

ADVERSE SELECTION AND MATURITY CHOICE IN CONSUMER CREDIT MARKETS: EVIDENCE FROM AN ONLINE LENDER*

ANDREW HERTZBERG[†], ANDRES LIBERMAN[‡], AND DANIEL PARAVISINI^{††}

ABSTRACT. This paper exploits a natural experiment to document adverse selection among prime consumer credit borrowers in the US. In our setting, some borrowers are offered only a short term loan while an observationally equivalent set of borrowers is offered the same short term loan as well as an additional long maturity option. We isolate adverse selection from the causal effect of maturity on repayment by comparing the ex post default behavior of borrowers who took the same short term loan in both settings. We show that when the long term option is available, borrowers who choose the short term loan default less and have higher future FICO scores. Thus, a longer loan maturity induces adverse selection by attracting unobservably less creditworthy borrowers. The difference in the default rate of borrowers who choose the pre-existing short maturity loan and those that self select to the new long maturity option is five percentage points, an economically large effect for prime borrowers whose average default rate is 8.6%. Our results highlight the potential for underprovision of long maturity consumer credit.

Keywords: Adverse Selection, Loan Maturity, Consumer Credit.

JEL codes: D82, D14.

[†] Columbia University. Email: ah2692@columbia.edu.

[‡] New York University. Email: aliberma@stern.nyu.edu.

^{††} London School of Economics. Email: D.Paravisini@lse.ac.uk.

* We thank Emily Breza and participants at the NYU-Columbia Junior Faculty Seminar. We thank Siddharth Vij for outstanding research assistance. All errors and omissions are ours only.

I. Introduction

Adverse selection can limit the effectiveness with which credit markets allow households to efficiently allocate consumption over time. Indeed, credit rationing can arise in equilibrium when borrowers possess private information about their own creditworthiness (Stiglitz and Weiss (1981)). The normal function of the price mechanism to equate supply and demand breaks down because lenders recognize that raising the loan interest rate will discourage the best quality borrowers from taking a loan.

But despite its importance in theory, there is little evidence of adverse selection in consumer credit markets.¹ The main reason for this lack of evidence is that adverse selection may be confounded with the causal effect of loan terms on repayment. The seminal methodological contribution of Karlan and Zinman (2009) was to show how a field experiment could be designed to distinguish adverse selection from moral hazard using random variation in anticipated and realized loan rates. The challenge that remains is to isolate the effect of adverse selection in other contexts where such intricate experimental variation in contractual terms is not feasible.²

This paper provides evidence of adverse selection in the choice of loan maturity among prime, unsecured consumer credit borrowers in the US. We exploit a natural experiment generated by a change in the menu of loan contracts offered to borrowers of an online lending platform, Lending Club (hereafter LC). The setting approximates the following ideal experiment. Some borrowers are offered a short maturity loan, while another set of observationally equivalent borrowers are offered the exact same short maturity loan and a longer maturity option. We measure the effect of selection by comparing the ex post default behavior of borrowers *who selected into the short maturity loan* in both groups. Since the short term loan is identical for both groups,

¹Evidence 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).

²Outside of the US, Rai and Klønner (2009) use a natural experiment involving a firm that organizes ROSCAs in South India to provide evidence of adverse selection. After the policy change borrowers are more constrained in their ability to bid on loans and they show this lowers the relationship between loan interest rate and default.

any difference in ex post default must be due to differences in the unobserved quality of borrowers who select into the long term loan. If default is lower at the short maturity loan when the long loan is offered then it must be that lower quality borrowers select into the long maturity option.³

The natural experiment occurred as follows. When a borrower applies for a loan at LC she is assigned according to a pre-specified rule to a narrowly defined risk category based on FICO score and other observable characteristics. Each borrower in a risk category is offered the same menu of loan choices that specifies the interest rate for any loan amount. Loans are available for amounts between \$1,000 and \$35,000 in either short–36 months–or long maturities–60 months. At the start of 2013 the long maturity loan was available only for amounts above \$16,000. During 2013 the menu of loan choices was expanded when the minimum size of long maturity loans was lowered two times, first to \$12,000 in March 2013 and then to \$10,000 in July 2013. Crucially for our experiment, the terms of all other previously available menu items were unchanged within each risk category. As a result, any change in the performance of borrowers whose loan amount was offered prior to the expansion must be due to ex-ante selection on unobserved borrower characteristics and cannot be attributed to changes in ex-post moral hazard or burden of repayment.

The experiment differs from the ideal setting because it introduced a new borrowing option to all borrowers, not only to a subset. However, we document that after the long maturity threshold was reduced, the number of short maturity loans at amounts affected by the reduction in the threshold (i.e., between \$10,000 and \$16,000, the “affected amounts”) fell by 20% relative to slightly lower and higher amounts. Thus, the bulk of selection into the new long maturity options is from borrowers who would have otherwise chosen a short maturity loan for the same amount. In similar tests we

³This approach to studying selection is similar to the experimental design employed by Beaman, Karlan, Thuysbaert, and Udry (2014) who study self-selection into micro-finance lending in Mali. They compare the marginal returns to a grant given to farmers who chose not take out a micro-finance loan to the returns of the same grant given to farmers where no selection into micro-finance was possible.

find no evidence of selection from short maturity loans above \$16,000 or below \$10,000, nor from long maturity loans above the \$16,000 threshold.

Thus, we can approximate the ideal experiment described above by studying the repayment of short maturity loans issued at affected amounts, before and after the reduction in the long maturity threshold. We control for time varying shocks to creditworthiness and credit demand by comparing borrowers of affected amounts with borrowers in the same month, LC risk category, and four-point FICO score range who took a short maturity loan at amounts just above \$16,000 and below \$10,000.⁴ In addition we exploit the staggered reduction of the long maturity amount threshold to use borrowers at amounts affected by one reduction of the threshold as additional controls in the other phase of the experiment. Importantly, *we do not* compare the performance of borrowers who took the short and long maturity loans because this comparison combines three effects that we are unable to distinguish: 1) selection on the intensive margin from the short to the loan maturity loan, 2) selection on the extensive margin from no loan to the long maturity loan, and 3) the impact of the different loan terms on default.

We find that the average default rate of short maturity borrowers decreases by one percentage point when a long maturity loan is available. This implies that borrowers who are less creditworthy choose the longer maturity loan. In economic terms, the expected default rate of the 20% of borrowers who self selected into the new long maturity loans is 5 percentage points higher than the remaining 80% of borrowers who chose to borrow short term. To get a sense of this magnitude, the average default rate of all short term borrowers within our sample before the menu expansion is 8.6%. We also find that the average FICO score of short maturity borrowers of amounts affected by the expansion increases by 1.8 points more than for borrowers of unaffected amounts. As with the result on default, this implies that borrowers who choose the 36 month loan when the long maturity was available expected their FICO score to be 9 points higher compared to borrowers who chose instead to take the long maturity loans.

⁴This also ensures that our results are not driven by a change in the pool of all potential applicants on the LC platform.

Our results demonstrate economically important adverse selection in the US unsecured consumer credit market. This finding is particularly striking as LC exclusively offers loans to prime US borrowers. On average, borrowers in our sample have a FICO score of 695 (and minimum of 660), 15 years of credit history, annual income of \$64,750, debt payments to income of 17%, and significant access to credit with a revolving balance of \$14,107 and utilization of 60.6%. Hence, we expect our estimate of adverse selection to under-represent its importance for the universe of all US borrowers and highlights that this phenomenon is important across the entire spectrum of observed consumer creditworthiness.

Our empirical setting has several important advantages that underline the robustness of our estimates. First, the data contains the entire set of variables that LC and its investors have about every borrower. By comparing borrowers offered exactly the same loan contract, within the same 4 point FICO score bin, and controlling for all other observable measures of creditworthiness, we can rule out that differences in loan performance are due to changes in the information that LC or investors possess at origination.⁵ Second, loans offered on the LC platform are funded by small individual 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 filled.⁶ Moreover we know that the algorithm did not change the price or any terms of the short term loans we study during the period of analysis. This rules out that selection is occurring based on supply side screening decisions or the possibility of that our results could be explained by reverse causality of loan terms impacting default. Third, LC's revenue is derived from an upfront origination fee that varies between 1.1 and 5% of a borrower's loan amount, which is subtracted from the amount borrowed. Thus, borrowers who took a short maturity loan could not costlessly refinance their loans and swap them for long maturity loans after the expansion.⁷ This ensures that the pool of borrowers who

⁵In theory the results could be driven by selection within each 4 point FICO score bin whereby higher (lower) FICO score borrowers within each bin choose short (long) term loans. However, by regressing default on FICO, we show that this effect cannot explain more than a 0.008% reduction in default.

⁶In rare occasions when a loan is not filled, LC provides the funds itself.

⁷We test for this possibility in the data and find no evidence of such behavior.

select the short maturity loan prior to the expansion is not impacted by the expansion itself. However, to the extent that this was occurring, it would bias our results towards finding no selection at all.

Fourth, the expansion was a fairly minor change in LC's overall lending platform. As far as we can tell it was not accompanied by any additional marketing campaign, nor did it coincide with a discontinuous jump in the total number of loans issued by LC. Importantly, it was not even contemporaneously advertised on the LC website.⁸ Borrowers would only notice the new options once they began applying for a loan. Thus it is highly unlikely that our results are driven by the expansion altering the types of borrowers applying for loans at LC. Finally, the data allows us to run a placebo test as if the change in the expansion occurred at different loan amounts. If our results were driven by time varying differences in creditworthiness at different loan amounts these results would detect similar effects. Instead, these placebo tests find no such evidence.⁹

We develop a stylized framework to explain how the choice of maturity may induce adverse selection. In our framework, information about a borrower's creditworthiness is slowly disseminated to credit markets. Thus, long maturity debt provides borrowers with insurance against shocks to their creditworthiness.¹⁰ If information were symmetrically distributed, all borrowers would choose long term loans to take advantage of this insurance benefit. When instead borrowers are better informed about their own creditworthiness, unobservably better borrowers would subsidize less creditworthy households by borrowing long term. Better types may instead choose to separate themselves and take short maturity loans to reduce their borrowing cost from pooling with low quality borrowers, thereby creating adverse selection into long maturity loans. This argument follows Rothschild and Stiglitz (1976) result that asymmetric information leads low risk types to distinguish themselves by taking less insurance. Our results are also

⁸According 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.

⁹In unreported results, we also find no difference in default rates if we run a placebo test at a later time. We interpret these results with caution because the data is right truncated and there may not be enough time to detect default in the sample.

¹⁰Athreya (2008) develops a model in which unsecured debt provides insurance through default. Evidence of the insurance motive of debt is present in Mahoney (2012) and Dobbie and Song (2014).

consistent with signaling theories of loan maturity by Flannery (1986) and Diamond (1991) who both argue that short maturity loans can be used by firms to credibly signal high creditworthiness. The framework we present differs from those papers by focusing on the insurance benefits of long term debt to risk averse household borrowers.

Our paper contributes to the literature that studies information asymmetries in credit markets (see Zinman (2014)). Ausubel (1999) and Agarwal, Chomsisengphet, and Liu (2010) study the results of large scale market experiments conducted by credit card issuers in the US in which new borrowers were sent solicitations with randomly selected credit terms. Both papers show that after controlling for observed borrower characteristics default rates are higher on solicitations with higher interest rates, which is consistent with the existence of adverse selection but also with the causal effect of higher rates. Adams, Einav, and Levin (2009) and Dobbie and Skiba (2013) provide evidence of strong adverse selection among high risk, low income subprime consumer credit borrowers in the US. While these papers provide important evidence, it is not altogether surprising that adverse selection would be high among subprime borrowers with limited credit history, low income and limited household assets. We complement their analysis by showing striking evidence of adverse selection among *prime* borrowers.

We also provide the first evidence of adverse selection stemming from a borrower's choice of loan maturity. To date, research has focused on the adverse selection stemming from a borrowers decision to accept a loan contract (Ausubel (1999), Karlan and Zinman (2009), Rai and Klonner (2009) and Agarwal, Chomsisengphet, and Liu (2010)) or their decision over how much to borrow (Adams, Einav, and Levin (2009), Dobbie and Skiba (2013)). Studying the selection response to maturity is important for a number of reasons. First, many common consumer loan products such as mortgages, auto loans, and personal loans offer a selection of loan maturities. In addition, by producing a different minimum monthly payment for any amount borrowed, the choice between different types of loans (for example credit card debt versus home equity loan) is in part a selection between loans of different maturities. Second, existing empirical work has demonstrated that credit demand elasticities with respect to maturity are much

higher than those interest rates both for borrowers in developing economies (Karlan and Zinman (2008)) and in the US (Attanasio, Koujianou Goldberg, and Kyriazidou (2008)). Since this dimension of a credit contract induces such a large demand response, the potential for adverse selection is also potentially important.

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. Section III provides a simple framework to understand how long maturity loans could induce adverse selection. Section IV describes the empirical strategy. Section V describes our main results on the extent of asymmetric information in this credit market. Section VI concludes.

II. Setting

A. Lending Club

LC is a publicly traded online peer-to-peer lender that operates in 45 US states. LC loans are unsecured amortizing loans for amounts between \$1,000 and \$35,000 (in \$25 intervals). LC loans are available for 36 months, which are available for all amounts, and 60 months, which are available for different amounts at different points in time. Using a proprietary credit risk assessment model, LC determines whether borrowers may apply for a loan and, if so, assigns the borrower to an initial credit risk category (one of 25 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 subgrade are weakly increasing in amount and maturity. Investors review the borrower's information, credit risk sub grade, and loan terms and choose to fund a portion of the loan. Loans typically get funded in less than 2 weeks: the median loan in our main sample is funded after 7 days, while the 90th percentile is funded after 14 days. LC charges a fee that varies between 1.1% and 5% depending on the credit score, which is subtracted at origination.

B. Staggered expansion of 60 month loans

Before March 2013, 60 month loans were only available for amounts above and including \$16,000. Figure 1 documents how after March 2013 the minimum threshold for a 60 month loan was lowered, first to \$12,000, and later to \$10,000. The figure shows the number of 60 month loans for amounts between \$5,000 and \$10,000, between \$10,000 and \$12,000, between \$12,000 and \$16,000, and between \$16,000 and \$18,000 (closed left and open right interval in all cases). The graph shows a clear break in the number of 60 month loans between \$12,000 and \$16,000 on March 2013, and between \$10,000 and \$12,000 on July 2013, relative to the other amount categories.

Importantly, the unaffected loan amounts do not exhibit any discontinuous jumps in the number of 60 month loans issued, which confirms that this expansion did not coincide with a marketing campaign or a large surge in demand for LC loans. In an Internet search we found no evidence of any contemporaneous changes in LC’s lending policy around the expansion of these loans. Further, using the Internet Archive website we verified that LC did not advertise this change, which suggests that the pool of applicants right before and after the expansion was fairly constant. 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

We select our main sample period so that: 1) LC’s lending policies remain constant during the period, and 2) to allow a reasonable pre- and post- period of time before and after the introduction of the 60 month loan options. Based on an Internet search and on our analysis of the data we found that LC changed the model it used to assign a borrower’s risk category (sub grade) in December 2012. Further, the model remained constant until the end of October 2013. Hence, we limit our sample period to all 36 month loans whose “list date” variable (`list_d`) is between and including these two

months for amounts between \$5,000 and \$18,000.¹¹ This amount interval includes loan amounts affected by the 60 month threshold reduction (\$10,000 to \$16,000) as well as amounts just above and just below this interval to control for time varying shocks to creditworthiness and credit demand. We use LC’s publicly available information to infer each borrower’s initial sub grade by reverse engineering LC’s risk model to obtain our final sample of 56,904 loans.¹²

Panel A of Table 1 presents summary statistics for the 11,250 36 month loans in our main sample that appear as listed during the pre-expansion period, that is, between December 4, 2012 and before March 2013. On average, loans for this sub-sample have a 16.4% APR and \$362 of monthly installment. Note that 87% of all loans were issued to refinance existing debt (this includes “credit card” and “debt consolidation” as stated purpose). As a comparison, 5.8% loans are issued for home improvement and 2% represent a major purchase. This suggests that the decision to borrow at LC is independent of intertemporal consumption decisions, and that the loan amount is most likely determined by the external borrowing that borrowers refinance. As of December 2014, 8.6% of these loans are late by more than 30 days.

Panel B shows statistics of borrower-level variables of this sample of loans. On average, LC borrowers in our sample have an annual income of \$64,750 and pay approximately 17% of their monthly income on other debts excluding mortgages. The average FICO score at origination is 695, and later credit report pulls show that the FICO score has on average decreased to 685 approximately one year later.¹³ LC borrowers have access to credit markets: 55% report that they own a house or have an outstanding mortgage. The average borrower has \$37,515 in debt excluding mortgage

¹¹In some placebo tests we shift our sample to loans issued between November 2013 and June 2014. We verify that the lending policy remained constant during this period as well. 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”.

¹²In the data, LC reports a borrower’s “final” credit risk sub grade, which starts from the initial sub grade (which is unobservable) and is modified to account for a borrower’s choice of amount and term. In the Appendix we detail how we reverse engineer the final subgrade using LC’s publicly available info on their credit risk model to infer the initial subgrade. We are able to assign an initial sub grade to 98.6% of all loans in the sample period.

¹³This “last FICO score” variable is updated every time LC discloses new information, which happens quarterly, expect for borrowers who have fully paid their loans or who have been charged off.

debt and \$14,107 in revolving debt, which represents a 60% revolving line utilization rate. Finally, LC borrowers have on average 15 years of credit history. The degree of information asymmetry among this group of borrowers, for which a large degree of information is already publicly known, probably leads to an underestimate of the level of adverse selection among the population of borrowers.

III. Framework

In this section we develop a framework to study the way that maturity choice in consumer credit can induce adverse selection. To do this, we build upon the fundamental model of adverse selection presented by Stiglitz and Weiss (1981). The simple addition is that over time some of a borrower's private information about her type will be revealed to the market. When a borrower rolls over a short term loan the market will use this new information to update the terms of future borrowing, whereas a long maturity loan will not adjust to this information and by doing so provides partial insurance against this news.

A. Setup

The time-line of the model is shown in Figure 2. At $t = 1$ there is a continuum of observationally equivalent households who need to borrow a fixed amount A .¹⁴ Each borrower anticipates receiving risky income at $t = 3$ that will be used to repay this loan. In the interim period $t = 2$ information about a borrower's creditworthiness is released in the form of a signal $S = \{G, B\}$ indicating either good or bad news, respectively. A borrower for whom good news is released ($S = G$) will earn income of $I = E > 0$ with certainty. By contrast a borrower for whom bad news is released will generate income of $I = E$ with probability $q \in (0, 1)$ and zero income otherwise.

¹⁴The evidence we present below suggests that the offer of a long term loan induced selection on loan maturity but not on loan size. This assumption can be motivated if borrowers have a fixed sized indivisible consumption need at $t = 1$ (or to refinance existing credit card debt). We could however relax this assumption to allow screening at $t = 1$ based on both loan size and maturity. Assuming loans at $t = 1$ are of a fixed size allows us to focus exclusively on the screen role played by maturity.

Income is not directly observed by the lender 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 $k \in \{H, L\}$. A borrower's type determines the probability with which each signal is released and hence the probability with which she will default: a borrower of type i will have good news released at $t = 2$ with probability $p_k \in (0, 1)$ with $p_H > p_L$. Let $\phi \in (0, 1)$ be the fraction of borrowers who are the high type. The model has the simplifying feature that the signal at $t = 2$ is a sufficient statistic for estimating a borrowers expected probability of default. Hence, conditional on receiving the bad signal, q is the same for both types.¹⁵

The supply of credit is perfectly competitive, the opportunity cost of funds and borrowers' subjective discount rate are normalized to zero, and lenders are risk neutral. We assume that lenders offer non-callable debt contracts with a single promised payment in the future.¹⁶ At $t = 1$ lenders can offer either a short or long maturity loan that promises a single payment at $t = 2$ or $t = 3$, respectively. Let $D_{t,\tau}$ denote the face value of a loan issued at $t \in \{1, 2\}$ that is due at $\tau \in \{2, 3\}$. Since the household earns no income at $t = 1$ a borrower who takes a short term loan will need to refinance by taking a second short term loan at $t = 2$ to repay at $t = 3$. 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 contingent on the signal.

When a loan is due a borrower can either repay the loan (using either income or refinancing) or default. All loans are uncollateralized so a creditor is unable to seize any household assets upon default. In the event of default the household incurs a

¹⁵This assumption is not necessary for the analysis but is made to eliminate the need to study the potential for additional screening at $t = 2$.

¹⁶The restriction to debt contracts could be motivated by costly state verification although we abstract from that here. In our empirical setting loan repayments amortize over the life of the contract. In essence, a long maturity amortizing loan is a bundle of short and long term loans. In order to focus on the difference in maturity between contracts we only consider loans with a short or long term promised payment.

utility cost of $\Omega > 0$ (e.g. the inconvenience of being contacted by collection agents and the (un-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. This allows the analysis to focus purely on the friction generated by ex ante asymmetric information. The exact condition to ensure this is provided below. Assume that $A < qI$, which as we show below is sufficient to ensure that any short term debt taken at $t = 1$ will always be refinanced at $t = 2$ regardless of what signal is released.

Households consume only at $t = 3$ and have a strictly increasing and concave utility function $u(c)$.¹⁷ A household will consume their income net of any debt payments due at $t = 3$. If the face value of debt due at $t = 3$ is greater than household income then the household will default and consume zero. Households will choose a loan contract at $t = 1$ to maximize their expected utility at $t = 3$. For simplicity we assume that a borrower can choose either a short or long term loan at $t = 1$ and not a combination of the two.

B. Symmetric Information

Consider the benchmark case in which information about a household's type is known by all agents at $t = 1$. Suppose that a borrower takes the long term loan at $t = 1$. This loan will be repaid if and only if the borrower has income of $I = E$, and hence default will occur with probability $(1 - p_k)(1 - q)$. Therefore, perfect competition in the supply of credit will ensure that the face value of the long maturity loan is:

$$D_{1,3} = \frac{A}{p_k + (1 - p_k)q}.$$

If at $t = 2$ the signal is bad ($S = B$) the household will continue to face uninsured consumption risk at $t = 3$. To insure against this they will borrow against all remaining

¹⁷Since the amount borrowed A is the same for all households we can ignore consumption at $t = 1$. We abstract from consumption at $t = 2$ to avoid considerations of inter-temporal consumption smoothing. Allowing for consumption at this interim period does not change the central results but does complicate the analysis.

unpledged expected income and save the proceeds.¹⁸ If $I = E$ then the household will have $E - D_{1,3}$ that can be pledged to creditors at $t = 2$. So the household will issue additional short term debt at $t = 2$ with face value $D_{2,3}^B = E - D_{1,3}$, which will raise $q(E - D_{1,3})$. No such precautionary savings motive exists when $S = G$ because all income risk disappears in that scenario. The resulting expected utility at $t = 1$ to taking a long term contract is:

$$U_1^{Long} = p_k u \left(E - \frac{A}{p_k + (1 - p_k)q} \right) + (1 - p_k) u \left(qE - \frac{qA}{p_k + (1 - p_k)q} \right) - (1 - p_k)(1 - q)\Omega.$$

Alternately, if a borrower takes a short term loan at $t = 1$ this loan will always be refinanced at $t = 2$ and hence $D_{1,1} = A$. There are two scenarios to consider for the refinancing of this loan, one for each possible realization of the signal S . If $S = G$ then the borrower will repay the new short term debt with certainty and faces no remaining income risk. Hence, she will issue just enough riskless debt to refinance her initial loan $D_{2,3}^G = A$. Alternately if $S = B$ then the borrower will face income risk at $t = 3$ and so will borrow against all potential income by issuing debt with face value $D_{2,3}^B = E$. Since this amount will be repaid with probability q , it will raise funds of qE at $t = 2$. Since we have assumed $qE > A$ this is sufficient to repay the initial short term debt and the remaining proceeds will be saved for consumption at $t = 3$. The resulting expected utility for taking the short term contract is:

$$U_1^{Short} = p_k u(E - A) + (1 - p_k) u(qE - A) - (1 - p_k)(1 - q)\Omega.$$

To complete the model we need to ensure that, as conjectured, default will only occur when $I = 0$ at $t = 3$. We show in the Appendix that this is assured under all contracts if:

$$(1) \quad \Omega \geq u(qE - A + E) - u(qE - A).$$

We assume that the cost of default Ω is high enough to ensure condition (1) holds.

¹⁸None of the results of the model rely on this motive for the household to take on additional debt at $t = 2$. It is included for completeness but the same qualitative results obtain if we assume that households can at most roll over existing debt at $t = 2$.

Comparing contracts under symmetric information, the long term loan is strictly preferred to the series of short term contracts. To see this note that the expected consumption at $t = 1$ under both contracts is the same: $E_1 [c] = [p_k + (1 - p_k) q] E - A$. This follows from the fact that lenders are risk neutral and the credit market is perfectly competitive. The difference between the two contracts comes from the insurance that is provided by the long term contract. To see this observe that:

$$E - A > E - \frac{A}{p_k + (1 - p_k) q} > qE - \frac{qA}{p_k + (1 - p_k) q} > qE - A,$$

which implies that consumption under the short term contract is a mean preserving spread of consumption under the long contract. The long term contract spreads the risk premium required to compensate for the possibility of default across all possible realizations of the intermediate signal S . It does so by being written before this uncertainty is resolved. By contrast, with a series of short term contracts the entire risk premium for default is paid only when the signal is bad $S = B$. Hence the higher cost of borrowing is paid exactly in the state of the world where a borrower's expected income is lower. Since creditors are risk neutral the efficient allocation of risk is for this uninsured income risk to be borne by them. A short term contract does not provide this insurance.

C. Asymmetric Information and Adverse Selection

Now suppose that a borrower's type is only known privately and so from a lenders perspective all borrowers are observationally equivalent at $t = 1$. It is easy to show that in any equilibrium the low type will choose the long term contract.¹⁹ The analysis then depends on which contract a high type will take at $t = 1$. We look for the conditions that will support a separating equilibrium in which the high types endogenously self select into short term loans at $t = 1$. Under this proposed separating equilibrium the

¹⁹Indeed, since low types strictly prefer this contract when their type is known, pooling with a high type would only make the long term contract more desirable. Moreover, there is no incentive for a low type to take a short term loan since the interest rate charged before S is revealed is zero independent of the borrower's type.

long term loan will be priced to break even assuming that it is taken only by borrowers of low creditworthiness. As a result the face value of the long term debt contract will be:

$$(2) \quad D_{1,3} = \frac{A}{p_L + (1 - p_L)q}.$$

From the perspective of a high type, the long term loan (2) is overpriced relative to the symmetric information case studied above. The price of the short term loan is unchanged from the symmetric information case (because the loan from $t = 1$ to $t = 2$ remains riskless and the loan from $t = 2$ to $t = 3$ depends only on the realized signal S). As a result the expected utility from taking the short term contract is the same as the symmetric information case. A separating equilibrium in which high type borrowers select the short term loan obtains if p_H is sufficiently larger than p_L . We show in the Appendix that this requires:

$$(3) \quad p_H \geq \bar{p}_H \equiv \frac{u\left(qE - \frac{qA}{p_L + (1-p_L)q}\right) - u(qE - A)}{u(E - A) - u(qE - A) - u\left(E - \frac{A}{p_L + (1-p_L)q}\right) + u\left(qE - \frac{qA}{p_L + (1-p_L)q}\right)}.$$

When borrowers are risk neutral $\bar{p}_H = p_L$, and is strictly higher than this when $u(\cdot)$ is strictly concave. The intuition for condition (3) is as follows. In order to achieve the insurance offered by the long term contract the high type must overpay for this loan. If the difference between types is sufficiently large (i.e. when p_H is sufficiently high) then the high type will forgo this overpriced insurance and opt instead for the series of short term contracts. Note that relative to the symmetric information case this results in an inefficient under provision of insurance.²⁰

The central intuition is directly analogous to the seminal paper of Rothschild and Stiglitz (1976), who show that adverse selection will result in inefficient risk sharing in insurance markets. Here, the under provision of insurance comes in the form of

²⁰In the Appendix we also characterize the conditions which support a pooling equilibrium in which both types choose a long term loan at $t = 1$. There exists a $\hat{p}_H \geq \bar{p}_H$ such that a pooling equilibrium exists if and only if $p_H \leq \hat{p}_H$. If the difference in creditworthiness between the two types is sufficient so that $p_H > \hat{p}_H$ then the separating equilibrium characterized here is unique.

an under provision of loan maturity in a competitive equilibrium. Moreover, suppose that at there is an initial period $t = 0$ in which borrowers are identical and their types are yet to be realized. The expected utility of all borrowers at $t = 0$ would be strictly higher if only the long term loan was available. This follows because, from the perspective of a borrower at $t = 0$, pooling of good and bad types simply represents a transfer from good to bad future realizations of their future self.

The key empirical prediction of our framework is that, in equilibrium, only high types will choose to borrow short term loans. In the next section we explain the empirical strategy we use to uncover this fact among LC borrowers. The model also predicts that long term loans will have higher interest rates to compensate for the increased risk of lending to low types. In our setting, this is precisely what we find, as long term loans at LC are on average 3% more expensive than short term loans of the same amount (for the same observed creditworthiness).²¹ The model also suggests why in equilibrium LC is able to offer interest rates lower than prevailing credit card interest rates. Credit card, with its low minimum monthly repayment and terms which cannot be revised on existing balances (outside of default) is in essence a very long maturity loan. LC is able to screen higher type borrowers who are prepared to forgo the insurance provided by the long term contract implicit in credit card debt to borrowers who are unwilling to pool with less credit worthy observationally equivalent borrowers.

IV. Empirical strategy

A. Ideal Experiment

Our goal is to study the way in which borrowers with different private information select among loans of different maturity. The methodology we employ is based on the following ideal experiment. Suppose that observationally equivalent borrowers are randomly assigned to one of two groups: A and B. Both groups are offered a menu of

²¹The contemporaneous spread between 3 and 5-year US Treasury bonds explains around 0.5% of this difference.

lending contracts. Each contract $C_{i,M}(Q_i, M)$ specifies a loan amount Q_i and a loan maturity M . In our empirical setting among observationally equivalent borrowers the loan interest rate is uniquely determined by the choice of loan amount and maturity, and is weakly increasing in both.²² For simplicity and to match the natural experiment we exploit at LC, suppose that loan maturity can be either short or long: $M \in \{S, L\}$. Group A is offered the following menu of contracts:

$$\Phi_A = \left\{ \begin{array}{ll} C_{1,S} \{Q_1, S\}; & - \\ C_{2,S} \{Q_2, S\}; & - \\ C_{3,S} \{Q_3, S\}; & C_{3,S} \{Q_3, L\} \end{array} \right\},$$

where each column represents the loans available at the two possible maturities and each row corresponds to a different available loan amount. Group A are offered short term contracts at three different loan amounts and a long maturity loan for the amount Q_3 .

Now suppose that group B is offered an expanded menu Φ_B that includes identically the same set of choices offered to group A as well as an additional loan maturity loan in amount Q_2 :

$$\Phi_B = \left\{ \begin{array}{ll} C_{1,S} \{Q_1, S\}; & - \\ C_{2,S} \{Q_2, S\}; & C_{2,S} \{Q_2, L\} \\ C_{3,S} \{Q_3, S\}; & C_{3,S} \{Q_3, L\} \end{array} \right\}$$

Borrowers can choose at most one option from the menu they face or decide to reject all offers and take no loan at all: C^\emptyset .²³

For any short term contract $C_{i,S}\{Q_i, S\}$, we compare the realized default rate of borrowers in group A and group B who took that menu item. Consider a borrower in group A who selected $C_{i,S}\{Q_i, S\}$ from Φ_A . By revealed preference, an identical

²²All other loan terms are the same. In our context all loans are ammortized equally over the life of the loan, have no prepayment penalties, and require no collateral.

²³At LC it is not possible for the same borrower to take multiple loans at the same time. Some borrowers may take up one more loan while another LC loan is active. Our data does not allow us to distinguish which borrowers have more than one loan at LC.

borrower in group B would either choose $C_{i,S} \{Q_i, S\}$ or the new long maturity option in Φ_B : $C_{i,S} \{Q_i, S\}$ (i.e., she would not choose 1) not to borrow or 2) any other menu option).²⁴ If the default rate on contract $C_{i,S} \{Q_i, S\}$ is lower (higher) in group B than in Group A, then it must be that the borrowers who selected to take the new long maturity loans were less (more) creditworthy than those who continue to take the same short maturity option.

Notice that this experiment captures adverse selection on the intensive margin—how hidden credit quality affects the decision *between* long and short maturity loans. Since we cannot measure the credit quality of borrowers who select no loan, we are unable to measure directly whether longer maturity loans induce lower quality borrowers to take a loan instead of no loan. Also, we do not compare the default rate of borrowers who took the new loans that were offered in Φ_B to the default rate on loans offered in Φ_A to borrowers on either group A or group B. Such a comparison would combine at least three effects that we cannot disentangle: 1) the selection between short and long maturity loans 2) the selection between no loan and the new long maturity choice, and 3) the direct effect of the different loan contract terms on default.

B. Application to Empirical Setting

We exploit the staggered reduction of the amount threshold for 60 month loans during 2013 to identify selection of borrowers into long maturity loans. As prescribed in the ideal experiment, the expansion offered new menu items at longer maturities for amounts already offered on short term contracts prior to the expansion. However, relative to the ideal setting, this expansion of the menu of loan choices affected the entire pool of potential applicants at once. In this sense, all borrowers who applied for a loan at LC after the expansion would be in group B of our ideal experiment. This is a problem for the econometrician if there are time-of-origination-varying shocks to borrower default rates coming from either changes in macroeconomic conditions or

²⁴Note that by experimental design the only new menu item offered to group B is a long maturity loan at an amount already offered in Φ_A at the short maturity. Thus, we isolate selection induced by maturity and keep the loan amount constant.

changes in the unobserved credit quality of potential borrowers. That is, there is no observationally equivalent group of loan applicants at the same time of origination that were prevented from borrowing at the new long maturities that could serve as a control group for these confounding shocks.

To solve this problem, we focus our analysis on loan amounts that allowed borrowing in *only* short maturity loans before the expansion and allowed borrowing short *and* long maturity loans after the expansion, i.e., loan amounts between \$10,000 and \$16,000. We refer to this set of amounts as the “affected amounts”. These are the empirical counterpart of $C_{2,S}\{Q_2, S\}$ in our ideal experiment. It is theoretically possible that offering $C_{2,L}\{Q_2, L\}$ in Φ_B would induce selection from other short term loans ($C_{1,S}\{Q_1, S\}$ and $C_{3,S}\{Q_3, S\}$). Thus, we first show that the expansion of the set of loan choices affected only borrowers who would have borrowed a short maturity loan within the set of affected loan amounts ($C_{2,S}\{Q_2, S\}$ in our ideal experiment). Figure 4 presents graphical evidence of this fact. The graph shows the time series evolution of the (log) number of 36 month loans originated each month across 4 amount categories: \$5,000 to \$10,000, \$10,000 to \$12,000, \$12,000 to \$16,000, and \$16,000 to \$18,000. The four curves move in parallel before March 2013, a condition that is consistent with the assumption that the trends of each of these loan categories is a valid counterfactual for each other. After March 2013 and July 2013 (the first and second dotted lines respectively), the relative number of loans issued in the \$12,000 to \$16,000 and \$10,000 to \$12,000 categories decreases relative to the other amount categories, respectively. Thus, the introduction of the option of borrowing at 60 months affected those borrowers who absent this option would have taken a loan of the *same* amount interval.²⁵

We formally test this fact by collapsing the data and counting the number of loans $N_{j,t,amount1000}$ at the month of origination $t \times$ sub-grade $j \times amount1000$ level for all loans issued during our sample period as defined above (December 2012 to October 2013). Note that *amount1000* includes all loan amounts in the same \$1,000 increment

²⁵Note that this suggests that credit demand is relatively inelastic to loan amount. This is consistent with the modal stated purpose of loans in our sample shown in Table 1, which is to refinance a predetermined level of existing debt.

(e.g., \$10,000 to \$11,000, \$11,000 to \$12,000, etc).²⁶ We define a dummy variable $D_{amount1000,t}$ for each loan issued in our sample period that takes the following values:

$$D_{amount1000,t} = \begin{cases} 1 & \text{if } \$16,000 > amount_{1000} \geq \$12,000 \text{ \& } t \geq \text{March2013} \\ 1 & \text{if } \$12,000 > amount_{1000} > \$10,000 \text{ \& } t \geq \text{July2013} \\ 0 & \text{in other cases} \end{cases} ,$$

and run the regression:

$$(4) \quad \log(N_{j,t,amount1000}) = \beta_{amount1000} + \delta_{j,t} + \gamma \times D_{amount1000,t} + \epsilon_{i,t}.$$

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). We include amount level fixed effects $\beta_{amount1000}$, which control for level differences in the number of loans for each \$1,000 amount increment. In turn, sub-grade \times month fixed effects $\delta_{j,t}$ control for the terms of the contract offers.²⁷

Table 2 shows the results of regression (4). Column 1 shows the baseline effect, estimated on the full sample of borrowers who took a 36 month loan between \$5,000 and \$18,000 during the sample period (December 2012 to October 2013). The coefficient γ implies that the number of borrowers who took a short term loan is 21% lower once the new long term loan option for the same amount becomes available. This is the selection at the intensive margin induced by the existence of another loan option: absent the option to borrow long term, we expect to see 20% more short term loans issued at the amounts where the long term loans were offered (Q_2).²⁸

To ensure that our results are not simply driven by secular trends in the demand for loans of different amounts, we run regression (4) on a sample shifted forward by 7 months to the period of time exactly after the expansion of 60 month loans to lower

²⁶Results are insensitive to other definitions of the amount interval.

²⁷Results are qualitatively and quantitatively unchanged by collapsing the data at the month of origination $t \times amount1000$ level instead and not including $\delta_{j,t}$ fixed effects.

²⁸These results are generally robust to the choice of interval size below \$10,000 and above \$16,000.

amounts is concluded. That is, we shift the definition of $D_{amount1000,t}$ forward by 7 months and run the regression on the sample of loans originated between July 2013 and May 2014. Column 2 of Table 2 shows the results. The coefficient on $D_{amount1000,t}$ equals -3.3% and is insignificant, and given the confidence interval we can reject the null that this coefficient equals our main estimate. This result strongly suggests that the expansion of the menu of borrowing choices affected the decision set of some borrowers who would have otherwise chosen to borrow a short term loan.

We study whether the expansion of the menu of loans had any differential effect in the number of loans issued contemporaneously at other loan amounts, both at the short and long maturities. To this extent we modify our regression slightly. We replace the staggered introduction dummy $D_{amount1000,t}$ with the interaction of the dummy $after_t$ which equals 1 one after March 2013 and $treated_{amount1000}$, where each \$1,000 amount interval is defined to be treated or not appropriately for each regression. The interaction term measures the differential change in the number of loans issued for amounts defined as treated relative to those not defined as treated in each regression, before and after March 2013. Column 3 of Table 2 replicates the result documented in Column 1 by defining the treated group as those with loan amounts between \$10,000 and \$16,000: the coefficient is -21.6% and highly statistically significant.

Next, in Column 4 we shift the sample to all loans originated during our main sample period but for amounts between \$16,000 and \$25,000, and we define the treated group for this regression for amounts between \$16,000 and \$20,000. The coefficient on the interaction term is 7.9% and not significantly different from zero. In Column 5 we shift the sample to all loans originated during our main sample period for amounts between \$1,000 and \$10,000. We define the treated group for this regression for amounts between \$6,000 and \$10,000. The coefficient on the interaction term is -5.2% and not significantly different from zero. These two results show that there was no noticeable reduction in the number of 36 month loans issued at amounts just below and just above the \$10,000 to \$16,000 interval in which the expansion of 60 month loans was concentrated.

Finally, in Column 6 of Table 2 we show the output of the same regression as in Column 4, that is, loans originated during our main sample period for amounts between \$16,000 and \$25,000, but this time for 60 month loans. We define the treated group for loan amounts between \$16,000 and \$20,000. Here we test whether there was selection into the new 60 month loan amounts from borrowers who otherwise would have chosen a 60 month loan for a larger amount. The coefficient is -8%, which suggests some of this selection may have occurred, but is not significantly different from zero. Taken together, these results suggest that selection into the new 60 month loan options only came from borrowers who otherwise would have chosen a 36 month loan at the same amounts at which these options were offered.

Having established that the bulk of selection into the long term loans occurred from the same loan amounts as the new long term loan options, we implement our methodology to test for adverse selection along maturity. We test for differences in ex post measures of creditworthiness of short term borrowers whose loan amount was affected by the expansion of the long term menu options (in our ideal setting, this amount is Q_2). We control for secular trends in creditworthiness and credit demand with the same difference in measures of creditworthiness but on loan amounts that never had the option to borrow long term (i.e., just below \$10,000, which in our ideal setting is depicted as Q_1) and amounts that always had the option to borrow long term (i.e., just above \$16,000, which in our ideal setting is depicted as Q_3).

Figure 5 shows our main results in a graphical manner. We label Treated loans as those for amounts between \$10,000 and \$16,000, and Control are other amounts in our main sample (\$5,000 to \$10,000 and \$16,000 to \$18,000). The figure shows that the 30 day default rate for Treated borrowers is lower after the 60 month expansion relative to Control borrowers. Given that our sample is right censored, the default rate exhibits a downward trend. Hence, we also shows the 30 day default rate for a period of 1 year after each month of origination, which confirms the graphical intuition. The last graph shows that Treated borrowers also have a higher FICO score as of December 2014. These graphs hint at our main result of adverse selection induced by the long maturity

loans: borrowers who chose short term loans at affected amounts seem to have been more creditworthy than those who chose a long term loan.

We implement a formal regression test on our sample of 36 month loan borrowers:

$$(5) \quad \text{outcome}_i = \beta_{\text{amount1000}} + \delta_{j,\overline{FICO},t} + \gamma \times D_{\text{amount100},t} + X_{i,t} + \epsilon_i,$$

where data is at the loan level i and the coefficient of interest is again γ , the change in the outcome variable for short maturity loans originated for affected amounts before and after the expansion of the menu options. We include granular month of origination $t \times$ sub-grade $j \times 4$ -FICO score at origination (\overline{FICO}) bin fixed effects $\delta_{j,\overline{FICO},t}$ that ensure we compare borrowers who took a loan on the same month, with the same contract terms (same sub-grade), and with very similar level of observed creditworthiness. We also include a vector of control variables observable at origination, $X_{i,t}$. In our baseline specification, $X_{i,t}$ includes US state address \times month fixed effects and annual income.²⁹ We also report results including the *full* set of variables that LC reports and that investors observe at origination. These variables include, for example, 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). This regression allows us to interpret our results as demonstrating selection on unobservables at origination.

Our outcomes include default30_i , a dummy that equals one if the loan is late by more than 30 days, default_i which includes all measures of default including if a borrower is late by less than 30 days, and FICO_i , the high end of borrower’s FICO score 4 point bin, all three variables measured as of December 2014.³⁰ We also report regression results for fullypaid_i , a dummy that equals one if the loan has been prepaid,

²⁹We implement using the methodology in Gormley and Matsa (2014) for regressions with two high-dimensional fixed effects using the REG2HDFE Stata command (see Guimaraes and Portugal (2010)).

³⁰Our 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.

and $chargeoff_i$, which equals one if the loan has been charged off, all as of December 2014. In all our regression, standard errors are clustered at the sub grade level.

V. Results

Table 3 reports the main regression results of our paper. Columns 1 through 4 show, using various measures of default, that borrowers who took a 36 month loan after the 60 month loan option was available for the same amount are significantly less likely to default relative to before the option was available and relative to borrowers of slightly larger and smaller amounts. Column 1 reports the result of our main outcome variable, $default30_i$: the coefficient on the regression results when the outcome variable is $default30_i$ equals one percentage point and is significant at the 1% level. Note that as per Table 2, the expansion of the 60 month loan option reduces demand for the short term loan by roughly 20%. Thus, the average default rate of the 20% of borrowers who chose to take a 60 month loan when it became available is $1\% \times 5 = 5\%$ higher than those who chose instead to take the 36 month loan. To get an idea of the economic magnitude of this effect, note that as per the summary statistics shown in Table 1, the average default rate of 36 month loans issued between December 2012 and February 2013 is 8.6%.

This result says nothing of the default rate of 60 month borrowers. Indeed, the default rate on 60 month loans aggregates three effects: 1) the average creditworthiness of borrowers who would have borrowed short term had the 60 month loan option been unavailable, 2) the average creditworthiness of borrowers who would have not borrowed at LC unless the 60 month loan option been unavailable, and 3) the causal effect of the 60 month loan on default. Our results allow us to measure the first effect, which reveals substantial adverse selection along the maturity dimension. Further, our result implies that borrowers who chose a 60 month loan but who would have chosen a 36 month loan of the same amount before the expansion would have had a higher default

rate on the 36 month loan. Presumably, choosing a longer loan may increase or reduce the rate at which these same borrowers default.

Next we study whether the lower default rate of borrowers who selected a short maturity loan could be predicted by variables available to investors at the time of origination. In that case, investors would presumably modify their required interest rates to counteract the effect of default on profitability. Note that our main result already controls in a very granular manner for month of origination by sub grade by 4-point FICO score bin fixed effects, as well as by state of residence by month fixed effects. However, in Column 2 of Table 3 we run the same regression as in Column 1 but adding every single variable known at origination that is available in LC's dataset as a control in the right hand side. The results are striking: the coefficient goes down from -1% to -0.9%, and remains statistically significant at the 5%. This suggests that the adverse selection we capture with our main result represents unobserved heterogeneity and cannot be priced in by the lender.

We do not measure the borrower's exact FICO score at origination. Instead, LC provides in its data a 4-point range—all our regressions and fixed effects are calculated using the high end of each range bin. This is potentially a problem if our default regression simply captures selection along FICO scores *within* each 4-point FICO bin. In that case, we could not interpret our results as evidence of adverse selection along unobservables but rather as selection along observables that the econometrician cannot control for. While this is a theoretical possibility, we find that the effect of FICO on default in our sample is not very large. Indeed, a regression of *default30* on the high end of the FICO range at origination within each sub grade by \$1,000 amount range by month gives a coefficient of -0.0000268 (i.e., a 1 point increase in FICO score at origination is correlated with a 0.003% decline in default rate, not statistically significant). Thus, variation in default rates within FICO score bins can at most account for a 0.008% difference in default rates, quantitatively irrelevant next to our estimated effect of 1% reduction in default.

Columns 3 and 4 of Table 3 run the same regression as Column 1 with alternative definitions of default. Column 3 includes as default every single measure of delinquency reported by LC, including borrowers who are late by one day. Column 4 instead focuses on loans that have already been charged off, which typically does not occur before 151 days after a loan stopped being current. As both columns report, the effects are negative and significant, consistent with our main effect.

All our results measure default as of December 2014, the last time the data was updated. Since our dataset is right censored as of December 2014, loans issued earlier in the sample have a longer time to default than loans issued later. This is not a problem for our empirical strategy because we compare loans issued in the same month at affected and unaffected amounts. However, we study the differential evolution of default since origination by limiting the definition of default to those loans whose last payment due was x months after origination, and varying x from 1 to 20 months. Figure 6 shows the regression coefficient and 90% confidence interval for each of these default variables. The figure shows that the default rate of borrowers who choose to borrow at 36 months when the 60 month option is available is always lower, and declines monotonically with months since origination. The result becomes statistically significant approximately 1 year after origination. Note that this effect is consistent with the main assumption underlying our theoretical framework: private information about a borrower's creditworthiness is slowly learned by market participants and does not lead to immediate default.

Finally, we study whether adverse selection can be measured with respect to FICO scores. Note that a borrower's FICO score aggregates repayment on a borrower's entire set of liabilities, including LC. Thus, if borrowers prioritize the repayment of some debts over others, lower default rates on LC debt would not necessarily translate into higher FICO scores. E.g., if long and short maturity borrowers prioritize repayment of different debts, the negative effect of defaults at LC on FICO scores would cancel out.

Column 5 of Table 3 uses as outcome a borrower's FICO score pulled at the time we downloaded our dataset (December 2014). The coefficient implies that on average,

borrowers who chose the short term loan have a 1.8 higher FICO score in subsequent data updates. In economic terms this means that the average FICO score of the 20% of the total pool of pre-expansion short term applicants who became long term borrowers is $1.8 \times 5 = 9$ points higher. This higher FICO score may, for example, result in better access to credit and labor markets. This result is consistent with our framework and with theories that emphasize borrowers own assessment of their ability to refinance short term debt in the future.

VI. Discussion

We have documented economically important adverse selection among prime borrowers in US consumer credit markets. This adverse selection is driven by the choice of maturity: long maturity loans attract unobservably worse borrowers. We provide a framework that rationalizes this finding in equilibrium based on the seminal contributions of Stiglitz and Weiss (1981) and Rothschild and Stiglitz (1976). In our framework, long maturity loans provide borrowers with insurance. But if information about their creditworthiness is privately known by borrowers, better borrowers may choose to separate and borrow short term to avoid paying large interest rates. In equilibrium this may lead to an under provision of maturity (insurance) in a competitive equilibrium, with associated efficiency costs. Here, the under provision of insurance comes in the form of an under provision of loan maturity in a competitive equilibrium.

Our results help rationalize upward sloping rates along maturity. Further, they may help explain the success of Internet sites such as LC or Prosper, who derive most of their business from lending to prime borrowers who want to refinance their consumer debt. Indeed, these sites may be capturing good borrowers who do not want to pay high credit card interest rates.

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.
- Athreya, Kartik B, 2008, Default, insurance, and debt over the life-cycle, *Journal of Monetary Economics* 55, 752–774.
- 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.
- Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry, 2014, Self-selection into credit markets: Evidence from agriculture in mali, Discussion paper National Bureau of Economic Research.
- Chiappori, Pierre-André, and Bernard Salanie, 2000, Testing for asymmetric information in insurance markets, *Journal of political Economy* 108, 56–78.
- 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.

- 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.
- Gormley, Todd A, and David A Matsa, 2014, Common errors: How to (and not to) control for unobserved heterogeneity, *Review of Financial Studies* 27, 617–661.
- Guimaraes, Paulo, and Pedro Portugal, 2010, A simple feasible alternative procedure to estimate models with high-dimensional fixed effects, *Stata Journal* 10(4), 628–649.
- 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.
- Mahoney, Neale, 2012, Bankruptcy as implicit health insurance, *NBER Working Paper*.
- 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.
- Stiglitz, J.E., and A. Weiss, 1981, Credit rationing in markets with imperfect information, *The American economic review* 71, 393–410.
- Zinman, Jonathan, 2014, Household debt: Facts, puzzles, theories, and policies, .

Appendix

Appendix A. Figures and Tables

FIGURE 1. Staggered expansion of 60 month loans

This figure shows the time series of the number of 60 month loans by listing month for each of the following four amount intervals: \$5,000 to \$10,000, \$10,000 to \$12,000, \$12,000 to \$16,000, and \$16,000 to \$18,000.

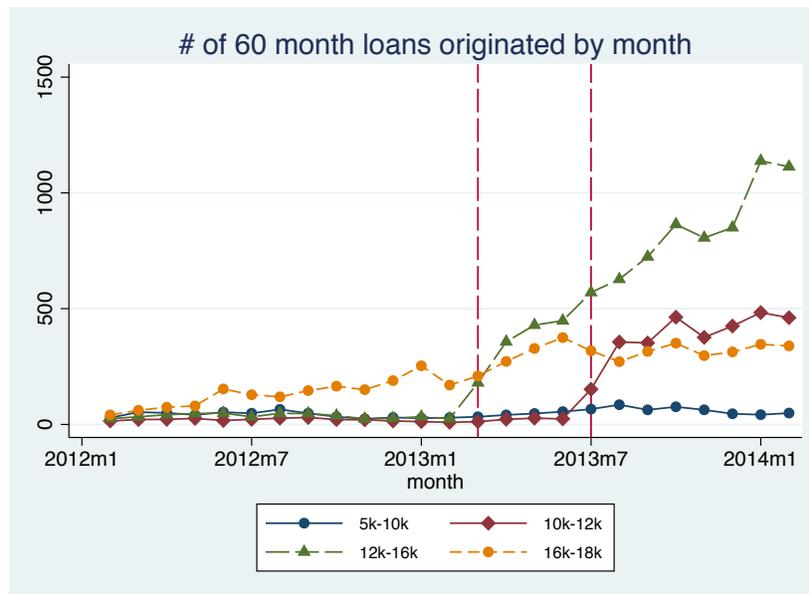


FIGURE 2. Model Time-line

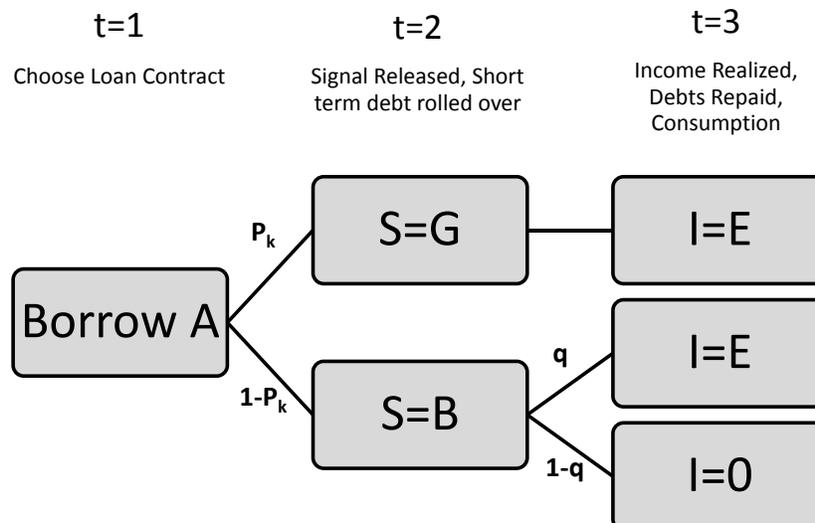


FIGURE 3. Total \$ amount issued by LC by month of listing

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.

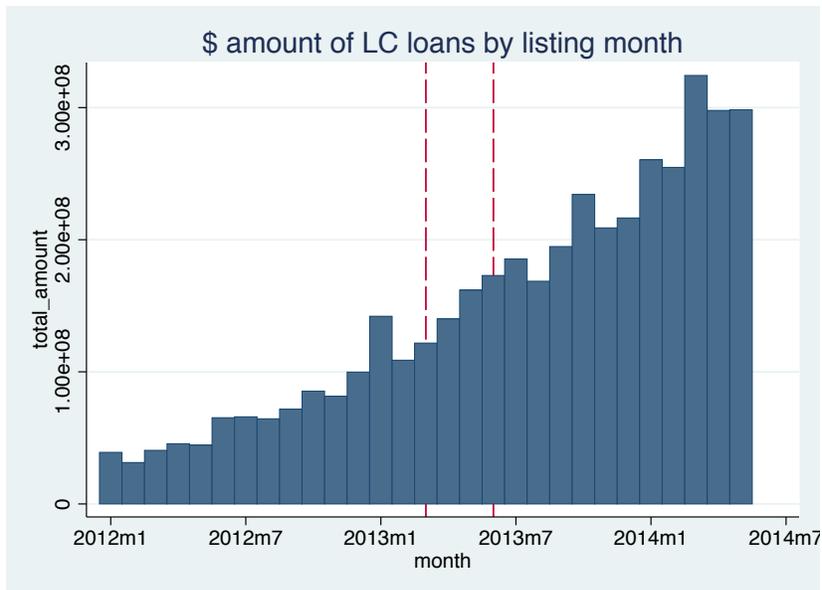


FIGURE 4. Log(number of 36 month loans) by month

This figure shows the time series of the logarithm of the number of 36 month loans by listing month in our main sample period for each of the following four amount intervals: \$5,000 to \$10,000, \$10,000 to \$12,000, \$12,000 to \$16,000, and \$16,000 to \$18,000.

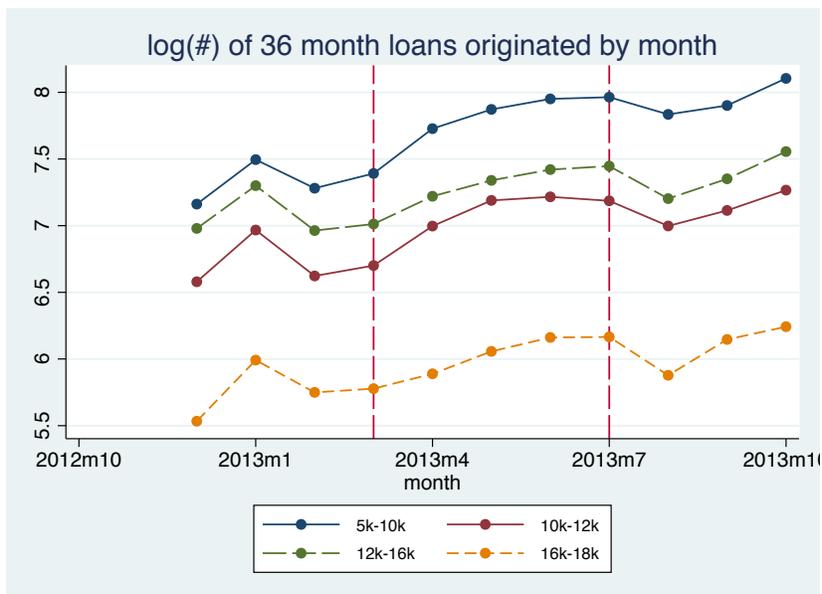


FIGURE 5. Default and FICO by month of origination

This figure shows the time series of the 30 day default rate, the 1-year forward 30 day default rate, and the December 2014 FICO score for borrowers in our main sample. In this graph, Treated (or Affected) amounts include loans for amounts between \$10,000 and \$16,000, while Control (or Unaffected) amounts include loans for \$5,000 to \$10,000 and \$16,000 to \$18,000. The dotted vertical bars represent March and July 2013, when the 60 month amount threshold was lowered to \$12,000 and \$10,000 respectively.

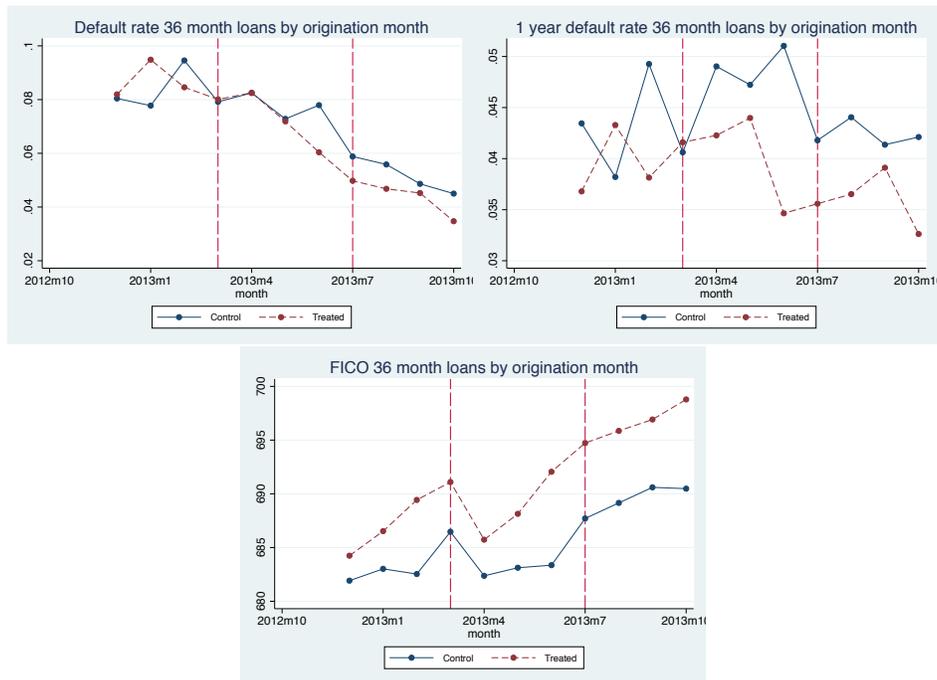


FIGURE 6. Default rate coefficient by number of months since origination

This figure shows the estimated coefficient and 90% confidence interval of the regression of default_{30} , a dummy that equals one if a loan is not current by more than 30 days as of December 2014, on $D_{\text{amount}1000,t}$, 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 subgrade level.

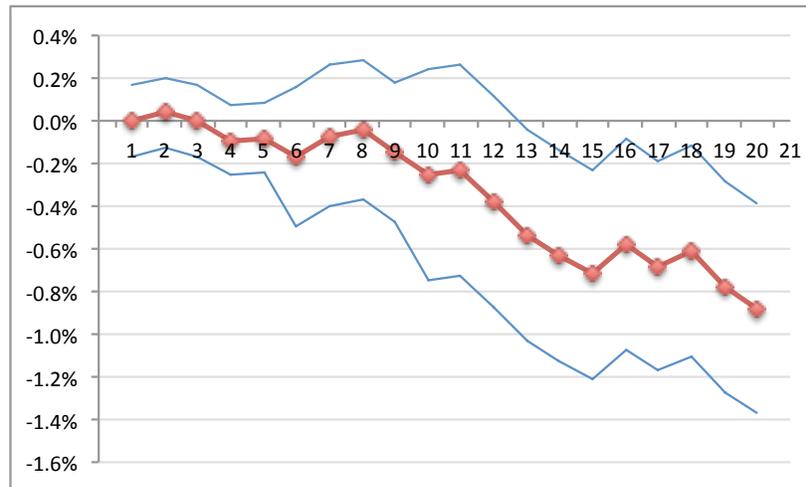


TABLE 1. Summary statistics

This table shows some summary statistics of our main sample of Lending Club borrowers for pre-expansion months, which includes all 36 month loans whose listing date is between December 4, 2012 and March 2013, for an amount between \$5,000 and \$18,000, and for which we estimate an initial subgrade based on LC's publicly available information.

	mean	p50	sd
Panel A: loan characteristics			
APR (%)	16.4	16.0	4.0
Installment (\$)	362	351	111
For refinancing (%)	86.6	100	34.0
Default 30 days (%)	8.6	0	28.1
Panel B: borrower characteristics			
Annual income (\$)	64,750	56,000	75,107
Debt payments / Income (%)	17.3	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	694	66
Home ownership (%)	55.0	100	49.8
Total debt excl mortgage (\$)	37,515	28,888	33,71
Revolving balance (\$)	14,107	11,216	12,430
Revolving utilization (%)	60.6	62.6	22.0
Months of credit history	180	163	84
N	11,250		

TABLE 2. Regression results: selection into long maturity loans

This table shows that selection into the new 60 month options was more important for borrowers who would have chosen to borrow a 36 month loan of the same amount. Columns 1 and 2 show 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 maturity loan was first not available and then made available, and zero otherwise. The sample corresponds to loan amounts between \$5,000 and \$18,000. Column 1 is the main sample, includes loans whose list date is between December 2012 and October 2013. Column 2 is the placebo sample, includes loans issued between July 2013 and May 2014. Column 3 restricts the sample to 36 month loans issued in the main sample period for amounts between \$16,000 and \$25,000. Column 4 restricts the sample to 36 month loans issued in the main sample period for amounts between \$1,000 and \$10,000. Standard errors are clustered at the initial credit risk sub grade (25 clusters). *, ** and *** represent significance at the 10%, 5%, and 1% respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	$\log(\#loans)$	$\log(\#loans)$	$\log(\#loans)$	$\log(\#loans)$	$\log(\#loans)$	$\log(\#loans)$
$D_{amount1000,t}$	-0.2022*** (0.026)	-0.0334 (0.027)				
$after \times treated$			-0.2156*** (0.038)	0.0793 (0.062)	-0.0521 (0.064)	-0.0805 (0.067)
Sample	Main	Placebo	Main	36m, 16k - 25k	36m, 1k - 10k	60m, 16k - 25k
Observations	3,226	3,410	3,234	1,826	2,374	1,953
R^2	0.831	0.885	0.835	0.789	0.761	0.714
# clusters	25	25	25	25	25	25

TABLE 3. Regression results: adverse selection along maturity dimension

This table shows our main result of adverse selection along maturity. 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 and July 2013. Outcomes include *default30*, a dummy that equals one if a borrower is late by more than 30 days; *default*, a dummy that equals one if a borrower is late by any number of days; *chargedoff*, a dummy that equals one if a loan has been charged off. All three default variables are measured as of the December 2014 LC data update. The last outcome is FICO, which measures a borrower's FICO score at the time of the data pull, or at the time of charge off or prepayment. The sample corresponds to loan amounts between \$5,000 and \$18,000 whose listing date is between December 4, 2012 and October 25, 2013. All regressions include sub grade \times 4-point FICO bin \times month, and US state \times month fixed effects. Column 2 includes all borrower level variables observed by investors at the time of origination. Standard errors are clustered at the initial credit risk sub grade (25 clusters). *, ** and *** represent significance at the 10%, 5%, and 1% respectively.

	(1)	(2)	(3)	(4)	(5)
	<i>default30</i>	<i>default30</i>	<i>default</i>	<i>chargedoff</i>	<i>FICO</i>
$D_{i,t}$	-0.0101*** (0.003)	-0.0091** (0.003)	-0.0081* (0.004)	-0.0071* (0.004)	1.8337** (0.850)
Observations	56,904	53,780	56,904	56,904	56,897
R^2	0.118	0.131	0.117	0.116	0.289
# clusters	25	25	25	25	25

Appendix B. Mathematical Appendix for Framework

A. Confirming Conjecture that Default Occurs if and Only if $t = 3$ and $I = 0$

We show here that condition 1 is sufficient to confirm the conjecture in our analysis that a borrower only defaults on her loan in the scenario where she has no income at $t = 3$.

A.1. Default When Borrower Takes Short Term Contract at $t = 1$

Start by considering the case where a borrower has taken the short term loan at $t = 1$. Suppose the borrower that has taken the short term loan at $t = 1$ has received the bad signal $S = B$ and has refinanced her loan and issued debt with face value $D_{2,3}^B = E$. After refinancing her loan the household will have savings of $qE - A$ available to fund consumption at $t = 3$. If the household realizes income of $I = 3$ at $t = 3$, not defaulting will result in consumption of $qE - A$, whereas defaulting will raise consumption by the face value of the debt to $qE - A + E$ but will invoke the utility loss of Ω . She will choose to repay her loan only if the utility loss from default outweighs the utility from the additional consumption from defaulting:

$$(6) \quad \Omega \geq u(qE - A + E) - u(qE - A)$$

which is the assumption we have made in 1. Since $u(\cdot)$ is concave, if 6 holds for $A > 0$ then it must also hold when $A = 0$. So when we consider a deviation in which a borrower does not roll over her loan at $t = 2$ this condition assures us that any new borrowing done at $t = 2$ with face value up to E will be repaid.

If we consider the scenario where the borrower has taken the short term loan and received the good signal $S = G$ the condition for repayment will be identical to (6) if we set $q = 1$:

$$(7) \quad \Omega \geq u(E - A + E) - u(E - A).$$

Observe that the difference in consumption from defaulting is the same in (6) and (7). And since $E > q$ the concavity of $u(\cdot)$ ensures that

$$u(qE - A + E) - u(qE - A) \geq u(E - A + E) - u(E - A)$$

and hence (6) is sufficient to ensure (7) holds.

Next we need to ensure that the household will always choose at $t = 2$ to repay the short term debt taken at $t = 1$. Suppose that the signal is bad $S = B$. If the borrower rolls over her existing debt her expected utility is

$$u(qE - A) - (1 - q)\Omega.$$

Alternately, if she defaults on her initial loan she will incur utility loss Ω and will take a new short term loan with face value E that will raise qE for consumption. We have already argued above that (6) ensures that this new loan will be repaid at $t = 3$ if $I = E$ and will be defaulted on if $I = 0$. The expected utility from this deviation is:

$$u(qE) - (2 - q)\Omega.$$

Thus the borrower will choose to repay her short term loan at $t = 2$ when $S = B$ if and only if

$$(8) \quad \Omega \geq u(qE) - u(qE - A).$$

Observe that (6) is sufficient to ensure (8) if and only if

$$u(qE - A + E) - u(qE - A) \geq u(qE) - u(qE - A),$$

which must hold since by assumption $qE > A$, which ensures $qE - A + E \geq qE$. This establishes that when the borrower takes the short term contract (6) ensures default only occurs when $I = 0$ at $t = 3$.

A.2. Default When Borrower Takes Long Term Contract at $t = 1$

Focus attention on the case where the borrower's type is presumed to be low and so the long term debt taken at $t = 1$ has a face value of

$$D_{1,3} = \frac{A}{p_L + (1 - p_L)q}.$$

The conditions needed to rule out default when the long maturity debt has a smaller face value will only be weaker than the conditions we consider here. Given this long term obligation the remaining pledgeable income (when $I = E$) is

$$E - \frac{A}{p_L + (1 - p_L)q}$$

and the household will fully borrow against this at $t = 2$ to insure against income risk. If $S = B$ this will generate savings of

$$qE - \frac{qA}{p_L + (1 - p_L)q}.$$

The only question is whether all debt will be repaid at $t = 3$ when $I = E$. By construction total debt due at $t = 3$ is E and so if the borrower repays she can consume from her savings. Alternately defaulting will afford an additional E in consumption but will incur the utility cost Ω . The borrower will choose not to default if and only if

$$(9) \quad \Omega \geq u\left(qE - \frac{qA}{p_L + (1 - p_L)q} + E\right) - u\left(qE - \frac{qA}{p_L + (1 - p_L)q}\right).$$

An analogous condition rules out default when $S = G$ but with $q = 1$. Due to the concavity of $u(\cdot)$ this condition is satisfied whenever (9) is. Finally we can show that (6) is sufficient to assure (9) holds as well. To see this note that the difference in consumption, E , from defaulting is the same in (6) and (9). However the level of consumption of consumption is strictly lower in (6). To see this compare the consumption achieved when the the loan is repaid in each scenario:

$$qE - A < qE - \frac{qA}{p_L + (1 - p_L)q} \iff 0 < p_L(1 - q)$$

where the right hand side is generated by simplifying the left hand side and is assured to be true by assumption. It follows directly that (6) is necessary and sufficient to ensure that default only occurs under either the short or the long term contract when $I = 0$ at $t = 3$.

B. Conditions for Separating Equilibrium

The expected utility for a high type taking the short term contract is the same as in the symmetric information case:

$$U_1^{Short,H} = p_H u(E - A) + (1 - p_H) u(qE - A) - (1 - p_H)(1 - q)\Omega.$$

If the high type borrowers deviates from the proposed separating equilibrium and takes the long term loan at $t = 1$ that is priced for the low types as per (2) her expected utility from this deviation will be

$$\tilde{U}_1^{Long,H} = p_H u\left(E - \frac{A}{p_L + (1 - p_L)q}\right) + (1 - p_H) u\left(qE - \frac{qA}{p_L + (1 - p_L)q}\right) - (1 - p_H)(1 - q)\Omega.$$

Note that this involves the same pattern of additional financing for consumption smoothing purposes at $t = 2$ as outlined in the symmetric information case in the text. Such a deviation will not be taken, and hence a separating equilibrium can only be sustained, when $U_1^{Short} \geq \tilde{U}_1^{Long}$, which holds when (3) holds.

C. Conditions for Pooling Equilibrium

We now look for conditions under which an equilibrium in which both types take the long term loan. Begin by observing that if a borrower funds A using short term debt then none of the loan terms she is offered depend on the market's assessment of her type. As a result we do not need to specify out of equilibrium beliefs for any borrower who deviates from the proposed equilibrium and funds A with short term debt. Hence the expected utility to any borrower who uses short term debt is identical

to the symmetric information case studied in the paper:

$$U_1^{Short} = p_k u(E - A) + (1 - p_k) u(qE - A) - (1 - p_k)(1 - q)\Omega.$$

Under a pooling equilibrium the perfectly competitive credit market will price the long term debt to have a face value of

$$D_{1,3} = \frac{A}{\bar{p} + (1 - \bar{p})q} \text{ where } \bar{p} \equiv \phi p_H + (1 - \phi)p_L.$$

Similar to the symmetric information case, the household will borrow as much additional funds as possible at $t = 2$ if $S = B$ to fully insure against the remaining income risk.

The resulting expected utility to type k for taking the long term loan is then

$$U_1^{Long} = p_k u\left(E - \frac{A}{\bar{p} + (1 - \bar{p})q}\right) + (1 - p_k) u\left(qE - \frac{qA}{\bar{p} + (1 - \bar{p})q}\right) - (1 - p_k)(1 - q)\Omega.$$

Since the low type strictly prefers the long term contract even in a separating equilibrium it follows immediately that they have no incentive to deviate from a pooling equilibrium.

A pooling equilibrium therefore can only exist if the high type is unwilling to deviate.

This requires $U_1^{Long,H} \geq U_1^{Short,H}$ which holds if and only if

$$p_H u\left(E - \frac{A}{\bar{p} + (1 - \bar{p})q}\right) + (1 - p_H) u\left(qE - \frac{qA}{\bar{p} + (1 - \bar{p})q}\right) \geq p_H u(E - A) + (1 - p_H) u(qE - A)$$

which holds if and only if

$$p_H \leq \hat{p}_H \equiv \frac{u\left(qE - \frac{qA}{\bar{p} + (1 - \bar{p})q}\right) - u(qE - A)}{u(E - A) - u(qE - A) - u\left(E - \frac{A}{\bar{p} + (1 - \bar{p})q}\right) + u\left(qE - \frac{qA}{\bar{p} + (1 - \bar{p})q}\right)}.$$

If $u(\cdot)$ is risk neutral then $\hat{p}_H = \bar{p}$ and hence pooling can only be sustained if there is no difference between the types. When $u(\cdot)$ is strictly concave we will have $\hat{p}_H > \bar{p}_H$.

Appendix C. inferring initial credit risk subgrade from data

LC assigns each loan's interest rate depending on the credit risk sub grade. In the data, the variable subgrade 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 subgrade 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.³¹ 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 subgrade"), 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 subgrade is then modified depending on the requested loan amount and maturity. For example, the initial subgrade of 36 month loans was not modified, while the initial subgrade 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:

³¹See <https://web.archive.org/web/20121204222850/http://www.lendingclub.com/public/how-we-set-interest-rates.action>

	Initial subgrade		
	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 subgrade of a borrower who requests a loan for \$10,000 is the same as her final subgrade before the modification for maturity. But a borrower who was ranked initially as C1 (equivalent to an 11) who requests a \$16,000 loan will see her grade modified two steps to a C3 (13).

Borrowers who share the same initial subgrade will have very similar risk characteristics as assessed by LC's lending model, while their interest rate will only vary according to their choice of amount and maturity. Thus, our analysis above uses the initial subgrade before amount and maturity modifications to construct fixed effects. But this variable not observable in the data. Instead, LC only provides the credit risk subgrade after all modification 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 subgrade that we cannot assign an initial subgrade from our reverse engineering procedure:

% of loans in our main sample period not assigned an initial subgrade

Final subgrade	Mean	Number of loans	Final subgrade	Mean	Number of loans
A1	0.007	1,511	D1	0.008	2,388
A2	0.002	1,160	D2	0.002	2,217
A3	0.004	1,259	D3	0.007	1,746
A4	0.008	1,458	D4	0.003	1,431
A5	0.016	2,157	D5	0.003	887
B1	0.004	5,075	E1	0.003	312
B2	0.008	5,592	E2	0.000	414
B3	0.005	5,979	E3	0.000	196
B4	0.006	5,908	E4	0.006	155
B5	0.006	2,908	E5	0.000	118
C1	0.092	3,772	F1	0.879	33
C2	0.016	2,939	F2	0.941	51
C3	0.005	3,112	F3	1.000	9
C4	0.005	2,676	F4	1.000	8
C5	0.005	2,218	G1	1.000	1

By construction, almost all loans below an F1 rating (26) will not have an initial subgrade because LC's model states that only 25 initial grades are issued. Second, we succeed in matching a borrower's initial subgrade for more than 98% of the loans of each final subgrade in 24 out of the 25 top subgrades. Grade C1 (grade 11) is slightly problematic as the success rate drops to 91%. The reason is that we should not observe C1 loans between \$15,000 and \$20,000, but LC categorizes 329 of these loans. All our results are robust to eliminating loans issued in final grade C1.