub-grade level (25 clusters), the standard error of the coefficient γ' in Column 1 of Table 2 is 0.028.

selection away from 36-month loans in amounts above or below the range of amounts in which the menu expansion occurred. We compare the evolution of the number of 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 implement this by estimating difference-in-differences regressions with a modified definition of the treatment dummy D_{kt} to equal 1 one after March 2013 or July 2013 for different loan amounts as explained below.

First, using the subsample of 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 loans in the control group were affected by the expansion. The coefficient on the interaction term is 5.9% and not significantly different from zero (Table 2, column 2). This result confirms 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.

We repeat the exercise at the \$10,000 amount threshold. 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 3). 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 Column 1, 2 and 3 confirm our conjecture that the bulk of any selection to longer maturity loans induced by the expansion of the menu was in the treated amounts.

Our identification strategy also rests on the assumption that in the absence of the menu expansion there would be no difference in the change in origination of 36-month loans between treated and control amounts after March 2013 and July 2013. We test for differences in pre-trends by running an amended version of (2) using a series of dummies that become active τ months after a 60-month loan is offered at each amount. Formally, we define:

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

and we run the following regression:[26]

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

Figure 6 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 (2) 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 4 of Table2 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.

B. Selection and Repayment Behavior

We now measure how the default probability of 36-month loans changed due to the selection induced by the menu expansion to include longer term loans. Table 3, columns 1 and 2, reports results of regression (1) using the default dummy, $default_i$, as the dependent variable. 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% lower than the default rate of the non-selected borrowers (column 1), and the magnitude is robust to the including as additional controls every single variable observable at origination in LC's dataset (column 2).

This decline in the default probability is due to the borrowers that self-select into the long term loan, which we estimated to be roughly 14.5% of the borrowers in the not-selected sample (Table 2, column 1). Combining the two estimations 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,

[26]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).

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 sub-grade 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.

The results using the future FICO score of borrowers as the dependent variable (Table 3, columns 3 and 4) echo the ones on default: the positive γ estimates indicate that borrowers that self-select into the 36-month loan when the 60-month loan is available have a lower future expected default probability than those that self-select into the 60-month loan. The coefficient estimate implies that, on average, borrowers who self-select into the short term loan have a future FICO score that is 2.7 points higher. 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. Since a borrower's FICO score aggregates repayment on a borrower's entire set of liabilities, the results indicate that borrowers are not able to undo the damage caused by defaulting on LC on their future credit rating by changing their repayment behavior elsewhere.

C. Robustness

We present in Table 4 several tests that demonstrate the robustness of our results. First, Columns 1 and 2 of Table 4 present counterparts to our main results in Table 2, Columns 1 and 2, 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 estimates using default as the outcome are 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 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 treated amounts. The results in column 3 and 4 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 (e.g., on a 36-month loan).

D. 60-month Loan Performance

In order to isolate the effect of selection our analysis has so far focused on the propensity to default holding the terms of the contract constant, on 36-month loans. Thus, our analysis tells us what the default probability of borrowers that 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 given that they chose a 60-month loan. One of the main reasons for this is that 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.

Notwithstanding this problem we can provide suggestive evidence by comparing the average default rate 36-month and 60-month loans that have the same measured expected default risk (initial sub-grade 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. These stylized facts are suggestive that borrowers who select longer maturity loans are riskier even under the long term contract. This stylized evidence and conclusion are consistent with the findings in 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.

E. Screening on Maturity

Our empirical analysis benefited from the natural experiment created by LC's decision to expand the availability of long term loan contracts without changing any of the characteristics of the short term 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.[27] 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 into long maturity loans.

This provides suggestive evidence that LC, through offering a menu of short and long term maturity loans, managed to sort borrowers into high and low default risk without relying on observable borrower characteristics. In equilibrium, low risk borrowers are better off because they pay a lower interest rate on their credit. By revealed preference, high risk borrowers cannot be worse off since they could have chosen the short term loan at a lower rate and they did not. According to our suggested interpretation of the results, riskier borrowers' willingness to pay for the long term contract is higher thanks to the insurance service it provides. In the next subsection we provide further evidence consistent with this interpretation.

F. Private Information about What?

We now turn to the question of what is the nature of the private information borrowers hold that drives the selection decision. We argue that borrowers self-select into longer maturity loans when they privately observe that they are more exposed to shocks to their ability to repay. This could stem from a greater risk of income fluctuations or shocks to unavoidable expenses such as medical bills. There are several plausible alternative hypotheses. For example, borrowers may have private information about when in the future they will have income available to repay, about their intention

[27]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 sub-grade, 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.

to default strategically, or about the severity of their current liquidity constraints. We discuss these alternatives in more detail below.

It is worth highlighting the challenges involved in determining the sources of adverse selection. Indeed, by definition, the analysis cannot be based on variables that are observable by the lender at the time of origination, such as type of job. If certain jobs have a higher income volatility than others, one would expect lenders to take the job type information into consideration when pricing the loan. Adverse selection can only occur due to private information. In the absence of variables that are observable by the econometrician but not by the lender, the analysis must hinge on differences in borrowers' ex-post behavior. In the analysis below we focus on two such ex-post measures: the volatility of the borrower's FICO score and the timing of default.

Recall from the motivation in the introduction that long term credit provides insurance to borrowers from being reclassified as bad risk in the future due to exposure to shocks to their ability to repay. We can 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.[28] The cross sectional average and standard deviation of this measure for loans in our sample that were issued in the three pre-period months are 24.5 and 19.1, respectively.

The point estimate in Column 1 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 represents direct evidence of the insurance rationale behind adverse selection: borrowers who select long maturity loans are (unobservably) more exposed to reclassification risk. In contrast, this finding is difficult to reconcile with borrowers self-selecting on the basis of private information on when in the future they will have income available to repay or on their intention to default strategically.

Another dimension of the borrowers' ex post behavior that is useful to distinguish between the potential explanations is the timing of default. As we verify formally in the final section of the paper, long term loans provide greater insurance against income or consumption shocks, the further in the future these shocks may occur. Thus, borrowers who have private information about the volatility of income and consumption shocks are more likely to self-select into long-term loan term loans when these shocks occur further in the future. In contrast, borrowers that default strategically will tend to default as early as possible. Also, borrowers that choose long term loans because they know all their income comes in the future may be forced into early default when they take a short term loan (because short term loan have larger installments).

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

To explore the relationship between maturity self-selection and the timing of default, 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 2 and 3 of Table 5. Column 2 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 4), 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 3 shows that the differential propensity to default is present at the 24 month horizon from origination.

We present in Figure 9 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.[29] 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.[30]

As argued above, the results on the timing of defaults is consistent with self-selection occurring due to a desire by borrowers to insure against future shocks to their ability to repay. In contrast, it is not easy to reconcile with borrowers self-selecting on the basis of their intention to strategically default or their private knowledge about the timing of their income. The result also rules out interpretations of the result related to immediate liquidity constraints. Longer maturities imply lower installment amounts, which could in turn improve the repayment behavior of borrowers that are currently facing severe liquidity constraints. Constrained borrowers that are privately informed about the severity of their current constraints could choose to self-select into long term loans, and thus avoid being pushed into immediately into default by larger installments amounts. However, long term loans do not appear to alleviate immediate constraints for the self-selecting borrowers, as the difference between selected and non-selected borrower default rates occur over a year after origination. Note that this does not say anything about the effect of lower installments on the default probability of constrained borrowers. It only highlights that maturity selection in this market is not driven by private information on the severity of liquidity constraints. One potential explanation is that there is sufficient public information available to lenders about a the current severity of a

[29] 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 treatment and control amounts in the same way and does not affect the identification strategy.

[30] 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).

borrower's constraints (e.g., it is relatively easy to verify if a borrower is currently employed and their salary). If this is the case, then there is little scope for adverse selection along this margin.

IV. When is Screening on Maturity Optimal?

Our results imply that maturity choice *may* 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 these papers, screening is achieved because bad credit risks are unwilling to incur the higher transaction costs Flannery (1986) or increased chance of firm liquidation (Diamond (1991)) that comes with short-term debt. 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 amounts. The purpose of this final section of the paper is to explore these questions. 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 to allow lenders to screen borrowers using both loan maturity (as per Flannery (1986) and Diamond (1991)) and loan amount (as per Stiglitz and Weiss (1981)). In unsecured consumer credit markets these are the two primary contract dimensions available to creditors. This allows us to derive a testable condition for when maturity will be, in equilibrium, the optimal way to screen borrowers as opposed to loan amount. Then, we discuss evidence that supports this condition in our setting and provide additional evidence that LC is using maturity as a screening device.

A. Setup

The time-line of the model is shown in Figure 7. 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 $t = 1$.[31] 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 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

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

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 $x_L \geq x_H$. [32]

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: $\{A_1, D_{1,2}, D_{1,3}\}$. $A_1 \geq 0$ 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: $\{A_2, D_{2,3}\}$ 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.

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 captures a borrower's private information about their exposure to long-term shocks to their ability repay 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 information about their exposure to short term shocks to their ability to repay. The distinction between the two is central to the analysis which follows.

[32] 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 $t = 2$.

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 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 $t = 2$.[33] All households receive full insurance that is provided by the default feature of the debt contract.[34] 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 $t = 1$.[35] 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 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 numerically solutions to show that the findings are robust to a wider set of cases. Applying the logic of Rothschild and Stiglitz (1976) we know that 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.

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

We compare two extreme cases of the model analytically (all derivations in Appendix B). First, consider the case where there is only asymmetric information about a borrower's exposure to shocks to her ability to repay 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: $A_1^{*H} = E(1 - x_H)$, $D_{1,2}^{*H} + D_{1,3}^{*H} = E$,

---

[33]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 $A_1^{**k} = E(1 - p_k(1-q) - x_k)$, $D_{1,2}^{**} = 0$, and $D_{1,3}^{**} = E$.

[34]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).

[35]In the 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.

$D^{*H}_{1,2} \geq \bar{D}_{1,2} \in (0,E)$ and $D^{*H}_{1,3} \geq 0$. In words, optimal screening is achieved by any contract that has a sufficiently high repayment at $t = 2$ when there is asymmetric information about a borrowers ability to repay. 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 there is only asymmetric information about a borrowers ability to repay 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 $t = 2$: $D^{*H}_{1,2} = 0$, $D^{*H}_{1,3} \in (0,E)$, and $A^{*H}_1 = (1-(1-q)\,p)\,D^{*H}_{1,3} < (1-(1-q)\,p)\,E$. In this case screening is achieved by through quantity rationing - high type borrowers chose a long maturity loan that has no repayment die at $t = 2$ for an amount that is lower than the present value of their expected income. Contrasting both cases, we see that screening is optimally achieved by targeting repayment towards the horizon from origination when 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 borrowers 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.

C.2. Numerical Solution

Next, we solve the model numerically and conduct a comparative static exercise that varies the relative degree of information asymmetry about short and long-term creditworthiness. 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 $\Delta \in [0,(1-q)\,(p_L - \bar{p}_H)]$.[36] Our comparative statics exercise will show how the optimal contract offered to the high type varies with Δ. Increasing Δ lowers the degree of information asymmetry about a borrowers ability to repay at $t = 3$ $(p_L - p_H)$, and raises the degree of information asymmetry about a borrowers ability to repay at $t = 2$ $(x_L - x_H)$. 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.

Panel A of Figure 8 presents the comparative statics of the equilibrium lending contract varying Δ. The left axis measures the Macaulay duration of the contract offered to high type households at $t = 1$ (the solid line).[37] 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

[36] 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.

[37] Formally: $Duration = 1 \times \frac{D^{*H}_{1,2}(1-x_H)}{A^{*H}_1} + 2 \times \frac{D^{*H}_{1,3}(1-p_H(1-q)-x_H)}{A^{*H}_1}$.

when the household's private information has more power to predict their ability to repay 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 loans value is repaid at $t = 2$. When the degree of information asymmetry is higher with regard to the ability to repay in the short term credit 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.

The comparative statics presented in Figure 8 are robust to the parameter assumptions used to derive those numerical solutions. We demonstrate this in the Appendix in Figure 11. We solve the model allowing consumption at $t = 1$ by setting $\alpha < 1$ (Panel A), by varying the choice of q (Panel B) and to the choice of utility function (Panel C). The comparative statics are qualitatively the same in each case.

Our theoretical framework demonstrates that loan maturity can be used to screen borrowers thereby rationalizing our finding that borrowers who are more exposed to shocks to their creditworthiness will self select into long maturity loans. The novel insight of the model is to show that maturity rather than quantity will be the optimal screening device when the power of their private information to predict their ability to repay is increasing in the time from origination. We now use our empirical strategy to test this condition.

D. Is Screening on Maturity Optimal?

We finish by interpreting our empirical results under the light of the framework to address the question of whether, in the LC setting, screening on maturity is optimal. Our theoretical framework demonstrates that maturity is the optimal way to screen borrowers when the power of their private information to predict their ability to repay is increasing in the time from origination. We showed evidence consistent with this condition in the previous section: borrowers who self-select into long term loans have no differential propensity to default within the first year of the loan relative to those that self-select into short term loans. The entire difference in the default probability of these borrowers occurs after 12 months of origination (see Table 3). The linear pattern in Figure 9 suggests that there is a constant difference in hazard rate of default between high and low type borrowers.

We return to our model and ask which contract feature will be used for screening under this condition. To do this we hold constant the hazard rate of default for each type by setting $p_k = \frac{(1-x_k)x_k}{(1-q)}$. Panel B of Figure 8 presents the comparative statics of the equilibrium lending contract varying the difference in the constant hazard rate of default for each type. The results indicate that both maturity and quantity rationing are optimally be used to screen borrowers. More rationing of both is required as the degree of asymmetry increases. The numerical results suggest that the bulk of screening is optimally achieved through maturity. For example when the difference in the hazard rate of default between types is 5% the high type takes a loan that is 13.4% smaller than that taken by the low type and has two thirds of its value repaid at $t = 2$ with only a third repaid at $t = 3$.

V. Conclusions

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 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 allows thinking about some of the welfare implications of 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 on-line 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, Yan Chang, and Abdullah Yavas, 2012, Adverse selection in mortgage securitization, *Journal of Financial Economics* 105, 640–660.

Agarwal, Sumit, Souphala Chomsisengphet, and Chunlin Liu, 2010, The importance of adverse selection in the credit card market: Evidence from randomized trials of credit card solicitations, *Journal of Money, Credit and Banking* 42, 743–754.

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.

Attanasio, Orazio P, Pinelopi Koujianou Goldberg, and Ekaterini Kyriazidou, 2008, Credit constraints in the market for consumer durables: Evidence from micro data on car loans*, *International Economic Review* 49, 401–436.

Ausubel, Lawrence M, 1999, Adverse selection in the credit card market, Discussion paper working paper, University of Maryland.

Bester, Helmut, 1987, The role of collateral in credit markets with imperfect information, *European Economic Review* 31, 887–899.

Chiappori, Pierre-André, and Bernard Salanie, 2000, Testing for asymmetric information in insurance markets, *Journal of political Economy* 108, 56–78.

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.

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.

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.

Garmaise, Mark J, and Tobias J Moskowitz, 2004, Confronting information asymmetries: Evidence from real estate markets, *Review of Financial Studies* 17, 405–437.

Genesove, David, 1993, Adverse selection in the wholesale used car market, *Journal of Political Economy* pp. 644–665.

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

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

Karlan, Dean S, and Jonathan Zinman, 2008, Credit elasticities in less-developed economies: Implications for microfinance, *The American Economic Review* pp. 1040–1068.

Kelly, Bryan, and Alexander Ljungqvist, 2012, Testing asymmetric-information asset pricing models, *Review of Financial Studies* 25, 1366–1413.

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.

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*.

Rai, Ashok, and Stefan Klonner, 2009, Adverse selection in credit markets: Evidence from a policy experiment, Discussion paper.

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.

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.

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

Van Tassel, Eric, 1999, Group lending under asymmetric information, *Journal of Development Economics* 60, 3–25.

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

Zinman, Jonathan, 2014, Household debt: Facts, puzzles, theories, and policies, .

Appendix

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.

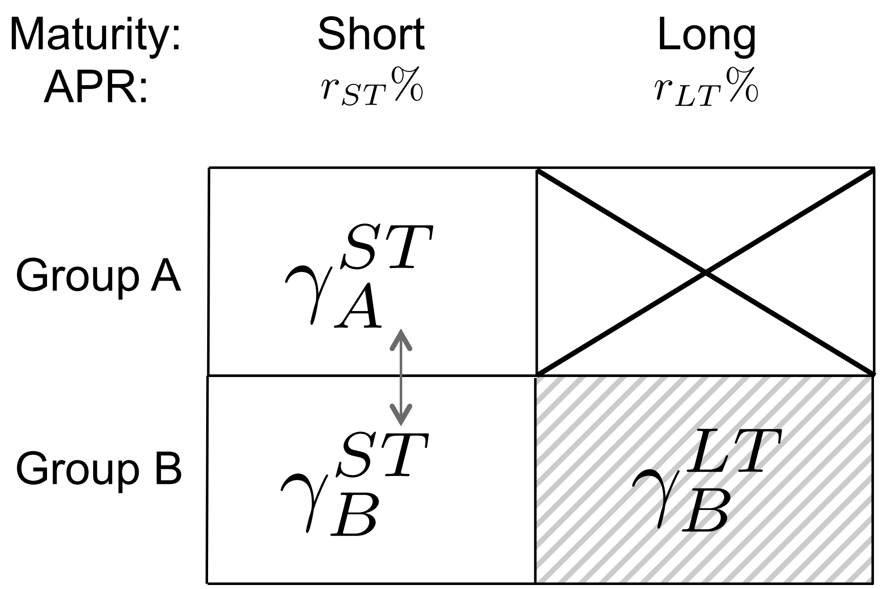


FIGURE 2. Staggered expansion of 60-month loans

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.

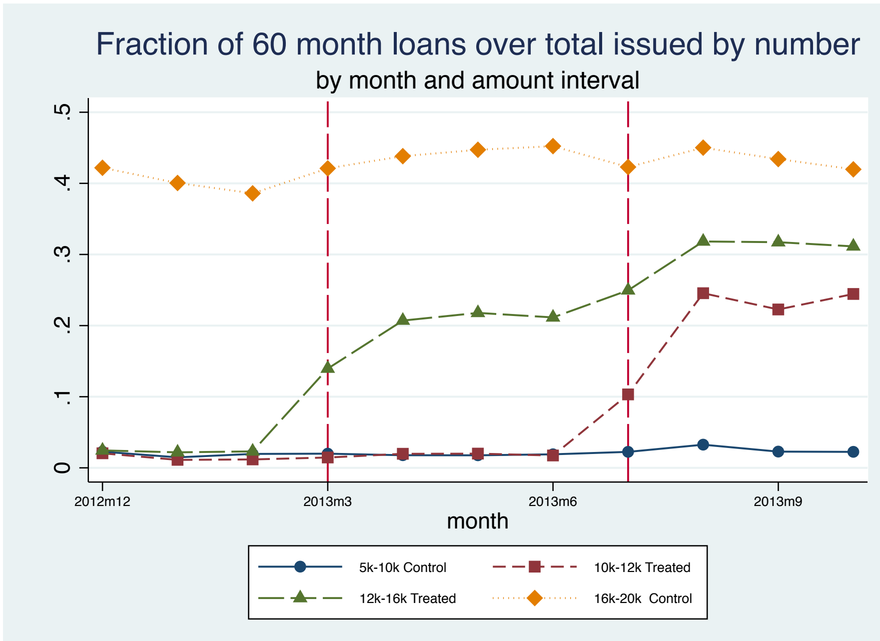


FIGURE 3. Total $ amount issued by LC by month of listing

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.

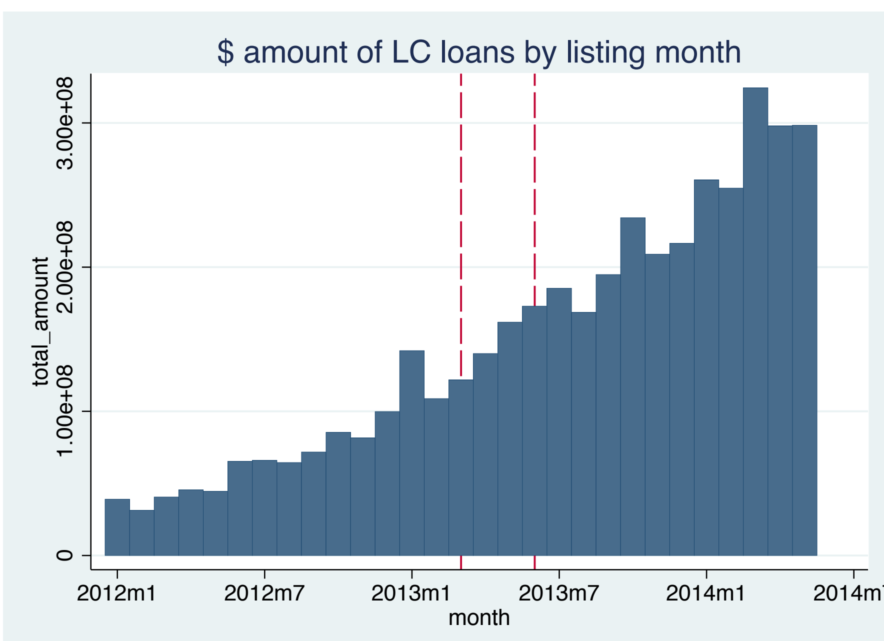


FIGURE 4. Hazard rate of default

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 $t-1$.

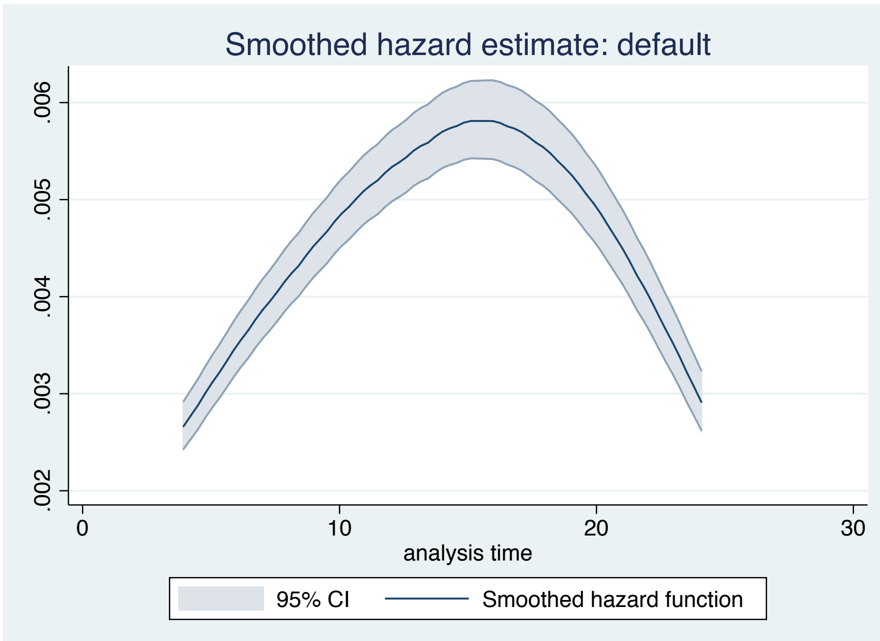
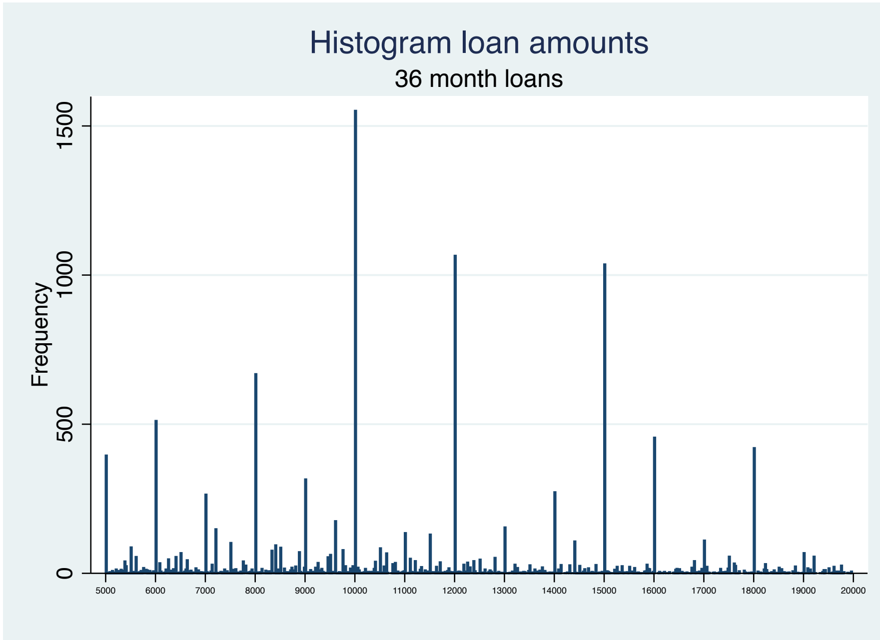


FIGURE 5. Pre-period loan amount histogram

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.



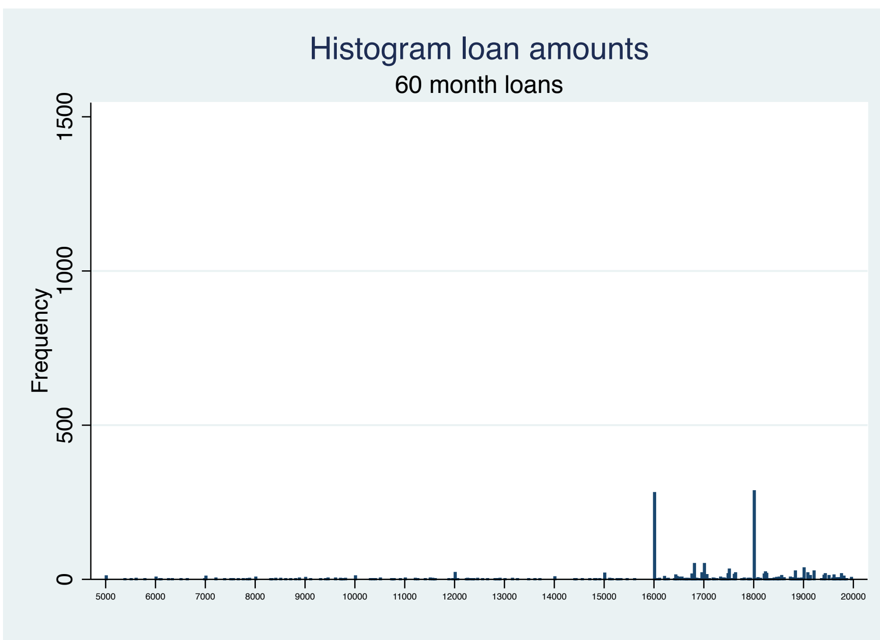


FIGURE 6. Pre-trends on number of loans originated

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 treated and control amounts τ months after the threshold expansion. Standard errors are robust to heteroskedasticity.

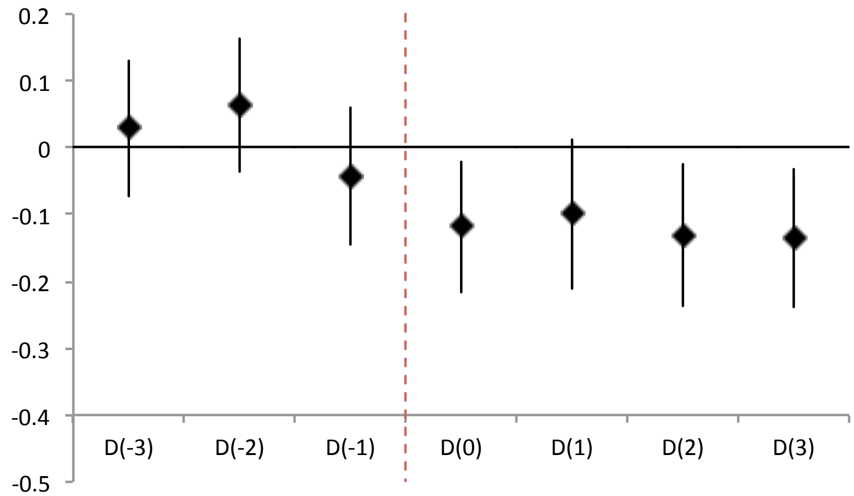


FIGURE 7. Model Time-line

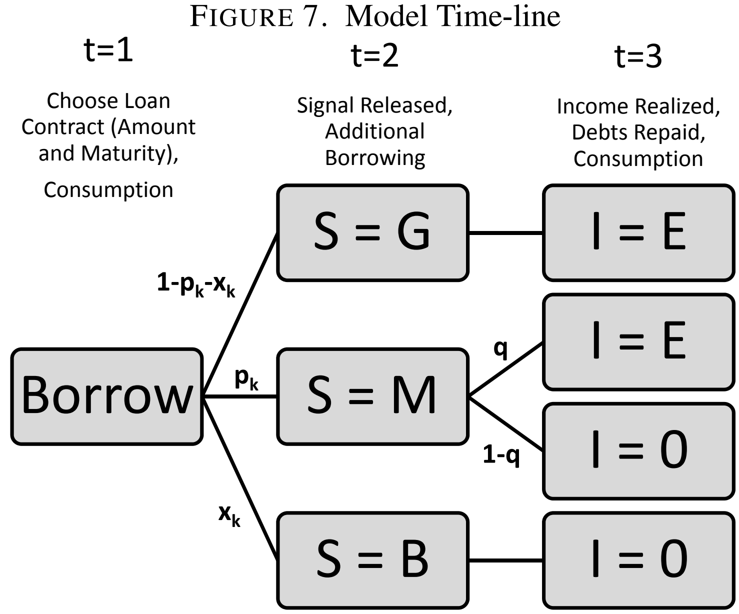
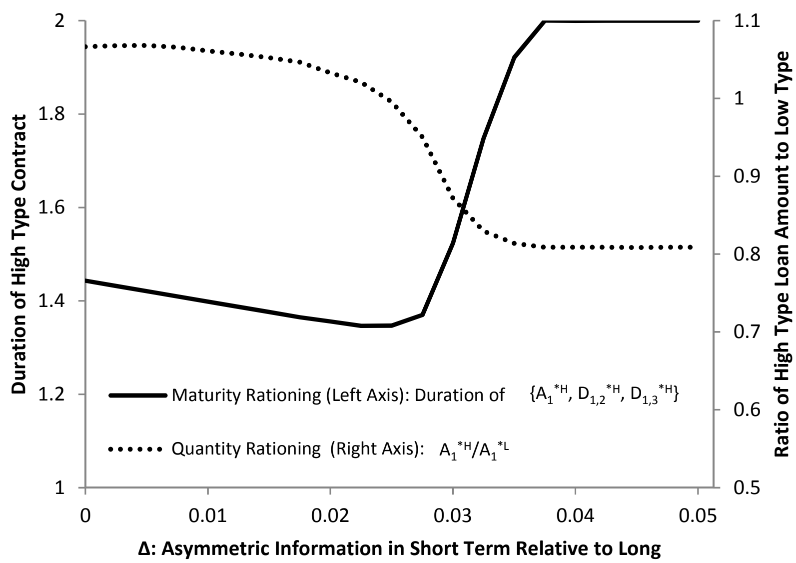


FIGURE 8. Model Comparative Statics

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. In Panel A $E = 100$, $p_L = 0.3$, $x_L = 0.1$, $q = 0.75$, $\bar{p}_H = 0.1$, $\eta = 0.1$ and$\alpha = 1$. The equilibiurm contract is shown for varying values of Δ. In Panel B $E = 100$, $x_H = 0.1$, $q = 0.75$, $\bar{p}_H = 0.1$, $\eta = 0.1$ and$\alpha = 0.5$. The equilibrium contract is shown from valrying values of x_L. For both types $p_k = \frac{(1-x_k)x_k}{(1-q)}$ so as to maintain a constant hazard rate of default at each horizon (i.e. equal to x_k). 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: $Duration = 1 \times \frac{D^{*H}_{1,2}(1-x_H)}{A^{*H}_1} + 2 \times \frac{D^{*H}_{1,3}(1-p_H(1-q)-x_H)}{A^{*H}_1}$. 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 $t = 1$: $\frac{A^{*H}_1}{A^{*L}_1}$.

Panel A: Varying the Degree of Information Asymmetry from Time of Origination



Panel B: Constant Hazard Rate of Default for Each Type

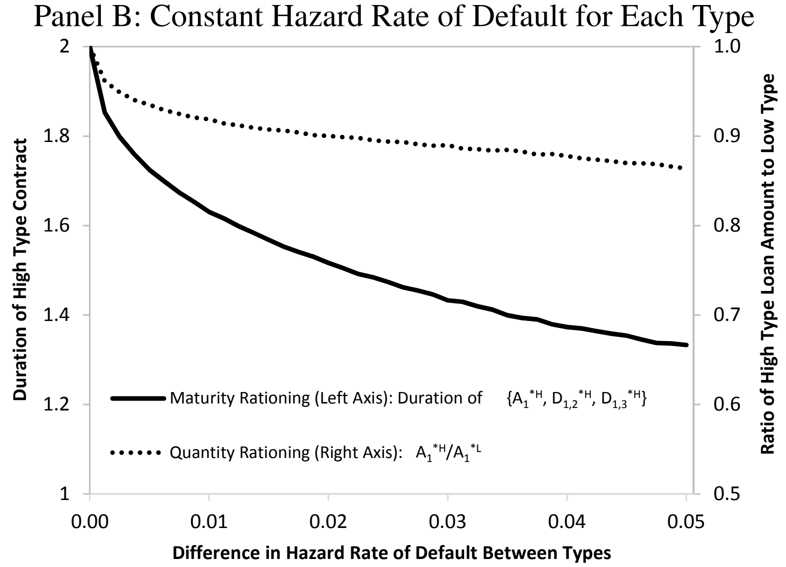


FIGURE 9. Default rate coefficient by number of months since origination

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.

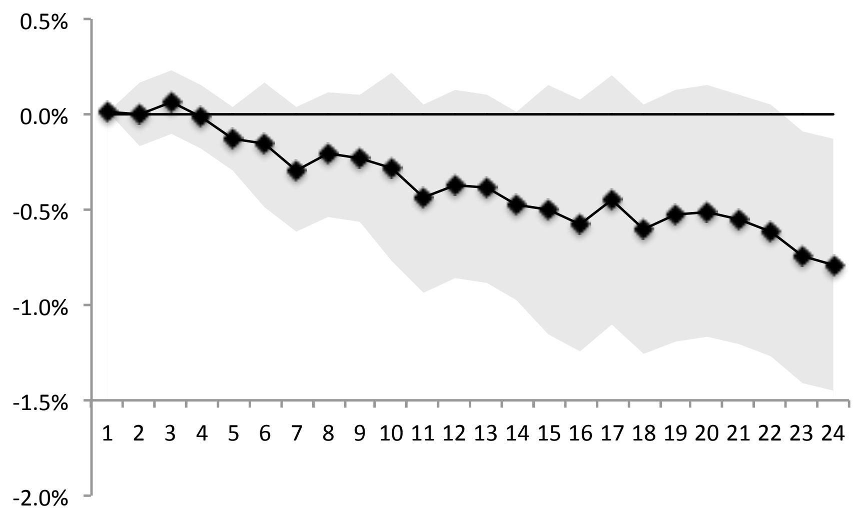


FIGURE 10. Reduction in APR

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.

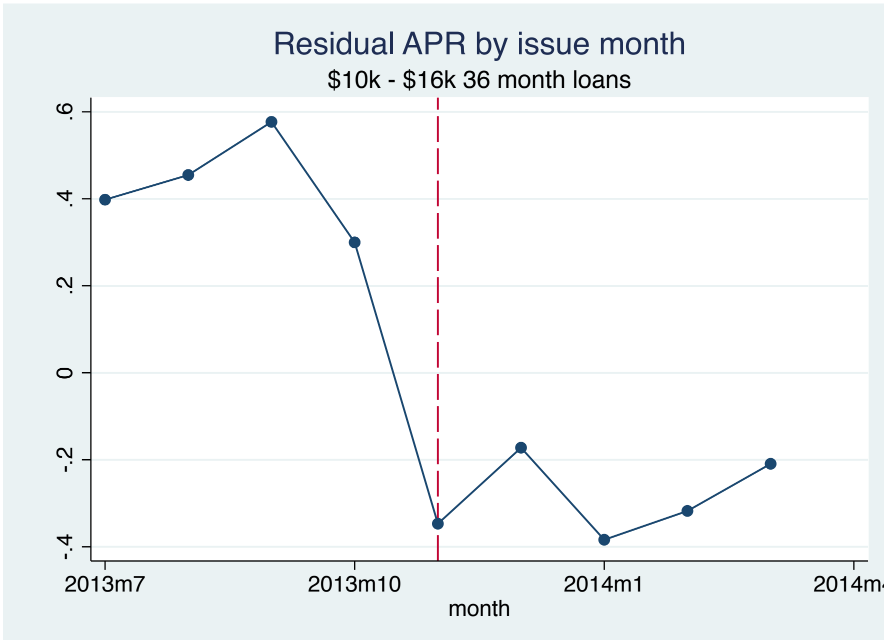


TABLE 1. Pre-period summary statistics

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 sub-grade 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 the logarithm of the number of loans at each month, credit risk sub-grade, 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 and 3 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 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 3 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 4 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 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)
	$log\,(\#loans)$	$log\,(\#loans)$	$log\,(\#loans)$	$log\,(\#loans)$	$log\,(\#loans)$
$D_{amount1000,t}$	-0.1451***	0.0586	-0.0355	-0.0825	-0.0441
	(0.033)	(0.064)	(0.048)	(0.065)	(0.028)
Sample	Main	36m, 16k - 24k	36m, 1k - 10k	60m, 16k - 24k	Placebo
Observations	3,663	1,637	2,374	1,738	3,861
R^2	0.817	0.802	0.761	0.724	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). Outcomes include $default$, a dummy that equals one if a borrower is late by more than 120 days;$FICO$, which measures a borrower's FICO score. Both outcome variables are 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 sub-grade $\times$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)	(3)	(4)
	$default$	$default$	$FICO$	$FICO$
$D_{i,t}$	-0.0081**	-0.0080**	2.7464**	2.6705**
	(0.004)	(0.004)	(1.032)	(0.999)
Sample	MAIN	MAIN	MAIN	MAIN
Observations	60,511	57,263	60,511	57,263
R^2	0.035	0.047	0.192	0.215
# clusters	45	45	45	45

Table 4. Robustness

The table shows the output of several robustness tests. Columns 1 and 2 replicate columns 1 and 2 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 3 and 4 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 3 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 4 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 sub-grade$\times$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)	(4)
	$default$	$FICO$	$default$	$default$
$D_{i,t}$	-0.0066*	2.80**	0.0018	-0.0106
	(0.004)	(1.049)	(0.010)	(0.007)
Sample	6k - 19k	6k - 19k	36m, 16k - 24k	36m, 1k - 10k
Observations	54,689	54,689	14,652	33,493
R^2	0.037	0.189	0.061	0.035
# clusters	45	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 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; $default12m$ 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. 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 sub-grade×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)
	$sd(FICO)$	$default12m$	$default24m$
$D_{i,t}$	-0.5764**	-0.0039	-0.0082*
	(0.266)	(0.003)	(0.004)
Observations	60,511	60,511	60,511
R^2	0.027	0.024	0.032
# clusters	45	45	45

Appendix 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

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

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 $A_1^{**k} = E\left(1-p_k(1-q)-x_k\right)$. 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.[38] 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 designed for this high type household. Formally $U^{*L'}(A_1, D_{1,2}, D_{1,3})$ is the maximized objective of

[38] 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.

the following problem:

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

subject to

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

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

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

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:

$$\mathscr{L} = \max_{D_{1,2}, D_{1,3}} (1 - p_H - x_H)\, u\,(E - x_H D_{1,2} - (x_H + (1-q)\, p_H)\, D_{1,3}) \tag{28}$$

$$+ p_H u\,(qE - x_H D_{1,2} + ((1-q)(1-p_H) - x_H)\, D_{1,3}) \tag{29}$$

$$+ x_H u\,((1 - x_H)\, D_{1,2} + (1 - x_H - (1-q)\, p_H)\, D_{1,3}) \tag{30}$$

$$+ \lambda_1 u\,((1 - p_L (1-q) - x_L)\, E) \tag{31}$$

$$- \lambda_1 (1 - p_L - x_L)\, u\,(E - x_H D_{1,2} - (x_H + (1-q)\, p_H)\, D_{1,3}) \tag{32}$$

$$- \lambda_1 p_L u\,(qE - x_H D_{1,2} + ((1-q)(1-p_H) - x_H)\, D_{1,3}) \tag{33}$$

$$- \lambda_1 x_L u\,((1 - x_H)\, D_{1,2} + (1 - x_H - (1-q)\, p_H)\, D_{1,3}) \tag{34}$$

$$+ \lambda_2 D_{1,2} + \lambda_3 D_{1,3} \tag{35}$$

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

$$\begin{aligned} D_{1,2} :\quad & -(1 - p_H - x_H)\, x_H u'\left(c_3^G\right) - p_H x_H u'\left(c_3^M\right) + x_H (1 - x_H)\, u'\left(c_3^B\right) \\ & + \lambda_1 \left[(1 - p_L - x_L)\, x_H u'\left(c_3^G\right) + p_L x_H u'\left(c_3^M\right) - x_L (1 - x_H)\, u'\left(c_3^B\right)\right] + \lambda_2 = 0 \end{aligned} \tag{36}$$

$$\begin{aligned} D_{1,3} :\quad & -(1 - p_H - x_H)(x_H + (1-q)\, p_H)\, u'\left(c_3^G\right) + p_H ((1-q)(1-p_H) - x_H)\, u'\left(c_3^M\right) + x_H (1 - x_H - (1-q \\ & + \lambda_1 \left[(1 - p_L - x_L)(x_H + (1-q)\, p_H)\, u'\left(c_3^G\right) - p_L ((1-q)(1-p_H) - x_H)\, u'\left(c_3^M\right) - x_L (1 - x_H - (1-q)\, p_H \right. \end{aligned} \tag{37}$$

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

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

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

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 $\bar{D}_{1,2}\in\left(0,E\right)$ 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

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

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 $\Upsilon\left(D_{1,3}\right)$ then

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

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

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

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 $\left\{A_1,D_{1,2},D_{1,3}\right\}$ as

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

where

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

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 8.

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:

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

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 8 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.

Appendix C. inferring initial credit risk 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.[39] 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 subgrades 1 to 5), 5 grades for B borrowers (initial subgrades 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).

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

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

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

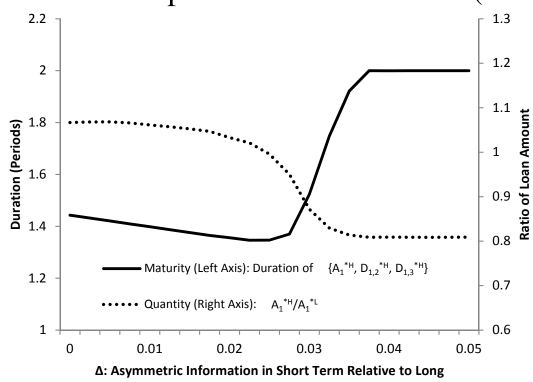
First, by construction, almost all loans below an F1 rating (26) will not have an initial sub-grade because LC's model states that only 25 initial grades are issued. Second, we succeed in matching a borrower's initial sub-grade for more than 98% of the loans of each final sub-grade in 24 out of the 25 top subgrades. 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.

Appendix D. Supplementary Tables and Figures

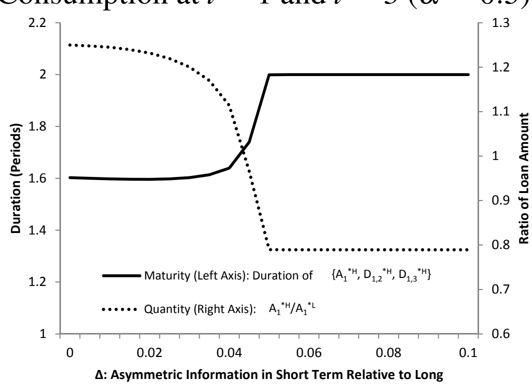
FIGURE 11. Model Comparative Statics - Robustness

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 8. The following parameters are used (identical to Figure 8): $E = 100$, $p_L = 0.3$, $x_L = 0.1$, $\bar{p}_H = 0.1$. 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 8) 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 8: $q = 0.75$ $\alpha = 1$. 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