Real Effects of Search Frictions in Consumer Credit Markets*

Bronson Argyle† Taylor Nadauld‡ Christopher Palmer§

July 2017

Abstract

We document significant price dispersion in the market for car loans, provide direct evidence that it persists in part because of search frictions, and estimate how search frictions in credit markets distort consumption and modulate interest-rate shocks. Using rich microdata from millions of auto-loan applications and originations by hundreds of financial providers, we isolate plausibly exogenous variation in interest rates due to institution-specific rule-of-thumb pricing rules. These discontinuities lead to substantial variation in the benefits of search, which we find affect physical search behavior and distort extensive- and intensive-margin loan and car choices through quasi-random interest-rate markups. We further show that these discontinuities are more consequential in areas we measure as having high search costs. Overall, our results provide evidence of the real effects of the costliness of shopping around for credit, the continued importance of proximate bank branches, and how search frictions inhibit the transmission of monetary policy to durable goods purchases.

Keywords: price dispersion, search, auto loans, durables, regression discontinuity

*We thank our discussants Paul Calem, Anthony DeFusco, and Neale Mahoney; seminar and workshop participants at Berkeley-Haas, BYU, the 2016 CFPB Research Conference, the NYU Stern Salomon Center Conference on Household Finance, Ohio State, MIT, the WFA, and Duke; and John Campbell, Claire Celerier, Jan Eberly, Brigham Frandsen, Peter Ganong, Kyle Herkenhoff, Lars Lefgren, Andres Liberman, Brigitte Madrian, Adrien Matray, Adair Morse, Holger Mueller, Hoai-Luu Nguyen, Andrew Paciorek, Brennan Platt, Rodney Ramcharan, David Scharfstein, Aaron Schroeder, Amit Seru, David Sraer, Bryce Stephens, Johannes Stroebel, Stijn Van Nieuwerburgh, Stephen Zeldes, and Jonathan Zinman for helpful conversations. Tommy Brown and Sam Hughes provided excellent research assistance. An anonymous information-technology firm provided the data.

†Brigham Young University; bsa@byu.edu

‡Brigham Young University; taylor.nadauld@byu.edu

§Massachusetts Institute of Technology; cjpalmer@mit.edu

1 Introduction

Some of the most important open questions in household finance center around how credit-market imperfections affect consumption, including the role of adverse selection in consumer credit markets (Adams, Einav, and Levin, 2009), the importance of credit constraints in explaining high marginal propensities to borrow and consume out of credit (Gross and Souleles, 2002), and identifying the inhibition of credit expansions to the household sector (Agarwal et al., 2017a). In this paper, we provide evidence that costly search represents an additional friction in consumer debt markets that not only leads to interest-rate dispersion among similar loans but can distort extensive- and intensive-margin loan and consumption choices.

Using administrative data on 4 million auto loans extended by 326 different financial institutions in all 50 states and loan application data on 1.5 million potential loans from 41 institutions, we establish four main empirical facts. First, there is significant price dispersion for the same credit product across providers—60% of borrowers in our data could access significantly dominating loan offers if they could costlessly query all nearby financial institutions (see Figure 1). Second, such search is costly, and borrowers' propensity to search for loans with better terms is lower in areas likely to have higher search costs.[1] Third, the segment of the auto lending market we study does not feature pure risk-based pricing; we observe large loan-rate and loan-term discontinuities at various institution-specific FICO thresholds. Fourth, consumer purchasing and financing decisions are distorted by the resulting interest rate dispersion around these lending thresholds. Taken together, we argue that consumers fail to consistently identify optimal financing terms because of costly search in the retail auto loan market; this distorts financing decisions at the extensive and intensive margin as well as consumption decisions of durable goods.

We focus on the market for automobile-secured loans for several reasons. Auto loans

[1]Nationally representative survey evidence points to the apparent costliness of consumer search in credit markets. According to the 2013 Survey of Consumer Finances, one in five people self-report doing "almost no searching" when taking out a new loan. While such behavior could be driven by expected benefits of non-costly search being low, our results provide evidence that the benefits of search are likely substantial for many borrowers.

are ubiquitous and play an important role in the consumer credit complex. Over 85% of car purchases are financed, and there are over 0.8 outstanding car loans per U.S. household. Vehicles represent over 50% of total assets for low-wealth households (Campbell, 2006). Auto debt is the fastest-growing and the third-largest category of consumer debt, with over 106 million outstanding loans comprising $1.15 trillion in aggregate auto debt. From an empirical-design standpoint, auto loans are a relatively homogeneous credit product and can essentially be described completely by their interest rate, term, and amount. Finally, auto loan markets are quite local. The median borrower in our sample borrowing directly from a lender (as opposed to indirect loans originated via auto dealers) originates a loan from a branch that is within a 15 minute drive of her home, whereas the median worker in the United States commutes 26 minutes to work. This stylized fact that direct auto loan markets are more local than labor markets motivates our inquiry into the distortions that physical search frictions might cause in consumer debt markets.

Our empirical strategy features a setting where potential gains to search are high and quasi-randomly assigned. We document large discontinuities in offered loan terms around FICO thresholds across lending institutions. Lending policies that jump discontinuously at various FICO thresholds appear to exist in 173 of the 326 lending institutions in our sample. Notably, the location of the thresholds along the FICO spectrum varies across institutions; while some thresholds appear more popular than others, there is no consensus set of thresholds used by a plurality of lenders. Variation in the location of thresholds for lenders even in the same geography means that borrowers on the "wrong" side of a threshold at one institution could be on the "right" side of a threshold at another institution. We document in first-stage results that borrowers on the right of FICO thresholds are offered lower interest rates. On average, borrowers to the right of an institution's FICO threshold are offered loans with 1.37 percentage point lower interest rates as compared to otherwise similar borrowers just below a FICO threshold. As a result, borrowers just to the left of a lender's threshold would benefit from searching for loan offers from institutions with either

no discontinuity in offered rates or from institutions where borrowers would be on the right side of a given threshold.

Figure 2 provides examples of such interest-rate discontinuities for three different credit unions in our data with detected discontinuities using the lending policy rule estimation procedure described in Section 5.2. As discussed in Section 5.3 below, the observed FICO thresholds isolate supply-side changes in loan characteristics from demand-driven factors under the assumption that demand-side factors (e.g., preferences, income, financial sophistication) are not likely to also change discontinuously at quasi-random FICO thresholds that vary across institutions in the same geography. We support this assumption with evidence that ex-ante borrower characteristics (including age, gender, ethnicity, application DTI, application loan size, and the number of loan applications per FICO bin) are balanced around FICO thresholds.

What impact does sharp variation in loan pricing for otherwise identical borrowers have on borrower outcomes? Borrowers quasi-randomly offered expensive credit on average purchase cars that are 4 months older, spending an average of $901 less. The similarities mentioned above in borrowers across FICO thresholds suggest that borrowers arriving quasi-randomly on the expensive side of an arbitrary FICO threshold have similar preferences to those on the low interest-rate side of a pricing discontinuity and would thus presumably also like to purchase a more expensive and newer car had they not been assigned higher interest rates. Given that treatment (high markups) is as good as randomly assigned, we further ask whether there is selection into take-up by examining ex-post borrower outcomes. Subsequent changes in credit scores and ex-post loan performance do not change differentially by cutoffs, which we interpret as evidence that borrowers who take up dominated loan offers are not disproportionately likely to be low-quality borrowers, allowing us to interpret conditional-on-origination effects on second-stage consumption outcomes as causal. By using above-threshold borrowers as a counterfactual for below-cutoff borrowers, we are the first paper to our knowledge to quantify how search frictions distort consumption.

Potential explanations for equilibrium differences in offered interest rates around lending thresholds include adverse selection, measurement error, and search costs. Identifying search costs as a meaningful friction requires direct evidence that explains variation in measures of consumer search using measures of the costliness of consumer search. We show that borrowers on the expensive side of FICO thresholds reject high-interest-rate loans most often when the number of nearby alternative lenders is high.[2] Using the physical branch locations of every bank and credit union in the United States, we calculate the number of financial institutions within a 20 minute drive from each borrower as a proxy for search costs. We find that differences in loan take-up rates across FICO thresholds are smaller for borrowers in high search-cost areas. Borrowers that would presumably have to exert more effort to search for a loan with better terms are more likely to accept the loan pricing they are offered even though these terms are strongly dominated by nearby alternatives. Using a subsample of our data that allows us to link borrowers across loan applications to different lenders, we verify that borrowers are more likely to submit multiple loan applications when our search-cost measure is low. Finally, we show that search frictions have aggregate consequences for the transmission of monetary policy. When we examine how the interest rates of new originations change in response to contemporaneous changes in five-year Treasury yields, we find that high search cost areas have 10% less pass-through.

We also consider a series of robustness tests to address potential omitted variables that could be correlated with our physical measure of search costs. A Bartik strategy based on the 1990 network of lender branches in the United States allows us to address potential time-varying endogeneity in the number of proximate lenders (e.g., that banks close branches in response to hyperlocal economic conditions that also determine interest-rate elasticities). Because shift-share instruments essentially rely on the exogeneity of preexisting conditions, we also present a difference-in-differences strategy that allows us to rule out time-invariant

[2]Importantly, while loan take-up rates are lower on the expensive-side of FICO thresholds, borrowers do not *apply* for loans at differential rates across the FICO thresholds, consistent with our assumption that demand-side factors do not change at cutoffs.

explanations for our results (e.g., the concern that areas with low financial sophistication jointly have low interest-rate sensitivity and fewer nearby branches). Taken together, our evidence suggests that search costs represent a meaningful market friction that enables the persistence of equilibrium price dispersion and ultimately distorts consumption in the retail auto loan market.

The remainder of the paper proceeds as follows. After contextualizing our work in several related literatures in section 2, section 3 details the administrative data we use throughout the paper, including an analysis of its representativeness. Section 4 documents price dispersion in the market for auto loans. Section 5 presents results detecting discontinuities in lender price rules and introduces our regression-discontinuity identification strategy. In Sections 6 and 7, respectively, we present evidence that consumers propensity to search is correlated with measures of search costs, and we estimate the effects of costly search on loan and durable-purchase outcomes. To demonstrate the aggregate importance of search frictions beyond the sample of borrowers we consider here, Section 8 examines the differential transmission of interest-rate shocks to areas with high and low search costs. Section 9 concludes.

2 Related Literature

In this section, we motivate our work in connection with the literature on search frictions, auto loans, and FICO-based regression discontinuities.

Theories of costly search (e.g., Stahl, 1989) suggest that when some agents find it too costly to solicit the full menu of offered prices, equilibrium prices will reflect the distribution of offered prices and the random draw that each agent acquires from the offered price distribution. Lenders can expect to make loans in the presence of search costs despite not offering the lowest rates among their competitors because of the possibility that a randomly arriving customer will not exert the effort required to find better rates. Consider a financial institution that offers an interest rate on auto loans that is high relative to competitors,

conditional on borrower quality. If search is costly, consumers that arrive randomly to solicit a loan are more likely to accept the offered rate despite the existence of better available rates. Similarly, entrants cannot profitably undercut overpriced competitors because of entrants' inability to inform and attract consumers. Lowering search costs should therefore result in lower price dispersion as consumers increase their propensity to search and are more likely to be informed about the complete distribution of available prices. In equilibrium, if consumer search costs are reduced, lenders offer more competitive rates, essentially facing a decline in market power. For a comprehensive treatment of the history of thought in the theoretical and empirical search and price dispersion literature, see Baye, Morgan, and Scholten (2006).

Multiple empirical papers establish the existence of equilibrium price dispersion (necessitating ruling out product heterogeneity as a driver of price variation) and connect it to evidence that consumer search is costly in a given domain. For example, Sorenson (2000) documents dispersion in prices of prescription drugs that are driven by proxies for likely search intensity. In consumer finance, Hortaçsu and Syverson (2004) find large dispersion in the fees charged by very similar mutual funds that are driven by information/search frictions. Woodward and Hall (2012) document that mortgage borrowers overpay for mortgage broker services due to a reluctance to shop for mortgages. Survey evidence also confirms the costliness of consumer search. In addition to documenting price dispersion in mortgage rates, Alexandrov and Koulayev (2017) provide survey evidence indicating two key findings; first, close to half of consumers did not shop for a mortgage before origination and second, consumers are unaware of price dispersion. Zinman & Stango (2015) use a self-reported measure of shopping intensity to explain variation in price dispersion in the credit-card market. All of these results are consistent with questions on search intensity in the 2013 Survey of Consumer Finance wherein many borrowers self-report doing very little shopping around for a loan.

Relative to the literature on price dispersion and search intensity that often models consumers as having inelastic unit-demand for a final good, our setting allows for measurement

of distortions in consumption that can result from costly search, in particular how search frictions in credit markets have real effects on purchasing decisions. Analogous to supply-side frictions that modulate the pass-through of monetary policy (e.g., Scharfstein and Sunderam, 2016 and Agarwal et al., 2017a), we further show that search frictions have the potential to temper the efficacy of monetary policy if consumers are unwilling or unable to search out the distribution of available credit to find the rates that have responded to declining risk-free rates.

Recent work by Agrawal, Grigsby, Hortacsu, Matvos, Seru, and Yao (2017) shows that in the cross-section, intensive loan search is correlated with higher interest rates, running counter to the standard prediction that search and selected prices are inversely correlated. Agarwal et al. (2017b) explain this with a model of borrower private information about the returns to search—low credit-worthy borrowers search until they find a lender who offers them an advantageous interest rate, albeit higher than the rates offered without search to (observably) high-quality borrowers. In our setting, the quasi-random assignment of our regression-discontinuity design effectively allows us to abstract away from cross-sectional variation in private information and rely the conceptual argument that for a given borrower, the relationship between search and interest rates should be negative.

We also note that we are not the first paper to exploit FICO-based discontinuities in treatment variables. Keys et al. (2009 and 2010) find that the probability of securitization (and thus loan screening) change discontinuously at a FICO score of 620. Bubb and Kaufman (2014) provide evidence for other discrete FICO thresholds in the mortgage underwriting process, including detailing the likely genesis of threshold-based policies. More recently, Agarwal et al. (2017a) use sharp FICO-based discontinuities in credit limits to estimate heterogeneity in marginal propensities to borrow and lend, and Laufer and Paciorek (2016) evaluate the consequences of minimum credit-score thresholds for mortgage lending. Building on this collection of papers that either use FICO-based discontinuities as natural experiments or explicitly study their consequences, we are the first to identify

credit-score–based discontinuities in pricing rules and to link those discontinuities to price dispersion, costly consumer search, and distortions in consumption.

Finally, we contribute to a growing literature studying the automobile loan market and the frictions therein, including Attanasio, Goldberg, and Kyriazidou (2008), Adams, Einav, and Levin (2009), and Einav, Jenkins, and Levin (2012 and 2013).

3 Data

We analyze the loan contract terms and auto purchasing decisions of just under 4 million individual borrowers in the United States from 326 retail lending institutions between 2005–2016. The loan data are provided by a technology firm that provides administrative data warehousing and analytics services to retail-oriented lending institutions nationwide. Roughly two thirds of the lending institutions represented in the data set are credit unions ranging between $100 million and $4 billion in asset size. The remainder are non-bank finance companies of unknown total asset size, although the vast majority (98.5%) of the loans in our data are originated by credit unions.[3] Borrowers from all 50 states are represented in the data, but the five largest states in the data are Washington (465,553 loans), California (335,584 loans), Texas (280,108 loans), Oregon (208,358 loans), and Virginia (189,857 loans).

The dataset contains information capturing all three stages of a loan's life: application, origination, and ex-post performance, although we have loan application data for only approximately 1.5 million loans from 41 different institutions. The available loan application data report borrower characteristics (ethnicity, age, gender, FICO scores, and debt-to-income (DTI) ratios at the time of application), whether a loan application was approved or denied, and whether it was subsequently withdrawn or originated. For originated loans, the data additionally include information on loan amounts, loan terms, car purchase prices, and whether the loan came through a direct or indirect origination channel.[4] We restrict our sample to

[3]Our results are unchanged if we exclude loans from finance companies, which are generally of lower credit quality.

[4]The terms direct and indirect loans refer, respectively, to whether the borrower applied for a loan

direct loans in an effort to address concerns that indirect loans are potentially endogenously steered to specific financial institutions (perhaps because car dealers become aware of lenders' pricing rules over time). Finally, to measure ex-post loan performance, we observe a snapshot of the number of days each borrower is delinquent, whether each loan has been charged off, and updated borrower credit scores as of the date of our data extract.

Panels A, B, and C of Table 1 present summary statistics on loan applications, loan originations, and measures of ex-post performance, respectively. As reported in Panel A of Table 1, the median loan application in our data seeks approval for a five and a half year $19,824 loan at a median interest rate of 3.75%.[5] Borrowers applying for loans in our data have an average credit score of 646 and an average DTI ratio of 28.3%. The percentage of loans approved is 50.2%, with 78.4% of the approved borrowers subsequently originating a loan. Throughout the paper we refer to the number of loans originated divided by the number of applications approved for a particular group as the loan take-up rate. We exploit variation in the loan take-up rate in Section 6.2.

Panel B of Table 1 reports summary statistics on loan originations, revealing several interesting patterns. Compared with loan applications, originated loans have smaller average sizes, higher interest rates, shorter terms, and are from more creditworthy and less constrained borrowers. Average monthly payments for originated loans are $338 per month with an interquartile range of only $200.

Panel C tabulates measures of ex-post loan performance. While the average loan is 36 days delinquent, most loans are current; the 75th percentile of days delinquent is zero and only 2.1% of loans have been charged-off (accounted as unrecoverable by the lender). Defining default as a loan that is at least 90 days delinquent, default rates average 2.2%. In untabulated results, default rates for borrowers with sub-600 FICOs average 6.8%, compared

directly to the lending institution or through an auto dealership that then sent the loan application to lending institutions on the buyer's behalf.

[5]Application interest rates are strongly right skewed with a mean interest rate of 17.3% and a 75th percentile of 12.7%. These risky outliers appear to be rejected, as they are not in the originated loan sample in Panel B.

to a default rate of 2.6% for borrowers with FICOs between 600 and 700 and 1.6% for over-700 FICO borrowers. Lending institutions periodically check the credit score of their borrowers subsequent to loan origination, creating a novel feature of our data. Summary statistics for ΔFICO represent percent changes in borrowers' FICO scores from the time of origination to the lender's most recent (soft) pull of their FICO score.[6] Updated FICO scores indicate that borrowers on average experienced a 1% reduction in FICO score since origination, although borrowers with FICO scores below 600 on average realized a 5.7% increase in FICO score.

3.1 Data Representativeness

The bulk of our auto loan data come from credit unions, prompting questions about the representativeness of the data. Our data confirm popular perception is that credit union usage is concentrated among older consumers. Over 41% of borrowers in our sample were between 45 and 65 years old at loan origination whereas 34% of the adult U.S. population are between the ages of 45–65. Borrowers in our sample are also less racially diverse than the general public. Over 73% of our sample are estimated to be white (as of 2015), compared to an estimated 65% of adults in the general population as recorded by the 2015 American Community Survey.[7] Borrowers in our data report median FICO scores at origination of 715 (Table 1, Panel B) over the full 2005-2016 sample period. The NY Federal Reserve Consumer Credit Panel (CCP), a representative 5% sample of U.S. borrowers, reports median FICO scores for originated auto loans of 695 during the period our sample was collected.[8] In summary, our sample contains borrowers that are slightly older, less racially diverse, and of a higher average credit quality than national averages. These sample biases should not limit our ability to draw inference given that much of our inference relies on a regression

[6]The time between FICO queries varies by institution, but institutions that provide updated FICO scores do so at least once a year such that conditional on having an updated FICO score, the amount of time between the original FICO recording and the current FICO is roughly equal to loan age.

[7]Borrowers do not report race at the time of loan origination but most lenders in our sample estimate race ethnicity in an effort to comply with fair lending standards.

[8]The CCP data report quarterly median FICO scores over our sample period. The reported 695 median FICO is actually the median of the quarterly medians that span our sample period.

discontinuity (RD) design that leans crucially on an assumption of smoothness in borrower demographics across discontinuities at a given institution.

A second data validity issue involves the distribution of loan originations through time. Over 70% of loan originations in our sample occurred between 2012 and 2015, despite a sample period that runs from 2005–2016. The large increase in loans through time reflects the increase in the client base of our data provider through time rather than auto credit origination in general. Auto loan originations in the general population have increased through time, from an aggregate outstanding balance of $725 billion in Q1 2005 to just over $1.15 trillion in Q4 2016, but not at the rate reflected in our dataset. We view the non-representative time series of our data as less relevant to our inference given our cross-sectional identification strategy.

A third data validity issue is whether credit unions capture a meaningful fraction of the auto loan market. Experian data from 2015 indicates that credit unions originated 22% of all used car loan originations and 10% of new car originations in the United States. The Experian data do not differentiate direct lending from indirect lending, but in the auto loan data made available to our data provider by its clients, roughly two-thirds are direct loans. Finally, we note that data on the performance of auto loans as reported in the CCP suggests that auto loans originated by credit unions and banks have substantially lower default rates as compared to loans originated by auto finance companies.[9]

4 Documenting Price Dispersion

Diagnosing a market with dispersed prices requires ruling out any product differentiation, i.e. establishing that differences in prices truly represent identical goods being sold for different prices in the same market. For any given borrower with an observable set of attributes, we estimate the spread between the borrower's loan note rate and the lowest available interest rate at another lender in our data for another borrower with very similar attributes. To cal-

[9]The CCP does not separate auto loans made by credit unions from those made by banks.

culate this spread, we group borrowers in the same Commuting Zone, six-month transaction date window, five-point FICO bin, $1,000 purchase-price bin, same loan maturity, and ten percentage-point DTI bin. The $1,000 auto purchase bins are non-overlapping, beginning from $2,000 to $2,999, up to a maximum purchase amount of $100,000. Five-point FICO bins are also non-overlapping. We consider loans originated to borrowers within the same CZ $\times$ time $\times$ price $\times$ FICO $\times$maturity$\times$DTI cell to be effectively identical.[10] Owing to the strictness of this criteria, many borrowers in our data are in their own cell, limiting our ability to calculate a spread. Moreover, because we do not observe interest-rate offers from lenders that are not clients of our data provider, these spreads are lower bounds (having the universe of interest rates offered to a given cell could only weakly decrease the best available rate).[11] Albeit incomplete, because of the richness of our data coverage across hundreds of providers, we have thousands of cells with multiple borrowers.

Figure 1 plots the density of the spread to the best available rate in percentage points for the 60% of borrowers who did not attain the best rate in their cell. The mean and median of this distribution are 173 and 119 basis points, respectively. Including the 19% of borrowers who are getting the best available rate given their discrete borrower type, the average borrower in our data is thus paying 1.4 percentage points more than an observationally equivalent borrower at the same time in the same place. We provide further evidence of price dispersion in section 6.

Exclusivity of Credit Unions By definition, a credit union is a member-owned cooperative financial institution that requires membership to receive financial services. Often, credit unions' membership requirements restrict eligibility to well-defined groups. Because most of the loans in our sample were originated by credit unions, one concern is whether a given

[10] Although there may be some degree of residual heterogeneity within a cell, the magnitude of the variation we find is sufficiently large that it would be difficult to explain with borrower-level heterogeneity alone. Moreover, in 43% of the cells, the best rate in the cell is achieved by a borrower with a lower FICO and higher DTI than other borrowers in the cell. Nevertheless, our RD design below establishes the existence of large pricing disparities for identical credit risks.

[11] We discuss the particular case of digital lenders in section A.3 below.

borrower could have joined the credit union with the best available rate used to demonstrate the existence of dominating loan opportunities for that particular cell. For example, if the lowest available interest rate that we assume could have been obtained by a borrower was offered by a firefighters' credit union, then borrower search costs would not only involve the effort required to find the low rate but also the effort required to become a firefighter. To address this concern, we re-calculate the spread-to-lowest-available rate measures using a sample comprised entirely of credit unions whose primary membership requirement is residence in a specified geographic area. In other words, all borrowers in our CZ-based matched portfolios are eligible to become a member at any of the credit unions included in their cell by virtue of living in the same CZ as others in their cell. Our results are nearly identical after making this restriction. We also note that the finance companies in our sample have no membership requirements.

5 Estimating the Effects of Search Costs

In this section we introduce an empirical strategy designed to identify the role of costly search in the price discovery process. As previously noted, an accurate diagnosis of price dispersion can be hindered by measurement error and/or unobserved heterogeneity, particularly in a credit market where soft information can be at play. We exploit exogenous variation in the data that creates a setting where the potential gains to search are high for some borrowers, yet variation in the gains to search is quasi-randomly assigned across borrowers. Exploiting the variation in gains to search, we estimate whether borrowers propensity to search is correlated with proxies for the cost of search. The setting also has the benefit of limiting the possibility that measurement error and/or unobserved heterogeneity explain the patterns we observe in the data. In subsequent sections, we use the RD laboratory to document distortions in auto purchases that can be attributed to financing terms that persist on account of search frictions.

5.1 Detecting Discontinuities

Lending institutions make underwriting decisions about whether to approve a loan application using a combination of hard and soft information on borrower credit quality. Hard information generally consists of quantifiable credit metrics provided by credit bureaus or verified with paystubs and tax statements such as FICO scores, debt-to-income ratios, bankruptcy history, and annual earnings. Soft information, loosely defined as information that cannot be easily quantified related to the likelihood of a borrower's future willingness or ability to repay a loan, is by definition unobservable to the econometrician.[12] Any econometric analysis that specifies loan outcomes as the dependent variable is subject to the critique that equilibrium loan outcomes are influenced by unobservable soft information, complicating inference related to factors causing an outcome of interest. Our setting is no exception. While our dataset consists of millions of equilibrium lending outcomes, our ability to draw inference is hindered by the possibility that unobserved soft information plays a role in jointly determining selection into application and origination, observed loan terms, and subsequent loan performance. Because our sample consists of direct auto loans, soft information in our setting would most likely be generated from the relationship between credit unions and their long-term customers, observable to a loan officer.

We address this possibility, and other potential omitted variables, with an RD design that exploits observed discontinuities in offered loan terms across several FICO thresholds. Unlike the 620 FICO heuristic in mortgage *underwriting* first exploited by Keys et al. (2009 and 2010) that affects screening at both origination and securitization (Bubb and Kaufman, 2014), we focus on discontinuities in loan *pricing*, i.e., the interest rate offered to a borrower conditional on having a loan application approved by underwriting. Moreover, no industry standard set of thresholds exist in auto lending as opposed to mortgage lending. Still, while auto loan lending institutions do not adhere to a common set of FICO cutoffs, the use of a given threshold at some point across the FICO spectrum is prevalent for most

[12] See Petersen (2004) for a careful treatment of hard and soft information in financial markets.

lenders in our data. See Bubb and Kaufman (2014) for a model of a particular credit risk processing cost structure and Al-Najjar and Pai (2014) for a model of overfitting that could each rationalize binning risk types. Anecdotally, lending institutions have confirmed that their pricing functions explicitly incorporate discrete FICO thresholds to set interest rates and loan terms.[13] Also in contrast to Keys et al. (2010), FICO thresholds observed in our data have little to do with secondary markets given that many auto loans are retained by the lending institutions in our dataset. Rather than reflecting demand for securitization or a loan's subsequent marketability on a secondary market, FICO discontinuities may have been incorporated into software systems as a holdover from a time when pricing was done via rate sheets instead of automated algorithms and could persist in part because costly consumer search prevents more accurately risk-based pricers from gaining market share.[14]

To illustrate the effect of FICO thresholds on equilibrium interest rates, we estimate lender-specific interest-rate and loan-term policies nonparametrically. For each lender c in our data, we characterize their lending policies across FICO bins with a set ψ of parameters $\{\psi_{ck}\}$ where k indexes FICO bins denoted $\mathcal{F}_k$. Pooling loan-level data from institution c, we estimate ψ by regressing an origination outcome y_{ic} (interest rates or loan maturity) on a set of indicator variables for each 5-point FICO bin $\mathcal{F}_k$

$$y_{ic} = \sum_k \psi_{ck} \mathbb{I}(FICO_i \in \mathcal{F}_k) + \varepsilon_{ic} \tag{1}$$

where ε_{ic} includes all other factors that influence loan pricing. The 5-point FICO bins begin at a FICO score of 500 where the first bin includes FICO scores in the 500-504 range, the second bin includes 505-509, etc., up through FICO scores of 800. The estimated coefficients

[13] As an example, one executive pointed to a FICO score of 610 as the explicit cutoff that determines the loan terms offered to prospective borrowers at that executive's credit union. Applicants with a FICO score just below 610 were offered higher rates and loan terms below 60 months in contrast to applicants with FICO scores above 610.

[14] In the mortgage industry, Bubb and Kaufman (2014) write that "Though [Automated Underwriting Systems] calculate default risk using smooth functions of FICO score, they also employ a layer of 'overwrites' which trigger a 'refer' recommendation when borrowers fall into certain categories—for instance, borrowers with FICO scores below 620." See Hutto & Lederman (2003) for a history of the incorporation of discrete credit score cutoffs into automated underwriting systems for mortgage lending, such as those created by Fannie Mae and Freddie Mac.

on each FICO bin represent the average interest rate for loans originated to borrowers with FICO scores in that bin relative to the estimated constant (the omitted category is loans outside this range—we focus on relative magnitudes for this exercise).

Figure 2 presents interest-rate plots for three different financial institutions. The estimated $\hat{\psi}$ point estimates represent how that lender's pricing rules appear to vary with borrower FICO score, and the accompanying 95% confidence intervals provide a sense of how reliant on FICO scores was each lender's pricing rule. Panel A of Figure 2, estimated on one institution in our data with approximately 12,000 borrowers (rounded to preserve lender anonymity), illustrates breaks in average interest rates for borrowers with FICO scores around FICO cutoffs at 600, 660, and 700. The breaks in interest rates at the FICO cutoffs are large (representing jumps of over 2 percentage points). Average interest rates for borrowers in the 595-599 FICO bin are 2.5 percentage points higher than the average interest rate for borrowers in the 600-604 FICO bin, and the difference in average interest rates between the two bins are statistically significant at the .001 level. Panels B and C illustrate similar rule-of-thumb FICO breaks for unique institutions with approximately 6,000 and 25,000 loans, respectively. One important observation arising from these anecdotal plots is the fact that the breaks occur at different FICO scores across different institutions, consistent with our understanding that the discontinuities are reflective of idiosyncratic pricing policies across institutions.

In order to standardize our analysis to include every institution that employs discontinuous pricing rules, we empirically identify the existence of discontinuities at each institution (if they exist at all) in our sample through the following criteria. We first estimate the interest-rate FICO bin regressions following equation (1) for each institution in our sample separately. To establish the existence of a economically and statistically significant interest-rate discontinuity, we require that interest rate differences across consecutive bins be larger than 50 basis points and be estimated with p-values that are less than 0.001. We further refine the set of discontinuities by requiring that differences between leading and following

FICO bin coefficients ψ_{ck} have a p-value of at least 0.1 and that an identified discontinuity not lie within 20 FICO points of another identified discontinuity at the same institution. This restriction limits any potential contamination that could occur if borrowers simultaneously fall into a treated sample at one observed threshold but serve as a control for a sample at a different threshold. We further examine each potential threshold visually to ensure that the identified discontinuities are well behaved around the candidate thresholds. Finally, in an effort to maximize the statistical power in our RD design, we require that each candidate threshold contain 100,000 loans within the span of 38 FICO points around the candidate threshold, forming a discontinuity sample (Angrist and Lavy, 2004). The 38 FICO points represent 19 points on either side of a threshold that do not bump up against a different threshold that could exist within 20 FICO points. Implementing each of these restrictions ultimately results in large and meaningful discontinuities in interest rates and loan terms at FICO scores of 600, 640, and 700 across 57 institutions and 378,521 loans.[15] Table 2 reports summary statistics for our ultimate estimation sample (the set of loans within 19 points of one of our thresholds). A comparison of the full sample summary statistics (Table 1) with the threshold-specific sample (Table 2) reveals that the threshold sample is similar to the full sample along observables. All of the RD estimates reported in the paper use the discontinuity sample.

5.2 First-Stage Results

To validate our RD design, we present a series of diagnostics designed to test whether our data meet the two main identifying assumptions underlying RD estimation. First, the RD approach assumes that the probability of borrower treatment (i.e., offered interest rates) with respect to loan terms is discontinuous at FICO thresholds of 600, 640, and 700. Second,

[15] We reiterate that not all institutions have thresholds at 600, 640, and 700—these are merely the most popular detected discontinuities satisfying our criteria. Relaxing the requirement of 100,000 loans within 38 FICO points around the threshold results in a larger set of identified thresholds. The two most populated thresholds outside of our selected three thresholds are at 680 and 660 which contain approximately 90,000 and 80,000 loans, respectively.

valid RD requires that any borrower attribute (observed or unobserved) that could influence loan outcomes change only continuously at interest-rate discontinuities. This smoothness condition requires that borrowers on either side of a FICO threshold are otherwise similar, such that borrowing outcomes on either side of a threshold would be continuous absent the difference in treatment induced by policy differences at the threshold.

In our remaining specifications, we normalize FICO scores to create a running variable $\widetilde{FICO}_{ict}$ that measures distance from a interest-rate discontinuity. For example, for loans near the 600 FICO score threshold, $\widetilde{FICO}_{ict} = FICO_{ict} - 600$. Panel A of Figure 3 plots average interest rates against normalized borrower FICO scores for a sample restricted to loans with borrower FICO scores between 581 and 619. The plots demonstrate smoothness in the conditional expectation function except for the points corresponding to a FICO score of 599 and 600, where interest rates jump discontinuously. We repeat the plot using similar 38 point FICO ranges for the 640 and 700 FICO thresholds in panels B and C of the same figure. These plots confirm the existence of large interest-rate discontinuities at these thresholds. The magnitude of the discontinuities appears to be smaller at higher FICO thresholds, which might arise from smaller relative differences in credit quality at high FICO score levels.

To establish statistical significance and introduce our RD design, we estimate first-stage regressions of the form

$$y_{ict} = \beta_1 \widetilde{FICO}_{ict} + \beta_2 \mathbb{I}(\widetilde{FICO}_{ict} \geq 0) + \beta_3 \widetilde{FICO}_{ict} \cdot \mathbb{I}(\widetilde{FICO}_{ict} \geq 0) + \alpha_{cz(i)} + \delta_t + \varepsilon_{ict} \quad (2)$$

where y_{ict} is the outcome for loan i originating from lending institution c in quarter t, $\mathbb{I}(\widetilde{FICO}_{ict} \geq 0)$ is in indicator variable equal to one if the normalized FICO score $\widetilde{FICO}_{ict}$ is above the threshold, and $\alpha_{cz(i)}$ and δ_t are Commuting Zone and and quarter fixed effects, respectively. In practice, we conservatively estimate equation (2) using the Robust RD estimator of Calonico, Cattaneo, and Titiunik (2014), estimating the effect of the running variable $\widetilde{FICO}$ above and below the cutoff at $\widetilde{FICO} = 0$ using local linear regression (as opposed to the unweighted linear specification we provide for intuition in equation (2)) and

a local quadratic bias correction.[16] Our baseline regression specification pools each of the three discontinuities into one dataset using the FICO normalization described above. We cluster our standard errors by normalized FICO score.

Table 3 presents results of this exercise. Interest rates for borrowers with FICO scores immediately above one of our thresholds are estimated to be 1.37 percentage points lower than borrowers just below (column 1). Column 2 reports that loan maturities for borrowers just above a FICO threshold are 1.47 months longer than otherwise similar borrowers below the threshold. Given an average interest rate in our estimation sample of 6.8% (Panel B of Table 2, the magnitude of this coefficient is economically meaningful and shows that landing on the so-called wrong side of a interest rate discontinuity has material consequences on the cost of credit. For the remainder of the paper, to facilitate brevity and ease of interpretation, we refer to left-of-threshold borrowers, those on the expensive side of thresholds or below thresholds as LOT (left-of-threshold) borrowers. Right-of threshold borrowers, borrowers above thresholds, or borrowers on the lower-interest rate side of thresholds are referred to as ROT (right-of-threshold) borrowers.

5.3 Testing Quasi-Random Assignment Assumption

To test whether other observables beside the treatment variables (interest rate and loan maturity) also change discontinuously at our detected FICO thresholds, in Figure 4, we pool loans in the neighborhood of all three FICO thresholds and plot the average value of other borrower characteristics around these FICO thresholds along with the Calonico et al. (2014) estimated RD function and associated confidence intervals. Importantly, these graphs are constructed with loan application data in order to ensure that borrowers are similar at FICO thresholds along characteristics at the time of application. Panels A–E plot borrower debt-to-income ratios, loan amounts, borrower age in years, borrower gender (an indicator for male), and borrower ethnicity (an indicator for white), respectively. These plots

[16] While our reported results use a uniform kernel with a bandwidth of 19, our results are robust to alternative kernels and a wide range of bandwidths.

indicate smoothness in ex-ante borrower characteristics around FICO thresholds. Borrowers on either side of FICO thresholds do not appear meaningfully different in terms of their debt capacity, their willingness to borrow, or along demographics. Finally, Panel F plots the number of applicants within each normalized FICO bin, along with the McCrary (2008) test for bunching in the running variable, showing that borrowers do not appear to select into applying for a loan based on their FICO score. Such manipulation of the running variable—a discontinuity in the propensity to apply for a loan at a FICO threshold—would raise selection concerns but would be difficult to accomplish given the uncertainty applicants face about their own credit scores (owing to the volatility of FICO scores and uncertainty about which credit bureau(s) a lender will query) and the low likelihood that prospective borrowers are aware of the precise thresholds used by a given lender.

Table 4 reports the magnitude and significance of the discontinuity coefficients using the loan-application data, available for a subset of lending institutions. The estimates indicate no statistical difference in requested loan amounts for borrowers on either side of the threshold (column 1). In column 2 we present estimates of differences in debt-to-income ratios around the thresholds. Ex-ante debt-to-income ratios of borrowers on either side of the thresholds are statistically indistinguishable. Finally, we count the number of applications received from borrowers of each normalized FICO score and examine these counts at the FICO-score level using our RD estimator. Column 3 shows that the number of borrowers applying for loans is also not statistically different on either side of as threshold.

Our empirically detected discontinuities in loan pricing at specific FICO thresholds are large (nearly 150 basis points) and are unaccompanied by similar discontinuities in borrower composition supporting our reliance on a regression discontinuity design.

6 Evidence on Loan Search and Search Costs

6.1 Measuring Potential Gains to Search

The documented discontinuities represent exogenous variation in the benefits to search because they estimate a counterfactual interest rate should a near-threshold borrower choose to search for a different loan. This is particularly salient in our setting because of variation in the location of FICO thresholds across institutions, even within the same commuting zone (CZ). Borrowers that find themselves on the expensive side of a lending threshold, knowingly or unknowingly, could plausibly find a lower interest rate (all else equal) were they to search.

We quantify the magnitude of better available rates for LOT borrowers in Table 5, which tabulates the average spread-to-the-best-available rate for borrowers with FICO scores from 595 to 599, 635–639, and 695–699 as being 3.4 percentage points (pp), 2.5 pp, and 1.6 pp, respectively. That is, for borrowers with FICO scores between 595 and 599 there was a loan with a 3.4 percentage point lower interest rate originated to someone with the same FICO and DTI in the same CZ at the same time and to secure a similarly priced car. The standard deviation of the spread across these cells is 2.5%, 2.1%, and 1.4% with an average number of borrowers in the cell of 2.46, 2.99, and 3.99, respectively. Figure 5 provides visual evidence of better available rates by plotting the spread-to-lowest-available-rate for LOT and ROT borrowers. Dotted lines in each plot are for LOT borrowers, solid lines for ROT borrowers. LOT borrowers, those offered exogenously higher rates, exhibit substantially higher average spreads from the lowest available rate and larger variance in those spreads. This pattern holds for each of the three thresholds. Taken together, these plots confirm that price dispersion is largest for borrowers exogenously offered higher interest rates. We further note that Figure 5 plots do not just confirm higher average rates for LOT borrowers, they also confirm that a larger fraction of LOT borrowers, as compared to the ROT counterfactual borrowers, accepted interest rates that were clearly dominated by nearby available loans.

6.2 Measures of loan search and search costs

Why did borrowers treated with expensive rates not avail themselves of better lending opportunities? In this section we evaluate whether proxies for search costs can explain borrowers reluctance to shop. This analysis requires the construction of two measures: a measure of search propensity and a measure of search costs.

Our first measure of search propensity is the fraction of borrowers that accept an offered loan, a measure we refer to as the loan take-up rate. The loan take-up measure relies on the assumption that borrowers who applied for, but then reject an offered loan reject the loan in favor of searching for a different loan. While borrowers could reject loans for different reasons, including the decision not to originate any loan, we view differences in the decision to accept an offered loan around lending thresholds as a reasonable measure of differences in borrowers' propensity to engage in further loan search. Using reported FICO scores at the time of loan application, we estimate differences in take-up rates around FICO thresholds.

To proxy for search costs, we use FDIC and NCUA data to identify the physical location of every bank branch and credit union branch in the United States for each year in our application data. We then create a measure we call *proximity to financial institutions* (PFI hereafter), which measures the driving-time density for an individual borrower by geocoding and counting the number of physical branch locations within a 20-minute drive of the borrower, similar to the approach employed by Degryse and Ongena (2005). The calculation of actual driving times relies on posted speed limits along current driving routes and abstracts from traffic conditions and any changes to the road network between the time of loan origination and 2016 (the date of our driving-time data). We calculate driving distances and trip durations for only those institutions which existed at the time of loan origination.

The driving-time density measure, PFI, is designed to capture the effort, proxied by travel time and physical distance, for each borrower to search out a lending institution that is within a reasonable distance from their home.[17] Our PFI measure finds external validity

[17]Recall that our sample consists of direct auto loans originated through a lending institution, as opposed

in work by Goodstein and Rhine (2017), who show that households that live within close proximity to bank branches are more likely to have a bank account and thus less likely to use nonbank transaction products. We also note that there are many other dimensions over which search is costly besides time and distance that we do not measure, for example the disutility of filling out financial paperwork and potential concerns that numerous credit-registry queries negatively impacts credit scores (see Liberman, Paravisini, and Pathania, 2016, for a discussion of the impact of this characteristic of credit scoring).

If PFI is correlated with search costs, and thus influences the propensity to search, differences in loan take up-rates around FICO thresholds should be larger in areas with higher PFI (i.e. more lenders live close by, thus lower search costs). If a borrower faces higher search costs (lower PFI, i.e. lower driving-time densities), LOT borrowers in that area, i.e. those offered unfavorable loan terms, would be less likely to reject the unfavorable loan terms in favor of searching for better terms elsewhere. Importantly, in a differences-in-differences spirit, our empirical specification measures *differences* in loan take-up rates around FICO thresholds. We then compare differences in loan take-up rates for borrowers in high versus low PFI areas.

The median borrower in our application data lives within a 20 minute drive of 72 lending institutions. Borrowers in the 25th percentile of driving distance live less than a 20-minute drive from 23 institutions, as compared to 168 institutions for borrowers in the 75th percentile. Just under 4% of applicants in our sample live in an area with one or fewer lending institutions within a 20-minute drive. In contrast, 45% of applicants live within a 20-minute drive of at least 100 different lending institutions. Our baseline results categorize borrowers with fewer than 10 lending institutions within a 20-minute drive as facing high search costs (lowest 15% in the density distribution) to capture nonlinearity in the effect of additional

to indirect auto loans where dealers undertake a digital search across multiple lenders at the time of the auto purchase. We discuss the question of digital search in A.3. Distance can also proxy for soft-information producing relationships (see Nguyen, 2017), although we do not believe auto-loan lending to be a particularly relationship-intensive credit product.

nearby lenders on search costs.[18] In Section 6.2.1 below, we verify robustness of our results to the definition of high search cost area.

Using our standard RD framework described by equation (2), we specify the dependent variable as an indicator equal to one for accepted loans and estimate differences in the take-up rate around FICO thresholds. In Table 6, we estimate RD regressions of equation (2) using the Calonico et al. (2014) estimator for the full sample and separately for borrowers with more than and less than 10 lending institutions within a 20-minute drive. For the full sample, column 1 documents differences in take-up rates of 7.6 percentage points around the threshold; that is, borrowers randomly drawing high markups from a dispersed distribution of prices are 7.6 percentage points less likely to accept the offered loan. We estimate the same regression in high and low search cost areas and report results in columns 2 and 3, respectively. In high search cost areas, LOT borrowers demonstrate no differences in loan take-up rates as compared to ROT borrowers. In contrast, differences in take-up rates around FICO thresholds are 9.4 percentage points in low search cost areas (column 3). In other words, borrowers in low search-cost areas are much less likely to take-up a loan if they are assigned a high interest-rate markup, whereas high search-cost borrowers are likely to accept loan terms regardless of where they fall in the distribution of prices.[19] Finally, column 4 shows that the difference between the discontinuity coefficients in the high and low search-cost samples is statistically significant.

[18] In unreported results we split the sample based on bottom vs. top quartile and again by above and below-median driving density. Top quartile vs bottom quartile driving densities offer qualitatively similar results to our reported results. Above and below-median splits produce similarly signed, but a smaller magnitude of differences in the difference of take-up rates across thresholds.

[19] One threat to our interpretation of Table 6 is that the magnitude of the first-stage may be different in high and low search cost areas. Note, however, that for this to explain our results, low search-cost areas would have to have *larger* discontinuities in rates. Appendix Table (A1) finds a small and insignificant difference in the size of the average interest-rate change at a FICO threshold. See Appendix Section A.1 for further discussion.

6.2.1 Robustness to Definition of High Search Cost Area

Figure 6 probes whether our take-up results are sensitive to the choice of cutoff for high search cost area. Each point and associated confidence interval plots $\hat{\gamma}_2$ from estimating

$$takeup_{ict} = f(\widetilde{FICO}_{ict}) + \gamma_1 HighSearchCost_i + \gamma_2 HSC_i \cdot \mathbb{I}(\widetilde{FICO}_{ict} \geq 0) + \alpha_{cz(i)} + \delta_t + \varepsilon_{ict} \tag{3}$$

via OLS where $f(\widetilde{FICO})$ is an RD quadratic in the normalized FICO score running variable ($f(\cdot)$ includes interactions between each term and an indicator for $\widetilde{FICO} \geq 0$) and *High Search Cost* is a dummy equal to one when the number of lenders within a 20-minute drive is in a particular range. We estimate (3) varying the cutoff for high search cost to be 5 institutions, 10, 15, etc. Two interesting patterns emerge from the coefficients. First, most of the point estimates are negative. Recall that the main effect of a FICO discontinuity on take-up rates is positive; ROT borrowers are much more likely to accept an offered loan because they are being offered much more favorable interest rates than LOT borrowers. Being in a high search cost area blunts the effect of the FICO discontinuity on take-up. Second, while borrowers with five or less PFIs are rare and so the resulting take-up discontinuity is estimated imprecisely, the estimates for borrowers who have less than 20 PFIs are significantly less than zero. Moreover, estimates for definitions near our baseline cutoff of 10 PFIs (15th percentile) are similar, suggesting that our conclusions are fairly insensitive to the exact definition of high search cost. Finally, consistent with our intuition that from a search-cost perspective, the value of an additional PFI diminishes (i.e., each lender matters a lot when there are only a handful of nearby options few lenders compared to when there are, for example, over 50 lenders nearby), the point estimates diminish towards zero in the number of PFIs used in the definition of high search cost.

6.2.2 Explaining Variation in Price Dispersion with Search Costs

Our PFI measure of search costs also has power to explain some variation in price dispersion. One prediction of search theory is that lowering search costs should result in lower price

dispersion, as consumers facing low search costs are more likely to be informed about the complete distribution of available prices. In Figure 7, we plot the cumulative density function of the spread-to-lowest available rate variable. The solid line is the CDF for borrowers facing low search costs and the dotted line is for high search cost borrowers. The CDF plots indicate that a smaller fraction of high search cost area borrowers accept loans with lower spreads-from-the-lowest-available-rate, i.e. price dispersion is larger in high search cost areas. Although the magnitude of the difference between the two distributions is likely attenuated by the limited coverage of our data relative to the entire local auto-loan market, Kolmogorov-Smirnov tests confirm that the distributions are statistically different at the 1% confidence level.

6.3 Direct Evidence of Loan Search

In this section, we present a more direct measure of loan search by attempting to link loan applications across financial institutions for the same borrower to evaluate whether the propensity of an individual borrower to search for an auto loan, as measured by the number of filed loan applications, varies with our PFI proxy for search costs. We assume that loan applications originating from the same nine-digit zip code from a borrower with the same birthdate are from the same individual and that loan applications occurring within 90 days of each other are for the same prospective purchase. As before, we divide borrowers into high and low search cost areas as those with less than and more than 10 lending institutions within a 20-minute drive, respectively. Table 7 reports that applicants in high search cost areas (column 1) apply for an average of 1.07 loans per vehicle purchase. In contrast, applicants in low search cost areas (column 2) apply for an average of 1.12 loans/vehicle purchase. Column 3 shows that this difference is statistically significant at a 1% significance level. Although small, these averages are a lower bound as we do not observe applications to any lender not in our data (the vast majority of potential lenders). Moreover, in the sequential search equilibrium of Stahl (1989), consumers always buy from the first seller in equilibrium.

Regressions of loan applications per purchase on the count of lending institutions within a 20-minute drive confirms a positive and significant relationship. These results are consistent with the more comprehensive indirect evidence presented above that borrowers facing high search costs search less and accept worse rates than borrowers facing relatively lower search costs.

6.4 Comments on Identification

See the Appendix for a discussion of identification issues related to our measure of search costs.

7 Distortions to Consumption

To estimate how search costs can have real effects on consumption we next establish that being treated with a higher interest rate causally distorts loan and purchase decisions. Whether a given credit-market imperfection is a binding constraint on optimal consumption is difficult to ascertain because it requires estimating counterfactual consumption in the absence of the alleged constraint. However, we can determine the existence of potential distortions by evaluating the auto purchasing decisions of borrowers on either side of the documented FICO thresholds. Given the empirical result that, ex-ante, borrowers are similar around FICO thresholds, we start with the null hypothesis that borrowers around FICO thresholds would also have similar demand for cars, conditional on obtaining the same set of financing terms. Exploiting our ability to observe the exact amount that each borrower spent on a car, we test whether borrowers spend differently around the observed FICO thresholds.

Figure 8 plots car purchase amounts around the normalized FICO threshold. Purchase amounts are smooth leading up to the FICO threshold and then jump discontinuously at the threshold. Using the same RD design used in our first-stage analysis presented above, we formally test for statistical differences in purchase amounts. As before, we estimate

equation (2) by controlling for lending institution fixed effects, quarter-of-origination fixed effects, allowing for a local linear function of the running variable, bias-correction using the local quadratic approach of Calonico et al. (2014), and using a bandwidth of 19 around the normalized FICO threshold with a uniform kernel. Table 8 presents these reduced-form results. Borrowers quasi-randomly offered more expensive loans spend an average of $901 less on the cars they purchase. Column 2 presents results with loan amounts as the dependent variable. Originated loan sizes are an average of $1,424 lower on the expensive side of a detected FICO discontinuity. The fact that loan sizes increase by larger amounts around the threshold than purchase amounts indicates that, ex-post, borrowers on the right side of the cutoff are allowed and choose higher loan-to-value ratios. This result is confirmed in column 3 of Table 8, which indicates that ex-post LTV ratios are an average of 2.8 percentage points higher for borrowers to the right of FICO thresholds.

Detailed data on loan amounts and loan terms allow us to calculate the implied monthly payment of borrowers on either side of the thresholds. In column 4 of Table 8 we test whether ex-post monthly payments are different around the thresholds. On average, monthly payments increase by $9.64 per month for ROT borrowers. Shorter terms and higher interest rates cause LOT borrowers to purchase less expensive cars and use less financing in their purchase than ROT borrowers, essentially purchasing less car and less credit for only a $10 reduction in monthly payment.

Still, higher ex-post LTV ratios and slightly higher monthly payments, implying higher ex-post DTI ratios, warrant further consideration. One concern is the possibility that borrowers on either side of the FICO thresholds are different ex-ante in their ability to service debt, violating the smoothness/exogeneity conditions required for valid RD. Given that ex-ante DTI ratios in the loan application data are continuous around the thresholds (Table 4), we interpret these results as further evidence of the easing of credit terms for ROT borrowers. That is, ex-post, ROT borrowers are offered lower rates, longer terms, and apparently allowed higher ex-post LTV and DTI ratios.

This evidence of otherwise similar borrowers spending different amounts on the cars they purchase as a result of the financing terms they are offered is consistent with consumption distortion. An alternative explanation is the possibility that dealers somehow price discriminate and exploit borrowers' ability to service larger debt amounts by charging more for the exact same car than otherwise similar LOT borrowers.[20] We address this possibility by controlling for year-make-model (e.g., 2013 Honda Accord) fixed effects in our RD regressions.[21] Column 1 of Table 9 reports results when controlling only for make-model fixed effects. Even within a make and model category, borrowers quasi-randomly assigned expensive credit continue to spend $805 less on cars, suggesting that the purchasing behavior we observe in Table 8 is not driven by people choosing to purchase different model cars as a result of their assigned credit. Contrasting the coefficients in columns 1 and 2 provides indirect evidence as to the exact nature of the substitution patterns in this market. When we include year-make-model fixed effects in column 2, we find no statistical differences in the car purchase price, suggesting that it is not that borrowers with more affordable credit are paying more for the same car by choosing extra add-on features or simply being overcharged by the seller. Reconciling these two results, column 3 provides direct evidence with vehicle age at purchase in months as the dependent variable (controlling for make-model fixed effects since vehicle age would be collinear with year-make-model fixed effects). Borrowers with access to easier credit purchase 4.15 months newer cars, on average.

How do borrowers respond to being arbitrarily offered more expensive credit than their creditworthiness would warrant? The evidence presented in Tables 8 and 9 indicate that borrowers offered expensive credit spend less on their car purchases by selecting an older car than they would have otherwise, originating smaller loans in the process and having slightly smaller monthly payments. The lack of a difference in car pricing outcomes with year-make-model fixed effects is also evidence that there is little private information or selection around

[20] Again, our decision to restrict our sample to direct loans mitigates this concern somewhat, although buyers may share details of their financing terms with dealers.

[21] Year-make-model fixed effects are made possible by vehicle identification numbers (VINs) provided in our data set.

buyers' savvy or skill in car buying decisions based on assignment to one side of a FICO threshold. We view this as evidence that borrowers' inability to find the best available loan terms distorts consumption away from efficient levels.

7.1 Evaluating Alternative Explanations: Adverse Selection

In this section we address the possibility that LOT borrowers who take-up loan offers are different on unobservable dimensions from ROT car buyers. While we have already demonstrated that borrowers seem similar on observable dimensions at the ex-ante application stage, an alternative explanation for our results is that (unobservably) high credit-quality borrowers who are arbitrarily offered expensive interest rates withdraw their loan applications and look elsewhere for credit. Under this scenario, borrowers who follow through with the origination of expensive loans are those who know they are of poor credit quality and unlikely to do better given their unfavorable soft attributes. Lending institutions could also recognize that borrowers that choose to accept the unfavorable terms are indeed lemons, as anticipated, and so the arbitrary thresholds reinforce an equilibrium that separates high credit-quality borrowers from low credit-quality borrowers, with the appropriate pricing differences offered to each borrower type.

We test for the possibility that adverse selection drives the observed equilibrium outcomes in our data by comparing ex-post borrower performance around the FICO thresholds. If an unobservable selection process guides differences in who accepts expensive loan offers, this should be revealed by ex-post credit outcomes as lower credit-quality borrowers eventually default relatively more. To test this hypothesis, we first specify as a dependent variable in our RD setting the number of days a borrower is subsequently delinquent on their car loan. The coefficient in column 1 of Table 10 estimates that ROT borrowers are an average of 1.25 fewer days delinquent than LOT borrowers, indicating that borrowers on either side of the threshold do not exhibit economically meaningful or statistically significant differences in delinquency. Similarly, ROT borrowers are 0.11% more likely to have their loan charged

off (written off as a loss by the lender) and 0.01% less likely to be in default (over 90 days past due), both of which estimates we view as relatively precise zeroes.[22]

A novel feature of our dataset allows for a second test of adverse selection as an explanation for our observed results. As a means of monitoring borrowers, many lending institutions in our dataset pull credit scores on borrowers after loan origination. Ex-post credit score queries occur as frequently as every six months, and, in a few cases, as infrequently as once post-origination. The most common convention for the subset of institutions that pull credit ex-post is to pull credit scores once a year. Ex-post credit scores allow us to calculate changes in credit scores over time, capturing broad changes in borrower distress and financial responsibility. Any unobserved heterogeneity driving selection into loan take-up should impact credit scores over time if low credit-quality borrowers for whom the below-threshold expensive interest rate is "fair" are the only ones to originate such loans. Using the sub-sample of institutions that collect updated FICO scores after origination, we use the percentage change between credit scores at origination and the most recently observable credit score as the dependent variable in our RD framework. Results presented in column 4 of Table 10 show no meaningful differences in credit score changes for borrowers around the threshold.

Taken together, the evidence on borrower delinquency, defaults, and ex-post changes in credit scores indicate that LOT borrowers do not represent meaningfully different credit risks as compared to otherwise similar ROT borrowers. While adverse selection is undoubtedly a motivator of many features of retail car loan markets (Adams, Einav, and Levin, 2009), information asymmetries do not appear to be a primary determinant of the acute differences in lending behavior around the observed FICO thresholds.[23]

[22]The sample size differs across columns in Table 10 because of inconsistent data coverage of all variables across lenders.

[23]Adverse selection is not the only alternative explanation for our observed results around thresholds. For example, FICO thresholds could promote the steering of financially unsophisticated borrowers into higher rate loans. However, as any such borrower naïveté is not manifest in differences at loan application, more expensive car purchase prices paid, differential ex-post default rates, or differences in ex-post credit scores, it is unlikely to be a driving factor for the phenomena we document here.

8 Interest-rate Passthrough

While the evidence above supports our conclusion that costly search significantly distorts consumption among borrowers seeking loans from lenders in our data with pricing discontinuities, do search frictions affect origination interest rates more broadly? In this section, we examine the relationship between comparable-maturity risk-free interest rates $Treas5_t$ and the origination interest rates of car loans. To do so, we estimate a pass-through regression explaining individual interest rates in our data r_{igt} in location g in month t

$$r_{igt} = \beta_1 Treas5_t + \beta_2 Treas5_t \cdot HighSearchCost_g + \beta_3 HSC_g + X_i'\gamma + \xi_{FICO(i)} + \alpha_{cz(g)} + \delta_t + \varepsilon_{igt} \tag{4}$$

where $High\,Search\,Cost$ is an indicator equal to 1 if location g has 10 or fewer lenders within a 20-minute drive, X_i is a vector of loan-level controls including loan amount, loan maturity, and an indicator for high-cost loans (loans with rates above 6% that are likely evaluated as high credit risks and less responsive to risk-free rates). The terms $\xi_{FICO(i)}$, $\alpha_{cz(g)}$, and δ_t are fixed effects for FICO bin, Commuting Zone, and month, respectively. Table 11 reports the results of estimating specifications resembling equation (4). Column 1 shows that for a 100 basis point decrease in the 5-year Treasury note, car loan interest rates are expected to decrease by 61 basis points that month. The interaction term between high search cost and the 5-year note yield indicates that high search cost areas decrease by only 39 basis points for every 100 basis point decrease in the 5-year, passing through substantially less of innovations to the risk-free rate. Column 2 controls for month fixed effects, absorbing the main effect of the level of the 5-year note and any other macroeconomic changes that affect interest rates nationwide. Column 3 additionally controls for loan amount and maturity. In both of these additional specifications, the interaction term is negative and significant, indicating attenuated pass-through of interest rate shocks to areas with high search costs. The R^2 of estimating equation (4) indicates that this parsimonious model explains about half of the variation in interest rates across time and space. We interpret these results as

consistent with the idea that in areas with high search costs, originated car loan interest rates are less reactive to monetary policy because even if some lenders do decrease rates in response to changes in risk-free rates, borrowers are less likely to discover those prices.

9 Conclusion

Mounting evidence indicates that credit-market imperfections influence household debt outcomes. In this paper, we make three main points. First, we presented evidence that substantial price dispersion exists in retail auto lending markets. About 80% of borrowers in our sample for whom a comparable borrower exists did not originate a loan with the lowest available interest rate.

Second, we demonstrated the importance of considering search costs as a credit market imperfection with material effects on loan characteristics. In a setting where the gains to search are potentially very high, we show that borrowers' acceptance of dominated loan terms is related to measures of search costs. Because arbitrary pricing schedules vary across lending institutions within the same commuting zone, borrowers on the expensive side of FICO discontinuities in loan pricing at one institution would likely find themselves on the favorable side of a pricing threshold at a different institution. Absent search frictions, borrowers would be much less likely to accept dominated loan terms. However, we find that for borrowers randomly offered high interest rates in our sample, a lower interest rate loan was frequently originated by an equally creditworthy borrower purchasing a similarly priced car on a similar date, suggesting that borrowers are either unwilling or unable to search for more favorable loan terms. Proxying for the costliness of acquiring information on loan terms with the number of institutions within 20 minute drive from a borrower, we show that in areas with higher search costs, borrowers are more likely to accept inferior loan terms, limiting their access to marginal prices in credit markets. Although we have focused on probing the interaction between our detected discontinuities in pricing rules and search costs, the evidence

suggests that the broader retail car loan market is subject to costly search and the resulting price dispersion.

Third, we estimated distortions in consumption that are associated with borrowers' acceptance of dominated loan terms. Borrowers that accept randomly-offered higher interest rates purchase cars $901 lower in value, that are 4 months older on average, and borrow $1,424 less than otherwise similar borrowers that accept randomly-offered lower rates. This highlights the importance of well-functioning consumer credit markets in determining durable goods consumption patterns. Moreover, when consumer search is costly, the pass-through of credit expansions (e.g. from easier monetary policy) may be muted. Relative to consumer search models with inelastic unit demand where dispersed prices have no associated deadweight loss and just represent a transfer from buyers to sellers, there are welfare consequences of costly search in the real world. Given demand with at least some degree of elasticity, consumers facing markups from a market characterized by price dispersion may choose not to purchase the given good at all or, conditional on purchasing it, may adjust the characteristics of the purchased good away from their first-best level of consumption.

Even with a well-developed financial sector including secondary markets for many forms of consumer debt, household consumption is still distorted by credit market imperfections such as costly search. At least one answer to Zinman's (2014) query as to why efficient risk-based pricing is still not ubiquitous in the era of big-data-based credit modeling appears to be demand-side obstacles finding lowest available prices. While the existence of price dispersion has been documented in prior empirical work, we provide novel estimates of the cost of search at the extensive and intensive margin. Even with the possibility of shopping for interest rates online, searching for consumer credit products remains an opaque, local, and costly process for many borrowers. This relationship between costly search and distortionary credit market imperfections extends our understanding of equilibrium price dispersion to credit markets and could motivate extra regulatory attention on so-called banking deserts.

References

Adams, W., L. Einav, and J. Levin (2009). "Liquidity constraints and imperfect information in subprime lending." *The American Economic Review* 99.1, pp. 49–84.

Agarwal, S., S. Chomsisengphet, N. Mahoney, and J. Stroebel (2017). "Do Banks Pass Through Credit Expansions to Consumers Who Want to Borrow?" *Quarterly Journal of Economics* Forthcoming.

Agarwal, S., J. Grigsby, A. Hortaçu, G. Matvos, A. Seru, and V. Yao (2017). "Search and Screening in Credit Markets: Evidence from the US Mortgage Market." Working Paper.

Alexandrov, A. and S. Koulayev (2017). "No Shopping in the U.S. Mortgage Market: Direct and Strategic Effects of Providing More Information." Working Paper.

Angrist, J. D. and V. Lavy (1999). "Using Maimonides' rule to estimate the effect of class size on scholastic achievement." *The Quarterly Journal of Economics* 114.2, pp. 533–575.

Attanasio, O. P., P. Koujianou Goldberg, and E. Kyriazidou (2008). "Credit constraints in the market for consumer durables: Evidence from micro data on car loans." *International Economic Review* 49.2, pp. 401–436.

Baye, M. R., J. Morgan, and P. Scholten (2006). "Information, search, and price dispersion." *Handbook on Economics and Information Systems*. Ed. by T. Hendershott. Vol. 1. Citeseer, pp. 323–77.

Bertrand, M. and A. Morse (2011). "Information disclosure, cognitive biases, and payday borrowing." *The Journal of Finance* 66.6, pp. 1865–1893.

Bubb, R. and A. Kaufman (2014). "Securitization and moral hazard: Evidence from credit score cutoff rules." *Journal of Monetary Economics* 63, pp. 1–18.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82.6, pp. 2295–2326.

Campbell, J. Y. (2006). "Household finance." *The Journal of Finance* 61.4, pp. 1553–1604.

Degryse, H. and S. Ongena (2005). "Distance, lending relationships, and competition." *The Journal of Finance* 60.1, pp. 231–266.

Einav, L., M. Jenkins, and J. Levin (2012). "Contract pricing in consumer credit markets." *Econometrica* 80.4, pp. 1387–1432.

— (2013). "The impact of credit scoring on consumer lending." *The RAND Journal of Economics* 44.2, pp. 249–274.

Gross, D. B. and N. S. Souleles (2002). "Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data." *Quarterly Journal of Economics* 117.1, pp. 149–185.

Hortaçsu, A. and C. Syverson (2004). "Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds." *The Quarterly Journal of Economics* 119.2, pp. 403–456.

Hutto, G. W. and J. Lederman (2003). *Handbook of Mortgage Lending*. Mortgage Bankers Association of America.

Keys, B. J., T. Mukherjee, A. Seru, and V. Vig (2009). "Financial regulation and securitization: Evidence from subprime loans." *Journal of Monetary Economics* 56.5. Carnegie-Rochester Conference Series on Public Policy: Distress in Credit Markets: Theory, Empirics, and Policy November 14-15, 2008, pp. 700–720.

— (2010). "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans." *The Quarterly Journal of Economics* 125.1, pp. 307–362.

Laufer, S. and A. Paciorek (2016). "The Effects of Mortgage Credit Availability: Evidence from Minimum Credit Score Lending Rules." FEDS Working Paper No. 2016-098.

Liberman, A., D. Paravisini, and V. S. Pathania (2016). "High-Cost Debt and Borrower Reputation: Evidence from the UK." Working Paper.

McCrary, J. (2008). "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics* 142.2, pp. 698–714.

Al-Najjar, N. I. and M. M. Pai (2014). "Coarse decision making and overfitting." *Journal of Economic Theory* 150, pp. 467–486.

Nguyen, H.-L. (2017). "Do Bank Branches Still Matter? The Effect of Closings on Local Economic Outcomes." Working Paper.

Petersen, M. A. (2004). "Information: Hard and Soft." Working Paper.

Scharfstein, D. S. and A. Sunderam (2016). "Market Power in Mortgage Lending and the Transmission of Monetary Policy." Working Paper.

Sorensen, A. T. (2000). "Price Dispersion in Retail Markets for Prescription Drugs." *Journal of Political Economy* 108.4, pp. 833–850.

Stahl, D. O. (1989). "Oligopolistic pricing with sequential consumer search." *The American Economic Review*, pp. 700–712.

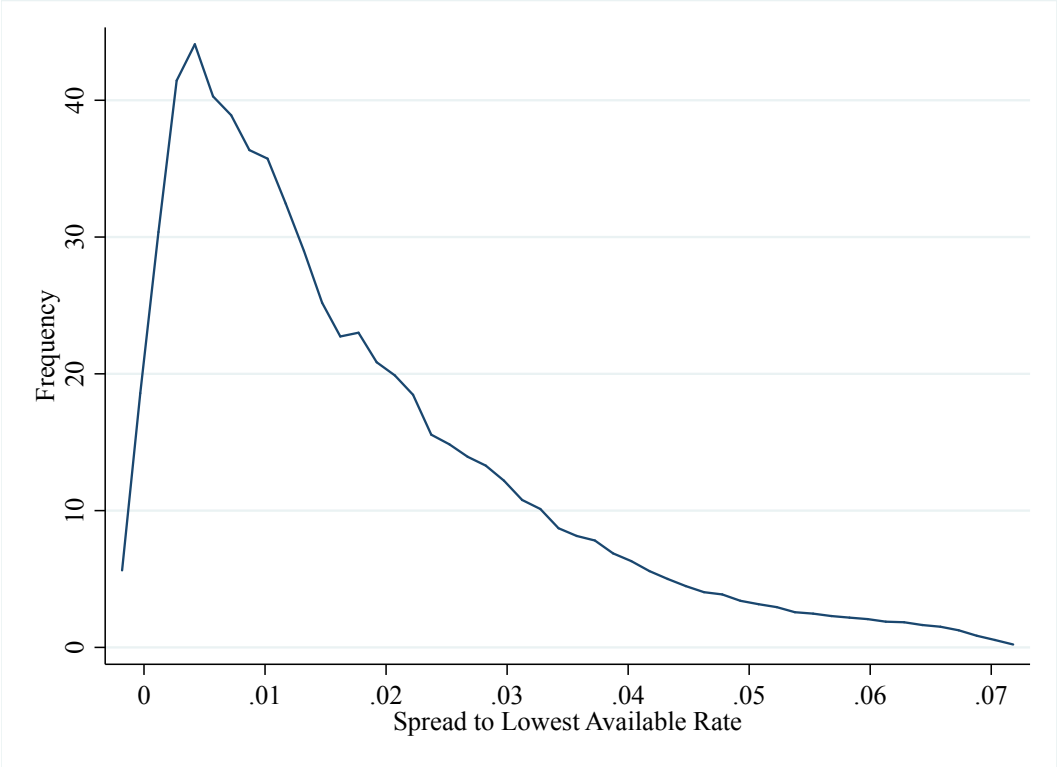
Stango, V. and J. Zinman (2009). "Exponential growth bias and household finance." *The Journal of Finance* 64.6, pp. 2807–2849.

— (2015). "Borrowing high versus borrowing higher: price dispersion and shopping behavior in the US credit card market." *Review of Financial Studies*, hhv072.

Woodward, S. E. and R. E. Hall (2012). "Diagnosing consumer confusion and sub-optimal shopping effort: Theory and mortgage-market evidence." *The American Economic Review* 102.7, pp. 3249–3276.

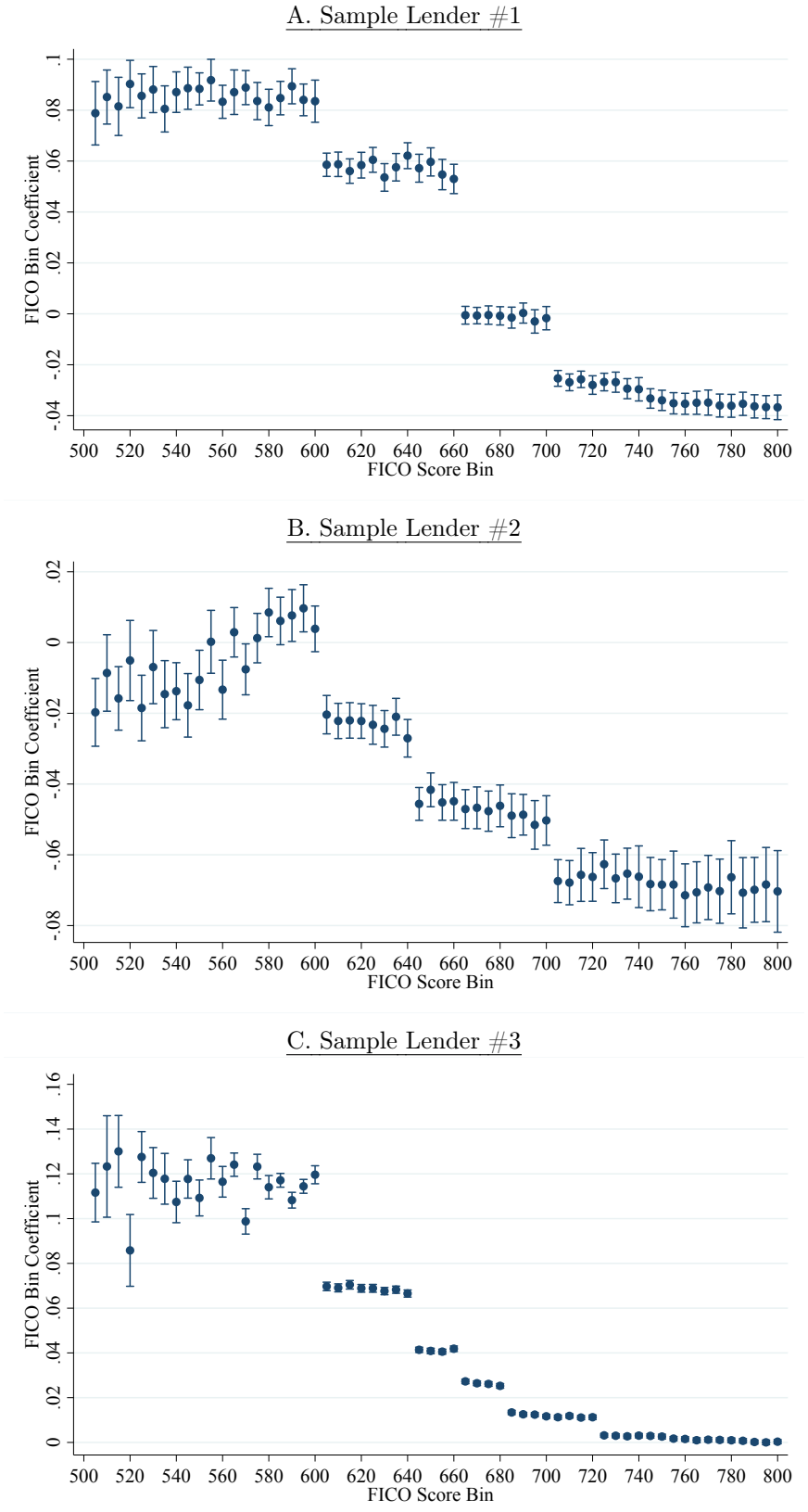
Zinman, J. (2014). "Consumer Credit: Too Much or Too Little (or Just Right)?" *The Journal of Legal Studies* 43.S2, S209–S237.

Figure 1: Density of Spread to Lowest Available Rate



Notes: Figure reports the kernel densities of the spread (in percentage points) to the lowest available rate for borrowers not receiving the best available rate in their cell. A cell is defined as as all borrowers in the same commuting zone taking out a loan in the same $1,000 collateral-value bin, 5-point FICO bin, 10-point DTI bin, 6-month time period, and loan maturity.

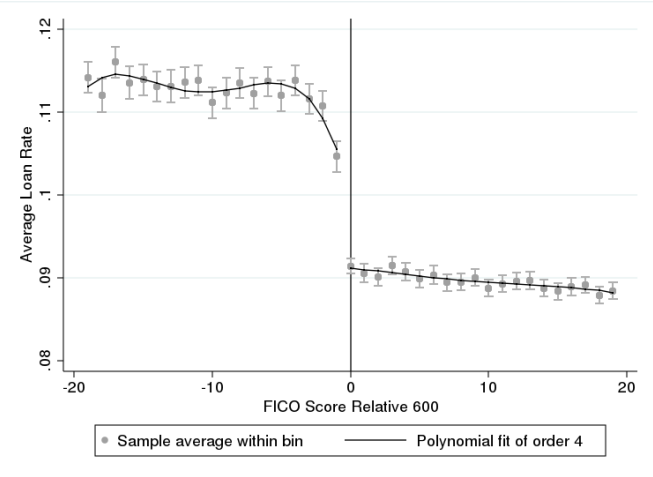
Figure 2: Examples of FICO-Based Discontinuities in Interest-Rate Policies



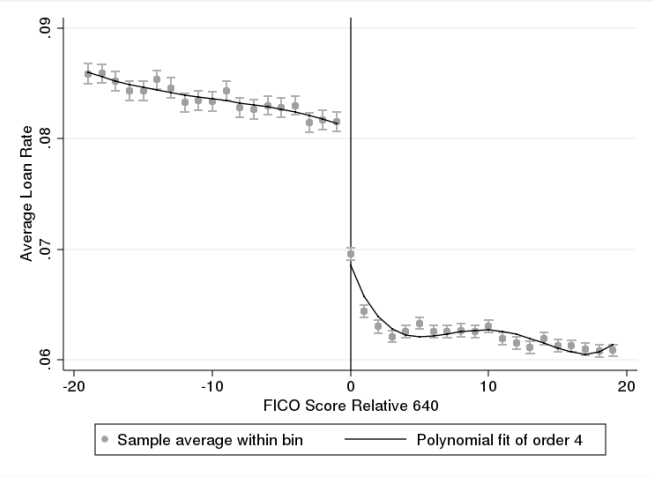
Notes: Each panel plots estimated interest-rate rules (with 95% confidence intervals) for a lender in our sample. Loan rates in percentage points are regressed on 5-point FICO bin indicators as in equation (1).

Figure 3: FICO-Based Lending Policies - Interest Rates

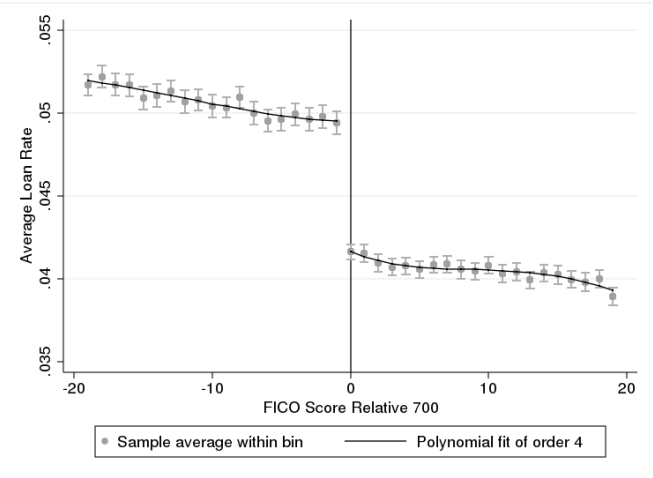
A. Interest Rates Around FICO = 600 Discontinuities



B. Interest Rates Around FICO = 640 Discontinuities

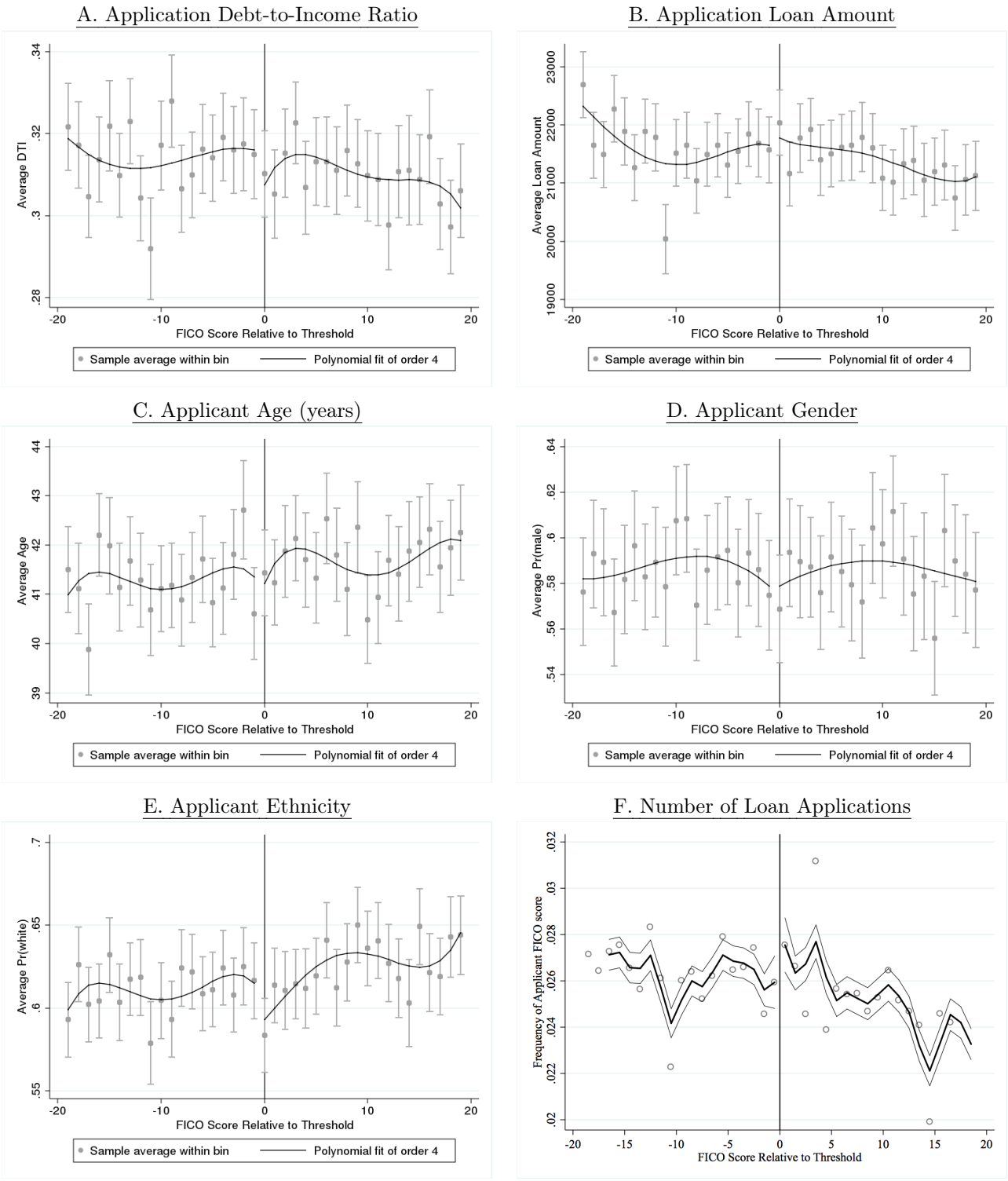


C. Interest Rates Around FICO = 700 Discontinuities



Notes: Figures plot average interest rates on the vertical axis against borrower FICO scores normalized to each threshold along the horizontal axis for institutions with pricing discontinuities detected at FICO scores of 600, 640, and 700, respectively.

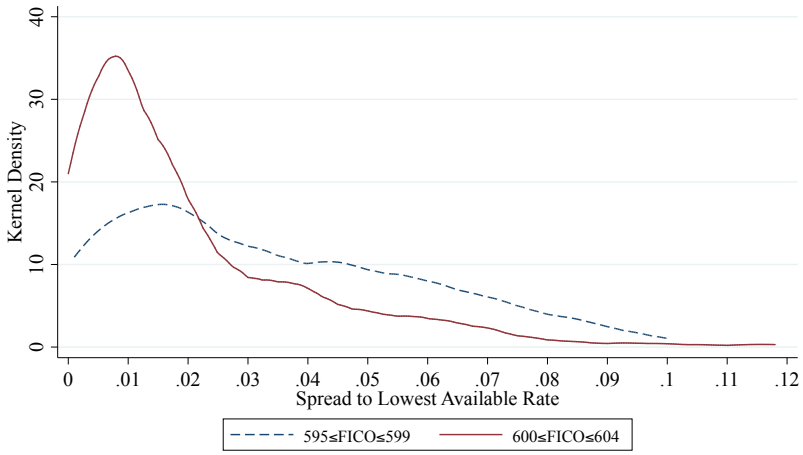
Figure 4: Balance of Borrower Characteristics Across FICO Thresholds



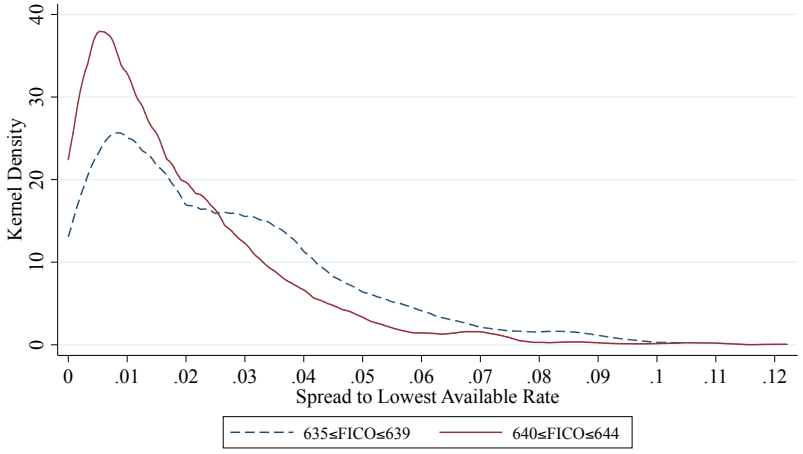
Notes: Figures plot average values of ex-ante borrower characteristics around FICO thresholds for institutions with detected discontinuities with Calonico et al. (2014) local linear estimates and confidence intervals. Applicant gender in panel D is an indicator for male, and ethnicity in panel E is an indicator for whether the applicant is estimated as white by the lender. Panel F plots the application count within each normalized FICO bin along with the estimated McCrary (2008) test.

Figure 5: Density of Spread to Lowest Available Interest Rate Around FICO Thresholds

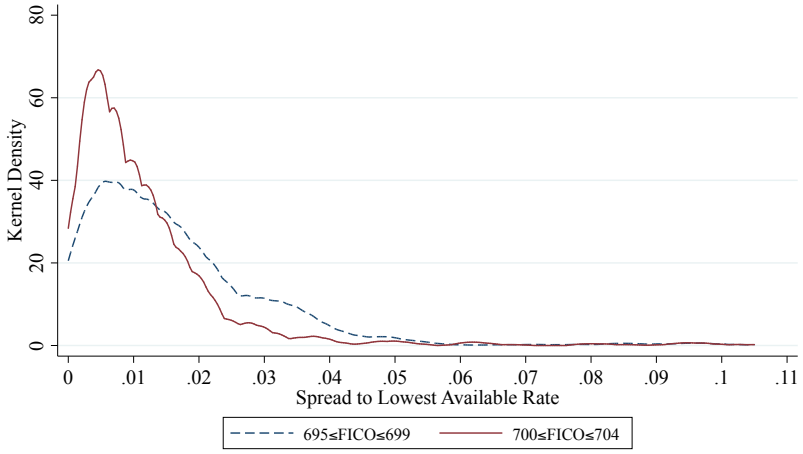
A. Borrowers Around a FICO = 600 Threshold



B. Borrowers Around a FICO = 640 Threshold

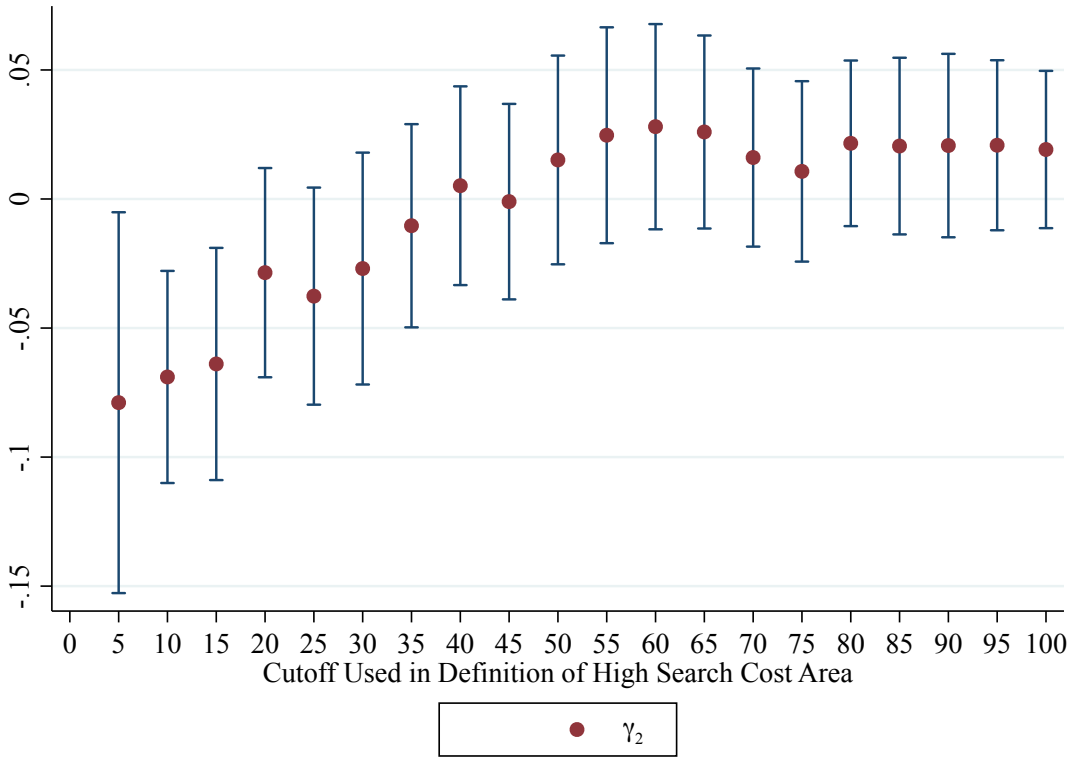


C. Borrowers Around a FICO = 700 Threshold



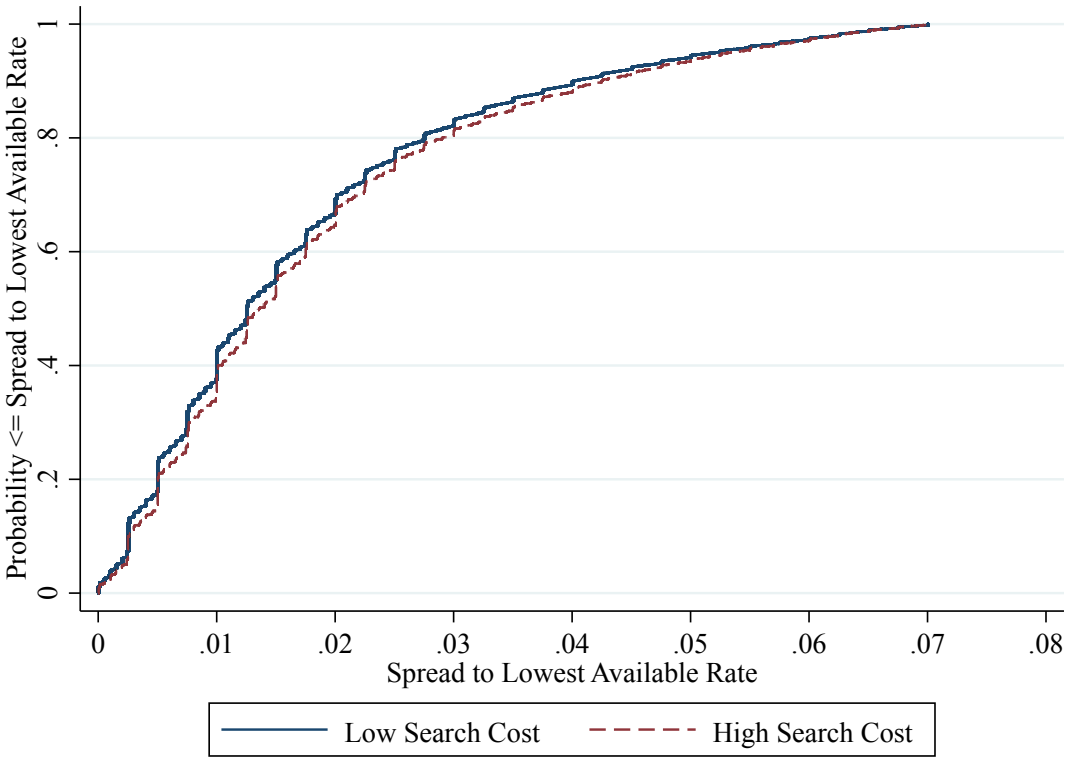
Notes: Figure reports the kernel densities of the spread (in percentage points) to the lowest available rate for borrowers with FICO scores just to the left of a threshold that borrowed from institutions with lending thresholds of 600, 640, and 700, respectively.

Figure 6: Takeup Rate Discontinuities Differences by Search Cost Definition



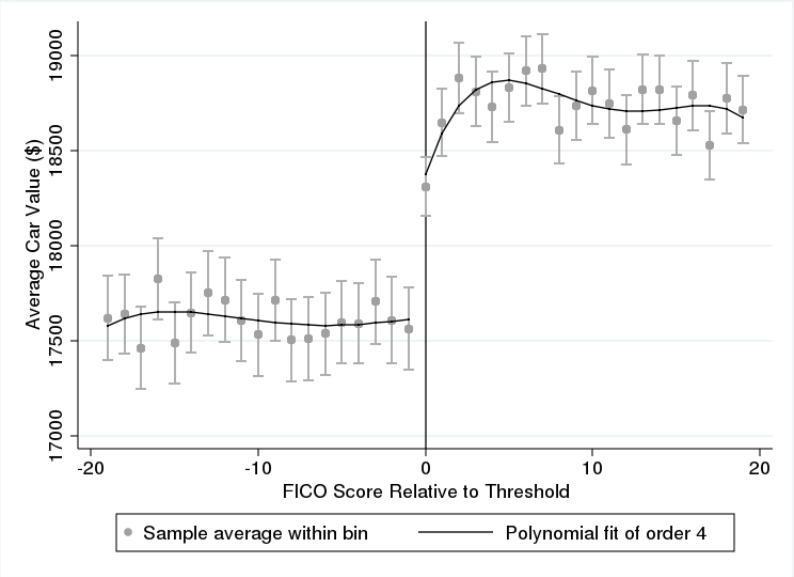
Notes: Figure reports RD estimates of γ_2 in equation (3) of take-up on an interaction of the discontinuity with a High Search Cost dummy defined as less than or equal to indicated cutoff.

Figure 7: CDFs of the Spread to Lowest Available Rate by Search Costs



Notes: Figure plots estimated cumulative density functions of the spread to the lowest available rate (the difference between a loan's interest rate and the best rate among similar borrowers in its cell) by our measure of search costs based on the number of nearby lenders. See notes to Figure 1.

Figure 8: Effect of FICO Threshold on Value of Car Purchased



Notes: Figure plots average car prices by normalized FICO score for loans originated by lenders with detected interest discontinuities at FICO thresholds of 600, 640, and 700.

Table 1: Summary Statistics

				Percentile		
	Count	Mean	Std. Dev.	25th	50th	75th
	A. Loan Applications					
Loan Term (months)	1,152,729	65.4	25.3	60	66	72
Loan Amount ($)	1,356,197	21,335	11,738	12,550	19,824	28,382
Loan Rate	1,356,197	0.0446	0.0374	0.0199	0.0375	0.0625
FICO	1,297,642	646.4	110.7	594	656	718
Debt-to-Income	1,266,666	.283	.205	.131	.272	.401
	B. Originated Loans					
Loan Rate	3,968,914	0.049	0.029	0.028	0.040	0.060
Loan Term (months)	3,968,914	64.52	20.45	60	61	72
Loan Amount ($)	3,968,914	19,441	11,196	11,211	17,393	25,366
FICO	3,526,409	712.25	73.4	664	715	769
Debt-to-Income (%)	1,862,758	0.264	0.357	0.089	0.300	0.370
Collateral Value ($)	3,968,914	20,697	11,071	12,781	18,650	26,405
Monthly Payment ($)	3,968,914	337.74	162.49	222.8	310.5	422.2
	C. Ex-Post Loan Performance Measures					
Days Delinquent	2,848,807	35.63	1054.2	0	0	0
Charged-off Indicator	3,968,914	0.021	0.142	0	0	0
Default Indicator	3,968,914	0.022	0.147	0	0	0
Current FICO	2,759,442	704.42	83.53	653	713	771
%ΔFICO	2,727,173	-0.010	0.089	-0.044	-0.002	0.032

Note: Panels A–C respectively report summary statistics for loan applications, originated loans, and ex-post loan performance. Loan Rate is the annual interest rate of the loan. Loan Term is the term (in months) of the loan. Debt-to-Income is the ratio of debt service payments to income. Collateral Value is the value of the car at origination. Days Delinquent is the number of days that a borrower has missed one or more monthly payments. Charged-off Indicator is a dummy for whether a loan has been written off the books of the lending institution. Default is an indicator for whether a borrower has been delinquent for at least 90 days. Current FICO is an updated FICO score for each borrower as of the date of our data extract. ΔFICO is the change in FICO score since origination as a fraction of the FICO score at origination.

Table 2: Summary Statistics for Estimation Sample with Identified FICO Discontinuities

				Percentile		
	Count	Mean	Std. Dev.	25th	50th	75th
	A. Loan Applications					
Loan Rate	38,236	0.052	0.054	0.038	0.042	0.060
Loan Term (months)	42,568	70.6	16.3	61	72	84
Loan Amount ($)	73,516	24,923	11,924	16,000	23,474	32,294
FICO	54,715	663	42.1	623	683	700
Debt-to-Income	51,257	0.270	0.195	0.130	0.270	0.390
	B. Originated Loans					
Loan Rate	378,521	0.068	0.0307	0.0445	0.0625	0.0854
Loan Amount ($)	378,521	17,777	105,79	10,026	15,812	23,286
Loan Term (months)	378,521	63.2	21.2	49	60	72
FICO	378,521	652	36.8	626	647	688
Debt-to-Income	212,472	0.267	0.315	0.106	0.272	0.38
Collateral Value	378,521	18,311	9,824	11,350	16,477	23,337
Monthly Payment	378,521	327	152	219	302	405
	C. Ex-Post Loan Performance Measures					
Current FICO	285,379	647	69	608	652	693
%ΔFICO	285,379	-0.00838	0.0929	-0.0512	0	0.0458
Days Delinquent	266,804	41.3	275	0	0	0
Charged-off Indicator	378,521	0.0342	0.182	0	0	0
Default Indicator	378,521	0.0295	0.169	0	0	0

Note: Table reports summary statistics for the discontinuity sample (restricted to a 19-point bandwidth around detected FICO discontinuities in lender pricing rules). Panels A, B, and C describe loan applications, loan originations, and ex-post loan performance, respectively. See notes to Table 1 for further details.

Table 3: First-Stage Effects of FICO Discontinuity on Loan Rate and Loan Term

	(1)	(2)
	Loan Rate	Loan Term
Discontinuity Coefficient	-0.0137***	1.47***
	[-18.15]	[6.75]
Commuting Zone FE	✓	✓
Quarter FE	✓	✓
Number of Observations	339,308	339,308

Notes: Table reports regression discontinuity estimates of equation (2), pooling the three discontinuities shown in Figure 3 by normalizing FICO scores around each threshold and using the estimator of Calonico et al. (2014). All specifications include lending institution fixed effects and quarter-of-origination fixed effects. Robust t-statistics reported in brackets are clustered by normalized FICO score.

Table 4: Loan Application Covariate Balance Regressions

	(1)	(2)	(3)
	Loan Amount	Debt-to-Income	Number of Loan Applications
Discontinuity Coefficient	-83.23	-1.94	62.7
	[-0.26]	[-0.43]	[0.677]
Commuting Zone FE	✓	✓	✓
Quarter FE	✓	✓	✓
Number of Observations	58,022	40,323	17,853

Notes: Table reports reduced-form RD results for the subset of institutions for which we have detailed loan application data. See notes to Table 3 for more details. Each observation in the data used for column 3 represents a normalized FICO score. Robust t-statistics reported in brackets are clustered by normalized FICO score.

Table 5: Spread to Lowest Available Rate Summary Statistics

FICO Range	# of Cells	Mean # Borrowers in Cell	Mean	Std. Dev.	Percentile 25th	50th	75th
$595 \leq FICO \leq 599$	124	2.46	0.034	0.025	0.01	0.03	0.05
$635 \leq FICO \leq 639$	1,006	2.99	0.025	0.021	0.01	0.02	0.04
$695 \leq FICO \leq 699$	671	3.92	0.016	0.014	0.006	0.01	0.02

Notes: Table reports summary statistics for the spread between a just-to-the-left-of-the-threshold borrower's interest rate and the best available interest rate for borrowers in the same cell. Cells are defined as borrowers with in the same commuting zone, 5-point FICO bin, $1,000 purchase-price bin, 5 percentage point DTI bin, maturity, who take out loans in the same six month window. Within each of the matched bins, we then calculate the average difference between the lowest interest rate in the cell and each individual loan in the cell. Summary statistics are reported for only those cells that contain at least 2 borrowers.

Table 6: Effect of Search-Cost Proxies on Loan Offer Take-up Decisions

Search Costs	Full	High	Low	Difference
	(1)	(2)	(3)	(4)
Discontinuity	.076	-.0007	.094	-0.095
Coefficient	[3.01]	[-0.02]	[4.87]	[-1.89]
Quarter FE	✓	✓	✓	
Commuting Zone FE	✓	✓	✓	
Number of Observations	23,828	4,480	19,348	

Notes: Table reports results for reduced-form RD regressions of loan take-up (conditional on being offered a loan) separately for the full sample (column 1) and for borrowers in areas with low and high search costs in columns 2 and 3, respectively, using the specification in equation (2) (see text for definition of high and low search costs). Search costs are estimated using the number of lending institutions within a 20-minute drive. Robust t-statistics reported in brackets are clustered by normalized FICO score. See notes to Table 3 for estimation details.

Table 7: Number of Loan Applications per Vehicle Purchase by Search Cost Group

	High Search Costs (1)	Low Search Costs (2)	Difference (1) - (2)
Mean	1.073	1.117	-.044
Standard Deviation	(.298)	(.399)	[-18.85]
Number of Observations	20,066	157,249	

Notes: Table reports average number of applications per vehicle purchase for applications with reported birthdates and nine-digit addresses. Standard deviations are reported in parentheses. Column 3 calculates the difference in means, along with the robust t-statistic in brackets for the statistical significance of the difference between columns 1 and 2.

Table 8: Reduced-Form Effects of FICO Discontinuity on Origination Outcomes

	(1)	(2)	(3)	(4)
	Purchase Price	Loan Amount	LTV	Monthly Payment
Discontinuity Coefficient	901.4***	1,423.5***	0.028***	9.64***
	[10.98]	[14.49]	[5.17]	[8.22]
Commuting Zone FE	✓	✓	✓	✓
Quarter FE	✓	✓	✓	✓
Number of Observations	339,308	339,308	339,308	339,308

Notes: Table reports reduced-form RD estimates of equation (2) using the estimator of Calonico et al. (2014). Columns 1, 2, and 4 are measured in dollars. Robust t-statistics reported in brackets are clustered by normalized FICO score. See notes to Table 3 for more details.

Table 9: Reduced-Form Effects Robustness to Vehicle Heterogeneity

	(1)	(2)	(3)
	Purchase Price	Purchase Price	Car Age (months)
Discontinuity Coefficient	805.4***	76.43	-4.15***
	[8.75]	[1.44]	[-9.13]
Commuting Zone FE	✓	✓	✓
Quarter FE	✓	✓	✓
Make-Model FE	✓		✓
Year-Make-Model FE		✓	
Number of Observations	308,946	308,938	308,946

Notes: Table reports reduced-form RD estimates of equation (2) on car purchase prices (columns 1–2) and car age in months (column 3). Columns 1 and 3 include make × model fixed effects and column 2 includes year × make × model fixed effects. Robust t-statistics reported in brackets are clustered by normalized FICO score. See notes to Table 3 for more details.

Table 10: Effect of FICO Discontinuities on Ex-Post Credit Outcomes

	(1)	(2)	(3)	(4)
	Days Delinquent	Charge-off	Default	%ΔFICO
Discontinuity Coefficient	-1.25	0.00112	-0.0001	-.0008
	[-0.42]	[-0.419]	[-0.056]	[-0.36]
Commuting Zone FE	✓	✓	✓	✓
Quarter FE	✓	✓	✓	✓
Number of Observations	241,038	339,308	339,308	254,179

Notes: Table reports RD estimates of equation (2) on ex-post loan and borrower outcomes. Days delinquent is the number of days a borrower is delinquent as of our data extract. Charge-off is an indicator for whether a loan has been written off the books of the lending institution. Default is an indicator for whether a borrower has been delinquent for at least 90 days. ΔFICO is the change in FICO score since origination as a fraction of the FICO score at origination for the sub-sample of institutions that report credit scores after loan origination. Robust t-statistics reported in brackets are clustered by normalized FICO score. See notes to Table 3 for more details.

Table 11: Pass-through of Rate Shocks to Origination Interest Rates by Search Costs

	(1)	(2)	(3)	(4)
5-year Treasury	0.61		0.15	
	[49.57]		[18.69]	
5-year Treasury × 1(High Search Cost Area)	-0.012	-0.026	-0.018	-0.026
	[-2.36]	[-5.48]	[-4.90]	[-6.77]
1(High Search Cost Area)	0.001	0.001	0.001	0.001
	[11.22]	[12.06]	[10.75]	[11.32]
Commuting Zone Fixed Effects	✓	✓	✓	✓
FICO Bin Fixed Effects	✓	✓	✓	✓
Month Fixed Effects		✓		✓
Loan-level Controls			✓	✓
Observations	3,416,575	3,416,575	3,416,575	3,416,575
R-squared	0.55	0.58	0.75	0.76

Notes: Table reports OLS estimates of equation (4) of origination interest rates on the level of five-year Treasury yields, an indicator for high search cost areas, and their interaction. Robust t-stats in brackets clustered by month. Loan-level controls include loan amount, maturity, and a high-cost loan indicator.

A Identification Appendix

In this appendix, we discuss some additional robustness checks that support the interpretation of our results in the main body of the paper.

A.1 First-stage Heterogeneity

Our take-up estimates represent unbiased estimates of borrowing elasticities at the extensive margin to the extent that the discontinuities satisfy the identifying assumptions required of valid RD estimation, with one important caveat. The proximity of financial institutions (PFI) could influence first-stage estimates. It would not be surprising to find that below-threshold borrowers search more in high PFI geographies if differences in offered rates across thresholds are correlated with PFI. Appendix Table A1 reports estimates of differences in offered interest rates around thresholds for high and low search cost geographies. High search cost areas (low PFI) have estimated average differences of 1.35 percentage points across thresholds, compared to differences of 1.39 percentage points in low search cost areas (high PFI). The differences in the first-stage estimates between high and low PFI areas are not statistically significant, suggesting that our measure of search costs does create bias through differences in first-stage effects.

A.2 Potential Omitted Variables Correlated with High Search Costs

A more immediate challenge to our inference is the interpretation of differences in borrowing elasticities across high and low search cost geographies. The number of PFIs is plausibly correlated with a host of unobserved variables, including the financial sophistication of borrowers, the advertising of credit (which of course influences search costs), or hyper-local (within-CZ) economic or demographic conditions, to name a few. Candidate omitted variables could be time varying (e.g. hyper-local economic conditions) or time invariant (e.g. financial sophistication).

A.2.1 Addressing Time-varying Omitted Variables

If a given location has few nearby lenders because lenders anticipated forthcoming local economic shocks, then our proxy for high search costs may be correlated with a growing inelasticity of local demand that could generate our results even if the number of PFIs had remained high. We address possible bias from such time-varying omitted variables in the following way. We create a Bartik measure of PFI, calculated as the PFI for each borrower in our sample using 1990 financial institution locations. Using NETS data to calculate the location of every financial institution in the U.S. as of 1990, we calculate PFI for our sample of borrowers relative to 1990 financial institution locations. We then grow the 1990 PFI measure using the national growth rate in financial institutions from 1990 through the year of each observation in our sample period, calculating the national growth rate for each location excluding the contribution of that location to the aggregate changes. Variation in the Bartik-PFIs is thus driven by local branching concentration as of 1990 and aggregate variation in national branching trends, neither of which is likely to be correlated with time-series variation in local economic conditions during our sample period of 2005–2015. We sort our sample based on high and low-Bartik PFI and estimate take-up elasticities. Results are reported in Appendix Table A2. These results indicate that borrowing elasticities in high-Bartik PFI geographies (low search costs) are higher than in low-Bartik PFI geographies (high search costs), with high search cost areas being relatively insensitive to interest rates and low search cost areas have strong and statistically significant reactions. These results suggest that time-varying economic shocks that occur during our sample period do not appear to be a compelling explanation for differences in loan take-up rates across high and low PFI geographies.

A.2.2 Addressing Time-invariant Omitted Variables

A second challenge to our inference is the issue of time-invariant omitted variables, such as financial sophistication. Ultimately, the strength of the Bartik instrumentation strategy

relies on the exogeneity of the 1990 branch network, which may itself be correlated with local characteristics. For example, borrowers with low financial sophistication could exhibit lower borrowing elasticities if they do not appreciate the costs associated with differences in loan interest rates (see, e.g., Stango and Zinman, 2009 and Bertrand and Morse, 2011). While our regression-discontinuity design uses random assignment to one side or the other of a FICO threshold and thus conceptually holds such unobservables fixed, when we compare the effect of discontinuities in high and low PFI areas, we need to concern ourselves with possible time-invariant correlates of the number of nearby lenders. In general, our specifications above attempted to address bias caused by time-invariant omitted variables with commuting-zone fixed effects. However, while commuting-zone fixed effects address time-invariant omitted variables at the metropolitan-area level, they do not address likely variation in sophistication correlated with high or low PFI within a commuting zone.

Having shown above that potentially endogenous changes to the branch network cannot explain our results, we use temporal variation in the branch network in a difference-in-differences setting that looks at changes in take-up rates within a very narrow location (9-digit zip code). In doing so, we set aside the discontinuity sample and instead focus on ascertaining the reaction of shopping behavior of locations g that become high search cost areas after not being high search cost areas previously. We first restrict our attention to only those areas g in our data that eventually transition from having ten or more PFIs to fewer than 10. This results in a small sample of 754 observations from locations that we observe before and after becoming high search cost areas (note that 9-digit zip codes can often be a single address). We estimate the effect γ of becoming a high-search cost areas in a panel setting as follows

$$takeup_{igt} = \eta_g + \delta_t + \gamma High\,Search\,Cost_{gt} + \beta FICO_{igt} + \varepsilon_{igt}, \tag{5}$$

where $HighSearchCost$ is a dummy for whether location g was a high search cost location in

quarter t. The virtue of this specification is that it absorbs fixed differences across extremely local areas (e.g., financial sophistication) and thereby identifies the effect of search costs solely off of the *timing* of changes in search costs, which we argued in Appendix Section A.2.1 seem unrelated to unobservables driving variation in demand elasticities across space. Column 1 of Appendix Table (A3) reports estimates of equation (5), showing that when borrowers face a reduction in nearby potential lenders, they are 5.3 percentage points more likely to accept a given loan offer.

To leverage locations which did not change search-cost status, we also estimate a version of equation (5) in first-differences. For each location g in our sample that we observe more than once, we calculate the change in take-up rates $\Delta Takeup_{gt}$ and regress them on commuting-zone and quarter-pair fixed effects (a fixed effect for the pair of quarters in which location g was observed to calculate $\Delta Takeup_{gt}$), the change in the location's average FICO score, and an indicator $\Delta High\,Search\,Cost_{gt}$ equal to one if location g was a high search cost area in quarter t but not in the last period location g's take-up rate was observed before period t.

$$\Delta Takeup_{gt} = \eta_{cz(g)} + \delta_{t,\Delta t} + \gamma \Delta High\,Search\,Cost_{gt} + \beta \Delta FICO_{gt} + \varepsilon_{gt}.$$

Again, the identifying assumption behind this specification is that the timing of decreases in the number of PFIs is unrelated to counterfactual trends in take-up rates, supported by the Bartik results above. Including commuting-zone fixed effects in the differences specification allows for each metropolitan area to change its propensity to shop for a loan differentially. The prediction of our search-cost explanation for high take-up rates in high search-cost areas for this setting is that cross-sectional *changes* in take-up rates, having differenced out time-invariant factors, will be positively related to changes in our search cost measure. Column 2 of Appendix Table A3 shows that the estimated $\hat{\gamma}$ is similar to the levels specification $\hat{\gamma}$, again positive and significant and consistent with the idea that when search costs go up, take-up rates increase as borrowers shop around less.

A.3 Digital Search

Many consumers now search for loans on the internet (including using such information aggregators as Bankrate.com), potentially limiting the relevance of lender density and driving distances as a proxy for 21st-century search costs. An anecdote from one of the larger lenders in our sample suggests that formal digital search (actually filing out an application on line) is less prominent than might be expected: only 8% of their total applications are digital. While credit-union clientele skew slightly older than the general population, another potential explanation for the ability of physical search measures to explain variation in search propensity is that, although borrowers can be easily pre-approved on the internet, the actual closing of loans (signing documents, transfer of title, etc.) still most frequently occurs at physical branch or dealer locations, even for direct loans. Ultimately, however, we view price dispersion results as lower bounds given the possibility of digital lending, perhaps making our results with respect to PFI and loan search as even more noteworthy given trends in online shopping for credit products.

A.4 Competition and Measuring Search Costs

Proxying search costs with driving-time density may not uniquely measure borrower search costs. Driving-time density, as constructed, might also be a correlate of other local factors such as the degree of price competition among lenders. Indeed, costly search is a fundamental source of imperfect competition. In an effort to differentiate between search costs and a pure competition story, we construct empirical measures of lending competition within CZs. We calculate the share of originated mortgage loans by each HMDA lenders within a given CZ and use the origination shares to construct a CZ-level Herfindahl index and divide loans into low and high competition areas based on our constructed Herfindahl index. We then reestimate the specification of Table 6 for all four combinations of high and low search cost areas crossed with high and low competition areas.

The results of this exercise, presented in Appendix Table A4, highlight that even within

a competition category, there are statistically significant differences by search costs in the difference of take-up rates across FICO thresholds. Even for areas with a highly competitive banking sector, borrowers in high search cost areas are much more likely to accept dominated loan terms. For low-competition CZs, in low search-cost areas, the difference in take-up rates around lending thresholds is 25 percentage points (treated borrowers are 25 pp more likely to walk away from an expensive loan than control-group borrowers). In comparison, borrowers in high search-cost areas in the same competition bin are only 11 percentage points more likely to walk away when offered an expensive loan. For the high competition bin, we find similar results with higher search costs (lower search costs) resulting in a take-up differential of 8 percent (13 percent). This result shows that regardless of the overall level of competition, borrowers in areas we expect to have high search costs are much less sensitive to interest rates in their extensive-margin loan take-up decisions.

Table A1: Effects of FICO Discontinuity on Origination Interest Rates by Search Costs

Search Costs Sample	High	Low	
	Loan Rate	Loan Rate	Difference
Discontinuity Coefficient	-.0135***	-.0139***	0.00084
	[-20.78]	[-16.926]	[0.382]
Commuting Zone FE	✓	✓	
Quarter FE	✓	✓	
Number of Observations	58,718	280,590	

Notes: Table reports reduced-form RD estimates of first-stage equation (2), splitting the sample into high and low search cost areas using the number of institutions within a 20 minute drive as a proxy. Robust t-statistics reported in brackets are clustered by normalized FICO score. See notes to Table 3 for more details.

Table A2: Effect of Bartik Search-Cost Proxies on Loan Offer Take-up Decisions

Bartik Search Costs Sample	High	Low	Difference
Discontinuity Coefficient	0.024	.0932	-.0689
	[0.65]	[4.91]	[-1.66]
Commuting Zone FE	✓	✓	
Quarter FE	✓	✓	
Number of Observations	4,327	18,973	

Notes: Table reports results for reduced-form RD regressions of loan take-up (conditional on being offered a loan) separately for borrowers in areas with low and high search costs in columns 2 and 3, respectively, using the Bartik instrument discussed in the text. Robust t-statistics reported in brackets are clustered by normalized FICO score. See notes to Table 3 for estimation details.

Table A3: Effects of Search Costs and Competition on Take-up Decisions

	(1)	(2)
High Search Cost Area	.053	
	[2.00]	
ΔHigh Search Cost Area		.066
		[3.00]
FICO	.00004	
	[2.30]	
ΔFICO		-.0002
		[-5.97]
Geographic Fixed Effects	Zip9	CZ
Time Fixed Effects	Quarter	Quarter Pair
Number of Observations	754	31,588
R-squared	0.54	0.04

Notes: Table reports difference-in-differences regression results relating local take-up rates to a given location's high search cost status. Column 1 estimates this relationship in levels at the individual × quarter level, restricting the sample to only those locations that transition to being high search cost. Column 2 estimates this relationship in changes, specifying the dependent variable as the take-up rate at the location × quarter level, and includes all locations for which we observe originated loans in multiple quarters. "Quarter Pair" fixed effects are time fixed effects for each pair of quarters over which the differences in column 2 are calculated. ΔFICO is the change in the average FICO score of a given location between observations for that location. Robust t-statistics in brackets are clustered by quarter (column 1) and quarter pair (column 2).

Table A4: Effects of Search Costs and Competition on Take-up Decisions

		Competition	
		LOW	HIGH
Search Costs	LOW	0.25	0.13
		[5.33]	[3.44]
	HIGH	0.11	0.08
		[2.29]	[1.65]

Notes: Table reports results for reduced-form RD regressions of loan take-up for borrowers in each combination of areas with low and high search costs and high and competition, (see text for definition of high and low search costs). Search costs are estimated using the number of lending institutions within a 20-minute drive. Competition is measured using CZ-level lender mortgage market shares in HMDA data. All regressions include lending institution fixed effects and quarter fixed effects. Robust t-statistics reported in brackets are clustered by normalized FICO score. See notes to Table 3 for estimation details.