

Borrowing High vs. Borrowing Higher:
Price Dispersion and Shopping Behavior
in the U.S. Credit Card Market

Victor Stango
Graduate School of Management
University of California, Davis

Jonathan Zinman
Department of Economics
Dartmouth College, IPA & NBER

July 2015

Abstract

We document substantial cross-individual dispersion in U.S. credit card borrowing costs, even after controlling for borrower risk and card characteristics. That remaining dispersion arises because cross-lender pricing heterogeneity generates dispersion in APR *offers* to borrowers, and borrowers vary in shopping intensity. Our empirics match administrative data to self-reported card shopping intensity, and use instruments suggested by Fair Lending law to account for the endogeneity between APRs and search. The results show that shoppers and non-shoppers pay APRs as different as those paid by borrowers in the best and worst credit score deciles. We discuss implications for policy and practice.

JEL codes: D14, G2

Keywords: credit cards, price dispersion, risk-based pricing, loan shopping, search costs

vstango@ucdavis.edu, jzinman@dartmouth.edu. Thanks to Anna Lunn and Gordon Vermeer for research assistance, and to John Driscoll, Charles Sprenger, David Silberman, Justin Sydnor, Jeremy Tobacman, and conference/seminar participants at the Consumer Financial Protection Bureau, American Economic Association Meetings, Berkeley (Haas), Boston University School of Management, Cornell (Dyson School), the Federal Reserve Bank of Philadelphia, the Federal Reserve Board/George Washington Financial Literacy Seminar Series, ITAM, the NBER (Behavioral Economics, Household Finance), the UC-Boulder Conference on Consumer Financial Decision Making, UC-Davis, and Wharton (Decision Processes) for comments. Special thanks to Bob Hunt, Dubravka Ritter and Steph Wilshusen of the Payments Cards Center of the Federal Reserve Bank of Philadelphia, and to Geng Li of the Federal Reserve Board, for sharing data and expertise. All errors and opinions are our own.

Funding: This work was supported by the Filene Research Institute; the Federal Bank of Chicago; the Rockefeller Center at Dartmouth College; and the Graduate School of Management at UC-Davis.

Borrowing High vs. Borrowing Higher:
Price Dispersion and Shopping Behavior
in the U.S. Credit Card Market

Abstract

We document substantial cross-individual dispersion in U.S. credit card borrowing costs, even after controlling for borrower risk and card characteristics. That remaining dispersion arises because cross-lender pricing heterogeneity generates dispersion in APR *offers* to borrowers, and borrowers vary in shopping intensity. Our empirics match administrative data to self-reported card shopping intensity, and use instruments suggested by Fair Lending law to account for the endogeneity between APRs and search. The results show that shoppers and non-shoppers pay APRs as different as those paid by borrowers in the best and worst credit score deciles. We discuss implications for policy and practice.

JEL codes: D14, G2

Keywords: credit cards, price dispersion, risk-based pricing, loan shopping, search costs

In the wake of the financial crisis, policymakers and practitioners are working to help consumers minimize their borrowing costs. On the policy side, rules and enforcement actions by the new Consumer Financial Protection Bureau have constrained a variety of lender practices related to borrowing costs, and also mandated new disclosures intended to help borrowers make better shopping and borrowing decisions. On the industry side, online personal financial management services and specialized search engines offer information and referrals. But which policies make the most sense, and which third-party products can best serve consumers? Answering such questions requires understanding both lender and consumer behavior.

We identify several key aspects of consumer and lender behavior related to borrowing cost minimization – or lack thereof – in the U.S. credit card market, the second-largest market for unsecured consumer debt in the world.¹ Rich transaction-level administrative data, credit bureau data, and survey data on a panel of 4,312 consumers from 2006-2008 grant us a uniquely comprehensive view of the choice sets that borrowers face from lenders, the details related to such choices (APRs, fees, and so on), borrower-level card and debt holdings, borrower-level risk in the cross-section and over time, shopping decisions and outcomes.

Our first contribution is factual: we show that credit card borrowers pay substantially different borrowing costs (APRs). The balance-weighted APR interquartile range across cardholders is 800 basis points, even after discarding introductory “teaser rates” and omitting “transactors” who never borrow. Adding fees to our measure of borrowing costs

¹ The U.S. student loan market has higher outstanding balances as of 2013.

increases dispersion slightly, because high-APR borrowers tend to incur more fees. Our evidence on equilibrium price dispersion in credit cards is novel in the context of prior work on credit cards, which sought to explain the *lack* of APR dispersion in the 1980s and 1990s.²

Standard explanations for why APRs vary across borrowers – card issuer (lender) pricing of borrower risk, and confounding variation in other card characteristics such as rewards and fees – together fit less than half of cross-borrower APR dispersion. Default risk as measured by credit scores, late payments, borrowing, credit limits, utilization and so on explains roughly 40 percent of cross-sectional variation. Card features such as rewards, fees, fixed/variable rate pricing, and so on, and demographics (age, income and education), add only slightly to the fit. The bottom line is that even similarly risky borrowers, holding cards with similar characteristics and debt levels, pay substantially different APRs.

Our second contribution is identifying channels through which supply-side and demand-side behavior can yield significant equilibrium price dispersion. A key supply-side detail – one previously undocumented in the modern credit card market – is that consumers face substantial price dispersion in APR offers they receive, due to cross-issuer heterogeneity in risk-based pricing models. Direct mail solicitations during our sample period show that APRs offered by different issuers to the *same individual during the same month* often differ by several hundred basis points. In our administrative transaction- and account-level panel data, we confirm that different issuers systematically

² See, e.g., Ausubel (1991), Stango (2000; 2002), Knittel and Stango (2003).

price the same observable risk factors – even factors as coarse as credit scores – differently enough to yield dispersion of several hundred basis points in the APRs faced by a given individual. These cross-issuer differences are not a function of variation in fees or non-price terms (both of which we observe).

Within-consumer offer dispersion can generate cross-borrower variation in APRs if borrowers differ in shopping intensity, a possibility we can examine empirically because a subset of the borrowers in our sample take online surveys in which they self-report whether they “keep an eye out for better credit card offers.” Merging self-reported search intensity with our administrative data provides a rare opportunity to directly estimate the relationship between prices (APRs) and search intensity, conditional on all other observable borrower and card characteristics.

The link between APRs and shopping is likely endogenous, so in estimating that relationship we use selected demographic “protected characteristics” as instruments for search intensity. Fair Lending law prohibits card issuers from discriminating based on such protected characteristics – notably, gender, marital status and race – in any way that produces “disparate treatment.” In other words, by law lenders must meet the exclusion restriction in our IV model. The other criterion for a useful instrument – that it be correlated with the endogenous regressor – is plausible in our setting as well, because research in marketing shows that characteristics such as gender and marital status are correlated with shopping behavior. This implies that any reduced-form correlation between protected characteristics and APRs should be due to consumer behavior

conditional on receiving card offers, and not to lender pricing models that incorporate protected characteristics into offer terms or approval criteria.

Putting the IV model to work, we find that the most active shoppers pay borrowing costs several hundred basis points lower than do non-shoppers – a difference comparable to that paid by individuals in the best vs. worst decile of credit score. This echoes earlier work on the importance of search/shopping costs in the credit card market, although those studies did not link shopping to the cross-section of borrowing costs or benchmark its importance against that of borrower risk.³

We further explore the implied magnitude of search costs by estimating the lowest APRs individuals face in the market conditional on observables, and then using that and debt loads to estimate how much individuals could save annually by paying the lowest rate in the market. If we aggregate up, the estimated partial equilibrium savings from imposing even moderately more intense shopping on all U.S. credit card borrowers would be \$36 billion per annum. That figure is three times larger than the (also partial equilibrium) savings estimated by Agarwal et al (forthcoming) in their study of the CARD Act’s effects on credit card fees and APRs. It also exceeds estimates of “money on the table” from misallocation of debt across cards with different interest rates,⁴ or from payment of penalty fees in financial services due to limited attention/salience.⁵

³ See, e.g., Agarwal et al (2013); Berlin and Mester (2004); Bucks et al (2009, A12), Calem and Mester (1995); Calem et al (2006); FINRA (2009); Kerr and Dunn (2008); Stango (2002); Woodward and Hall (2012).

⁴ Ponce et al (2014), Stango and Zinman (2015).

⁵ See, e.g., Agarwal et al. (2013); Agarwal et al. (2009); Alan et al. (2015); Stango and Zinman (2009; 2014).

We scrutinize the external validity of our data and conclude that our results are not an artifact of the sample. Our observed APR dispersion is in line with that observed in other, nationally representative data sources (see the Data Appendix). One might worry that the shopping-related dispersion we identify is due to an over-representation of cardholders who don't shop, or shop poorly. But this worry seems inconsistent with the fact that our sample is relatively higher-income and higher-education. And on many observables our sample shows more homogeneity than in the general population, even on dimensions – such as doing some financial management online – that are plausibly correlated with shopping behavior. Finally, several of our inferences rely on *issuer* pricing behavior, and the large issuers are well-represented in our data.

Our results inform policy and practice in the credit card market and other consumer financial markets characterized by substantial price dispersion.⁶ Although recent legislative and enforcement activity has focused on contingent charges/fees and whether they are “shrouded” in ways that harm consumers, our findings suggest that improving borrowers’ shopping, on a “base” price like a contract APR, could be more impactful in some markets.⁷ Our findings also highlight the potential for market “infomediaries” and

⁶ See Zinman (forthcoming) for a review of evidence on price dispersion in consumer loan markets, and Choi et al. (2010), Hortascu and Syverson (2004), and Sirri and Tufano (1998) on mutual funds.

⁷ On the heels of the CARD Act and rulemaking and enforcement actions focused on borrowing costs in the mortgage, student loan, and auto loan markets, the Consumer Financial Protection Bureau (CFPB) may now be turning its attention to contract APRs and shopping in credit cards. In September 2014 it warned issuers “against deceptively marketing promotional offers.” This harks back to the beginning of modern consumer protection efforts in debt markets, which typically focused on improving comparison shopping (National Commission on Consumer Finance 1972). The CFPB also provides consumers with “Know Before you Owe” data on credit

other delegates to provide (customized) search for consumers by shedding light on the potential money left on the table by incomplete borrower loan shopping.⁸ Similarly, our results inform the business case for policy efforts like “Smart Disclosure” that seek to facilitate infomediation by promoting data standards and access.⁹

1. Data

A. Data Contents and Sample Characteristics

Our data come from Lightspeed Research (formerly Forrester Research). Individuals in our sample are members of the “Ultimate Consumer Panel,” which is one of many such panels maintained by Forrester/Lightspeed.¹⁰

The credit card data collected by Lightspeed have four main components. The first component is transaction-level and comes from monthly credit card account statements. The set of transactions includes all credits (payments, refunds) and debits (purchases, fees, interest charges, etc.) on the account.

The second component is account-month level and contains data on account terms: APR, cash advance APR, bill date, due date, ending balance on bill date, summaries of credits and debits during the month, and so on. We restrict the analysis to general purpose

card terms on its website, although this data is not readily tailored to individual consumer preferences or credit characteristics and is updated only quarterly or semi-annually.

⁸ A noteworthy infomediator in the credit card market is NerdWallet, which started soon after our sample ends.

⁹ <https://www.data.gov/consumer/smart-disclosure-policy> .

¹⁰ We also use the data in Stango and Zinman (2009; 2014). Other Forrester/Lightspeed panels track consumer behaviors such as the use and purchases of new technology. Those panels are widely used by industry researchers and academics; see, e.g, Goolsbee (2000; 2001), Kolko (2010), Prince (2008).

credit cards. These are much more common than store cards and more amenable to the research questions because store cards are more clearly differentiated products and have all-in prices (net of discounts, etc.) that are more difficult to measure.

The third component is credit report data from one of the major bureaus, “pulled” at around the time of the cardholder’s registration. The credit report data include data on “trades” (current and past loans of all kinds), delinquency, loan balances, and a credit score on the standard 850-point scale.

Finally, Lightspeed solicits and collects survey data from panelists. All panelists complete a registration survey in which they report demographics and financial characteristics. Lightspeed also periodically invites panelists to take online surveys. The data we use later in the paper regarding credit card shopping come from one of those periodic surveys.

Table 1 summarizes the data. Our data span 2006-2008, and in this paper our main sample consists of the 4,312 cardholders who enroll at least one credit card account and for whom we observe credit bureau data. We stratify cardholders by their quartile of average “revolving” (i.e., interest-accruing) debt to facilitate analysis that conditions on debt levels, to understand how heavy and light borrowers differ, and because our research questions are most salient for heavier borrowers. Within-cardholder revolving debt levels are quite persistent, with a month-to-month serial correlation of 0.96.

Seventy percent of cardholders enroll one or two accounts, and the remaining thirty percent enroll three or more. Roughly half of our sample enrolls a “complete” set of

credit cards, meaning that the number of accounts enrolled matches the number of “active credit card lines” on the cardholder’s credit report. In a working paper version of this study, we show that complete-set cardholders look quite similar to the full sample in terms of descriptive statistics and all of the analyses we conduct below.¹¹ This alleviates selection/measurement concerns, and suggests cardholders with “incomplete” sets register the cards that they use regularly.

The 1st quartile of revolving debt contains many “transactors” who essentially never revolve balances but use their cards for purchases. Consumers in the 3rd and 4th quartiles are heavier “revolvers” who consistently carry balances. For these revolvers interest charges are 81% and 92% of total borrowing costs.¹²

Perhaps surprisingly, there are many similarities between individuals in the highest and lowest quartiles of revolving debt. Purchase volume, credit scores, and education are each U-shaped with respect to revolving balances. Income increases with revolving debt. We also see the expected life-cycle pattern, with those in the middle of the age distribution carrying more debt.

Perhaps the most noteworthy overall pattern in Table 1 is the substantial heterogeneity, both within and across revolving quartiles, in every variable. “Who borrows?” is not easily explained by observable individual characteristics.

¹¹ See Stango and Zinman (2013).

¹² The remainder of borrowing costs comes from annual, late, over-limit, cash advance, balance transfer and other fees. See Stango and Zinman (2009) for further detail on fees in these data. We control for fees incurred in the analysis below.

The Data Appendix provides many additional details on panel construction, variable definitions, and sample characteristics.

B. Representativeness and External Validity

Our credit data benchmarks reasonably well against various other data sources (the Data Appendix provides detailed comparisons). Our sample is similar to the U.S. population in terms of cardholding, purchases, creditworthiness, APR distribution, and interest costs relative to total borrowing costs. The one key difference is that our cardholders have outstanding balances that are about half the national average. Given that our analysis focuses on identifying borrowing cost dispersion *conditional on debt amounts*, any “missing debt” will lead us to understate the potential impact of borrowing cost dispersion on saving rates.

In terms of demographics, our cardholders are younger, more educated, and higher income (conditional on age) than national averages. The online nature of the panel might affect inferences about the broader population of cardholders, if “being online” is correlated with shopping or allocating debt efficiently. To the extent that our sample is more homogeneous than the population by dint of being younger, “online” and willing to participate in the panel, our results could easily *understate* the level of diversity in borrowing costs and shopping behavior in the population.¹³

¹³ Panel participants are also relatively willing to share financial information, raising questions about whether they might be unrepresentative in other, unobservable ways. But the same could be said about any data source – including household surveys – that relies on opt-in from subjects.

The time period under consideration here, 2006-2008, is also noteworthy. We do not observe a decline in borrowing cost dispersion in the early stages of the financial crisis. Nor do we know of any reason to expect that our results—which are mostly about dispersion—would differ in calmer times, but this is clearly something worth exploring in future research.

2. Cross-sectional Variation in Borrowing Costs

A. Measuring Credit Card Borrowing Costs, With and Without Float and Teasers

We measure borrowing costs for each cardholder as the average balance-weighted annual percentage rate (APR) over our sample period. Balances accrue interest charges if they are “revolving”: not fully repaid after the due date of the bill. We focus on APRs because they constitute >80% of borrowing costs for heavier borrowers (Table 1), but all of our results hold if we include fees in our measure of borrowing costs. In fact, borrowing cost dispersion increases slightly when we include fees, because those with higher APRs tend to incur more fees.

The first rows of Table 2 show APR dispersion over revolving and non-revolving (zero-APR) balances. Our primary focus is on revolving and “steady-state” APRs, so the next rows exclude the 627 cardholders (15% of the sample) who never revolve balances during our sample period, and the account-months we classify as paying “teaser” (introductory) rates. Dropping teaser rates from the data has little effect on dispersion.¹⁴

¹⁴ Teaser rates have a negligible effect on cross-individual dispersion for three reasons: 1) teaser rates typically last only six months or so, and represent only a small proportion of account-

APR dispersion is substantial within every borrowing quartile and similar across the top three, with interquartile ranges of 800-900 basis points (bp), and 10th/90th percentile ranges of 1500-1600bp.

Because cardholders can have multiple cards with different APRs and credit limits, it is possible that dispersion in borrowing costs could arise because some borrowers allocate debt to their cheapest cards and others do not. Other work shows that such misallocation is substantial among borrowers in Mexico (Ponce, Seira, and Zamarripa 2014). In our case, however, misallocation is much less common and does not drive borrowing cost dispersion. The last rows of Table 2 illustrate this by presenting dispersion in cardholders' "best weighted" APRs – ones that would apply if all of cardholders' revolving debt were always allocated to the lowest-rate cards in their wallets, up to the credit limits of each card. The distribution of these best APRs is nearly identical to the actual APR distribution in our data, showing that differential misallocation across cards within borrowers' wallets is not driving borrowing cost dispersion across panelists.

The last row shows that regressing cardholder-month-level APRs on just a set of cardholder fixed effects yields a fit of nearly eighty percent. Most APR variation is in the cross-section of cardholders rather than within-cardholder over time.

months; 2) people have multiple cards, and hence a mix of teaser and non-teaser rates, at any point in time; i.e., even though one can sort *account-months* into teaser vs. non-teaser, the extent to which this sorting aggregates to the individual level is muted; 3) we do not actually find a significant tradeoff between introductory and post-teaser APRs. Anecdotally, most teaser rate offers recoup their lower APRs via 2-5% balance transfer fees rather than higher post-teaser APRs.

B. Scaling the Magnitude of APR Dispersion

Credit card APR dispersion could matter a lot economically. Take a borrower at the medians for income, interest costs, and revolving debt in our top revolving debt quartile – keeping in mind that outstanding balances for such an individual in our data are equal to the outstanding balances of the *median* cardholder in the U.S. That individual's savings rate could rise by 1.2 percentage points if borrowing costs fell from the 75th percentile of APRs to the 25th, or by 1.8pp if borrowing costs fell from the 90th percentile to the 10th. Alternatively, the same individual could hold total interest costs constant with \$4,000 (\$10,000) in additional debt or consumption, by moving from the 75th to the 25th (90th to the 10th) percentile of APRs. Our APR dispersion seems representative, so these magnitudes should be relevant for U.S. cardholders more broadly. That said, they are slightly smaller than losses incurred by individual investors due to excessive trading in Barber et al. (2009); they are larger than the 75th percentile of losses from investment mistakes among asset holders in Campbell, Calvet and Sodini (2007); they are similar to losses from sub-optimal 401(k) account contributions in Choi et al (2011); and they are larger than (amortized) losses from insufficient mortgage shopping in Woodward and Hall (2012).

Of course, these savings could well reduce savings/investment elsewhere in the market or economy (by operating through bank profits), or induce general-equilibrium changes in APRs, credit availability or other features of the market. We present the figures merely to indicate that they are economically meaningful.

3. Borrowing Cost Dispersion: Explanations and Empirical Strategies

What might explain the substantial cross-sectional APR dispersion in Table 2?

Consider two borrowers, Gretchen and Mary. Assume each has two credit card accounts with a total credit limit of \$10,000, and each revolves an average of \$6,000 across those two cards. If we find that Gretchen pays 22% APR on average and Mary pays 14%, what might explain the difference?

Broadly, there are two classes of explanation. One class holds that Gretchen and Mary face different prices from the market because they are *differently risky*, or use cards that are *differentiated products*; if so, their choice sets cannot be compared apples-to-apples. The other broad class of explanation is that Gretchen and Mary in fact face similar prices from the market (or similar distributions of prices), but *make different choices* given the same choice set.

Disentangling those explanations, and assessing their relative importance, requires rich data. One must observe measures of customer default risk and card characteristics, and be able to infer how those things are related to the APRs that customers face in the market (and hold in their wallets). The analysis would be enriched with observations of consumer choices about shopping for and using cards. Our data satisfy these criteria.

4. Borrowing Cost Dispersion, Borrower Risk and Product Heterogeneity

A. Specifications: Models Explaining the Cross-Section of Borrowing Costs

The most natural explanations for cross-sectional variation in APRs are default risk and card-level product differentiation.

Our data include much of the information used by issuers when setting and adjusting APRs, as well as significant details about card characteristics. We observe credit score,

supplementary credit bureau data (e.g., the number of current and past “lines” of credit of varying kinds), purchase volume and revolving balances, in-sample late/missed payments, credit limits and utilization, demographics (age, income and education categories), fees (annual, balance transfer, cash back, others), measures of rewards and affinity links, and fixed/variable rate pricing. We also observe geographic data (state of cardholder’s residence), but those data have no explanatory power conditional on the other covariates, so we omit them from the models. The Data Appendix provides additional details on variable construction.

To assess how well these covariates explain APRs we estimate a series of cardholder- and account-month-level models with APRs as the dependent variable, using all of our available data regarding risk factors and product characteristics as flexibly parameterized covariates. In cardholder-level regressions we include cardholder-level aggregates, as well as characteristics of cardholder’s primary card by average revolving balances. The cardholder-level models include balance-weighted issuer fixed effects, accounting for average APR differences across issuers stemming from omitted card characteristics, omitted level differences in pricing customer risk, and other unobservables. The account-month models include issuer and month-year fixed effects. All models include indicators for cardholders’ first and last months in the data, accounting for any systematic time-varying factors that affect APRs, and for cardholders’ different sample entry/exit dates.

B. Results and Robustness

Table 3 reports the fit of the APR regressions. The broad takeaway is that observable risk and card/issuer characteristics explain 30-40% of cross-sectional variation in

borrowing costs. Credit scores alone explain 5-20% of cross-sectional variation in APRs. Including in-sample risk measures adds substantially to the explanatory power of the model, in most cases allowing the model to explain 25-40% of cross-sectional variation. This compares favorably to analogous work predicting credit card delinquency (Gross and Souleles 2002; Allen, DeLong, and Saunders 2004). Card characteristics and demographics add very little to the explanatory power of the model.¹⁵ Reading across columns, the models do a better job fitting APRs for heavier borrowers than for “transactors.” And the cardholder-level models generally have better fit than the account-month models.

In Stango and Zinman (2013) we compare model fit early in the sample vs. late, highlighting the tradeoffs between informativeness of our credit bureau data (observed at the beginning of the sample period) and informativeness of our in-sample risk metrics (which grow more comprehensively backward-looking by the end of the sample period). The overall fit improves by 0.05 by the end of the sample, primarily because our in-sample measures become richer.¹⁶ Credit scores are almost equally informative at the end

¹⁵ Although we do not measure the dollar value of rewards, we do have indicators for whether a card has pecuniary rewards (such as cash back or airline miles), and whether the card has an affinity link (such as an affiliation with a professional sports team). Pecuniary rewards are slightly (-45bp) negatively related to APRs in the account-month models, while affinity links are slightly positively (+45bp) linked to APRs. As mentioned above, we also observe the panelist’s state of residence. Including state fixed effects leaves the r-squared of the models unchanged, and the p-value of joint significance for the state effects is 0.65. We have experimented with finer specifications that interact state fixed effects with other covariates such as credit scores; those interactions are also insignificant.

¹⁶ Our account-month models show that credit card pricing is sensitive to increases in consumer risk as measured by variables such as utilization and late fees. Overall during our sample period the number of “large” (>500bp) changes within-account averages 0.80, and the number of total APR changes per account averages 3.02. Some of these figures reflect shifts from teaser pricing

of the sample compared to the beginning, which dovetails with the stylized fact that credit scores are very stable within-person, over time. The interested reader can see in Appendix Table 1 a complete set of regression coefficients from our best-fitting cardholder-level specification (the fifth r-squared row in the last column of Table 3).

Even our richest model leaves more than half of APR dispersion unexplained. Figure 1 illustrates this, showing both the raw (de-meanded) variation in borrowing costs and the residual variation. The inter-quartile range in residual variation is 500 basis points, and the 10th/90th range is 1000bp.

A natural concern is that some of the residual dispersion is driven by variables that are commonly priced by all issuers but not (fully) observed by us. We further explore this issue in Section 5, and for now note several robustness checks suggesting that additional variables like geography (see footnote 15), account age, intra-household card sharing, and relationship banking do not contribute meaningfully to cross-individual price dispersion.¹⁷

as well as variable-rate pricing pegged to other market interest rates. The CFPB estimates that 16-24% of accounts per annum were subject to “risk-based repricing” or “penalty repricing” in the 2007-2008 period. See http://files.consumerfinance.gov/f/201309_cfpb_card-act-report.pdf.

¹⁷ We observe account ages (years since opening) for a subset of panelists, and in that sub-sample do not find account age to be significant correlated with APR. This is unsurprising given that: 1) many panelists have a mix of older and newer accounts; 2) we do observe panelist age, which is correlated with account age at the panelist level; 3) issuers can reprice accounts over time (implying that we wouldn’t expect to see APRs that were initially low due to macro conditions “stick” over time); 4) consumers can close any sticky-high APR accounts over time, or move balances out of them (implying that the effects of any sticky-high accounts would be muted). Another possible issue might be cards shared across individuals, but restricting the sample to “single, never married” panelists leaves the results unchanged. Only 7% of respondents report registering a card belonging to someone else. It is also possible that “relationship banking” – benefits granted to cardholders because they also hold, e.g., a deposit account or mortgage with

Another natural concern is that our models' functional forms might not capture the true, commonly-priced relationship between commonly-observed risk and product characteristics and APRs. But we have estimated models with even more flexible functional forms – to the point of over-fitting – without any improvement in explanatory power.¹⁸

Yet another contributor to unexplained variation in APRs could be randomization by issuers (see, e.g., Day 2003; Shui and Ausubel 2004). We cannot empirically distinguish between randomized pricing and the omitted credit risk story, but intuition argues against randomized prices as a *primary* driver of borrowing cost dispersion, given the considerable resources that issuers expend in developing proprietary internal risk models.

Overall, our finding of substantial cross-sectional dispersion in borrowing costs seems robust to various ways of controlling for credit risk and product differentiation. Nevertheless, we grant that our models fitting the cross-section of borrowing costs might be imperfectly specified. We therefore pursue a complementary approach, one focusing directly on the possibility that different issuers price the same risk characteristics differently, leading similarly risky borrowers to face different prices in the market and hold cards at different APRs.

5. Offer Dispersion in the Market and APR Dispersion in the Wallet

the same bank – could affect APRs (Agarwal et al. 2009). But in survey responses only 3% of our panelists report paying a lower credit card APR due to relationship banking.

¹⁸ To take one specific example, we have estimated models that interact all of our RHS variables with credit score decile – allowing for the possibility that characteristics such as borrowing levels or age might have differential relationships to APRs in different parts of the distribution of creditworthiness. In that model the unadjusted r-squared rises to 0.43 (compared to 0.39) – a small increase given that the number of parameters is essentially ten times higher. More important, the adjusted r-squared falls from 0.35 to 0.34.

A. Offer APR Dispersion

Our first evidence of within-individual offer dispersion comes from a separate dataset on the terms of credit card mailers from Mintel Comprimedia.¹⁹ The Mintel data allow us to measure dispersion in offers received for a particular individual *in a specific month*; this is a lower bound on dispersion measured over a longer time period. Looking at within-month offer APR dispersion eliminates any confounding effect of time-varying credit risk at the individual level. We focus on January 2007 in particular: January because it is a peak month for mail solicitations by credit card issuers, and 2007 because it sits in the middle of our Lightspeed sample period. We condition on having received more than one credit card offer during January 2007, dropping roughly 25% of individuals and leaving us with 1,211 people who received a mean (median) of 4 (3) credit card offers.

To illustrate within-individual dispersion in offers, Table 4a shows the distribution of within-individual differences between the highest and lowest APR offers, calculated two ways.²⁰ The first APR is the contract or “goto” APR – the APR after any teaser period expires (column 3). The second APR is an estimated “net-of-teaser” APR, which is the 24-month weighted average of the teaser and goto APRs (column 6).²¹ The median within-individual and within-month high-low goto (net-of-teaser) rate spread is 434 (750)

¹⁹ We are extremely grateful to Mintel, and to Geng Li at the Federal Reserve Board of Governors, for sharing summary statistics from these data. A paper by Li and coauthors (Han, Keys, and Li 2013) contains more detail about these data.

²⁰ The distribution of APRs shown here lies below that in our data, because these are initial offers and do not reflect the upward shift in APRs that occurs in the group of cardholders who are repriced or incur a penalty rate after accepting the initial offer.

²¹ If, for example, the teaser APR is zero for six months and the goto rate is 2000 basis points, the net-of-teaser APR equals $(6/24)*(0)+(18/24)*2000=1500\text{bp}$.

basis points, and the seventieth percentiles are 725 (986) basis points. These measures of dispersion must, mechanically, be at least weakly larger over longer time periods – longer time periods that are still short enough such that within-person variation in creditworthiness is trivial for nearly all consumers. In short, it is common for an individual to receive credit card offers at very different APRs.

B. Implied Offer Dispersion in Our Data: Issuer Heterogeneity in Risk-Based Pricing

Our second analysis of within-individual dispersion in choice sets uses the Lightspeed data to estimate cross-issuer heterogeneity in risk-based pricing. Relatively little is known about such heterogeneity, and whether it leads to significantly different APR offers for a particular individual, in part because issuers invest considerable resources in their internal modeling and view their models as valuable trade secrets. In some sense, of course, the fact that issuers expend significant resources is *prima facie* evidence that different internal models yield different “optimal” APRs for a given individual; otherwise, why invest in the models? Nonetheless, we know of no academic work documenting or estimating the magnitude of this heterogeneity.

Figure 2 illustrates cross-issuer heterogeneity by plotting distributions of the credit score/APR relationship for each of five large (anonymized) issuers in our data, and also for a sixth “all other issuers” group.²² The plots illustrate three sorts of heterogeneity across issuers, all of which are substantial. The first is that, even within a credit score decile, different issuers can have *APR levels* that differ by several hundred basis points

²² Credit scores are defined based on the entire distribution of APRs, so “decile 1” for different issuers captures exactly the same range of scores.

(e.g., compare the horizontal lines denoting the median rate across issuers for the same credit score decile). Another type of cross-issuer heterogeneity is in the credit score-APR *gradient*: the decline in APRs from the worst to best decile. The third type of cross-issuer heterogeneity is in the extent of APR *variance* within credit score deciles. At the least, these types of differences indicate differential emphasis on credit scores vs. other information (such as late payments) in pricing risk.

More formally, in order to quantify the potential impact of these differences on cross-cardholder borrowing cost dispersion, we take the simplest or richest account-month model in Table 3 and allow for issuer-specific coefficients on risk factors for each of the largest six issuers in our sample (which collectively make up 85% of cards in our sample and 75% nationally).²³

We use the coefficients from these issuer-level pricing models to predict implied APRs for every cardholder, in each month, from each of the six issuers. The hypothetical is “what would the set of APRs from these six issuers be, given cardholder X’s characteristics and the month-year of the data?” We then calculate the gap between the highest and lowest of these implied APRs, for every cardholder in every month. Note that because our models include month-year effects and estimate within-month high-low differences, time series variation in issuers’ pricing does not contribute to our estimate of within-individual price dispersion. All of the largest issuers market cards to the entire

²³ The unit of observation is an account-month in these models. Across all months we have roughly 120,000 observations (for an average of 20,000 per issuer), and within January 2007 we have 3900 observations or roughly 650 per issuer. We estimate both a fully parameterized model for richness and a model including only credit score decile on the RHS to avoid over-fitting.

range of credit scores (though with different emphases), so we are not extrapolating outside the range of the data for any issuer.²⁴

A useful feature of this approach is that it is quite conservative. It treats all smaller issuers as pricing identically, and we actually exclude the “all other issuers” category from our dispersion calculations below. It treats each larger issuer as applying a single pricing model, when in fact many large issuers employ different models, even internally, for a variety of reasons, one being legacy effects from acquisitions of other issuers with different models. It ignores residual variation, using only the conditional mean of the data to generate dispersion.²⁵ Finally, it is possible that our specification is less flexible than that actually employed by a given issuer, which makes our fitted APRs less dispersed than the ones an issuer would actually set.

Table 4b shows the key results of this exercise: the implied high and low APRs, and the high-low spread. We show data from January 2007 to facilitate comparison with the Mintel data in Table 4a. Dispersion from our predictions (Table 4b) is even greater than that in the Mintel data (Table 4a), perhaps because heterogeneity in ex-post repricing compounds heterogeneity in ex-ante pricing. In any case, the central takeaway is that both prediction model specifications in Table 4b – “all covariates” and “credit score decile only” – imply substantial price dispersion based simply on differential treatment of

²⁴ Every large issuer has customers in all credit score deciles during our sample period, and in all but six of the 60 decile-issuer bins (6 issuers, 10 deciles of credit score) the share of account-months in that bin is greater than 4 percent. Our results do not change if one excludes predictions into issuer-deciles where that issuer has less than four percent of its cardholder-months.

²⁵ For our large issuers, the standard deviation of APRs predicted by our model ranges from 50-75% of the standard deviation of actual APRs in the data for those issuers.

identical customer characteristics by the largest six credit card issuers. Even the 10th percentile of the high-low difference is an estimated 500 or 600 basis points. The 90th percentile is estimated at about 1300 basis points in both specifications.

In all, the evidence in Tables 4a-4b strongly suggests that any given individual receives offers at very different APRs from different issuers. Borrowers facing offer dispersion in the market might pay different APRs based on heterogeneity in shopping behavior.

6. Borrower Shopping Behavior and Borrowing Costs

A. Shopping/Search Behavior and Borrowing Costs: Descriptive Data

We now examine the link between card shopping and borrowing costs. For a subset of panelists (n=603), we observe agreement (on a 10-point scale) with the statement “I always keep an eye out for better credit card offers.” Cardholders supplied responses via one of the periodic and voluntary surveys emailed to panelists by Lightspeed; the survey was administered in the first quarter of 2007. Of those cardholders, 476 are borrowers for whom we also observe demographics such as marital status, gender and race – the instruments in our empirical models below.

Table 5 summarizes shopping responses grouped into four categories. Reading across columns 34% of cardholders report 1-3 on the 10-point scale (“non-shoppers”), 30% report 4-6 (“medium shoppers”), 26% report 7-9 (“high shoppers”), and 10% report 10, the strongest agreement (“super-shoppers”).

The top rows of Table 5 show that self-reported shopping intensity correlates sensibly with other variables that might reflect shopping: current cards held, previous

(now closed) accounts, and recent credit card applications. For example, only 15% of non-shoppers hold 5 or more cards, while 33% of super-shoppers hold 5+ cards; 16% of non-shoppers have 15+ past cards, while 40% of super-shoppers have 15+ past cards; and 5% of non-shoppers report having applied for 2+ cards recently, while 26% of super-shoppers report the same thing.

The next sets of rows provide descriptive evidence previewing our instrumental variables results below: shoppers pay lower APRs. In this table and the empirics that follow, we focus on the “best” APR a cardholder could pay if debt were allocated perfectly to the lowest-rate cards in the wallet conditional on credit limits. This makes the focus purely about which cards are *held* – i.e. based on shopping – rather than about both which cards are held, and how they are used in the wallet.

The last columns compare survey respondents to non-respondents. Non-respondents have fewer current/past cards and recent applications, are less creditworthy, borrow slightly less and pay higher APRs. These differences caution against extrapolating our results below from respondents to non-respondents.

B. Shopping Behavior and Borrowing Costs: Regressions

Can individual-level differences in shopping behavior explain meaningful differences in borrowing costs? We examine this question by adding the 10-point shopping intensity variable, in linear form, to our main cardholder-level specification from Table 3.²⁶ The

²⁶ We have experimented with other functional forms (e.g., fewer ordinal categories, dummies for “shopper” vs. “non-shopper” at different thresholds) with similar results.

key identification issue is that shopping may be endogenous; e.g., a high APR “shock” might increase shopping effort and thereby attenuate naïve estimates of the relationship between reported shopping intensity and borrowing costs.

To deal with endogeneity we instrument for the shopping variable with two cardholder characteristics: marital status and gender. We choose marital status and gender as instruments because they satisfy the exclusion restriction by law: the Equal Credit Opportunity Act (ECOA) prohibits lenders from discriminating based on marital status or gender, regardless of intent. Fair Lending examiners monitor compliance by testing lenders for “disparate treatment” and “disparate impact”: conditional within-issuer correlations between protected characteristics and credit outcomes. Lenders have strong incentives to pass these tests and thereby ensure they satisfy our exclusion restrictions.²⁷ Exogeneity also requires that APR shocks do not change the IVs themselves. This almost certainly holds for gender and marital status.

The bottom rows of Table 5 shed light on the first stage for both instruments in the raw data: single cardholders and male cardholders search more. There are many possible explanations for these correlations; we simply note that gender and marital differences in shopping have long been observed in marketing research.²⁸ ECOA suggests two other protected characteristics as instruments: age and race. We use age as an included

²⁷ See http://www.federalreserve.gov/boarddocs/supmanual/cch/fair_lend_over.pdf for Fair Lending guidance pertinent to our sample period (and today). Skanderson and Ritter (2014) review ECOA policies and practices as they apply to the credit card market in particular and discuss how credit card issuers work to comply with Fair Lending law.

²⁸ See, e.g., Laroche et al (2000) and references therein.

regressor rather than an instrument because in practice credit card issuers are allowed to use age in their risk models.²⁹ We discard race as an instrument because it is uncorrelated with borrowing costs in the reduced form, conditional on our other RHS variables.

One might wonder, given ECOA monitoring, how any of our instruments could satisfy that second criterion for usefulness: a significant reduced form correlation between the IVs and the LHS variable, conditional on all other RHS variables. But recall that our LHS measure of borrowing costs is (a) *ex post* of any shopping decisions, and (b) weighted across all balances from different issuers. Thus, while individual issuers will not set APRs using information on gender and marital status (and indeed, issuers will strictly speaking lack data on those variables), a reduced form correlation can arise if borrower choices about which cards to select and use affect borrowing costs.³⁰ That link is precisely the one we exploit with the IV strategy.³¹ We also confirm that within-issuer reduced form correlations between our instruments and account-month-level APRs are essentially zero (see Appendix Table 2), implying that the reduced-form correlations between IVs and APRs in Table 6 reflect borrower-level behavior (i.e. about shopping and usage), rather than differential treatment in the market.³²

²⁹ Issuers are permitted to use age as an approval criterion in “soft” decisions that involve human judgment so long as such criteria do not treat the elderly unfairly. Lenders are also permitted to use age in their automated credit scoring models. See Skanderson and Ritter (2014) for a discussion.

³⁰ We are grateful to an anonymous referee for clarifying our thinking on this point.

³¹ Indeed, issuers will not typically even have readily available data on gender and marital status because they do not ask for this information on applications; nor is it contained in credit reports.

³² Regulators monitor within-issuer compliance only, and do not conduct cross-lender comparisons.

Table 6 presents our estimates of the link between shopping and borrowing costs. The first column reports OLS results, while the second column shows IV results. In the IV specification we report the standard IV point estimate and standard error, the p-value of the CLR/AR test for whether the test of the null is robust to weak instruments, and the associated corrected-confidence interval for the coefficient on the endogenous regressor.³³ We also show coefficients on the instruments in both the first stage and the reduced form (in both cases, conditional on all other RHS variables). Both gender and marital status are significant in both the first stage and reduced form, and the coefficients comport with the pattern in the raw data of Table 5.

The OLS results reveal no strong relationship between shopping and APRs, but the IV specifications suggest a large, negative effect of search intensity on APRs. The IV point estimates suggest a roughly 100bp reduction in borrowing costs per “point” of shopping intensity. The IQR of shopping intensity in the sub-sample is 5 points – from 2 to 7 – implying a reduction in borrowing costs of 500bp by moving from the 25th to 75th percentiles of shopping intensity if we focus on the point estimates (the confidence intervals include both much larger, and very small, effects). The 500bp estimate is fairly consistent with our estimates of offer dispersion (the low-high spreads in Table 4). The implied gains from shopping are comparable to APR variation generated by cross-sectional variation in observable and commonly priced default risk; e.g., moving from the

³³ See Finlay and Magnusson (2009) for a discussion of the weak instrument problem and the Stata routine we use to deal with the issue.

25th to 75th percentile of credit score is also correlated with a 500-600bp reduction in borrowing costs.

The contrast between our OLS and IV results suggests strong endogeneity in borrowers' shopping behavior: higher APRs induce more active shopping. It is possible, at least in partial equilibrium, that APR dispersion would be even greater absent this endogeneity. Regardless, our results suggest a strong role for search cost heterogeneity in explaining APR dispersion. The strong correlations between our instruments and search are also noteworthy, in that (assuming that issuers do not violate ECOA) we show a strong relationship between search and prices in a setting where unobserved heterogeneity is arguably unable to drive the results.

D. Bounding the Distribution of Search Costs

The magnitude of the effect of shopping effort on borrowing costs prompts two related questions. First, what do our results imply about the distribution of search costs? Second, why don't more credit card users shop intensively?

On the first question, we can infer the upper bound of search costs using our data. In unreported results, we have estimated the "lowest rate available" in the market for each cardholder-month, using the two models described in Table 4b, which model cross-issuer variation in pricing based on observables.³⁴ We then compare the actual APR for each

³⁴ These models regress account-month APRs on panelist-, panelist-month- and month-level variables, letting the coefficients vary by issuer for the largest issuers in the market. Coefficients from these models yield a set of panelist-month-level implied offers from these large issuers. We estimate that the average APR savings in our sample from always paying the lowest APR available is roughly 400bp.

cardholder-month to the lowest rate, and scale by revolving balances to quantify dollars left on the table via imperfect shopping. Annualized dollars on the table are \$0-30 at the 25th percentile (the range reflects the two different methods of inferring the lowest rate available from Table 4b), \$60-150 at the 50th percentile, \$250-600 at the 75th percentile and \$900-1500 at the 90th percentile. This distribution is similar to those found in mortgage and mutual fund markets, for example.

Why do some people behave as though they face large search costs? One view is that some individuals simply find shopping very costly, for rational and/or behavioral reasons, in pecuniary and/or non-pecuniary terms. Another, complementary view is that other frictions offset the benefits of search. For example, switching costs are well-documented in credit cards (Stango 2002, Calem and Mester 2006).

Another complicating factor is that APRs are not necessarily exogenous to search costs. Issuers observe signals about shopping behavior such as applications, credit/trade lines (present and past), balance transfers, and cardholder attempts to renegotiate APRs. Issuers may then tailor offers based on search/switch costs. Coupled with the possibility that the Internet has changed both the mean and variance of search costs in the cross-section of borrowers, this raises the question of whether search costs could help explain long-term trends in credit card pricing as well as cross-sectional variation during our time period.

Regardless of the underlying model, the aggregate partial equilibrium savings consumers would receive from more comprehensive shopping would be substantial, even conservatively. The Federal Reserve estimates that during our sample period, credit card

debt stood at roughly \$900 billion dollars.³⁵ This implies reduced borrowing costs of \$9 billion annually per 100 basis points of savings; moving from the 75th to the 25th percentile of shopping intensity in our data implies, given our empirics above, a more than 400 basis point reduction in borrowing costs and \$36 billion in annual savings. That figure is three times larger than the estimated consumer savings from the CARD Act's restrictions on penalty fees and repricing Agarwal et al. (forthcoming). Of course, our figures should be interpreted cautiously given that our sample may not be representative, but the numbers would remain large even if (for example) one halved them.

7. Conclusion

We document cross-consumer dispersion in credit card borrowing costs that remains substantial even after controlling for debt levels, credit risk, and product characteristics. Our results suggest that dispersion is generated by the intersection of heterogeneity in issuer pricing and heterogeneity in consumer contract choice: different issuers offer different APRs to the same individual, and differences in consumer shopping behavior lead otherwise identical consumers to choose contracts at widely differing rates.

Our estimates of borrowing cost levels and dispersion probably err on the conservative side. Our sample borrowers carry less credit card debt than the broader population. Moreover, we observe only credit card borrowing costs, and not costs in other, even larger debt markets: mortgages, auto loans, and student loans.

³⁵ See http://www.federalreserve.gov/releases/g19/HIST/cc_hist_r_levels.html.

Our results may inform interventions designed to help improve credit market outcomes. If credit shopping is more malleable than creditworthiness (credit scores are quite sticky), then helping people shop for cards may be a relatively effective focus for interventions. This is not to say that our results support any particular policy, programmatic, or business tack: they are silent, for example, on how or how cost-effectively one could affect search behavior, and on what the general equilibrium effects of any such innovation would be. The difference between our OLS and IV estimates also suggests that shopping behavior is highly endogenous; in contrast to inert consumer behavior in other settings, our results suggest that shocks in the form of higher APRs do induce behavior change in the form of more active shopping. Perhaps such APR shocks are highly salient to some borrowers.

Identifying what drives search behavior – heterogeneity in time costs, preferences for leisure, in one or more behavioral factors, in skills/endowments, etc. – would inform modeling, policy, and practice. Our instrumental variables approach may translate to other settings, making links between customer characteristics and shopping patterns easier to identify, and thereby informing policy.

Our results also highlight that while research on wealth accumulation often focuses on the asset side, the liability side can matter too: borrowing cost dispersion is substantial enough that even conditional on debt levels it could explain cross-sectional dispersion in savings rates and wealth accumulation. Perhaps a price channel helps explain debt loads that seem puzzlingly high, and savings rates that seem puzzlingly low, even by the

standards of behavioral models.³⁶ We emphasize that this is merely a suggestion: our results do not identify whether people are “over-paying” for credit card debt in a behavioral sense or simply making optimal tradeoffs between shopping costs and benefits. Nor do we actually estimate the relationship between borrowing costs and savings rates. But we hope that our paper will provoke inquiry along these lines.

³⁶ E.g., even with quasi-hyperbolic discounting, the calibrated life-cycle model in Angeletos et al (2001) substantially underpredicts credit card borrowing. Zinman (2014) reviews other work on links between borrowing behavior and behavioral biases in preferences, expectations, and/or price perceptions.

REFERENCES

- Agarwal, Sumit, Souphala Chomsisengphet, Chunlin Liu, and Nicholas Souleles. 2009. "Benefits of Relationship Banking: Evidence from Consumer Credit Markets." Working paper.
- Agarwal, Sumit, Souphala Chomsisengphet, Chunlin Liu, and Nicholas Souleles. 2013. "Do Consumers Choose the Right Credit Contracts?"
- Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel. forthcoming. "Regulating Consumer Financial Products: Evidence from Credit Cards." *Quarterly Journal of Economics*
- Agarwal, Sumit, John Driscoll, Xavier Gabaix, and David Laibson. 2013. "Learning in the Credit Card Market."
- Agarwal, Sumit, Paige Skiba, and Jeremy Tobacman. 2009. "Payday Loans and Credit Cards: New Liquidity and Credit Scoring Puzzles?" *American Economic Review Papers and Proceedings* 99 (2): 412–17.
- Alan, Sule, Mehmet Cemalcilar, Karlan, Dean, and Jonathan Zinman. 2015. "Unshrouding Effects on Demand for a Costly Add-on: Evidence from Bank Overdrafts in Turkey."
- Allen, Linda, Gayle DeLong, and Anthony Saunders. 2004. "Issues of Credit Risk Modeling in Retail Markets." *Journal of Banking and Finance* 28: 727–52.
- Angeletos, George-Marios, David Laibson, Jeremy Tobacman, Andrea Repetto, and Stephen Weinberg. 2001. "The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation." *Journal of Economic Perspectives* 15 (3): 47–68.
- Ausubel, Lawrence M. 1991. "The Failure of Competition in the Credit Card Market." *American Economic Review* 81 (1): 50–81.
- Barber, Brad M., Yi-Tsung Lee, Yu-Jane Liu, and Terrance Odean. 2009. "Just How Much Do Individual Investors Lose by Trading?" *Review of Financial Studies* 22 (2): 609–32.
- Berlin, Mitchell, and Loretta Mester. 2004. "Credit Card Rates and Consumer Search." *Review of Financial Economics* 13 (1-2): 179–98.
- Bucks, Brian, Arthur Kennickell, Traci Mach, and Kevin Moore. 2009. "Changes in U.S. Family Finances from 2004 to 2007: Evidence from the Survey of Consumer Finances." *Federal Reserve Bulletin* February: A1–56.
- Calem, Paul, Michael Gordy, and Loretta Mester. 2006. "Switching Costs and Adverse Selection in the Market for Credit Cards: New Evidence." *Journal of Banking and Finance* 30 (6): 1653–85.
- Calem, Paul S, and Loretta J Mester. 1995. "Consumer Behavior and the Stickiness of Credit-Card Interest Rates." *The American Economic Review* 85 (5): 1327–36.
- Calvet, L., J.Y. Campbell, and P. Sodini. 2007. "Down or out: Assessing the Welfare Costs of Household Investment Mistakes." *Journal of Political Economy* 115 (5): 707–47.

- Choi, James, David Laibson, and Brigitte Madrian. 2011. "\$100 Bills on the Sidewalk: Suboptimal Saving in 401(k) Plans." *Review of Economics and Statistics* 93 (3): 748–63.
- Choi, James, David Laibson, and Brigitte Madrian. 2010. "Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds." *Review of Financial Studies* 23 (4): 1405–32.
- Day, G.S. 2003. "Creating a Superior Customer-Relating Capability." *MIT Sloan Management Review* 44 (3): 77–82.
- Finlay, Keith, and Leandro Magnusson. 2009. "Implementing Weak-Instrument Robust Tests for a General Class of Instrumental-Variables Models." *The Stata Journal* 9 (3): 1–24.
- FINRA Investor Education Foundation. 2009. *Financial Capability in the United States: National Survey – Executive Summary*.
- Goolsbee, Austan. 2000. "In a World without Borders: The Impact of Taxes on Internet Commerce." *The Quarterly Journal of Economics* 115 (2): 561–576.
- . 2001. "Competition in the Computer Industry: Online versus Retail." *Journal of Industrial Economics* 49 (4 (Symposium on E-Commerce)): 487–99.
- Gross, David, and Nicholas Souleles. 2002. "An Empirical Analysis of Personal Bankruptcy and Delinquency." *The Review of Financial Studies* 15 (1): 319–47.
- Han, Song, Benjamin Keys, and Geng Li. 2013. "Unsecured Credit Supply over the Credit Cycle: Evidence from Credit Card Mailings."
- Hortascu, Ali, and Chad Syverson. 2004. "Product Differentiation, Search Costs, and the Welfare Effects of Entry: A Case Study of S&P 500 Index Funds." *Quarterly Journal of Economics* 119 (2): 403–56.
- Kerr, Sougata, and Lucia Dunn. 2008. "Consumer Search Behavior in the Changing Credit Card Market." *Journal of Business and Economic Statistics* 26 (3): 345–53.
- Knittel, C., and V. Stango. 2003. "Price Ceilings, Focal Points, and Tacit Collusion: Evidence from Credit Cards." *American Economic Review* 93 (5): 1703–29.
- Kolko, Jed. 2010. "A New Measure of US Residential Broadband Availability." *Telecommunications Policy* 34 (3): 132–43.
- Laroche, Michel, Gad Saad, Mark Cleveland, and Elizabeth Browne. 2000. "Gender Differences in Information Search Strategies for a Christmas Gift." *Journal of Consumer Marketing* 17 (6): 500–522.
- National Commission on Consumer Finance. 1972. *Consumer Credit in the United States*. Washington, DC: U.S. Government Printing Office.
- Ponce, Alejandro, Enrique Seira, and Guillermo Zamarripa. 2014. "Borrowing on the Wrong Credit Card: Evidence from Mexico."
- Prince, Jeffrey. 2008. "Repeat Purchase amid Rapid Quality Improvement: Structural Estimation of Demand for Personal Computers." *Journal of Economics and Management Strategy* 17 (1): 1–33.
- Shui, Haiyan, and Lawrence Ausubel. 2004. "Time Inconsistency in the Credit Card Market."
- Sirri, Erik, and Peter Tufano. 1998. "Costly Search and Mutual Fund Flows." *Journal of Finance* 53 (5): 1589–1622.

- Skanderson, David, and Dubravka Ritter. 2014. "Fair Lending Analysis of Credit Cards." *FRB of Philadelphia Payments Cards Center Discussion Paper No. 14-2*.
- Stango, Victor. 2000. "Competition and Pricing in the Credit Card Market." *Review of Economics and Statistics* 82 (3): 499–508.
- . 2002. "Pricing with Consumer Switching Costs: Evidence from the Credit Card Market." *Journal of Industrial Economics* 50 (4): 475–92.
- Stango, Victor, and Jonathan Zinman. 2009. "What Do Consumers Really Pay on Their Checking and Credit Card Accounts? Explicit, Implicit, and Avoidable Costs." *American Economic Review Papers and Proceedings* 99 (2): 424–29.
- . 2013. "Borrowing High vs. Borrowing Higher: Sources and Consequences of Dispersion in Individual Borrowing Costs." *NBER Working Paper 19069*.
- . 2014. "Limited and Varying Consumer Attention: Evidence from Shocks to the Salience of Bank Overdraft Fees." *Review of Financial Studies* 27 (4): 990–1030.
- . 2015. "Pecuniary Costs of Bounded Rationality: Evidence from Credit Card Debt Allocation."
- Woodward, Susan E., and Robert E Hall. 2012. "Diagnosing Consumer Confusion and Sub-Optimal Shopping Effort: Theory and Mortgage-Market Evidence." *American Economic Review* 102 (7): 3249–76.
- Zinman, Jonathan. forthcoming. "Household Debt." *Annual Review of Economics*
- . 2014. "Consumer Credit: Too Much or Too Little (or Just Right)?" *Journal of Legal Studies* 43 (S2 Special Issue on Benefit-Cost Analysis of Financial Regulation): S209–37.

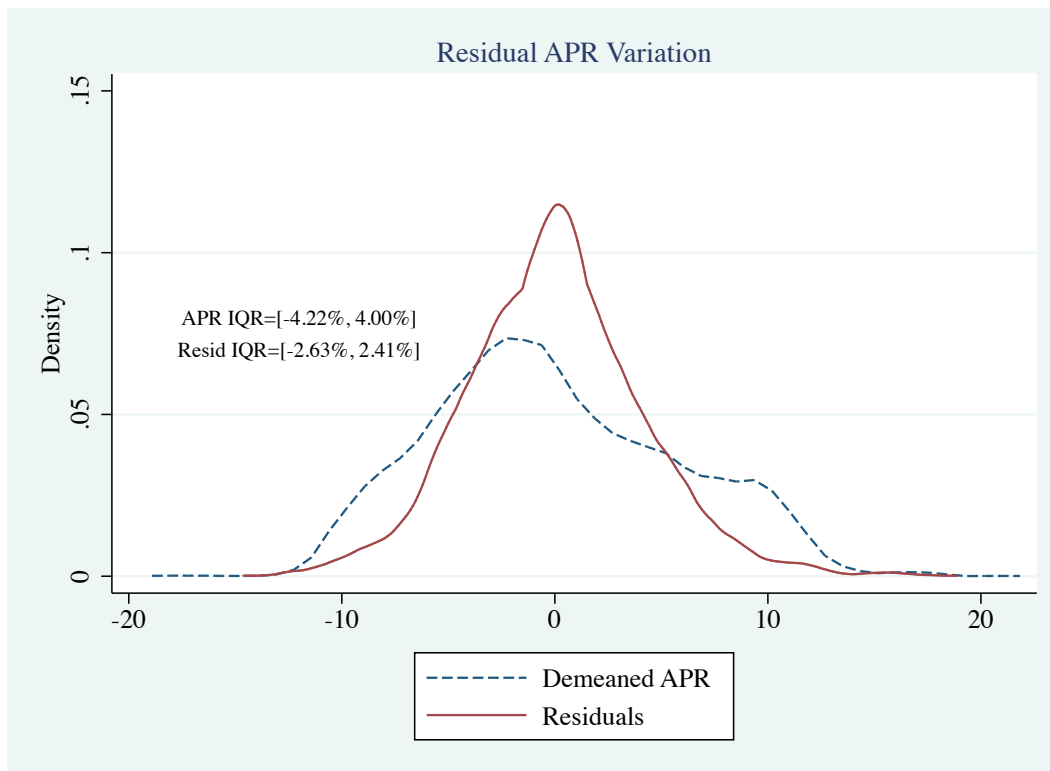


Figure 1. Raw and residual variation in weighted APRs.

Notes: “Demeaned APR” shows the distribution (kernel density) of panelist-level average weighted APRs on all revolving balances during the sample period, demeaned so that they are centered on zero. “Residuals” shows the distribution (kernel density) of residuals from the fullest specification of the panelist-level “above plus demographics” regressions described in Table 3. Fitted values and residuals are calculated using the quartile-specific coefficients in the first four columns of Table 3 (fifth row down).

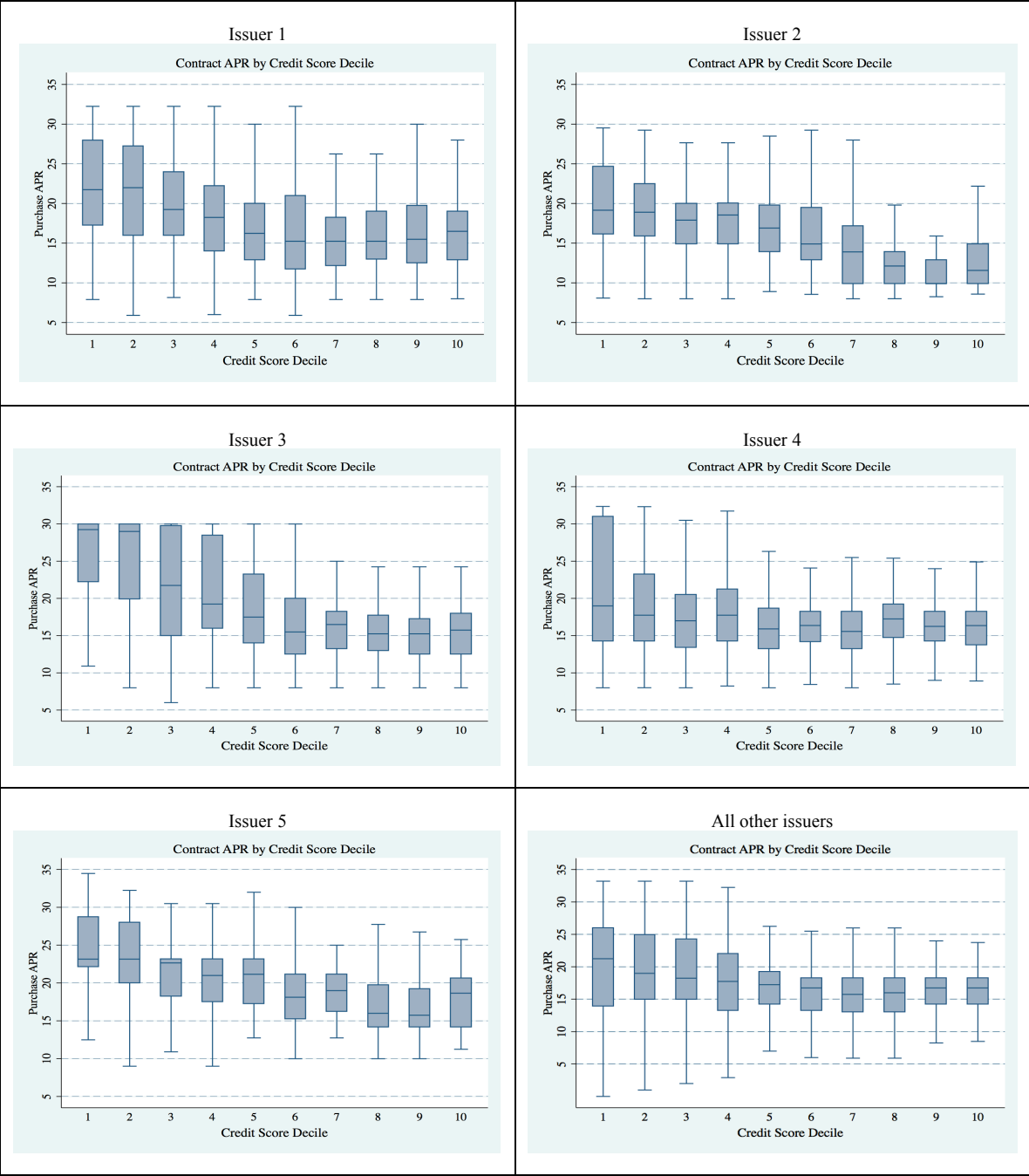


Figure 2. Credit scores and APRs for five large issuers and all other issuers.

Notes: Each pane shows the relationship between credit score decile and the within-decile distribution of contract APRs, for five largest issuers in sample and the remaining smaller issuers (the latter appearing in the bottom right pane). Each box-and-whisker plot shows the median APR as a solid horizontal line within the box, the 25th/75th percentiles as the top and bottom of the box, and the 5th/95th percentiles as the whiskers.

Table 1. Cardholder-Level Summary Statistics

Quartiles [revolving balances, \$]	Revolving Balance Quartile				
	1 [0, 499]	2 [499, 1534]	3 [1534, 4586]	4 [4586, 62515]	All [0, 62515]
Cards held	2.02	1.92	2.24	2.94	2.28
Average purchases per month, \$	730	393	499	740	591
Average revolving balances, \$	31	570	2199	11223	3505
Annualized interest costs, \$	6	113	412	1998	632
Interest costs/total borrowing costs	0.48	0.66	0.81	0.92	0.75
Annualized interest costs/annual income	0.00	0.00	0.01	0.03	0.01
Credit score	737	643	669	697	687
Income:					
under \$45k	0.42	0.40	0.33	0.26	0.36
\$45k-\$125k	0.51	0.54	0.57	0.63	0.56
\$125k+	0.07	0.06	0.10	0.11	0.08
Education:					
HS or less	0.08	0.12	0.10	0.08	0.10
Some college	0.23	0.34	0.31	0.28	0.29
College degree +	0.69	0.53	0.59	0.64	0.61
Age:					
Under 30	0.27	0.27	0.26	0.21	0.25
30-49	0.46	0.49	0.50	0.54	0.49
50+	0.27	0.24	0.24	0.26	0.24
Cardholders	1,078	1,078	1,078	1,078	4,312
Accounts	2,079	1,987	2,247	2,994	9,307
Cardholder-months	18,561	19,761	21,030	21,960	81,312
Account-months	29,438	29,681	35,117	47,851	142,087

Notes: All variables measured at cardholder level. Sample size is 4312 for all variables except income, which has 206 missing values due to item-nonresponse on registration survey. Cells show sample averages across cardholders, where each cardholder-level variable is evenly weighted across all cardholder-days in the sample. "Cards held" is the maximum number of distinct cards (accounts) observed on any one day in the Lightspeed data, at the cardholder level. Interest costs are calculated using daily balances and APRs for all card/days in the sample, and annualized. "Total borrowing costs" include interest costs, annual fees, late and over-limit fees, cash advance fees and balance transfer fees. Credit score is from one of the three major bureaus on an 850-point scale, observed upon entry into the panel. Income, education and age are self-reported upon entry into the

Table 2. Borrowing Costs in the Cross-Section of Cardholders

	Revolving Balance Quartile				Total
	1	2	3	4	
Quartile cutoffs (revolving balances)	[0, 499]	[499, 1534]	[1534, 4586]	[4586, 62515]	[0, 62515]
Cardholder-level weighted actual APR, all balances, all panelists (N=4312)					
10th	0.00	3.04	6.38	8.80	0.00
25th	0.00	8.21	11.21	11.91	3.45
50th	0.00	15.96	16.18	16.13	13.17
75th	1.08	21.11	21.68	20.77	19.53
90th	7.57	25.14	25.90	25.42	24.38
Cardholder-level weighted actual APR, revolving balances, no teaser rates (N=3629)					
10th	12.24	12.90	11.90	11.51	11.96
25th	14.99	15.74	15.24	14.01	14.99
50th	17.80	19.46	18.90	17.78	18.36
75th	21.07	24.03	23.78	22.31	23.21
90th	26.32	28.29	28.15	26.83	27.84
Cardholder-level weighted "best" APR, revolving balances, no teaser rates (N=3629)					
10th	9.90	10.89	9.87	9.17	9.90
25th	13.38	14.66	13.50	12.12	13.42
50th	16.99	18.66	17.97	16.55	17.59
75th	19.80	23.53	23.04	21.17	22.49
90th	24.24	28.09	27.85	26.02	27.19
R-sq.: monthly borrowing costs on panelist FEs	0.78	0.76	0.78	0.76	0.77

Notes: Weighted APR is at cardholder level across all card/days (or card/days without teaser APRs) in sample, weighted by total balances or only revolving balances. Balances that are non-revolving have an APR of zero. "Teaser rates" are defined by the authors as any APR below 7.99%. "Weighted best APR" is the lowest APR the cardholder could have paid in the sample period if balances were allocated to lowest-rate cards, conditional on credit limits and contract APRs. R-squared is from a regression of cardholder-month-level weighted APRs on revolving balances on a set of cardholder fixed effects; the r-squared therefore identifies the share of variation in cardholder-month-level APRs that is identified by time-invariant differences in APRs across cardholders (i.e., the pure cross-section).

Table 3. Explaining Borrowing Costs Using Observable Risk, Card Characteristics/Effects, Demographics, and Issuer/Time Fixed Effects

		Revolving Balance Quartile				Total
		1	2	3	4	
Cardholder-level models:		R-squared (unadjusted R-squared)				
	Credit score decile	0.06 (0.07)	0.14 (0.15)	0.20 (0.21)	0.12 (0.13)	0.15 (0.15)
	Above plus in-sample risk	0.09 (0.25)	0.24 (0.30)	0.34 (0.39)	0.30 (0.35)	0.27 (0.29)
	Above plus "issuer effects"	0.13 (0.38)	0.39 (0.46)	0.38 (0.47)	0.35 (0.45)	0.34 (0.37)
	Above plus card fees/characteristics	0.17 (0.42)	0.39 (0.47)	0.39 (0.48)	0.36 (0.46)	0.35 (0.38)
	Above plus demographics	0.16 (0.46)	0.38 (0.48)	0.39 (0.50)	0.37 (0.49)	0.35 (0.39)*
	N	448	1062	1061	1058	3629
Account-month-level models:		R-squared (unadjusted R-squared)				
	Credit score decile	0.07 (0.07)	0.16 (0.16)	0.17 (0.17)	0.06 (0.06)	0.12 (0.12)
	Time-invariant cardholder-level variables	0.27 (0.27)	0.38 (0.38)	0.32 (0.32)	0.26 (0.26)	0.27 (0.27)
	Time-varying usage and risk, issuer and time effects	0.24 (0.24)	0.36 (0.36)	0.29 (0.29)	0.21 (0.21)	0.25 (0.25)
	Time-varying usage and risk, issuer and time effects, card chars.	0.24 (0.25)	0.36 (0.37)	0.30 (0.30)	0.21 (0.22)	0.26 (0.26)
	All covariates above	0.32 (0.33)	0.41 (0.42)	0.34 (0.35)	0.38 (0.49)	0.30 (0.30)
	N	28375	28708	33480	44804	135367

Notes: Each cell reports the r-squared (unadjusted r-squared) from a regression of APRs on the set of listed covariates. Cardholder-level models use as the dependent variable the cardholder-level APR paid on all revolving balances, excluding teaser rates, weighted by balances across all accounts and days in sample period. Account-month-level models use the account-month-level APR as the dependent variable. Covariates are listed below and described in fuller detail in the Data Appendix. Full results from asterisked specification * are shown in Appendix Table 1. Models in demographics may have slightly fewer observations due to missing values for some demographics.

"**Credit score decile**" is a full set of indicator variables for the cardholder-level credit score. Base model also includes indicators for sample entry/exit timing.

"**In-sample risk**" (or "**time-invariant cardholder-level variables**") include the number of cards held (indicators up to 5+), cardholder-level average daily total credit line across all cards (decile indicators), cardholder-level indicators for quintile of total late fees in-sample and quartile of total over-limit fees in-sample, cardholder-level credit utilization decile indicators, average monthly purchase volume quartile indicators, and average monthly revolving balance quartile indicators.

"**Issuer effects**" in cardholder-level regressions are a vector measuring for each cardholder the average shares of revolving balances allocated to each distinct issuer in the data.

"**Card fees/characteristics**" include average fees paid per year (annual, balance transfer and cash advance) across all cards, and indicators for whether the cardholder's primary card (the one with the highest level of revolving balances, on average) has an annual fee, has a variable rate, and is a rewards card.

"**Demographics**" include indicators for income category, age and education category (see Table 1). Race, gender, and marital status are not included in these models because ECOA law prohibits lenders from pricing based on these protected categories. We discuss these variables (and show their impact on pricing in the reduced form) in Tables 5 and 6.

"**Time-varying usage and risk**" include cardholder-month level indicators for utilization decile, credit line decile, total late fees to date in sample and total over-limit fees to date in sample.

"**Issuer and time effects**" in account-month models include issuer fixed effects and month-year fixed effects.

"**Card characteristics**" in account-month models include card-level indicators for whether the card has an annual fee, is a rewards card or has a variable rate, and interactions between the variable rate indicator and month-year fixed effects.

Table 4. APR Dispersion in Choice Sets, in the Market and in the Wallet

Table 4a. Offer rate dispersion in Mintel data, within-individual, within-month, January 2007.

Percentile	Goto APR			Net-of-teaser APR		
	High (1)	Low (2)	Difference (3)	High (4)	Low (5)	Difference (6)
10th	10.99	8.99	0.00	9.90	3.75	0.46
30th	13.99	9.90	2.25	12.24	4.95	4.25
50th	16.15	9.99	4.34	14.24	6.45	7.50
70th	18.24	10.99	7.25	18.24	9.31	9.86
90th	19.50	14.90	9.25	18.99	13.39	13.95

Notes: Sample includes all reported credit card direct mail offers for 1211 individuals in the Mintel Comperemedia database from January 2007. "Goto" APR is the rate at which balances incur interest charges after expiration of the introductory "teaser" period (if any). "Net-of-teaser" APR is the average of the teaser and goto APRs over the first 24 months of the offer.

Table 4b. Estimated offer rate dispersion in Lightspeed data: within-cardholder (within-month) high APR "offer," low APR "offer" and high-low APR "offer" spread, January 2007.

Percentile	All covariates			Credit score decile only		
	High (1)	Low (2)	Difference (3)	High (4)	Low (5)	Difference (6)
10th	16.71	9.25	5.05	17.73	9.26	5.89
25th	18.52	10.91	6.42	18.74	10.67	7.11
50th	21.38	12.72	8.19	21.73	12.42	8.73
75th	24.12	14.75	10.43	24.10	14.55	10.84
90th	26.67	16.79	13.20	26.62	16.08	13.50

Notes: Estimated APR "offers" are calculated by first estimating OLS APR regressions for each of the largest six issuers, letting the relationship between cardholder characteristics and APR differ by issuer and using data from the entire sample period. Each model includes a full set of cardholder-month-level and card-month-level covariates described in Section V and the Data Appendix, (Columns 1-3 above), or just credit score decile and month-year fixed effects (Columns 4-6 above). We use the coefficients from each model to predict six fitted APRs for each panelist in each month - a hypothetical set of "offers" from the largest six issuers. This allows us to estimate a high APR, low APR, and high-low spread for each cardholder. The data above show the fitted APRs for January 2007 to facilitate comparison with Table 4a.

Table 5. Self-Reported Search Intensity, Other Shopping-Related Variables, Credit Characteristics, and ECOA-based Instruments

Variable	10-point scale (10 highest), "I always keep an eye out for better credit card offers"				All respondents	Non-respondents
	[1, 3]	[4, 6]	[7, 9]	10		
Current (in-sample) credit card accounts	2.73	2.96	3.12	2.82	2.92	2.04
Previous (closed) credit card accounts	2.87	2.95	2.97	3.49	2.98	2.66
Self-reported recent credit card applications	0.38	0.56	0.78	0.97	0.60	0.50
Credit score (median)	707	705	693	680	700	671
Average weighted "best" APR, no misallocation (median)	16.81	16.29	16.30	15.48	16.38	18.22
Revolving credit card balances (\$, average):	3182	4908	5257	3511	4313	4042
Marital status:						
Single, never married	0.16	0.21	0.22	0.22	0.20	0.18
Married/divorced/separated	0.81	0.78	0.72	0.69	0.77	0.77
Other	0.03	0.01	0.05	0.09	0.03	0.05
Female	0.71	0.68	0.58	0.47	0.63	0.70
N (panelists)	154	147	130	45	476	3079

Notes: Sample of "respondents" includes those who self-reported search intensity as part of responses to a longer email survey sent to all panelists in the first quarter of 2007. Survey content was not announced prior to the decision to take the survey. Sample excludes those who never borrow in-sample, and those with missing ECOA-based instruments. Search intensity is self-reported agreement with the statement "I always keep an eye out for better credit card offers," on a 10 point scale with 1 being "Does not describe me at all" and 10 being "Describes me perfectly." . In-sample credit card accounts is defined as in Table 1. Previous accounts is the number of previously held but closed credit card accounts from the cardholder's credit bureau file. "Applications" are the sum of affirmative responses to survey questions asking "Have you applied for any new credit cards in the last 12 months?" Surveys were emailed to panelists in 2004Q4, 2005Q1 and 2006Q1. Only those cardholders taking each survey (751 for the first, 972 for the second, and 1354 for the last) could have provided an affirmative response. "Non-respondent" column shows data for panelists who did not take the survey containing the search question, among those with non-zero average revolving balances and non-missing survey responses for ECOA instruments.

Table 6. Self-Reported Search Intensity, Instruments/Demographics, and Borrowing Costs.

	Dependent variable: weighted best APR (mean=16.35)	
	OLS	IV
Coefficient: Self-reported search intensity (10-point scale)	-0.083 (0.078)	-1.146** (0.490)
N	497	476
r-squared	0.59	0.42
full set of control variables?	yes	yes
CLR/AR test robust to weak instruments (p-value)	n/a	0.005
95% CI, robust to weak instruments	n/a	[-2.76, -0.35]
Sargan test (p-value)	n/a	0.81
Instruments in first stage:		
Marital status (1=Married/divorced/separated)		-0.677*
Gender (1=Female)		-0.797**
Instruments in reduced form:		
Marital status (1=Married/divorced/separated)		0.889*
Gender (1=Female)		0.831*

*, **, *** - significant at 10%, 5% and 1%.

Notes: Models are estimated at the cardholder level. "Weighted best APR" is the lowest APR the cardholder could pay in that month, averaged across days, if balances were allocated to lowest-rate cards conditional on credit limits. Search intensity is self-reported agreement on a scale of 1-10 with the statement "I always keep an eye out for better credit card offers," with 1 meaning "Does not describe me at all" and 10 meaning "Describes me perfectly." All models include the full set of regressors described in Tables 3 and A1. CLR/AR test is for significance of the endogenous regressor (search intensity), given that the instruments may be weak, and 95% CI is calculated using standard errors robust to the presence of weak instruments. Sargan test is for exogeneity of instruments (where rejection of the null indicates endogeneity), and is only applicable when the model is over-identified. See <http://econ.tulane.edu/kfinlay/pdf/FinlayMagnusson2009.pdf> for a discussion of the weak instrument problem and the Stata routine we use to deal with the issue. "Instruments in first stage" rows show coefficients on the instruments, from the first stage of the IV regression, with search intensity as the dependent variables (and all other covariates included). "Instruments in reduced form" rows show coefficients on instruments when they (rather than search intensity) are included as OLS covariates in the pricing model. The omitted category relative to "married/divorced/separated" is "single, never married."

Data Appendix

A. Panel Construction and Maintenance

Panelists enter the Ultimate sample by providing Lightspeed with access to at least two online accounts (checking, credit card, savings, loan or time deposit) held within the household. Panelists have typically participated in other Forrester/Lightspeed panels; the incremental payment for enrolling in the Ultimate panel averages \$20. After initial enrollment, panelists need take no action to maintain membership in the panel, and a panelist may request to leave the panel at any time.

Enrollment of new panelists occurs consistently throughout our sample period, as Lightspeed attempts to keep panel size constant by balancing enrollment against attrition. Our sample size falls over time, however, because later panelists tend not to have matched credit report data. Appendix Table 5 shows some data on how the number of panelists and their characteristics evolve over time. Because we focus on cross-panelist differences and generally employ panelist-level *time-invariant* variables in the analysis, those dynamics are not a focus of the analysis. Where appropriate, we do account for panelists' sample entry/exit dates in the empirics. We also check that our results are robust to using only individual months, or the first six months, of data.

B. Measuring Credit Risk

Our data include much, if not all, of the information used by issuers when setting and adjusting APRs:

1. *Credit scores:* A credit score from one of the major three bureaus is probably the single best summary source of information about credit history and risk. We observe one credit score for each panelist at entry into the sample, which is generally in January 2006, but occasionally later. The score, on the standard 850-point scale, summarizes risk by incorporating information about total debt, debt utilization, default history ranging several years into the past, and the number of “pulls” or applications for new credit.
2. *Supplementary credit bureau data:* We also observe other information from the report including total debt, the number of active credit cards, total credit available, the number of

active auto and mortgage “lines” (loans), the total number of past (closed) credit card accounts, and a few other variables.¹

3. *The number of credit cards held:* For each panelist on each day, we observe the number of registered credit card accounts. We define for each panelist the number of cards held as the maximum number of cards held on any one day. We have defined the variable other ways without any difference, because the number of cards held is very stable for a panelist over time.
4. *Purchase volume and revolving balances:* For each panelist we calculate average monthly purchase volume and average monthly revolving balances (these can be very different depending on whether the panelist revolves). We then bin each panelist into one of four quartiles based on each variable.
5. *In-sample late/missed payments:* A late or missed payment can trigger a “default” APR on the account in question, and is also in many cases reported to the credit bureau, leading other issuers to incorporate the late/missed payment history into APRs on new offers or existing cards. The credit score mentioned above should capture information about late/missed payments leading up to the panelist’s enrollment in Lightspeed, and once the panelist enters our data we directly observe late/missed payments. We measure running late payment counts for each account, a running count of late payments at the panelist level across all accounts, and several panelist-level and time-invariant measures of “total late fees,” “average late fees per month” and “any late fee in-sample.”
6. *Limits and utilization:* Issuers generally consider utilization (the ratio of balances to available credit) as a signal about risk. Cardholders may face higher APRs or offers either by having what an issuer considers “high” utilization, or by exceeding their credit line (going “over-limit”) on one or more cards. Again, the credit score we observe at panel enrollment incorporates all available information about utilization as of enrollment; after enrollment we

¹ Beyond the credit score itself, issuers may also incorporate this disaggregated information from the credit report into risk modeling for new account offers. In practice, adding such information non-parametrically to our models has little effect on the fit. This is partly because we use rich, disaggregated data on within-sample account performance, as described below. Customers may also self-report income, education, and other demographics on their applications, but an issuer generally does not directly observe those things. We include such demographics in our models and find that they do not improve the fit.

observe utilization levels (including both credit limits and card balances) and “over-limit” instances directly, at the card and panelist level. As with late/missed payments we here we calculate running utilization levels and over-limit instance counts, and also construct panelist-level time-invariant “Over-limit fees per month” and “Any over-limit fee in-sample” measures.

7. *Demographics*: We observe from the registration survey categorical variables measuring age, income and education. These may not be directly observed by issuers, but may proxy for variables (such as time since opening first credit card) that issuers incorporate into pricing.

Collectively, these variables are quite comprehensive. They likely compare favorably to the data observed by issuers on their own cards, although individual issuers may of course employ those data differently. They may dominate data observed by issuers on other cards (i.e., on accounts issued by other issuers).

C. Measuring Non-APR Account Attributes

We also observe a variety of card- and issuer-level characteristics:

1. *Annual fees*: For each card in the data, we observe annual fees incurred. We measure annual fees both as a cardinal number – the average annual fees paid per year, either at the card or panelist level – and using an indicator for “any annual fee incurred during the sample period,” again either at the panelist or account level.
2. *Other fees (balance transfer, cash advance, etc.)*: We observe balance transfer fees, cash advance fees, late fees, and over-limit fees as they are incurred, and include them as annual dollar costs per account or panelist. This is imperfect because we only observe fees that are incurred, rather than the contingent price that might be incurred. We have experimented with a variety of alternative approaches to this issue – inferring fees even when they are never incurred from data on actual fees paid by other panelists with the same card, for example – with little effect on the results.
3. *Rewards*: We observe for every card in the data its “card name” as a text string, which is the issuer’s name for the card. An example would be “MBNA CREDIT CARD.” The card name often reveals information about rewards or “affinity” links (e.g., “AMERICA WEST FLIGHTFUND CREDIT CARD,” “GREEN BAY PACKERS VISA”). We also observe its “account name,” which is an issuer- or panelist-defined name for the account and also

contains information about affinity/reward links (e.g., “NATIONAL WILDLIFE FEDERATION PLATINUM PLUS MASTERCARD,” “PLATINUM DELTA SKYMILES”). We do not directly observe rewards, but in practice the dollar value of rewards does not vary by much across cards. We have experimented with separate variables for rewards and affinity status, or a single combined indicator for the presence of either.

4. *Fixed/variable rate pricing:* A credit card APR may be “fixed,” meaning not pegged to another market rate, or “variable,” meaning that the APR moves monthly or quarterly with some market interest rate.² We construct an indicator measuring whether the rate is fixed or variable.
5. *Unobserved issuer-specific and state-specific effects:* We also observe the issuer (e.g., Bank of America, Capital One, Citi, etc.) for each credit card in the data. This allows us to construct a set of “issuer effects” measuring average APR differences across issuers, which might come from omitted card characteristics, or from systematic differences in pricing customer risk.³ Because a given panelist may have balances allocated across multiple cards from different issuers, our panelist-level regressions measure the average share of revolving balances held on cards of each issuer. In card-level regressions we simply include a fixed effect for the card issuer. (We’ve also estimated specifications with fixed effects for the card name, since, e.g., MBNA cards may be remain branded as “MBNA”, even after MBNA gets acquired by Bank of America. This alternative definition of issuer does not affect the results.) We also observe the panelist’s state of residence and in unreported specifications include fixed effects for state of residence; those effects might capture any number of omitted influences on state-level supply or demand for credit.
6. *Sample entry/exit dates:* Because panelists may be in the data for less than the entire sample period, we include a set of indicators for the panelist’s first and last months in the data. This corrects for variation in APRs generated by systematic time-varying APRs, combined with differential entry/exit dates by panelists.

² See Stango (2000) for a detailed discussion of fixed and variable rate pricing in credit cards.

³ Issuer effects are de-identified when we report the results, per confidentiality provisions of our data licensing agreement with Lightspeed.

D. Classifying “Teaser” Rates

Our data do not identify teaser rates as such, but they are fairly easy to classify empirically because they are significantly lower than even the lowest contract rate offered to the best credit risks during our sample period. We classify any APR below 7.99% as a teaser rate (source: tabulations from the Mintel data discussed in Section VI-A). This discards 5% of account-months, and 1% of panelists who always pay teasers in-sample.

E. Representativeness

Starting at the top of Table 1, our cardholding distribution matches up well with data from the 2007 Survey of Consumer Finances (SCF), particularly when one uses our “complete cards” sub-sample (Appendix Table 2) as the benchmark.⁴ Purchases also match up well with the SCF, in which the comparable 25th, 50th, and 75th percentiles of the weighted data are \$20, \$250, and \$1000.⁵ Comparisons of revolving debt are more problematic, given substantial underreporting in the SCF (Zinman 2009; Brown et al. 2011), and the lack of distinction between revolving and transaction balances in credit bureau data (and in the data that issuers report to regulators). But if we look simply at outstanding balances, we see about 50% less in our data than in the bureau (Brown et al Appendix Table 1).⁶ This suggests that our data may understate the *level* of debt and total interest costs, if not necessarily the degree of heterogeneity, relative to the broader population. For our panelists the share of total credit card costs from interest (vs. fees) is 74%, as compared to an estimate of 80% from 2007 issuer-side data (source: Cards&Payments).

Data from other sources on APR distributions is limited, but comparing our data to the SCF (which asks about a single APR, on the card used most often), we find similar dispersion; the interquartile range in the SCF is 800 basis points, which is comparable to what we observe, even if one restricts our data to the subsample of panelists’ “primary cards.” The dispersion we observe also looks similar to that in more recent administrative data from the OCC.⁷ How the central tendency in our data compares to other data is murkier. The APRs we observe are higher on average than the self-reported APRs in the 2007 SCF, but are similar to those in the OCC

⁴ Zinman (2009) shows that cardholding in the SCF matches up well with issuer-side data.

⁵ Zinman (2009) shows that card purchases in the SCF match up well with issuer-side data.

⁶ This may be explained in part by the life-cycle pattern of credit card balances (Brown et al Figure 4), coupled with the fact that our sample is relatively young.

⁷ Source: <http://files.consumerfinance.gov/f/2011/03/OCC-Presentation.pdf> .

administrative data from 2009. In short, we see little reason to believe that the dispersion we observe is uncharacteristic of the national population of U.S. credit cardholders.

In terms of demographics, our panelists are younger, more educated, and higher income (conditional on age) than national averages. The overall credit score distribution looks representative, conditional on demographics (source: tabulations from the Payment Cards Center of the Federal Reserve Bank of Philadelphia).

Appendix Table 1. Full Results from a Cardholder-Level Model (asterisked specification in Table 3, row 5, last column).

Dependent variable: cardholder weighted average APR on revolving balances, no teaser rates: mean (LHS)=17.48						
Variable	Coefficient	Standard error	Variable	Coefficient	Standard error	
Credit score: decile 2	-0.455	(0.322)	Total late fees >0: quintile 1	1.059**	(0.426)	
Credit score: decile 3	-0.531	(0.350)	Total late fees >0: quintile 2	1.795***	(0.457)	
Credit score: decile 4	-1.291***	(0.388)	Total late fees >0: quintile 3	2.569***	(0.452)	
Credit score: decile 5	-1.676***	(0.412)	Total late fees >0: quintile 4	3.899***	(0.473)	
Credit score: decile 6	-1.810***	(0.426)	Total late fees > 0: quintile 5	5.628***	(0.522)	
Credit score: decile 7	-2.040***	(0.454)	Total over-limit fees >0: tertile 2	-0.146	(0.403)	
Credit score: decile 8	-2.460***	(0.487)	Total over-limit fees > 0: tertile 3	0.032	(0.450)	
Credit score: decile 9	-2.651***	(0.514)	Average utilization: decile 2	-0.285	(0.492)	
Credit score: decile 10	-2.184***	(0.525)	Average utilization: decile 3	-0.074	(0.507)	
Two cards held	-0.294	(0.222)	Average utilization: decile 4	0.057	(0.512)	
Three cards held	-0.192	(0.294)	Average utilization: decile 5	-0.230	(0.539)	
Four cards held	-0.241	(0.380)	Average utilization: decile 6	-0.665	(0.567)	
5+ cards held	-0.433	(0.397)	Average utilization: decile 7	0.032	(0.585)	
Total credit line: decile 2	-0.309	(0.366)	Average utilization: decile 8	0.844	(0.612)	
Total credit line: decile 3	-0.953**	(0.450)	Average utilization: decile 9	0.883	(0.632)	
Total credit line: decile 4	-1.684***	(0.521)	Average utilization: decile 10	1.668***	(0.638)	
Total credit line: decile 5	-2.335***	(0.560)	Average monthly purchase volume: quartile 2	-0.134	(0.242)	
Total credit line: decile 6	-1.867***	(0.608)	Average monthly purchase volume: quartile 3	-0.304	(0.309)	
Total credit line: decile 7	-2.402***	(0.639)	Average monthly purchase volume: quartile 4	0.143	(0.396)	
Total credit line: decile 8	-2.292***	(0.694)	Average monthly revolving balances: quartile 2	-0.859**	(0.379)	
Total credit line: decile 9	-2.543***	(0.749)	Average monthly revolving balances: quartile 3	-1.003**	(0.483)	
Total credit line: decile 10	-2.671***	(0.837)	Average monthly revolving balances: quartile 4	-1.150*	(0.632)	
Average annual fees paid/year, all cards	0.027	(0.018)	Income \$35k-45k	-0.134	(0.290)	
Average balance transfer fees paid/year, all cards	-0.008	(0.036)	Income \$45k-87.5k	0.549**	(0.257)	
Average cash advance fees paid/year, all cards	0.035	(0.052)	Income \$87.5k-125k	1.092***	(0.360)	
Panelist's primary card: variable rate?	0.979***	(0.200)	Income \$125k+	1.423***	(0.402)	
Panelist's primary card: annual fee?	0.722***	(0.244)	Education: some college	-0.869***	(0.288)	
Panelist's primary card: rewards?	-0.358	(0.267)	Education: college+	-0.471	(0.289)	
			Age 30-39	-0.146	(0.244)	
			Age 40-49	0.124	(0.264)	
			Age 50-59	0.059	(0.289)	
			Age 60+	0.305	(0.377)	
Constant	21.949***	(1.179)				
Cardholder-level "issuer effects": p-value		0.000	N		3101	
Indicators for last month in sample: p-value		0.000	r-squared (unadjusted)		0.35 (0.39)	

*, **, *** - significant at 10%, 5% and 1%.

Notes: Coefficients are from an OLS regression at the cardholder level. Dependent variable is the cardholder-level weighted average APR on revolving balances during the sample period, excluding balances with "teaser rates." Sample begins with 4312 panelists from Table 1, dropping 627 who never revolve balances and 56 who borrow on teaser rates for entire time in sample. Remaining attrition occurs due to missing values for income and education. "Cards held" is the maximum number of different accounts open on any one day during the sample period. "Total credit line" quintile is measured using the average daily credit line on all cards. "Average utilization" is the average across all days in the sample of daily balances (revolving or not) divided by total credit line, across all cards. "Primary card" is the card on which a majority of balances are held during the sample period, across all days. Cardholder-level "issuer effects" are a vector measuring for each cardholder the average shares of revolving balances allocated to each distinct issuer in the data.

Appendix table 2. ECOA-based instruments in the reduced form, at the issuer level.

Instruments in reduced form, account-month-level models:	Dependent variable: account-month credit card APR					
	All	Issuer 1	Issuer 2	Issuer 3	Issuer 4	Issuer 5
Marital status (1=Married/divorced/separated)	-0.130	-0.046	-0.218	-0.273	0.175	-0.218
Gender (1=Female)	0.026	0.227	0.340	0.451	-0.451*	-0.214
Age (omitted = "under 30"):						
30-39	0.374**	0.066	0.373	0.984**	0.869**	0.325
40-49	0.550***	-0.219	0.060	0.851*	0.826*	0.608
50-59	0.597***	-0.305	0.171	1.279**	1.672***	0.378
60+	0.332	-1.205	0.693	0.906	1.052**	-0.351
N	126,785	15,116	24,612	26,640	20,109	19,585

*, **, *** - significant at 10%, 5% and 1%.

Notes: Coefficients are from OLS regressions of account-month-level credit card APRs on card-month-level covariates (as described in Table 3). Model includes month-year fixed effects. First column pools all issuers and adds issuer fixed effects. The next five models show coefficients from issuer-level regressions (i.e., allowing issuer-level variation in coefficients), for the largest five issuers in the sample. Model has slightly fewer observations than the account-month models in Table 3 due to missing values for some instruments and demographics.