Auto Dealer Loan Intermediation: Consumer Behavior and Competitive Effects†

Andreas Grunewald§, Jonathan Lanning*, David Low*, Tobias Salz‡

July 18, 2019

Abstract

This paper evaluates the intermediation of auto loans through auto dealers using new and comprehensive administrative data. The arrangement between auto dealers and banks gives dealers discretion to price the loans to consumers. Our project leverages detailed knowledge of these vertical relationships to demonstrate that consumers are substantially less responsive to finance charges than to vehicle charges. Dealers take this consumer-specific wedge into account when jointly pricing the car and the loan. The discretion that lenders give to dealers therefore leads to a form of price discrimination. Whether or not this price discrimination is beneficial or harmful for consumers depends on the nature of the wedge. If it is mostly driven by inter-temporal considerations, such as credit constraints, dealer discretion can be welfare improving. If instead the wedge is not an accurate representation of underlying preferences or constraints, this price discrimination may decrease consumer welfare. The data favors the latter view. In counterfactual computations we explore what happens to consumer welfare if dealers have no discretion in pricing loans. We find that the total price for consumers would drop by about $350.25, leading to a $1.78 billion increase in annual consumer surplus.

JEL classification: L13, L62, D90.

Keywords: Auto Loans, Add-On Pricing, Obfuscation, Oligopoly Competition.

†PRELIMINARY, COMMENTS WELCOME. The views expressed are those of the authors and do not necessarily reflect those of the Bureau of Consumer Financial Protection or the United States. Andreas Grunewald and Tobias Salz thank the Institute for Social and Economic Research and Policy at Columbia University and the Russel Sage Foundation for financial support. We thank Glenn Ellison, Matt Gentzkow, Kate Ho, Michael Riordan, Mike Whinston, and many seminar participants for helpful comments. Guy Aridor, Jasper Clarkberg, Thi Mai Anh Nguyen, and Yining Zhu provided outstanding research assistance.

§University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany. Email: gruni@uni-bonn.de

* Bureau of Consumer Financial Protection

‡ MIT Department of Economics, 50 Memorial Dr, Cambridge, MA 02142 10027. Email: tsalz@mit.edu
1 Introduction

Goods and services are often bundled with financial products, such as installment loans, deferred-interest financing, and extended warranties. Most importantly, this bundling is common in the auto market. American consumers have more than $1 trillion in auto debt, making it the third-largest debt market in the United States.¹ Consumers can unbundle vehicles from loans by obtaining financing “directly” from lenders, but more than 80% of auto loans are obtained “indirectly” through dealers as part of the vehicle sales process.² Dealers have considerable discretion to price loans, and so dealers and consumers bargain over both the car price (Ayres and Siegelman, 1995, Goldberg, 1996) and the interest rate on the loan.³ In this project we study how consumer demand responds to these two different prices and quantify the equilibrium implications. We pay particular attention to the institutional arrangements between lenders and dealers that enable this price discrimination.

First we show that consumers are substantially more sensitive to changes in the car price than the loan price. To identify the size of the wedge between buyers’ price elasticity and their finance charges elasticity, we use the contractual incentives facing dealers. To arrange a loan, dealers first request buy rates from lenders. The buy rate is the lowest interest rate at which a lender is willing to issue a loan to a particular consumer. After receiving buy rates, dealers choose a lender to provide the loan as well as a markup on the buy rate of the loan. Dealers’ compensation, called dealer reserve, typically consists of a fixed payment as well as a fixed share of the revenue generated by the markup. Our analysis makes use of new and rich transaction-level administrative data that includes buy rates, markups, and dealer reserve. This data allows us to directly back out the contractual incentives facing dealers, as well as the resulting loan outcomes.

Based on these vertical arrangements between dealers and lenders, the intuition for our approach to estimate price elasticities is straightforward. Dealers share profits from loan markups with lenders, but keep all profit from vehicle price markups. If consumers treat vehicle and finance charges equally, profit maximization would require that dealers charge the highest possible markup on the car and set the loan markup to zero. Conversely, if consumer demand is sufficiently less elastic to finance charges, then dealers should mark up loans. A dealer’s markup choice therefore reveals information on a consumer’s car price elasticity relative to loan price elasticity. Since we observe the markup for each contract in our data, we can derive individual-specific estimates of the wedge between buyers price elasticity and their finance charges elasticity.

We find that, on the margin and in equilibrium, consumers behave as if they would pay a dollar more in finance charges to reduce the vehicle price by $.80. Hence, consumer demand

²We estimate this number in Appendix B.1. See also Davis (2012).
³Consumers and dealers also typically negotiate over other terms of the deal, including vehicle add-ons and the price of a trade-in vehicle.
is substantially more elastic to vehicle price than loan price. We also derive a lower bound of $380 for the difference between actual finance charges and finance charges consumers appear to perceive. Since these measures are obtained for each consumer separately (i.e. non-parametrically), we can study heterogeneity in responsiveness to finance charges. We find that consumers with lower income, as well as consumers who live in areas with lower education levels and slower internet access, have the most attenuated response to an increase in the loan price.

We consider a number of explanations for these results, including intertemporal discounting and credit constraints, vehicle taxes, loan prepayment, and repeated dealer-lender interactions. These explanations can either qualitatively or quantitatively not account for our results. However, our findings are consistent with consumers systematically underestimating finance charges, which is then used by dealers to increase revenue.\footnote{In particular, consumers might lack financial literacy in general (Lusardi and Mitchell, 2014), the car price might be more salient to them (Chetty et al., 2009, Bordalo et al., 2013, Köszegi and Szeidl, 2012), dealers might exploit the complexity of financial contracts to shroud or obfuscate crucial information (Gabaix and Laibson, 2006, Ellison and Ellison, 2009), or consumers might neglect some dimensions of the price because their capacity to process information is limited (Sims, 2006, Mackowiak et al., 2018)}

To further explore how the joint pricing of cars and loans plays out in a competitive environment we estimate an equilibrium model in which dealers compete for customers, taking as given the contracts with lenders. This exercise allows us to explore the extent to which the wedge in consumers’ elasticities affects various market outcomes, e.g., dealer competition, lender pricing decisions, and consumer surplus. In addition to the administrative data, this part of the project uses a comprehensive dataset of dealer sales and loan terms that tracks most of the US car market.

In the model, dealers play a differentiated product-pricing game. To model consumer demand and competition we build on Berry et al. (1995), allowing dealers to set loan interest rates in addition to vehicle prices. We also model lender competition in the buy-rate auction. In this auction lenders win the contract if they quote the lowest buy-rate. However, relative to a standard first price auction, they also take into account the downstream pricing decision of the dealer and the resulting likelihood of sale, which affects profits through their share of the finance charges. In this modified auction game similar insights as in Guerre et al. (2000) prevail and allow us to recover lender cost as a function of observables and the primitives that govern dealer price competition.

Counterfactual computations from the model show that overall prices in the market would be 1.12 percent lower and consumer surplus 10.63 percent higher if consumers did not treat finance charges differently from vehicle charges. In equilibrium, dealers therefore substantially benefit from consumers attenuated response to an increase in finance charges. If dealers instead had to pass through the interest rates that banks quote to them, without additional discretion, prices would decrease by 1.25 percent and consumer surplus increase by 5.14 percent. In part, these changes are due to an informational advantage of the dealer. Dealers interact directly with customers and their markup choice likely depends on customer attributes that are
directly observable to them but not the lender. Without delegation lenders can not replicate the extent of price discrimination in which dealers were engaging.

We contribute to the literature studying consumer and car dealer decisions in the process of financing a car. Previous studies have focused on the market of subprime lending and how it is affected by down payment requirements (Einav et al., 2012), cash sensitivity (Adams et al., 2009), credit scoring (Einav et al., 2013), and the presence of binding credit constraints (Attanasio et al., 2008). We complement these papers by providing a detailed analysis of the vertical arrangements between auto dealers and lenders for the prime segment of the lending market. The corresponding insights also allow us to identify to what extent consumers are less sensitive to charges on the loan compared to charges on the car. While we conduct our analysis for prime consumers, our insights might not be limited to this segment of the market. In particular, price discrimination on the basis of the wedge between price sensitivities might be an important aspect for why subprime consumers sign loans with remarkably high interest rates.

There is a growing literature estimating empirical models incorporating some form of sub-optimal consumer decision making. Relying only on observational data, this type of work faces a pronounced identification challenge due to the necessity to uncover additional parameters that determine how much consumers depart from a fully standard model of decision making. It is oftentimes not clear how such parameters would be separately identified from those that capture preferences. Several studies in the literature have overcome this challenge in the context of health insurance plans (Handel, 2013), signing cell phone plans (Grubb and Osborne, 2015), and uptake of extended warranties (Abito and Salant, 2017). Complementing these insight we document for the car market that at least a part of the wedge in elasticities is stemming from sub-optimal decision making on the side of consumers. The existing studies have in common that they either leverage some form of exogenous variation or auxiliary data in addition to price and quantities. In our case we leverage instead our knowledge of the incentive structure that the supply side faces to derive insights about the demand side. An important advantage of this approach is that our estimates do neither rely on exact functional forms of consumer preferences nor do they require assumptions on the market structure.

To study the competitive effects of the joint pricing of cars and loans, we exploit our knowledge on the vertical arrangements between dealers and lenders. While dealer revenues generated by the intermediation of loans are substantial (Davis, 2012), the literature on competitive effects in the car market has largely abstracted from the joint pricing of cars and loans (Berry et al., 1995, Morton et al., 2001, 2003, Gavazza et al., 2014, Nurski and Verboven, 2016, Murry, 2017). We complement these studies by taking such strategic considerations

---

5Overviews on structural work in behavioral economics in general and within the field of industrial organization can be found in DellaVigna (2018) and Grubb (2015), respectively.
6For franchise dealers that sell vehicles to private costumers Davis (2012) estimates that more than half of their profits are generated through their F&I department.
into account when estimating demand and conducting counterfactual computations. The own price demand elasticities that we document turn out to be similar to the ones in the literature but slightly lower. Taking into account the joint pricing also allows us to evaluate policies that aim to change the bargaining process prior to purchasing a car, e.g., eliminating the discretion of the car dealer to price the loan.

While the auto market is important in its own right, this work also contributes to our understanding of other markets for durable goods that are commonly bundled with financial contracts. For example, household durables are often sold along with extended warranties, airline tickets with travel insurance, and car rentals with liability insurance. A common theme in these markets is that sales agents and intermediaries cater products to consumers that are either unsophisticated or find it costly to comparison shop (Egan, 2018, Egan et al., 2019, Guiso et al., 2018, Allen et al., 2014, Woodward and Hall, 2010). Our counterfactual analysis show that the discretion that sales agents hold when bargaining with consumers may lead to price discrimination that is harmful for consumer welfare. This insight may transfer to all of these markets as long as the intermediary holds the discretion to price the financial product as well as the durable.

Finally, there are several papers at the intersection of financial decision making and behavioral economics which inform our theory by documenting how consumers make financial decisions. For example, present bias has been used to explain why consumer are willing to hold high-interest debt (Ausubel, 1991) and why consumers hold large amounts of illiquid assets and at the same time borrow at high interest rates (Laibson et al., 2015). Stango et al. (2017) use RAND’s American Life Panel to relate various behavioral anomalies to each other and financial decision making. Stango and Zinman (2009) argue that consumer misperception of interest rates can be explained by a tendency of households to linearize exponential functions. These papers inform our approach to analyze the data. At the same time, we also contribute to this literature by showing that consumers are less sensitive to loan than to car prices prices. To the extent that this wedge in sensitivities is driven by suboptimal financial decision making, we also document how market outcomes depend on such anomalies. Our results suggest that some of the expansion in auto loan credit might be driven by a combination of car buyers’ inability to fully account for the price of the loan and dealers’ corresponding incentive shift markups. Since auto loans are a big contributing factor to the overall increase in household debt (Mian and Sufi, 2016, Schlagenhauf and Ricketts, 2012) these results might be of broader interest to the literature on household finance.

---

7For a recent survey on this literature see Madrian et al. (Forthcoming).
## 2 Institutional Details

Car prices and loan terms are negotiated between dealers and customers, so actual prices and interest rates often differ substantially from advertised ones. Typically, a sales agent and the consumer first negotiate vehicle specifics, e.g. trim level, options, and price.\(^8\) Afterwards, the dealer’s “Finance and Insurance” (F&I) agent and the consumer negotiate loan terms.\(^9\) To get rate quotes, the F&I agent typically submits the customer’s financial information into at least one of three major systems: DealerTrack, RouteOne, and/or Credit Union Direct Lending. Dealers may select specific lenders from which to solicit bids, or send the application to all lenders (DealerTrack advertises 1,500+ lenders). On average, dealers appear to maintain active relationships with more than 9 lenders, and are likely to request quotes from several of them for any given deal.

Each solicited lender submits a *buy rate*, which is the minimal interest rate at which it is willing to make the loan.\(^10\) The dealer may then add a discretionary markup to the buy rate.\(^11\) Most lenders allow markups of at most 200-250 basis points, but markups are otherwise discretionary.\(^12\) The revenue from this markup is split between dealer and lender according to pre-specified agreements. Details on these revenue sharing agreements are provided below.\(^13\)

Since dealers only act as intermediaries for financing, loans generally go directly on the balance sheet of the lender.\(^14\) Loans present both default risk and prepayment risk. Default risk is minimal for prime consumers. For subprime consumers, contracts often discipline markups and split default risk between dealers and lenders. In both the subprime and prime...

---

\(^8\)Sales agents are often paid based on commission and sales targets. The commission is often a function of the sale (e.g. a “flat” commission) and / or the simple profit to the dealer (i.e. price paid less direct vehicle cost). The sales targets are often discontinuous, and can result in direct bonuses and / or increases in the commission percent earned (either for additional sales, or occasionally for all sales in a month).

\(^9\)F&I agents are compensated by the dealer via a commission (or commission-like) mechanism based on the additional value of the deal above the sales price (often called the buy rate) (more on the profit netted, less on the number of deals done). F&I agents’ incentives are typically to sell add-ons and get the best “reserve” (net payment) from the lender.

\(^10\)This process can happen very quickly (seconds to minutes), or can take longer (hours to days). Super prime quotes (e.g. credit score 740+) are often fully automated, and take virtually no time. “High prime” deals (e.g. credit score 700+) are sometimes automated (depending on the lender) but are usually handled quickly. One bank that “manually” underwrites prime loans targets 2 minutes to a decision/rate. Non-automated rates typically vary by more than automated rates, so the likelihood of getting a lower rate is higher, though it can take more time. “Near/low prime” deals (credit score 620+) that require some manual pricing/underwriting are where lenders differ most. Some lenders try to move as quickly as possible, other lenders contact the dealer to discuss terms (figuring “hands on” approach is more likely to secure deal) or engage in a thorough assessment of risk (takes longer), but offer more favorable compensation to dealer. Subprime (credit score < 620) quotes are often available from only a few lenders, and often require phone calls and / or additional information and time. These customers are usually on the extensive margin of borrowing, and are unlikely to be particularly sensitive to rates (e.g. one subprime lender sets all rates to 23.99%, unless the customer has a competing offer).

\(^11\)These markups can be negative. “Markdown” is allowed by most lenders, where the dealer pays a fixed fee to decrease the contract rate below the lender’s buy rate. This is rare and usually happens when the buyer has a lower competing offer, and/or when the F&I agent has made enough on add-ons to be willing to subsidize the buyer’s rate.

\(^12\)Lender-imposed caps on markups arose after a series of class-action lawsuits against auto lenders that settled between 2003 and 2006. Before, many markups were even higher. See Cohen (2012).

\(^13\)http://www.elendsolutions.com/if-the-cfpb-axes-dealer-reserve-how-would-the-sales-financing-process-change/ provides some additional rationale for why the institution has arrived in this way. The dealer wants to present the customer a bundle, and both prices are negotiated at once. It is not clear how this could be done if the contract with the dealer entails a fixed payment with the bank. The dealer would not be able to quote the customer a price until the banks are contacted and the desired rate might not be feasible.

\(^14\)The story is more complex for captives and BHPH.
markets, prepayment risk is substantial.\footnote{Consumers may prepay an auto loan in cash, through refinancing, or through trading in their vehicle for a different one.} Dealers typically assume all prepayment risk for the first 3 to 6 months. If the loan is prepaid during this time period, the dealer often returns the entire dealer reserve to the lender. Lenders assume all prepayment risk after this time period.

Using the CCP, we estimate that at least 80\% of auto loans are “indirect”, i.e. obtained through auto dealers as described above. (For details of the estimation see Section ??). However, consumers can also get loan quotes directly from lenders. They can obtain these quotes either before negotiating with the dealer or afterward. If the consumer finances the vehicle directly, the dealer receives no revenue from the financing.

\section{Data and Descriptive Evidence}

This section first describes the datasets and then provides descriptive evidence. We pay special attention to the relationships between banks and dealers and the compensation that dealers receive for intermediating loans.

\subsection{Data}

The project combines three different datasets. First, and most importantly, we use a new administrative dataset of auto loans from various financial institutions. The data records car, loan, and buyer characteristics and includes several million transactions from 2010 to 2014. We observe the make and models of cars, whether they were new or used, their mileage and model year, and in some cases the price of add-ons. We observe a buyer’s zip code, income, and credit score. For each transaction, we also observe the encrypted numeric identifier of the lender. There are several lenders in our data, and an average of over 7,000 dealers per lender. The loan characteristics we observe include the APR, the term length, the down payment, and the trade-in value for the old car. Additionally, we observe whether loans are “subvented”, i.e. subsidized by car manufacturers to increase vehicle demand.\footnote{Subvented loans are typically from captives, but non-captive lenders do sometimes have agreements with car manufacturers to extend subvented loans.} Crucially, we observe the buy rate, the markup, and the dealer reserve. Recall that the buy rate is the rate at which the lender is willing to finance the transaction. The markup is the discretionary interest that the dealer adds to the buy rate, and the dealer reserve is the payment that the dealer obtains from the lender for originating and marking up the loan.

The administrative data provides detailed information on observed transactions, but it does not cover the entire market. We therefore use complementary data from AutoCount, from which we can observe market shares of lenders and dealers for the majority of states in the U.S. AutoCount also records whether dealers are franchised (i.e. associated with a

\footnote{Consumers may prepay an auto loan in cash, through refinancing, or through trading in their vehicle for a different one.}

\footnote{Subvented loans are typically from captives, but non-captive lenders do sometimes have agreements with car manufacturers to extend subvented loans.}
Figure 1: Markup Revenues and Dealer Reserve

![Figure 1: Markup Revenues and Dealer Reserve](image)

**NOTE:** The figure shows a bin-scatter plot where the x-axis is the distribution of revenues generated by dealers markups over the lifetime of the loan and the y-axis the distribution of dealer reserves. The graph shows that the relationship is well approximated by a linear contract.

car manufacturers) or independent. In the 2011 AutoCount Data we can look at the market shares of different types of lenders. Banks, captives, credit unions, finance companies, and Buy Here Pay Here companies had respectively 46.1%, 26%, 14.1%, 9%, and 4.8% of the market. However, the data from AutoCount does not include information on buy rates and mark ups.

Finally, we also use the CFPB’s Consumer Credit Panel (CCP), which is built from data from one of the three nationwide consumer reporting agencies. The CCP also does not contain data on buy rates or markups, nor does it include data on the vehicle securing the loan. However, we can use it to address questions on loan performance over time.

### 3.2 Contracts Between Banks and Dealers

Markup revenue is shared between the dealer and the lender according to a contractually-specified formula. Figure 1 plots dealer reserve against loan markup revenue in a binscatter plot. On average, dealers receive a fixed payment of $137 per loan plus about 66% of markup revenue. The dealer share of markup revenue varies slightly from lender to lender with a coefficient of variation of $0.079$.\(^{17}\) Adjustments of those terms across dealers and geographic regions appear to be rare. Two points are worth noting about the observed contracts. First,

\(^{17}\)For one lender we directly observe the dealers’ share. It varies somewhat across observations, so we take the median dealer share for this lender for the calculation of the coefficient of variation.
contracts are very nearly linear.\textsuperscript{18} Hence, dealers receive a pre-specified amount of every dollar of extra revenue they generate through loan markups. Second, the slope of the linear contract is substantially below unity. A dealer therefore receives only a fraction of markup revenue. We build on these observations in our theoretical considerations in Section 4.

3.2.1 Markups and Dealer Compensation

In our data 77.8 percent of loans are marked up.\textsuperscript{19} The average mark up in the dataset is 1.08, which at $\approx 43\%$ over the buy rate is substantial. While markups are common for all types of buyers, their size varies systematically with buyer observables. Table 1 shows how markup and dealer reserve vary with the buyer’s credit score, income, and the price of the vehicle. Markups are higher for buyers with low credit scores. As a consequence, dealer reserve is also higher on average for buyers with low credit scores. Markups are also higher for low income buyers. However, dealer reserve is lower for low income buyers, because they typically buy cheaper cars with smaller loans that generate less revenue for a given markup.

The results from Table 1 also inform the theoretical considerations below. In particular, the descriptive evidence suggests that dealers make extensive use of loan markups even though the corresponding profit must be shared with lenders. This is true regardless of buyer income, risk characteristics, and vehicle segment.

3.2.2 Subvented Loans

The core idea of Section 4 is to use dealers’ markup decisions to estimate buyers’ sensitivity to loan charges relative to vehicle charges. We will use dealers’ incentive structure and a model of optimal pricing to derive the conditions for identification of the consumer utility function. This analysis requires that dealers perceive the pricing of the car and the loan as a joint maximization problem.

Brown and Jansen (2019), Argyle et al. (2018) and Argyle et al. (2019) provide strong evidence that loan and vehicle prices are jointly determined. But since this is a critical assumption for our analysis, it is worth verifying in our data. To do so we focus on subvented loans. Recall that vehicle manufacturers subsidize interest rates on subvented loans, so markups on these loans are typically restricted and often prohibited. As a consequence, the mean markup for subvented loans is only 0.06 compared to 1.12 for non-subvented loans. Hence, dealers’ profit for these types of sales comes almost entirely from the markup on the vehicle price.

Table 2 shows the estimates from a linear regression model that predicts the effects of subvention on the price of the car and the loan. We include fixed effects for the lender that

\textsuperscript{18}If we do not pool lenders but instead plot the same graph for lenders separately, contracts still appear to very nearly linear. To protect the confidentiality of lenders in our data, we cannot show these graphs.

\textsuperscript{19}In very rare occasions dealers might also be allowed to mark down a loan. This is the case for 0.7 percent of loans we observe. Anecdotal evidence suggests that mark downs are typically not possible. However, lenders might allow large dealers a very limited number of mark downs per month if they are less than 0.25 percentage points.
<table>
<thead>
<tr>
<th>Quartile</th>
<th>Credit Score Mean</th>
<th>Credit Score Std Dev</th>
<th>Income Score Mean</th>
<th>Income Score Std Dev</th>
<th>Vehicle Price Mean</th>
<th>Vehicle Price Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Quartile</td>
<td>1.20</td>
<td>0.83</td>
<td>1.20</td>
<td>0.82</td>
<td>1.24</td>
<td>0.87</td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>1.14</td>
<td>0.82</td>
<td>1.13</td>
<td>0.81</td>
<td>1.13</td>
<td>0.82</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>1.09</td>
<td>0.82</td>
<td>1.09</td>
<td>0.81</td>
<td>1.08</td>
<td>0.80</td>
</tr>
<tr>
<td>4th Quartile</td>
<td>1.08</td>
<td>0.81</td>
<td>1.08</td>
<td>0.80</td>
<td>1.09</td>
<td>0.77</td>
</tr>
</tbody>
</table>

**Note:** Summary statistics of markups (upper panel), margin refers to mark ups as a fraction of the buy rate (middle panel) and reserve refers to the dealer reserve (lower panel) by credit score, income, and vehicle price.

4 Consumer Demand

Building on the institutional details described above, we now outline our empirical strategy to quantify consumer sensitivity to changes in finance charges relative to vehicle price. Our basic identification argument is the following. Dealers share a substantial fraction of loan markup revenue with lenders, but keep all vehicle markup revenue. If customers are indifferent between loan charges and vehicle charges, dealers should maximize the car markup and

issued the loan, the time the loan was issued, the purchased car model and make, and the buyer’s zip code. If car and loan prices are chosen independently in competitive markets, we would expect car prices not to be affected by loan subvention. Columns (1) and (2) show that this is not true. Subvention implies two changes in the prices. On the one hand, the price of the loan decreases by approximately $150. On the other hand, the price of the car increases by approximately $850. Hence, dealers seem to increase vehicle markups if their discretion to price the loan is reduced. In fact, overall transaction prices are estimated to be higher if loans are subvented.
Table 2: Predicted Effects on Price and Loan Outcomes

<table>
<thead>
<tr>
<th></th>
<th>Car Price</th>
<th>Finance Charges</th>
<th>Overall Charges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subvented</td>
<td>846.4**</td>
<td>-148.1*</td>
<td>780.1*</td>
</tr>
<tr>
<td>(2.13)</td>
<td>(-1.91)</td>
<td>(1.74)</td>
<td></td>
</tr>
<tr>
<td>Credit Score</td>
<td>123.7***</td>
<td>-577.1***</td>
<td>-443.1***</td>
</tr>
<tr>
<td>(3.21)</td>
<td>(-57.59)</td>
<td>(-9.93)</td>
<td></td>
</tr>
<tr>
<td>Log Monthly Income</td>
<td>0.0759***</td>
<td>0.0108***</td>
<td>0.0924***</td>
</tr>
<tr>
<td>(15.97)</td>
<td>(14.76)</td>
<td>(16.90)</td>
<td></td>
</tr>
<tr>
<td>Loan Term</td>
<td>181.3***</td>
<td>970.2***</td>
<td>1166.5***</td>
</tr>
<tr>
<td>(8.54)</td>
<td>(203.27)</td>
<td>(48.05)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Bank, month, model, and zip code fixed effects also included, but not shown. Standard errors clustered at zip code level.

keep all profit from the transaction. If instead dealers set positive markups, it must be because markups allow them to increase the total price of the loan and vehicle by at least the markup revenue remitted to the lender. This in turn can only be true if customers are less sensitive to loan charges than to vehicle charges.

This basic insight allows us to estimate consumers’ price sensitivity under very mild assumptions on the dealers’ pricing problem. We do not have to specify either the exact market structure under which dealers compete or the consumer model that generates the wedge between vehicle and loan price elasticities. Our approach is therefore reminiscent of recent structural work that does not require a full specification of the model (Pakes et al., 2015). It also shares certain advantages with the sufficient statistics approach (Chetty, 2009) in that it is less prone to errors due to model misspecification.

We start by developing the identification in a one-period model. Later, Section 5 considers a number of alternative explanations for our findings. In particular, we investigate to what extent time preferences (Section 5.1)\(^{20}\), prepayment risk (Sections 5.2 and 5.3) and relational contracts between dealers and banks (Section 5.4) can account for our results.

### 4.1 One-period model

Denote the buy rate for consumer \(i\) by \(b_i\) and the states sales tax by \(\tau_i\). The car dealer has a linear markup revenue sharing agreement with the lender (see Section 3.2). This agreement is characterized by slope \(\alpha\) and intercept \(\beta\), both of which are lender-specific. The dealer makes an offer to the consumer consisting of the vehicle price \(p_i\) and the loan interest rate

\(^{20}\)Since auto loans are typically fully amortizing loans, the allocation of payments between finance charges and car charges does not affect the inter-temporal allocation of payments. Therefore, our argument in this section is also valid in a multi-period model.
$r_i \geq b_i$, taking the downpayment $d_i \geq 0$ as given. The consumer can react in three different ways. First, she can take the loan and buy the car. Second, she can decline the entire offer and revert to some outside option with expected value $\bar{u}_i$, e.g., she does not buy a car or does so at a different dealer. We assume that $\bar{u}_i < 0$ and $|\bar{u}_i| > d_i$. Third, she can buy the car but get a loan directly from a lender. We denote the consumer’s interest rate in this scenario by $r_{L} > 0$, which she believes to be distributed according to the continuously differentiable density function $g_i(r_{L})$, which has compact support. Moreover, her expected cost from shopping is $s_i \geq 0$. Our main assumption is that consumer $i$’s disutility from finance charges $x$ is given by $M_i(x)$ where $M_i(x) \in C^2$.

**Assumption 1.** Consumer utility is additive in the disutility of vehicle charges $p$ and the disutility of finance charges $M_i(x)$, where $M_i(x) \in C^2$.21

Importantly, we do not impose any further structure on this function, i.e., we allow the consumer utility to follow any twice continuously differentiable pattern. If $M_i(x) = x$ the consumer treats financial and vehicle costs identically. If $M_i(x) \neq x$ for some $x$, the consumer treats the two dimensions differently.

The dealer’s maximization problem is given by:

$$\begin{align*}
\max_{r_i, p_i} & \quad p_i + (\tau_i \cdot p_i - d_i) \cdot (r_i - b_i) \cdot \alpha - c_i + \beta \quad (1) \\
\text{s.t.} & \quad -\tau_i \cdot p_i - M_i((\tau_i \cdot p_i - d_i) \cdot r_i) \geq \bar{u}_i \\
& \quad - M_i((\tau_i \cdot p_i - d_i) \cdot r_i) \geq -\int M_i((\tau_i \cdot p_i - d_i) \cdot r_{L}) \cdot g_i(r_{L})dr_{L} - s_i, \\
& \quad r_i \geq b_i, p_i \geq 0
\end{align*}$$

To understand the basic intuition of our argument, first consider the case of $M_i(x) = x$. For this case, the optimal policy for the dealer is a corner solution. If the incentives provided by the bank are not too high (formally, $\alpha < \frac{1}{(1+b_i)\tau_i}$), the dealer sets the interest rate as low as possible ($r_i = b_i$). The reason is that the consumer dislikes loan and vehicle price markups equally, but the dealer prefers vehicle price markups because he does not share them with the lender.

**Proposition 1.** Suppose $M_i(x) = x$ and $\alpha < \frac{1}{(1+b_i)\tau_i}$. Then any solution to the maximization problem in (1) features $r_i^* = b_i$.

We delegate proofs to Appendix A. Proposition 1 has an important implication: if $M_i(x) = x$, the dealer does not mark up loans with $\alpha < \frac{1}{(1+b_i)\tau_i}$. This inequality holds for 97.5 percent of observations in our data, and so the substantial markups documented in Table 1 are incompatible with $M_i(x) = x$, i.e. equal consumer sensitivity to loan price and car price.

---

21For our argument we implicitly assume that the dealer is aware of the consumer’s utility function $M_i(\cdot)$. However, our estimation procedure does not rely on this assumption. If the dealer does not know $M_i(\cdot)$ but has a point belief about it, our estimates reveal this point belief.
Using the size of markups in the data, we can go further and quantify the extent to which finance and vehicle charges are treated differently. Suppose next that the perceptions of finance charges by consumer \( i \) are given by \( M_i(x) \).

**Proposition 2.** For any optimal offer \((r^*_i, p^*_i)\) with \( r^*_i > b_i, \ p^*_i > 0 \) one of the following two statements holds:

(i) At the optimum only the first constraint is binding and the optimal solution satisfies

\[
M'_i(r^*_i(p^*_i - d_i)) = \frac{\alpha \cdot \tau_i}{1 - \alpha \cdot \tau_i \cdot b_i}.
\]

(ii) Both constraints bind at the optimum and

\[
M'_i(r^*_i(p^*_i - d_i)) < \frac{\alpha \cdot \tau_i}{1 - \alpha \cdot \tau_i \cdot b_i}.
\]

Proposition 2 allows us to use the prices set by the dealer to infer consumers’ sensitivity to additional charges at the equilibrium contract. We observe all variables on the right hand side in (2), as well as the argument \( r^*_i \cdot (\tau_i p^*_i - d_i) \) of function \( M_i(\cdot) \). Therefore, the data provides individual-specific estimates of the marginal perception \( M'_i(\cdot) \) at the observed contract. Furthermore, \( M'_i(r^*_i \cdot (\tau_i p^*_i - d_i)) \) has a transparent economic interpretation: it is the marginal dis-utility, in dollars, that a consumer associates with an increase in the loan price. If \( M'_i(\cdot) \) is smaller than unity, for example, the consumer is less sensitive to an increase in the loan price than to an equal increase in the car price. Whether (i) or (ii) holds for any given consumer depends on unobservables. Hence, our sufficient statistic establishes an upper bound for \( M'_i(\cdot) \) at the observed contract.

Our estimate for \( M'_i(\cdot) \) is an upper bound because of the second constraint in the dealer’s problem. When this constraint binds, it is because the dealer cannot charge a higher interest rate without causing the consumer to obtain the loan directly from a lender, even though the consumer would otherwise agree to pay a higher interest rate in exchange for a lower vehicle price. Hence search costs determine our estimates for \( M'_i(\cdot) \) in these cases, which are greater than actual \( M'_i(\cdot) \). This logic holds because our model does not allow dealers to condition the price of the vehicle on the source of financing. If instead we allowed dealers to charge higher prices for vehicles financed directly, the second constraint in the dealer’s problem would never bind, and our estimates for \( M'_i(\cdot) \) would be tight.\(^{22}\)

Our \( M'(\cdot) \) estimates measure consumers’ marginal valuation of finance charges. However, each marginal valuation is estimated at the specific contract terms in the data. As a result, two individuals with the same finance charges and the same estimated \( M'(\cdot) \) may still differ in their valuation of their finance charges. Next we calculate bounds on the perceived value of total finance charges.

\(^{22}\)Anecdotally, dealers often do condition the price of the vehicle on the source of financing.
Proposition 3. Suppose $\left(r^*_i, p^*_i\right)$ with $r^*_i > b_i$ and $p^*_i > 0$ is an optimal offer.

(i) If $M'(x) \leq 1 \forall x$, it holds that

$$
(r_i^* - d_i) \cdot (r^*_i - b_i) - [M_i(r^*_i \cdot (r_i^* - d_i)) - M_i(b_i \cdot (r_i^* - d_i))] \\
\geq (\tau_i \cdot p_i^* - d_i) \cdot (r^*_i - b_i) \cdot [1 - \tau_i \cdot \alpha \cdot (1 + b_i)] = B^M_i
$$

(ii) If $M''(x) > 0 \forall x$, it holds that

$$
(\tau_i \cdot p_i^* - d_i) \cdot r_i^* - M_i(r_i^* \cdot (\tau_i \cdot p_i^* - d_i)) \geq (\tau_i \cdot p_i^* - d_i) \cdot r^*_i \cdot \left[1 - \frac{\alpha \cdot \tau_i}{1 - \alpha \cdot \tau_i b_i}\right] = B^O_i.
$$

Proposition 3 establishes two different lower bounds on the difference between actual and perceived finance charges. Both can be estimated in our data. Part (i) of Proposition 3 derives an individual-specific lower bound $B^M_i$ for the difference between actual and perceived finance charges due to markup. Intuitively, this result leverages the fact that dealers choose the markups seen in the data rather than setting them to zero. This behavior can only be optimal if consumers’ disutility from the markups is not too high.

Part (ii) of Proposition 3 derives a lower bound $B^O_i$ for the difference between actual and perceived total finance charges. For this result, we have to make the additional assumption that $M_i(\cdot)$ is convex. If $M_i(\cdot)$ is convex, our estimates of $M'_i(\cdot)$ provide a natural lower bound on $M_i(\cdot)$.

While the two estimated bounds are highly correlated and the underlying ideas are similar, we still consider them to be important complements. $B^M_i$ is derived under milder assumptions, but only considers finance charges due to markups. Convexity of $M_i(\cdot)$ is a stronger assumption but allows us to consider all finance charges. $B^O_i$ is also important from the dealer’s perspective as it characterizes the extent to which he can take profit from differences in price sensitivities (see also Section ??).

4.2 Estimation

We now turn to estimation of the measures of consumer price sensitivities given by Propositions 2 and 3. Note that we can only estimate $M'(\cdot)$ and $B^O$ for the approximately 77.8% of our sample with positive markup; there are several potential explanations for zero markups (e.g. the consumer has a good outside option or low loan search costs, or the consumer is indifferent between vehicle and finance charges) that we cannot disentangle. However, we can estimate $B^M_i$ for the whole sample as it is zero for observations with zero markup.

Our procedure is non-parametric and yields individual-specific estimates. For each contract in our data, we obtain a specific $M'_i(\cdot)$, $B^M_i$, and $B^O_i$ estimate by applying equation (2), (3), and (4), respectively. We restrict our sample to prime consumers with a credit score above
Table 3: Summary Statistics of Estimates

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>P10</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>P90</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M'_i$</td>
<td>0.86</td>
<td>0.77</td>
<td>0.80</td>
<td>0.86</td>
<td>0.91</td>
<td>0.95</td>
</tr>
<tr>
<td>$B^O$ ($)</td>
<td>380.12</td>
<td>105.71</td>
<td>186.81</td>
<td>324.33</td>
<td>510.73</td>
<td>721.56</td>
</tr>
<tr>
<td>$B^M$ ($)</td>
<td>96.16</td>
<td>0.00</td>
<td>16.56</td>
<td>72.09</td>
<td>145.21</td>
<td>228.07</td>
</tr>
</tbody>
</table>

Note: Selected summary statistics of consumers’ sensitivity to finance charges and the dollar amounts that are not taken account of when making a purchase. For $M'_i$ and $B^O$, the estimation is conditional on the mark-up being positive, while the estimates $B^M$ are derived for the full sample.

Figure 2: Histogram of Estimated Bounds

Note: The left panel shows $B^M$ the bound on the difference between the consumer valuation of finance charges and finance charges under equal elasticities. The right panel shows this difference only for the finance charges implied by the dealers decision $B^O$.

720 to ensure that our results are driven by default risk.\textsuperscript{23,24} Table 3 provides a summary of our estimates, which reveal a substantial wedge between car and loan price sensitivities. In particular, the mean $M'_i$ is only 0.86. This implies that, at the contract terms seen in our data, an extra dollar of finance charges is treated as worth approximately 86 cents. For consumers at the 10th percentile, this value drops to 77 cents.

Figure 2 shows the estimated distribution of our estimates with respect to the two lower bounds $B^M$ and $B^O$. Two points are worth noting. First, the estimated bounds are substantial. On average, consumers act as if perceived finance charges are $380 less than actual finance charges. Similarly, on average consumers act as if perceived finance charges due to markup are $96 less than actual finance charges due to markup. This estimate corresponds to

\textsuperscript{23}In a sample of prime auto loans in the CCP originated in 2014 and 2015, .4% had resulted in either repossession, chargeoff, or bankruptcy by three years after origination. $M'_i$ is positively correlated with credit scores in our restricted sample but negatively correlated with credit scores in our full sample. This suggests that, for consumers with low credit scores, dealers may be incentivized or forced to restrict markups, which is consistent with anecdotal evidence.

\textsuperscript{24}For one lender, payments were not well described by a linear contract in the loan amount. In an earlier version of the estimation this led to very low $M'_i$ estimates and, hence, very low sensitivity to finance charges. To ensure that our results are not driven by these specific observations, we drop this lender entirely.
approximately 20.3 percent of total finance charges, and about 19.2 percent of the markup. Second, there is significant heterogeneity in our estimates. The difference between actual and perceived finance charges at the 90th percentile is seven times larger than at the 10th percentile. An obvious question of interest is therefore how these estimates correlate with other observables in our data. We answer this question in Section 6.

On a more speculative note, our estimates also allow us to explore the average curvature of the function \( M_i(\cdot) \). Specifically, we can investigate the effect of higher finance charges on our estimates on \( M'(\cdot) \). Finance charges are obviously endogenous to \( M_i(\cdot) \). However, the buy rate offered by the lender is strongly correlated with finance charges, and – because our data includes virtually all the information the lender has – exogenous to any unobservable consumer characteristics such as the consumer’s utility function \( M'(\cdot) \). It is therefore a valid instrument. The results of the corresponding 2SLS regression are summarized in Table 9. The estimated shape suggests that \( M_i(\cdot) \) is on average a convex function. Hence, consumers treat finance charges more accurately if more is at stake. This finding is reminiscent of recent work studying attention allocation (Bordalo et al., 2012, 2013, Kőszegi and Szeidl, 2012). A common theme of this work is that consumers pay more attention to dimensions with more pronounced differences across products. In our setting with increasing finance charges the difference of the outside option (not taking out a loan) and signing the proposed loan is scaled up. The proposed attention based theories would therefore predict more focus on loan terms if the implied finance charges rise, which is indeed what we find.

5 Explanations for Differential Price Responsiveness

5.1 Time Structure

The previous section illustrated our main identification argument in a one-period model. It may appear intuitive that in a multi-period model consumers who are impatient or liquidity-constrained may be willing to pay higher interest rates, and that this could partly explain our results. However, this intuition is largely incorrect, because auto loans have fixed monthly payments that fully amortize. For a given down payment, there is a one-to-one mapping between total transaction costs and monthly payments. This mapping does not depend on the allocation of transaction costs between loan and vehicle costs, and so if down payments are independent of this allocation, then intertemporal preferences play no role in our results.\(^{25}\) Section A.1 shows this formally.

However, if a larger car price causes a larger down payment, then intertemporal preferences could affect our results. Suppose, for example, the down payment is a fixed percentage of the car price. In this case, impatient or credit-constrained consumers may prefer to pay

\(^{25}\)Down payments will be independent of the allocation of charges if, for example, the down payment is determined before the price is determined, is at the discretion of the consumer, or is zero.
more for the loan to reduce the car price and hence the down payment. Figure 7 shows the
distribution of down payments as a fraction of the car price. This figure, together with our
discussions with industry experts, reveal that down payments for prime consumers are
generally not determined in this way. Variation of the relative size of the down payment is large
and a marginal increase in the car price does not appear to increase the down payment.

Still, we break down our estimates of $M'_i(\cdot)$ and $B^0_i$ by the size of the down payment in
the appendix in Tables 10 and 11, respectively. We find a considerable wedge between car
and loan price elasticities across the entire distribution of down payments. In particular, the
wedge is substantial for zero and negative down payments, for which time preferences cannot
explain any of our estimates. The median $B^0_i$ for consumers in the 10th to 25th percentile
of the down-payment distribution is $630 while it is $563 for consumers between the 75th
and 90th percentile. Overall, these results show that intertemporal preferences and credit
constraints cannot explain the substantial dealer markups in our data.

To further address this point, in Section A.2 we recompute $M'_i(\cdot)$ under the (evidently
counterfactual) assumption that dealers require a down payment that is a fixed percentage
of the car price, conditional on various discount factors for the consumer.\textsuperscript{26} The results are
summarized in Table 16. Recall that we are restricting our sample to individuals with credit
scores above 720. Hence, the sample consists of individuals that are unlikely to be either
impatient or credit constrained and should therefore have reasonably high discount factors.
Our results show quantitatively that even if the down payment is a fixed fraction of the price,
our estimates barely change for discount factors above 0.90.

\subsection{Early Prepayment Risk}

Consumers often prepay auto loans, either by trading in their vehicle for a new one, by pre-
paying in cash, or (less commonly) through refinancing. The potential effects of prepayment
on our results depends on when it occurs. Contracts between dealers and lenders include a
“clawback” period, typically the first three to six months of the loan, during which the dealer
bears prepayment risk. If the consumer prepays during the clawback period, the dealer re-
funds the entire dealer reserve to the lender. We define “early” prepayment as prepayment
that occurs during the clawback period, and “late” prepayment as prepayment that occurs
after it.

First, we explore the role of early prepayment. Early prepayment is not uncommon; in a
sample of prime auto loans in the CCP, 5.9\% are prepaid within the first 120 days. To account
for this, we add to our baseline model an early prepayment probability $(1 - \gamma)$. Hence, the

\textsuperscript{26}In doing so we first compute the fraction of the price that is equal to the observed down payment. We then assume that the dealer
would require this fraction for any car price.
dealer’s maximization problem becomes:

$$\max_{r_i, p_i} \ p_i + (\tau_ip_i - d_i) \cdot (r_i - b_i) \cdot \gamma \cdot \alpha - c$$  \hspace{1cm} (5)$$

$$s.t. \quad - \tau_i \cdot p_i - \gamma \cdot M_i((\tau_ip_i - d_i) \cdot r_i) \geq -\bar{u}_i$$
$$- \gamma \cdot M_i((\tau_ip_i - d_i) \cdot r_i) \geq -\gamma \cdot M_i((\tau_ip_i - d_i) \cdot r_i^E) - s_i,$$
$$r_i, p_i \geq 0$$

This problem can be solved along the same lines as the one in Proposition 2. On the one hand, consumers are now less sensitive to interest rates so dealers have an additional incentive to mark up the loan. On the other hand, the dealer values markups less because he may not keep the dealer reserve. The following corollary shows that the latter effect dominates. As a consequence, our estimates for $M'(\cdot)$ are upper bounds because they do not account for early prepayment risk.

**Corollary 1.** Suppose that consumer $i$ repays the loan early with probability $1 - \gamma$. Then, the statements from Proposition 2 hold true with the only change that:

$$M'_i(r_i^* \cdot (\tau_i \cdot p_i^* - d_i)) \leq \frac{\alpha \cdot \tau_i}{1 - \gamma \cdot \tau_i \cdot \alpha \cdot b_i}.$$  

### 5.3 Late Prepayment Risk

Lenders bear all prepayment risk after the end of the clawback period. If borrowers care about late prepayment risk, and if dealers can screen “late” prepayers from “early” prepayers, then the logic outlined in 5.2 does not hold, and borrowers may appear less sensitive to finance charges because they expect not to pay them. Late prepayment risk is substantial and so this is an important concern; 27% of prime auto loans (with term greater than two years) in the CCP are prepaid after 120 days but before two years.

There are two plausible kinds of late prepayment risk we need to examine. The first is risk that is observable to both the borrower and the dealer. To create a proxy for this risk, we run a logit regression predicting late prepayment in the CCP, using credit score, log loan amount, loan length, and state fixed effects. Coefficients are reported in Table 13. This is clearly a limited subset of the information available to both the borrower and the dealer, and yet it is remarkably predictive. 18.2% of those those in the bottom decile of predicted risk prepay late, while 42.9% in the top decile do.

While these variables are highly predictive of late prepayment and readily observable to lenders, recall lenders do not condition $\alpha$ on them. This is the first indication that borrowers with predictably higher prepayment risk do not pay higher markups.

Still, because lenders do not condition $\alpha$ on this information set, borrowers with predictably higher prepayment risk may agree to higher markups because the effective cost is lower. Tables 13, 12, and 14 provide percentiles of markup, $M'_i(\cdot)$ and $B^M_i$, respectively,
conditional on percentiles of predicted late prepayment risk. Table ?? provides estimates from a regression that controls for a number of consumer characteristics. If anything, the correlation between observable late prepayment risk and our elasticity wedge estimates is negative. Hence, we estimate higher $M_i'(\cdot)$'s and lower $B_i^M$'s for consumers with higher prepayment risk. This is strong evidence against the possibility that observable late prepayment risk drives our results.

One potential explanation for this result is that dealers cannot distinguish late prepayment risk from early prepayment risk. Indeed, the correlation between our measures of early and late prepayment risk in the CCP is .66. In this case, our finding that higher prepayment risk predicts higher $M_i'(\cdot)$ estimates suggests that the early prepayment result in Section 5.2 quantitatively dominates the opposing effect from predictable late prepayment. In this case, ignoring prepayment biases against our results in Section 4.2.

Another possibility with substantial empirical support is that, instead of considering the cost of a financed vehicle through the life of a loan, consumers instead care about monthly payments (Argyle et al. (2019)). Because there is a one-to-one mapping between monthly payment and the total cost of a loan if it is paid on schedule, this model is equivalent to ours if we ignore prepayment risk.

We still need to consider prepayment risk that is observable to the borrower but not to the dealer. Because we observe prepayment but not markup in the CCP, and markup but not prepayment in the supervisory data, we cannot test for this explicitly. However, recall that markup revenue sharing agreements between dealers and lenders are linear. This means that the dealer receives a fixed payment for every marginal dollar in finance charges he generates, assuming the loan is paid on schedule and not prepaid. If there were adverse selection on late prepayment risk in markups, then lender revenue would be concave in markups and this contract would not be optimal.

### 5.4 Relational Contracts between Banks and Dealers

Lenders and dealers interact repeatedly, so dealers may mark up loans to increase lenders’ profits in exchange for favorable treatment on later deals, such as financing for a subprime consumer who might otherwise not qualify for a loan. This argument is particularly plausible in light of our finding in Section 5.3 that markups are higher for consumers with lower prepayment risk. Fortunately, our AutoCount data includes important characteristics of dealer-lender relationships. This enables us to directly check the extent to which markups are higher in dealer-lender relationships that are more intense or that have lasted longer.

We proxy the intensity of the relationship between dealer $i$ and lender $j$ by the fraction $f$ of loans intermediated by dealer $i$ that are from lender $j$. Table 15 shows the results from an OLS regression where markups are the dependent variable and $f$ is the independent variable of interest. Importantly, the regression includes dealer fixed effects, so that the estimates
reveal if dealers choose higher markups when working with preferred lenders. We find that the intensity of a dealer-lender relationship has virtually no predictive effect for the markup. Markups for lenders that finance more than 20% of a dealer’s sales are only 3 basis points higher than markups for lenders that finance less than one percent of a dealer’s sales. Recall that the average markup in our data is 108 basis points. Even for lenders that finance less than five loans a year for a given dealer, we find that markups are on average 106 basis points. It is therefore implausible that dealers’ markup decisions arise from their relationships with lenders.

6 Heterogeneity Analysis

In this section we explore how $B^M_i$ and $B^O_i$ covary with other observables. While $B^M_i$ is estimated on the full sample, recall that $B^O_i$ is estimated only on the set of loans with positive markups.

This exercise is particularly useful because these bounds and much of our data are measured at the individual level. However, one potential caveat of this exercise is that the correlation between these bounds and the objects they bound may vary across segments of the population in a way that is unobservable but affects our results. In particular, the degree to which a bound $B_i$ is binding may depend on the curvature of $M_i(\cdot)$ and $i$’s unobservable search cost $s_i$. A regression of the bounds on observables may therefore suffer from omitted variable bias if the tightness of the bounds varies with observables. We now note assumptions under which this problem disappears. We then make these assumptions, and perform the empirical exercise.

As discussed in Section 4 the estimates on $M'(\cdot)$ are tight if dealers can choose the car price contingent on whether the consumer finances the car at their place or not. If we additionally impose structure on the functional form of $M(\cdot)$, we can ensure that the estimates on $B^O_i$ correspond to the ordering of the individual wedges between elasticities. For this purpose, we therefore make the following assumption for everything that follows:

**Proposition 4.** Suppose that for every consumer $i$ that $M_i(\cdot)$ can be written as $M_i(x) = x - h(\rho_i)\ln(x)$ for some positive, increasing, and continuously differentiable function $h(\cdot)$. Then $B^O_i \geq B^O_j \iff \rho_i \geq \rho_j \forall i, j$.

If $M(\cdot)$ can be represented by the family of functions described in Proposition 4, the parameter $\rho_i$ specifies the magnitude of the wedge between elasticities for consumer $i$, where larger $\rho$’s correspond to a larger wedge. Applying this parametric form, the ordering of the bounds $B^O_i$ induces an ordering on the space of all $\rho_i$’s. Hence, whenever $B^O_i \geq B^O_j$ for two

---

27We think that it is reasonable to assume that quoted car prices are contingent on the consumer taking the loan. In fact, in many conversations about this project people have pointed out to us that the dealer expected them to finance the car at the dealership with the quoted price.
consumers $i$ and $j$ it also holds that $\rho_i \geq \rho_j$. As a direct consequence, we can use the terms $B_i^O$ to conduct a heterogeneity analysis on the magnitude of the wedge in elasticities.

Figure 3 depicts an example of the functional form imposed in Proposition 4. While the additional structure imposed on $M(\cdot)$ is a caveat, the corresponding functional form seems plausible. Essentially it requires that the percentage of under-estimated finance charges is large for small amounts of financing. However, the marginal mistake that a consumer makes when increasing the loan, converges to zero for large amounts of financing. Hence, consumers reduce their mistake relative to the loan amount as the loan amount increases. This notion is reminiscent of the idea that consumers have scarce cognitive resources and allocate more of these to matters that are more important.  

6.1 Individual Characteristics and Regional Variation

In Column (1) and (3) of Table 4, we correlate our estimates for the bounds $B_i^O$ and $B_i^M$ with credit scores and income. We find that consumers with lower monthly income and lower credit score tend to exhibit a larger wedge between elasticities. Hence, our analysis suggests that consumers who face more difficulties to pay their loan rate on time are also more prone to exhibiting a wedge in their assessment of charges on the car and the loan.

28For papers that explore the behavioral implications of limited attention and the determinants of attention allocation see ?, Bordalo et al. (2012), Bordalo et al. (2013), and Mackowiak et al. (2018)
Table 4: Estimated Effect of Transaction and County Characteristics on Estimated Bounds

<table>
<thead>
<tr>
<th></th>
<th>$B_{jO}$</th>
<th>$B_{jM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Log Monthly Income</td>
<td>-9.277***</td>
<td>-8.688***</td>
</tr>
<tr>
<td></td>
<td>(0.262)</td>
<td>(0.261)</td>
</tr>
<tr>
<td>Credit Score, 100 points</td>
<td>-29.99***</td>
<td>-30.02***</td>
</tr>
<tr>
<td></td>
<td>(0.386)</td>
<td>(0.386)</td>
</tr>
<tr>
<td>Mileage, Tens of Thousands</td>
<td>5.054***</td>
<td>5.014***</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.081)</td>
</tr>
<tr>
<td>New Car</td>
<td>-4.931***</td>
<td>-4.915***</td>
</tr>
<tr>
<td></td>
<td>(0.411)</td>
<td>(0.411)</td>
</tr>
<tr>
<td>Log Loan Amount</td>
<td>393.8***</td>
<td>393.4***</td>
</tr>
<tr>
<td></td>
<td>(0.951)</td>
<td>(0.950)</td>
</tr>
<tr>
<td>Average Years of Education</td>
<td>-3.734***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.336)</td>
<td></td>
</tr>
<tr>
<td>Fraction with Internet Access</td>
<td>-8.425***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.469)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Table shows estimates from an OLS regression. Columns (1) and (3) control for transaction-specific observables. Columns (2) and (4) include county-level data from the ACS. Lender, model, and state fixed effects also included, but not shown. Standard errors clustered at zip code level.
Next we turn to the regional variation of our estimates. As a starting point Figures 8 and 9 in Appendix B show how our estimates for $B^O_i$ and $B^M_i$ differ across counties. To eliminate bank specific variation, we depict the residual variation in the bounds after taking out bank fixed effects. Lighter colors on the map, therefore, indicate that car buyers in this county exhibit larger wedges in elasticities compared to the inhabitants of other counties that acquire a loan from the same bank. The depicted regional variation of our estimates is systematically related to the average education within a county as well as its digital infrastructure. To delineate such patterns, column (3) and (6) add data from the American Community Survey (ACS) on the county level to the regression. To proxy education, we use the percentage of inhabitants in a county that has obtained a college degree.\textsuperscript{29} The quality of the digital infrastructure is proxied for by the speed and availability of internet access on a scale from 1 to 5, where a score of 5 represents counties in which internet access is fast and readily available, while a score of 1 means that internet access is either not available or slow. We find that education as well as digital infrastructure are negatively correlated with our estimates. Hence, lower educated consumers seem to be more prone to disregard finance charges as well as consumers that live in areas with low internet connectivity. Finally, we can also breakdown our estimates by the car model that a consumer purchases and finances. Table 19 depicts our average estimates for $B^M_i$ and $M'_i(\cdot)$ for the 20 most common models in our data, showing that there is some variation across the purchased models and makes.

7 Full Model Setup

The following section further explores how the wedge between consumers’ finance charges and price elasticities play out in a competitive environment. In particular, we are interested to quantify the effect of the wedge on consumer behavior and pricing strategies of dealers. We will then use the derived estimates to explore how two kinds of counterfactual experiments—consumers treat all charges identically and dealers are obliged to pass through the buy rate—change market outcomes. The following model captures the essential parts of the institutional details faced by lenders, car dealers, and consumers.

7.1 Consumer Utility and Dealer Market Shares

Consumers engage in a nested choice, deciding first on a dealer and then given the make selection of this dealer, which make to buy. There is a set of markets $\mathcal{M}$ and for every market $m \in \mathcal{M}$ a set of active dealers $\mathcal{D}_m$. Every dealer $d \in \mathcal{D}_m$ offers a specific set of makes $\mathcal{J}_d$, which we take to be exogenously given. Furthermore, each dealer cooperates with a set of banks $\mathcal{B}_d$ that extend loans to her customers. Dealers engage in differentiated product

\textsuperscript{29}The ACS also provides the percentage of inhabitants that obtained a high school degree. Our results also hold if we use this proxy for education or both proxies simultaneously.
Bertrand competition, setting both interest rates and car prices optimally, given consumer tastes, competing dealers’ behavior, their marginal cost of lending, and the revenue sharing agreement with banks.

To specify consumer \( i \)'s indirect utility function, we denote her travel costs to dealer \( d \) by \( g(d, i) \). Furthermore, consumer \( i \) has marginal utility of money \( \gamma_i \), which we allow to depend on her demographics such as her income. Importantly, consumers in our setting may disregard certain parts of the overall transaction price. To translate actual charges into perceived charges we parametrize the utility function of the consumers \( M_i(\cdot) \). In particular, we assume that

\[
M_i(x) = x - \rho_i \cdot \log(x).
\]

Hence, \( \rho_i \) is an individual specific parameter describing how elastic consumer \( i \)'s utility is to changes in finance charges. If \( \rho_i \) becomes large, the consumer utility becomes less responsive to changes in finance charges. If, on the contrary, \( \rho_i \) is close to zero there is no wedge between price and finance charges elasticity. Our estimates of \( M_i'(\cdot) \) from Section 4 allow us to back out \( \rho_i \) immediately. Summary statistics on the corresponding estimates can be found in Table 20. Given the observed car price \( p_{j,d} \) and interest rate \( r_{j,d} \) for a car \( j \) at dealer \( d \), the perceived overall transaction price \( \tilde{p}_{j,d} \) is then given by:

\[
\tilde{p}_{j,d} = p_{j,d} + p_{j,d} \cdot r_{d,d} - \rho_i \cdot \log(p_{j,d} \cdot r_{d,d}).
\]

Next to the travel distance and the car price, consumers’ indirect utility depends on observable loan and car attributes which we collect in the vector \( x_j \), whose relative attractiveness is captured by vector \( \beta \). Furthermore, there are two types of unobservable taste shocks in the model. First, the relative attractiveness of purchasing make \( j \) at dealer \( d \) is captured by a dealer-make specific scalar variable \( \xi_{j,d} \). Second, every consumer’s choice is affected by her individual specific taste shock \( \epsilon_{i,j,d} \) for a dealer-make combination, which are i.i.d type-1 extreme value distributed.

\[
u_{i,j,d} = g(d, i) - \gamma_i \cdot \tilde{p}_{j,d} + \beta \cdot x_j + \xi_{j,d} + \epsilon_{i,j,d}
\]

where

\[
\gamma_i = \beta_0 + \beta_y \cdot y_i + \nu_i \quad \text{where} \quad \nu_i \sim \mathcal{N}(0, \sigma) \quad (7)
\]

In practice, we collapse \( g(d, i) \) at the market level as a weighted average over the travel distances to dealer \( d \) from different zip-code centroids so that we have \( g(d) \). As is customary in the literature we refer to mean utility as \( \delta_{j,d} = g(d) - \alpha_c \cdot \tilde{p}_{j,d} + \beta \cdot x_j + \xi_{j,d} \). We refer to the vector of those mean utilities for an entire market as \( \delta_m \).

We now describe the customers’ choices, which determine the overall market share of a dealer-make combination. We are integrating over the unobserved customer type \( \nu \). Exploit-
ing the logit structure of our model, the share of make \( j \) at dealer \( d \) is given by:

\[
S_{jd}^{m}(\mathbf{p}_d, \mathbf{r}_d; \mathbf{p}_{-d}, \mathbf{r}_{-d}) = \frac{\int \exp(g(i, k) - \gamma_i \cdot \tilde{p}_{jd} + \beta \cdot \mathbf{x}_j + \xi_{jd})}{\sum_{d \in D_m} \sum_{k \in J_d} \exp(g(i, k) - \gamma_i \cdot \tilde{p}_{kd} + \beta \cdot \mathbf{x}_k + \xi_{kd})} f(\nu) \cdot d\nu
\]

7.2 Dealer Maximization Problem

Dealers decide on prices and interest rates for all models that they sell. The buy rates that dealers face are determined in equilibrium by the auction in which the set of bidders \( K_d \) for dealer \( d \) participates. The expected buy rate is a function of all the prices and interest that are charged in the market since banks will internalize that the buy rate (and the implied markup) will affect the dealers chance of selling the car. The optimal considerations of the bank are described in more detail in subsection 7.3. Given a buy rate and a downpayment \( \text{down}_{jd} \), dealer \( d' \)'s maximization problem is

\[
\max_{\mathbf{p}_d, \mathbf{r}_d} \sum_{j \in J_d} S_{jd}^{m}(\mathbf{p}_d, \mathbf{r}_d; \mathbf{p}_{-d}, \mathbf{r}_{-d}) \cdot \left( p_{jd} + \alpha_{jd} \cdot (p_{jd} - \text{down}_{jd}) \cdot (r_{jd} - b_{jd}) - c_{jd} \right) \tag{8}
\]

Denote the dealer’s profit exclusive the random component under bank \( l \)'s terms and the optimal price interest rate combination as \( \pi_{ld}^{l}(\mathbf{p}_d(b_l), \mathbf{r}_d(b_l); \mathbf{p}_{-d}, \mathbf{r}_{-d}) \).

7.3 Bank Maximization Problem

In modeling the banks maximization problem, we make a number of simplifying assumptions. First, we assume that there is one auction for each make-dealer combination. Hence, each of the banks in \( \kappa_d \) associated with dealer \( d \) bids once for all the sales of make \( j \). Second, banks are playing a symmetric game, i.e, all \( \alpha \)'s are identical for a dealer. Inspection of the data reveals that there is indeed relatively little variation in \( \alpha \) and allowing for heterogeneity would complicate the analysis considerably. Third, banks take all buy rates arising in auctions for other makes as given when making their bids for a particular contract.\(^{30}\)

**Assumption 2.** (i) All lenders \( k \in K_d \) are drawing from the same cost distribution \( r_k^{b} \sim F \) and have the same contractual arrangement \( \alpha, \beta \) with a given dealer \( d \). (ii) In each auction \( (j,d) \) bidders take the buy-rates arising in auctions for other makes as given.

The objective function is composed of the market share of the model for which the banks are bidding, the loan amount \( p_{jd} - \text{down}_{jd} \), the probability of winning the auction, and the

\(^{30}\)This simplification is reminiscent of the Nash-in-Nash approach that is frequently used in the analysis of bilateral oligopoly (see Horn and Wolinsky, 1988, Gaynor and Town, 2011, Collard-Wexler et al., 2014). In this literature, the bargains between a particular upstream supplier and a downstream firm take the outcome of all other bargains as given. Similarly, we assume that, when bidding for a particular contract, a bank takes all buy rates in other auctions as given.
profit margin, which is composed of a share of the markup and a direct payment. To estimate
the model, we extend the result of Guerre et al. (2000) and show that unobserved cost can be
expressed in terms of already known (after estimating the demand model) or observed objects.
Banks anticipate that they will issue the loan for a particular transaction if their buy rate is
lower than the offers of all opponents. Let \( N \) be the number of banks bidding for a contract
and \( r^j \sim F(.) \) the distribution of wholesale interest rates \( r \), which characterize the banks’
cost of financing a loan. Banks will anticipate the likelihood of a sale, which is given by
the downstream marketshare of the car, and depends on \( b \) through the dealer’s downstream
decision on all prices and interest rates. Bank \( k \)’s objective function when bidding for a
particular contract is therefore given by:

\[
\max_{b^k_{jd}} s_{jd}^m(p_d(b^k_{jd}), r_d(b^k_{jd}); p_{-d}, r_{-d}) \cdot \left( p_{jd}(b^k_{jd}) - \text{down}_{jd} \right) \cdot 
\left( 1 - F(\beta_{jd}^{-1}(b^k_{jd})) \right)^{N_{jd} - 1} \cdot \left[ (1 - \alpha_{jd}) \cdot m_{jd}(b^k_{jd}) + b^k_{jd} - r^k_{jd} \right]. \tag{9}
\]

For notational convenience we simply denote this function \( s_{jd}(b) \) here. For the derivation,
define \( \psi_{jd}(b) = s_{jd}^m(p_d(b_{jd}), r_d(b_{jd}); p_{-d}, r_{-d}) \cdot (p_{jd}(b) - \text{down}_{jd}) \).
Lastly, banks will anticipate that dealers mark up a specific model by \( m_{jd}(b) \). The markup function depends
on observable attributes and unobservable attributes of customers that purchase model \( j \)
as well as the other prices and interest rates in the market. Given our demand model, we can
simulate \( \hat{m}_{jd} \). We are applying the key inside from Guerre et al. (2000) that there is a relation-
ship between the observed distribution of buy rates and the unobserved distribution of
costs. In particular, \( G(b) = F(\beta^{-1}(b)) \) \( \forall b \Rightarrow g(b) = \frac{1}{\beta'(r)} \cdot f(r) \) \( \forall b \). This insight allows us to
substitute all unknown terms in the banks first order condition with either observed or already
estimated objects (based on the demand model and dealer pricing incentives given demand)
(see section A.2 for details).

**Proposition 5.** Banks cost can be recovered as:

\[
r^k_{jd} = (1 - \alpha_{jd}) \cdot m_{jd}(b^k_{jd}) + b^k_{jd} + \frac{1 + (1 - \alpha) \cdot m'_{jd}(b^k_{jd})}{\psi'(b^k_{jd}) - (N_{jd} - 1) \cdot g(b^k_{jd}) \cdot (1 - G(b^k_{jd}))^{-1}}. \tag{10}
\]

All expressions on the right hand side are observed or can be constructed from the demand
side estimates. We can therefore back out the cost distribution of banks in this market.

### 7.4 Discussion of Model Specification

Two of our modeling assumptions deserve discussion. First, the model treats the market as
a posted price market and thereby neglects the price dispersion that is due to haggling over
loan terms and vehicle prices. Retail markets for financial products are oftentimes analyzed
through the lens of search models, since they oftentimes feature large price dispersion despite being relatively homogeneous (Allen et al., 2013). In our case, we are studying a bundle of a retail financial product and an extremely differentiated durable. The latter requires us to account for unobservable product-specific components such as make cache. For this reason we have opted for a posted price demand model, which provides a well-established framework to deal with unobserved heterogeneity, at the expense of not being able to speak to the within product dispersion of prices, which exists in this market. Second, we assume that buyers never obtain loans from outside lenders. However, our assumption of no outside financing is in line with our observation from the consumer credit panel (Appendix section B.1), which shows that extremely few people obtain outside financing.

8 Estimation

For the estimation, we use data from 30 states, which amounts to 1134 counties. We consider each of such counties to be one market and estimate demand for the seventy most popular models, which account for the large majority of total sales. The remaining models are assigned to the outside good. We restrict our estimation to counties that have less than 45 dealers and in which at least one of the 70 most popular models is sold. After these data restrictions we are left with 917.0 markets each of which includes on average 6.916 dealers and a median of 5.0 dealers.

As described in Section 3, the AutoCount data covers the entire market for automobiles in the US while the supervisory data only covers a fraction of the market. We therefore, define the total market as all car purchases observed in the AutoCount data. Since the data are constructed from records of the Department of Motor Vehicles, they should be close to comprehensive. However, the AutoCount data lacks information on the dealer reserves and buy rates. To estimate the model we therefore impute buy rates in the AutoCount data from a saturated model that is informed by the supervisory data ($R^2 \geq 0.86$). While this procedure introduces some measurement error, we do not see this approach as a large shortcoming. In fact, the procedure is consistent with the timing of the sales process because dealers often offer loan terms before they have procured the loan from a lender, which means that they predict the buy rate themselves. Later on we will use the additional first order condition from dealers interest rate choice to recover the dealer profit shares $\alpha$. The $\alpha$ that we recover in this way then provides a robustness check of the imputation procedure as we can compare it to the observed values in the supervisory data.

---

31Empirical models of inter-modal competition, are, for example provided in Allen et al. (2014), where consumers obtain loan quotes from their home bank or run an auction among non-home banks. Another example is Salz (2017), where buyers can both search bilaterally or through brokers.

32The included states are Alabama, Arkansas, California, Colorado, Connecticut, Florida, Georgia, Idaho, Illinois, Indiana, Kentucky, Maine, Maryland, Michigan, Mississippi, Missouri, Montana, Nebraska, New Jersey, New Mexico, North Carolina, North Dakota, Ohio, South Carolina, Texas, Utah, Vermont, Virginia, Washington, Wisconsin
The estimation of the model further requires a proxy for the travel distance to each dealer. To acquire such a measure, we subdivide each county into zip-code tabulation areas and then query the travel distance from the centroid of each of those areas to the dealer using the Google distance matrix API. The overall travel distance to a dealer within a county is the population-weighted sum of the travel-distances from each of those centroids.

The price of the vehicle is potentially endogenous to the unobserved dealer-model specific demand shock $\xi_{jd}$. Failure to account for this potential endogeneity will result in biased price coefficients. We therefore interact $\xi_{jd}$ with the following set of instruments: the average miles per hour of other models at the same dealer, the vehicle length of other models at the same dealer, the mileage of other models at the same dealer, the average travel distance to other dealer in the same market, the buy-rate, and the average price of the same model in other markets.\(^33\)

\[
\delta^{h+1}_m = \delta^h_m + \log(s_m) - \log \left( \hat{s}(\delta^h, \sigma) \right) \tag{11}
\]

Following Berry et al. (1995) we use the observed aggregate market shares for a specific model-dealer combination, $s_m \ \forall \ m$ to compute a contraction mapping that recovers the mean utilities.\(^34\) We then regress those mean utilities on product attributes to uncover the linear parameters of the mapping (Nevo, 2000). Using $\delta_{jd}$ we can compute $\xi_{jd} = \delta_{jd} - \beta_x \cdot x$. Equation 12 provides the moment condition for estimation.

\[
G(\theta) = \sum_{d \in D_d} \sum_{j \in J_d} \xi_{jd}(\theta) \cdot z_{jd} \tag{12}
\]

9 Estimation Results

9.1 Demand Results

Table 5 shows the coefficients that we obtain from the demand model along with standard errors.\(^35\) The price interaction terms show that buyers with lower income are more price elastic. The coefficient on the price-income interaction is estimated somewhat imprecisely. Most other coefficients are estimated precisely and their sign lines up with intuition and what the literature has found previously. The coefficient on horsepower and vehicle length is positive but for both decreasing in income. The coefficient on miles per gallon, on the other hand, is negative but increasing in income, suggesting that higher income households value fuel efficient cars more. Consumers dislike travel distance to the dealer. We know discuss the size of the elasticities that are implied by those coefficients.

\(^33\)Except for the buy-rate and the distance, these instruments are standard in the demand estimation literature (Nevo, 2000).
\(^34\)Unlike the nested logit model in Berry (1994), the random coefficient model does not allow for an analytic inversion.
\(^35\)Table 21 further shows the demand coefficients associated with the different makes and their estimated standard deviation.
Table 5: Demand Model Coefficients

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>COEFFICIENT</th>
<th>STANDARD ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>-0.0772</td>
<td>0.00018</td>
</tr>
<tr>
<td>Horsepower</td>
<td>0.1765</td>
<td>0.00161</td>
</tr>
<tr>
<td>Vehicle Length</td>
<td>0.7627</td>
<td>0.00898</td>
</tr>
<tr>
<td>MPG</td>
<td>-1.5557</td>
<td>0.04973</td>
</tr>
<tr>
<td>Distance to Dealer</td>
<td>-0.162</td>
<td>0.00156</td>
</tr>
<tr>
<td>Price × Income</td>
<td>0.8213</td>
<td>0.02629</td>
</tr>
<tr>
<td>Horsepower × Income</td>
<td>-2.25044</td>
<td>0.24702</td>
</tr>
<tr>
<td>Income × MPG</td>
<td>15.6446</td>
<td>6.62776</td>
</tr>
<tr>
<td>Income × Car Length</td>
<td>-7.6912</td>
<td>1.34387</td>
</tr>
</tbody>
</table>

Note: The table shows the main coefficient estimates from the model. Travel time is in log-minutes and prices are in dollars. Income is measured in units of standard deviations.

Our model estimates consumers’ price elasticities for the overall transaction price including the car price and all finance charges. Over all market-dealer-make combinations we obtain about 35000 different elasticities, whose average is -3.73.

The own-price elasticities of the fifteen different makes that are included in the demand model vary from –6.2 for GMC to –2.9 for Hyundai (see Table 21 in Appendix B). The absolute value of our total price elasticities are close to what the literature has previously documented. Nurski and Verboven (2016), for example, find an average price elasticity of –3.14 for the Belgian market and Murry (2017) an average own-price elasticity of –4.9 for the US market. Note that is intuitively plausible that we estimate the demand to be slightly less elastic. In particular, our estimates refer to the overall price while the existing papers consider the car price only. as increases in the car price are accompanied by increases in loan charges, leaving aside loan charges should therefore lead to more negative estimates for demand elasticities. At an average of -0.16 consumers appear not to be too sensitive to changes in distance to different dealer locations.36

9.2 Dealer Cost

In order to conduct the counterfactual experiments, we use the first order conditions from a dealer’s pricing problem to recover their costs. In this process, we also estimate \( \alpha_{jd} \) that determines the contractually specified transfer (as a function of the markup) that dealer \( d \) receives from a bank when selling a car of make \( j \). A dealer maximizes her profits with regard to both interest rate and car price. We therefore obtain a set of \( 2 \times J_d \) first order

36Murry (2017) also documents distance elasticities using the distance in miles instead of average travel times as a measure for the dis-utility associated with traveling to a dealership. He finds that buyers are more elastic to distance with estimated elasticities ranging from –1.1 to –1.8. This might be a result of a different subset of markets that Murry (2017) focuses on.
Typically, it is possible to obtain a closed form matrix expression for the inversion to recover the costs. However, in our case, such a closed form does not exist for the joint set of costs and unobserved $c_{j,d}$-terms. We therefore conduct a non-linear search over this joint set to determine the solution of the first order conditions. The cost estimates that we obtain from this procedure are shown in Table 6. We estimate average cost of 17.6, ranging from 7.6 in the 10th percentile to 27.6 in the 90th percentile. The implied average Lerner index is 0.294.

Table 6: Lerner Index, and Cost, Summary Statistics

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>MEAN</th>
<th>P10</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>P90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lerner Index</td>
<td>0.294</td>
<td>0.102</td>
<td>0.147</td>
<td>0.228</td>
<td>0.383</td>
<td>0.606</td>
</tr>
<tr>
<td>Cost ($ 1000)</td>
<td>17.6</td>
<td>7.6</td>
<td>11.9</td>
<td>17.5</td>
<td>23.6</td>
<td>27.6</td>
</tr>
</tbody>
</table>

Note: The table shows summary statistics for the Lerner index across all estimated markets. In each market we weight the index according to the market shares of the respective model.

We use the second first order condition to back out the revenue sharing agreement $\alpha_{j,d}$. Since we also directly observe those agreements in the supervisory data the inferred values will serve as a validation of the model. From the inversion of the dealer FOC we recover an average revenue split of 0.68, very close to the 0.66 observed in the supervisory data.

9.3 Lender Cost Estimates

We use Equation 20 to recover the lenders cost. On average, we estimate cost as 0.072. Lenders are competing with other lenders that have an established relationship with the dealer. On average, a dealer conduct business with 4.35 and the median number of lenders is 4.0. If we condition only on lenders that have a significant market share with the dealer those numbers drop to 3.35 and 3.0. While the average buy-rate is, at 0.091, above cost the

\[^{37}\text{Remember that the dealer and lender need a revenue sharing agreement to split the additional surplus that the dealer markup generates.}\]
Figure 4: CDF’s of Interest rates, buy rates, and lender costs.

Note: This graph shows the CDF’s of the distribution of APR, buy rate, and lender cost.

The fact that for higher quantiles is lies above emphasizes that lenders expect dealers to mark up loans so that they can cover their cost. Table 7 shows means and quantiles of the lender cost distribution and compares them to the same summary statistics in the distribution of buy-rates and final interest rates. The lender cost distribution stochastically dominates the APR distribution over the entire domain. The same is not true for the comparison of lender cost and buy-rates.

Table 7: Lerner Index, and Cost, Summary Statistics

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>MEAN</th>
<th>P10</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>P90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lender Cost</td>
<td>0.072</td>
<td>0.026</td>
<td>0.044</td>
<td>0.063</td>
<td>0.08</td>
<td>0.148</td>
</tr>
<tr>
<td>Buy Rate</td>
<td>0.091</td>
<td>0.062</td>
<td>0.07</td>
<td>0.082</td>
<td>0.101</td>
<td>0.132</td>
</tr>
<tr>
<td>APR</td>
<td>0.139</td>
<td>0.086</td>
<td>0.106</td>
<td>0.13</td>
<td>0.162</td>
<td>0.204</td>
</tr>
</tbody>
</table>

Note: The table shows summary statistics for the Lerner index across all estimated markets. In each market we weight the index according to the market shares of the respective model.

10 Counterfactual Experiments

We present results from two different counterfactual experiments—one in which consumers treat all charges identically and one in which dealers are obliged to pass through the interest rate that the lender sets. We start by investigating the market equilibrium under the assumption that buyers are equally elastic to finance charges and vehicle charges. We refer to this as the No Wedge counterfactual. While this counterfactual has no direct policy relevance, it provides a useful benchmark of the overall effect of the differential price sensitivities in the
market as well as a comparison point against which the changes in the second counterfactual are to be interpreted. Our main outcome measures will be the effect on prices as well as the effect on consumer surplus and producer surplus.

10.1 Equal Sensitivities Counterfactual

To simplify the presentation of our results and to make interest rates comparable across contracts with different term length, we compute the total finance charges and show the implied two-period interest rate. For this counterfactual we hold buy-rates fixed.

Table 8: Overview Counterfactual Results

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Baseline</th>
<th>No Wedge</th>
<th>Δ%</th>
<th>No Dealer Discretion</th>
<th>Δ%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Total Price (p · (1 + r)) ($)</td>
<td>27181.0</td>
<td>26874.1</td>
<td>-1.13</td>
<td>26906.4</td>
<td>-1.29</td>
</tr>
<tr>
<td>Average Car Price ($)</td>
<td>24139.0</td>
<td>26683.9</td>
<td>10.54</td>
<td>23897.0</td>
<td>-1.26</td>
</tr>
<tr>
<td>Average APR (two period)</td>
<td>0.134</td>
<td>0.017</td>
<td>-87.14</td>
<td>0.12</td>
<td>-13.58</td>
</tr>
<tr>
<td>Total Cons. Surplus (p̂) ($Billion)</td>
<td>36.5</td>
<td>35.95</td>
<td>-1.38</td>
<td>38.0</td>
<td>4.12</td>
</tr>
<tr>
<td>Total Cons. Surplus (ρ = 0) ($Billion)</td>
<td>32.3</td>
<td>35.75</td>
<td>10.63</td>
<td>34.0</td>
<td>5.14</td>
</tr>
<tr>
<td>Total Dealer Profits ($Billion)</td>
<td>2.755</td>
<td>2.206</td>
<td>-26.6</td>
<td>2.19</td>
<td>-26.11</td>
</tr>
<tr>
<td>Total Lender Profits ($Billion)</td>
<td>3.5</td>
<td>-</td>
<td>-</td>
<td>3.36</td>
<td>-4.11</td>
</tr>
</tbody>
</table>

Note: This table shows counterfactual outcomes for two different scenarios. In scenario ρ = ̂ρ all markets are assigned the same ρ, which is the mean across all markets. In scenario No Dealer Discretion banks set interest rates directly and dealers compete downstream in prices taking them as given. The first three rows (Total Price, Car Price, APR) are averages across all markets. The last four rows are totals over all markets.

In No Wedge total transaction prices are on average $306.9 lower than in Baseline, which corresponds to a reduction of 1.12 percent. This overall price change is due to two opposing effects in the two dimensions of the pricing problem. On the one hand, dealers increase the prices of their cars on average from $24139.0 to $26683.9. On the other hand, dealers have a strong incentive to reduce interest rates because consumers now take the full amount of finance charges into account when deciding where to buy. The estimated decline in interest rates is from 0.134 to 0.017. The latter effect more than outweighs the former, which leads to the overall reduction in transaction prices. Intuitively, dealers are now competing for consumers that are overall much more price sensitive. Total prices fall as a result.

To analyze the corresponding change in consumer surplus, we compute average utility both under the initially estimated ̂ρ as well as under the paternalistic welfare criterion where we assume that ρ = 0. If utilities are computed with ̂ρ, consumer surplus decreases by -1.38 percent while it increases by 10.63 percent under the paternalistic welfare criterion. Total producer-surplus, aggregated across all markets, changes from 2.755 to 2.206 billion $, which amounts to a 26.6 percentage decrease. Hence, dealer profits would be drastically reduced if consumers were to treat finance charges and vehicle charges the same. Interestingly, there is a strong asymmetry in the change of the two different consumer surplus measures. Under the recovered consumer utility function, this counterfactual leads to a decrease in consumer
surplus. The change, however, is relatively small compared to the increase in consumer surplus under the no-wedge utility function.

Figure 5: Price Effects of the No Wedge Scenario $\rho$-decile.

![Graph showing price effects by decile of $\rho$.]

Note: The figure shows the estimated price effect of the No Wedge counterfactual by estimated $\rho$-decile. For individuals with a larger wedge, prices would fall by more in the counterfactual.

The overall counterfactual price change is largest for consumers who were previously exhibiting the largest wedge, for whom an extra dollar of finance charges now weights over-proportionately more. To illustrate the heterogeneity of the equilibrium outcome Figure 5 shows the change in total price by quantiles of our $\rho_i$-estimates. In the lowest quantile, the counterfactual leads to a price decrease of about $100 whereas in the highest quantile the decrease is more than $500.

10.1.1 The Effect of Competition

One question that arises naturally is to what extend competition can undo consumer harm when consumers are not fully aware of all components of the price (see also Gabaix and Laibson, 2006). To analyze how competition interacts with our results, we partition the response in the No Wedge counterfactual by number of dealers in the market. Figure 6 shows the results from this exercise. In markets with only two competitors the total price decreases by -537.7, whereas in markets with more than 15 competing dealers it is -262.5. Additional competition, therefore, strongly mitigates the detrimental effects of consumers’ reduced sensitivity to finance charges.

Similar to the market studied in this paper, many markets for products with add-ons are characterized by oligopolistic competition. Therefore, our results also provide an interesting benchmark for the interaction between competition and sub-optimal consumer behavior in other markets.
Figure 6: Price Effects of the No Wedge Scenario by Number of Dealers

**Note**: The figure shows the estimated price effect of the No Wedge scenario by the number of dealers in the market. The left panel depicts the corresponding scatter plot, the middle panel the histogram of estimated absolute price effects, and the right panel the histogram of estimated relative price effects.

### 10.2 No Dealer Discretion Counterfactual

In our key-counterfactual we explore what happens if dealers do not have discretion to set interest rates and lenders instead directly choose the interest rate, which the dealer then simply passes on to the consumer. As banks previously set buy rates in anticipation of a certain profit from the dealer markup, one would expect that the APR’s that banks quote lie significantly above the previous level of the buy-rates. We assume that lenders in this counterfactual bid according to the dominant strategy in a second price auction. This means that interest rates are set by the second lowest bid of \( N_d \) bids, where \( N_d \) is the number of lenders that contract with dealer \( d \).[^38]

There are two important forces that determine the outcome in the No Discretion counterfactual. On the one hand, the contractual incentives of dealers are changing as they no longer participate in finance charges. Holding the interest rate fixed, this leads them to increase price and reduce output. This effect is somewhat reminiscent of double marginalization: lender maximize their own profits in setting the interest rates and dealers simply take this as given. The difference is that the interest rate is not an input price to the dealers profit maximization problem. Instead it is one of the two prices that the consumer faces. On the other hand, there is an informational component. Lenders do not have the same information about consumers and, therefore, can not price consumer specific wedges optimally. This is, arguably, one of the main reasons for the existence of this institution is that dealers know better which customers are more prone to be less sensitive to finance charges.

To test whether that dealers in fact have access to pricing relevant information that lenders cannot observe, we estimate a fully saturated regression model with the optimal markup

[^38]: Note that while Proposition 5 provides a relatively straightforward way to compute lender cost, the reverse is not true. With the estimates in hand, it is computationally very hard to recompute an equilibrium, since the differential equation that results from the dealer bidding problem does not permit a closed form solution.
choice of the dealer on all variables that banks have access to for a given loan application. As shown in Table 22 in the Appendix even a fully saturated model exhibits an adjusted $R^2$ of less than 0.12. Hence, banks will be hardly able at all to condition their buy rate decisions on the wedge in elasticities in consumer preferences. A direct implication is that the effects of the counterfactual experiments should be heterogeneous across the population—consumer with higher $\rho$ should profit to a larger extent than their counterparts with lower $\rho$. The assumption that lenders bid according to the dominant strategy in a second-price auction, which means to simply bid the cost, automatically implies that they do not set interest rates with regard to the consumer-specific wedge.

An overview of the results from the first step can be found in Table 8, which shows that such a policy would lead to a decrease in overall transaction prices of about $-351.5$ which corresponds to -1.29 percent. In terms of consumer welfare the effect is more subtle. On the one hand, consumer welfare measured according to $\hat{\rho}$ is estimated to decrease by 4.12 percent as a response to such a policy. In particular, dealers cannot allocate charges to the interest rate dimension anymore which implies that they will increase the car price which in turn has a stronger impact on consumer utility. On the other hand, consumer welfare is estimated to increase on average by 5.14 according to the paternalistic welfare criterion with $\rho = 0$ for all consumers. While the effect would be positive, these results show that eliminating dealer discretion only generate a relatively small portion of the maximal attainable consumer surplus of the No Wedge counterfactual.

11 Conclusion

This paper explores the intermediation of auto loans through auto dealers. Based on the specific incentives that dealers face, we construct a measure that captures consumers’ differential responsiveness to loan and vehicle charges. We therefore leverage optimal supply side behavior to study demand side primitives, which is an approach that might find application in other settings. Our methodology neither relies on specific functional form assumptions for consumer utility nor on a fully specified model of competition. We find that consumers are significantly less responsive to finance charges compared to immediate charges on the car. The size of the wedge in responsiveness is negatively related to credit scores and education. To quantify the strategic implications of consumers’ attenuated responsiveness to finance charges, we use these results in a two-step estimator, which assumes differentiated product Bertrand competition among dealers. Computations from the model show that total vehicle prices would be lower and consumer surplus higher if consumers did not exhibit such a wedge in responsiveness to the two different prices in the market. Counterfactual computations also show that without dealer discretion prices in the market would fall and consumer

39 shows how the price affect of this counterfactual varies across markets populated by different numbers of dealers.
surplus increase.
References


A Proofs

Proof of Proposition 1

To solve the maximization problem in (1), ignore the second constraint and the constraint that \( p_i \geq 0 \) such that it becomes:

\[
\begin{align*}
\max_{r_i, p_i} & \quad p_i + (\tau_i p_i - d_i)(r_i - b_i) + \beta - c_i \\
\text{s.t.} & \quad -\tau_i p_i - (\tau_i p_i - d_i)r_i \geq \bar{u}_i, \\
& \quad r_i \geq b_i
\end{align*}
\]

Any solution to this problem clearly fulfills its constraint with equality. Therefore, we can rewrite this problem to be:

\[
\begin{align*}
\max_{r_i} & \quad -\frac{r_i d_i - \bar{u}_i}{(1 + r_i)\tau_i} + \left[ \frac{r_i d_i - \bar{u}_i}{1 + r_i} - d_i \right](r_i - b_i) + \beta - c_i \\
& \quad r_i \geq b_i
\end{align*}
\]

The derivative with respect to \( r_i \) is:

\[
\begin{align*}
\frac{d_i + \bar{u}_i}{(1 + r_i)^2 \tau_i} + \frac{d_i + \bar{u}_i}{(1 + r_i)^2}(r_i - b_i) + \frac{d_i + \bar{u}_i}{1 + r_i} \alpha & \leq 0 \\
\iff & \quad \left[ \frac{1}{r_i} + (r_i - b_i)\alpha - (1 + r_i)\alpha \right] \leq 0 \\
\iff & \quad 1 \geq \alpha(1 + b_i)\tau_i
\end{align*}
\]

Whenever the last equation is fulfilled, the derivative is negative such that the optimal solution features \( r_i = b_i \), which corresponds to a zero markup. Also note that the derived price is positive because \( |\bar{u}_i| \geq d_i \) is assumed for all \( i \). Hence, the constraint \( p_i \geq 0 \) is fulfilled. Finally, consider the second constraint. It will be fulfilled if:

\[-b_i(\tau_i p_i - d_i) \geq -(\tau_i p_i - d_i)\mathbb{E}[r^L] - s_i,\]

which is trivially fulfilled for \( b_i \leq \mathbb{E}[r^L] \). If in contrast, \( b_i > \mathbb{E}[r^L] \), we get that the second constraint is fulfilled if:

\[
\begin{align*}
P & \leq \left( \frac{s_i}{r_i - \mathbb{E}[r^L]} + d_i \right) \frac{1}{\tau_i} \\
\iff r_i d_i - \bar{u}_i & \leq \left( \frac{s_i}{r_i - \mathbb{E}[r^L]} + d_i \right) \frac{1}{\tau_i} \\
\iff s_i & \geq \left( \mathbb{E}[r^L] - b_i \right) \left( \frac{\bar{u}_i + d_i}{1 + b_i} \right).
\end{align*}
\]

As long as (15) holds, we will thus get that \( r_i^* = b_i \) for any optimal offer by the dealer. Finally, suppose that there is an optimal offer featuring \( r_i > b_i \) if (15) is not satisfied such that we have:
\( s < (\mathbb{E}[r^L] - b_i) \left( \frac{\bar{a}_i + d_i}{1 + r_i^L} \right) \). From the proof of Proposition 2, we know that for any solution with \( r_i^* > b_i \) the first constraint will be binding. Then from the arguments above it can not be the case that only the first constraint is binding, because \( r_i = b_i \) would be optimal. Hence, both constraints have to be binding. Then, we get

\[
p = \frac{r_i d_i - \bar{u}_i}{(1 + r_i) \tau_i} = \left( \frac{s_i}{r_i - \mathbb{E}[r^L]} + d_i \right) \frac{1}{\tau_i}
\]

\[
s_i = \frac{r_i - \mathbb{E}[r^L]}{1 + r_i} (d_i + \bar{u}_i)
\]

This however, is a contradiction to \( s_i < (\mathbb{E}[r^L] - b_i) \left( \frac{\bar{a}_i + d_i}{1 + r_i^L} \right) \), as \( |\bar{u}_i| > d_i \) and \( \frac{r_i - \mathbb{E}[r^L]}{1 + r_i^L} \) is increasing in \( r_i \). As a consequence, even if (15) does not hold, there cannot be an optimal offer featuring \( r_i^* > b_i \), which concludes the proof.

**Proof of Proposition 2**

First note that the optimal solution to the maximization problem in (1) will, in this case, satisfy at least one of its constraints with equality. Otherwise the dealer could increase \( p_i \) and thereby increase profits. In the first step of the proof we argue that if the second constraint is binding at the optimum also the first one is binding. For this purpose, suppose in contradiction that there exists an optimal solution \( (r_i^*, p_i^*) \) such that only the second constraint binds. Then \( (r_i^*, p_i^*) \) is a maximizer of:

\[
\max_{p_i, r_i} L = p_i + (\tau_i p_i - r_i^L)(r_i - b_i)\alpha + \beta - c_i + \\
\mu \left( M_i(r_i(\tau_i p_i - d_i)) - \int M_i((\tau_i p_i - d_i)\tau_i r^L)g_i(r^L)dr^L - s_i \right)
\]

Leading to the following first order conditions:

\[
\frac{dL}{dp_i} = 1 + \tau_i \alpha (r_i - b_i) + \mu \left( M_i'(r_i(\tau_i p_i - d_i))\tau_i r_i - \int M_i'((\tau_i p_i - d_i)\tau_i r^L)\tau_i r^L g_i(r^L)dr^L \right) = 0
\]

\[
\frac{dL}{dr_i} = \alpha (\tau_i p_i - d_i) + \mu M_i'(r_i(\tau_i p_i - d_i)) (\tau_i p_i - d_i) = 0
\]

Since both of these hold in equilibrium we get

\[
\frac{1 + \alpha \tau_i (r_i - b_i)}{M_i'(r_i(\tau_i p_i - d_i))\tau_i r_i - \int M_i'((\tau_i p_i - d_i)\tau_i r^L)\tau_i r^L g_i(r^L)dr^L} = \frac{\alpha}{M_i'(r_i(\tau_i p_i - d_i))}
\]

\[
M_i'(r_i(\tau_i p_i - d_i))(1 + \tau_i \alpha (r_i - b_i)) = \alpha (M_i'(r_i(\tau_i p_i - d_i))\tau_i r_i - \int M_i'((\tau_i p_i - d_i)\tau_i r^L)\tau_i r^L g_i(r^L)dr^L)
\]

\[
M_i'(r_i(\tau_i p_i - d_i))(1 - \tau_i \alpha b_i) = -\alpha \int M_i'((\tau_i p_i - d_i)\tau_i r^L)\tau_i r^L g_i(r^L)dr^L
\]

The last equality cannot be fulfilled since the left hand side is positive but the right hand side is negative, which establishes the contradiction.
Using the insights from the first step, there are only two possibilities left, either both constraints are binding or only the first one is. Suppose only the first one is binding such that the optimal solution to the maximization problem in (1) maximizes:

\[
\max_{p_i, \tau_i} \quad p_i + (\tau_i p_i - d_i)(r_i - b_i)\alpha + \beta - c_i \\
\text{s.t.} \quad -\tau_i p_i - M_i((\tau_i p_i - d_i)r_i) = \tilde{u}_i \\
r_i > b_i, p_i \geq 0
\]

The corresponding Lagrange function is given by:

\[
\max_{p_i, \tau_i} \quad L = p_i + (\tau_i p_i - d_i)(r_i - b_i)\alpha + \beta - c + \mu(\tilde{u}_i + \tau_i p_i + M_i(r_i(\tau_i p_i - d_i)))
\]

leading to the following first order conditions:

\[
\frac{dL}{dp_i} = 1 + \tau_i\alpha(r_i - b_i) + \mu\left(\tau_i + M'_i(r_i(\tau_i p_i - d_i))\tau_i r_i\right) = 0 \\
\frac{dL}{dr_i} = \alpha(\tau_i p_i - d_i) + \mu M'_i(r_i(\tau_i p_i - d_i))(\tau_i p_i - d_i) = 0
\]

Since both of these hold in equilibrium we get:

\[
\frac{1 + \tau_i\alpha(r_i - b_i)}{1 + M'_i(r_i(\tau_i p_i - d_i))\tau_i r_i} = \frac{\alpha}{M'_i(r_i(\tau_i p_i - d_i))} \\
\Leftrightarrow M'_i(r_i(\tau_i p_i - d_i)) = \frac{\alpha\tau_i}{1 - \tau_i\alpha b_i}
\]

In particular, any solution to the problem in which only the first constraint is binding will satisfy this equation. Next suppose that both constraints are binding. In this case, the corresponding Lagrange function is:

\[
L = p_i + (\tau_i p_i - d_i)(r_i - b_i)\alpha + \beta - c \\
+ \mu \left(\int M_i((\tau_i p_i - d_i)r^L)g_i(r^L)dr^L + s_i - M_i(r_i(\tau_i p_i - d_i))\right) \\
+ \lambda \left(-\tilde{u}_i - \tau_i p_i - M_i(r_i(\tau_i p_i - d_i))\right)
\]

Then from the Kuhn-Tucker Theorem, we know that there exists \( \lambda > 0 \) and \( \mu > 0 \) such that the following two equations are fulfilled:

\[
\frac{dL}{dp_i} = 1 + \tau_i\alpha(r_i - b_i) + \mu \left(\int M'_i((\tau_i p_i - d_i)r^L)\tau_i r^L g_i(r^L)dr^L - M'_i(r_i(\tau_i p_i - d_i)) r_i \tau_i\right) \\
+ \lambda \left(-\tau_i - M'_i((\tau_i p_i - d_i)r_i)\tau_i r_i\right) = 0 \\
\frac{dL}{dr_i} = \alpha(\tau_i p_i - d_i) - \mu M'_i(r_i(\tau_i p_i - d_i))(\tau_i p_i - d_i) - \lambda M'_i(r_i(\tau_i p_i - d_i))(\tau_i p_i - d_i) = 0
\]
Thus the second constraint is slack and the first one has to be binding. Hence, \( \tilde{\mu} \)

Substituting this into the first equation yields:

\[
1 + \tau_i \alpha (r_i - b_i) + \mu \left( \int M_i^\prime((\tau_i p_i - d_i)) \tau_i r_i g_i(r_i) dr_i - M_i^\prime((\tau_i p_i - d_i)) \tau_i r_i \right) + \frac{\alpha - \mu M_i^\prime((\tau_i p_i - d_i))}{M_i^\prime((\tau_i p_i - d_i))} (-\tau_i - M_i^\prime((\tau_i p_i - d_i)) r_i \tau_i) = 0
\]

\[
\iff 1 - \tau_i \alpha b_i + \mu \int M_i^\prime((\tau_i p_i - d_i)) \tau_i r_i g_i(r_i) dr_i \tau_i = \frac{\alpha - \mu M_i^\prime((\tau_i p_i - d_i))}{M_i^\prime((\tau_i p_i - d_i))}
\]

\[
\mu = \frac{\tau_i \alpha b_i + \frac{\tau_i \alpha}{M_i^\prime((\tau_i p_i - d_i))} - 1}{\tau_i + \int M_i^\prime((\tau_i p_i - d_i)) \tau_i r_i g_i(r_i) dr_i}
\]

since \( \mu \) is positive \( \tau_i \alpha b_i + \frac{\tau_i \alpha}{M_i^\prime((\tau_i p_i - d_i))} - 1 > 0 \) holds, which is equivalent to the statement that \( M_i^\prime((\tau_i p_i^* - d_i)) < \frac{\tau_i \alpha}{1 - \tau_i \alpha b_i} \).

**Proof of Proposition 3**

For the first part assume that \( M_i^\prime(x) \leq 1 \ \forall x \). Let \( r_i^* > b_i, p_i^* > 0 \) be the optimal offer. Denote by \( \tilde{p}_i \) the price that is optimal for the dealer given that he offers \( r_i = b_i \). Note first that \( \tilde{p}_i > p_i^* \) holds. In particular, both constraints in (1) have to be slack if \( r_i = b_i \) and \( p_i = p_i^* \). Thus, the dealer can profit from increasing the price and we get \( \tilde{p}_i > p_i^* \). As a consequence, the first constraint in (1) will also be binding at \( b_i, \tilde{p}_i \). To see this, recall that the first constraint is binding at \( r_i^*, p_i^* \) due to Proposition 2. We therefore get:

\[
\tilde{u}_i + \tau_i p_i^* = -M_i((\tau_i p_i^* - d_i)) r_i \geq -\int M_i((\tau_i p_i^* - d_i)) g_i(r_i) dr_i - s_i
\]

As \( \tilde{p}_i > p_i^* \),

\[
\tilde{u}_i + \tau_i \tilde{p}_i > \tilde{u}_i + \tau_i p_i^* \geq -\int M_i((\tau_i p_i^* - d_i)) g_i(r_i) dr_i - s_i \geq -\int M_i((\tau_i \tilde{p}_i - d_i)) g_i(r_i) dr_i - s_i
\]

Thus the second constraint is slack and the first one has to be binding. Hence, \( \tilde{p}_i \) and \( p_i^* \) are given by:

\[
-\tau_i \tilde{p}_i - M((\tau_i \tilde{p}_i - d_i) b_i) = \tilde{u}_i
\]

\[
-\tau_i p_i^* - M((\tau_i p_i^* - d_i) r_i) = \tilde{u}_i.
\]
The optimal offer has to imply a weakly higher profit than offering \( r_i = b_i \). Hence,

\[
p_i^* + (\tau_i p_i^* - d_i)\alpha(r_i^* - b_i) + \beta - c_i \geq \tilde{p}_i + \beta - c_i
\]

(16) implies that \( \tilde{p}_i \geq p_i^* \), we therefore get that:

\[
\begin{align*}
(\tau_i p_i^* - d_i)(r_i^* - b_i) & - [M_i((\tau_i p_i^* - d_i)r_i) - M_i((\tau_i p_i^* - d_i)b_i)] \\
= (\tau_i p_i^* - d_i)(r_i - b_i) - [M_i((\tau_i p_i^* - d_i)r_i) - M_i((\tau_i \tilde{p}_i - d_i)b_i)] \\
+ M_i((\tau_i \tilde{p}_i - d_i)b_i) - M_i((\tau_i p_i^* - d_i)b_i)] \\
\geq (\tau_i p_i^* - d_i)(r_i^* - b_i) - \tau_i (\tau_i p_i^* - d_i)\alpha(r_i^* - b_i) - [M_i((\tau_i \tilde{p}_i - d_i)b_i) - M_i((\tau_i p_i^* - d_i)b_i)] \\
\geq (\tau_i p_i^* - d_i)(r_i - b_i) - \tau_i (\tau_i p_i^* - d_i)\alpha(r_i^* - b_i) - ((\tau_i \tilde{p}_i - d_i)b_i - (\tau_i p_i^* - d_i)b_i) \\
= (\tau_i p_i^* - d_i)r_i - (\tau_i \tilde{p}_i - d_i)b_i - \tau_i (\tau_i p_i^* - d_i)\alpha(r_i^* - b_i)
\end{align*}
\]

(17) where the last inequality follows from the fact that \( M'(x) \leq 1 \ \forall x \). Moreover, (16) implies that

\[
\begin{align*}
p_i^* + (\tau_i p_i^* - d_i)\alpha(r_i^* - b_i) + \beta - c_i & \geq \tilde{p}_i + \beta - c_i \\
\implies b_i[\tau_i p_i^* - d_i + \tau_i (\tau_i p_i^* - d_i)\alpha(r_i^* - b_i)] & \geq [\tau_i \tilde{p}_i - d_i]b_i
\end{align*}
\]

Substituting this into (17), yields

\[
(\tau_i p_i^* - d_i)r_i - (\tau_i p_i^* - d_i)b_i - (M_i((\tau_i p_i^* - d_i)r_i) - M_i((\tau_i p_i^* - d_i)b_i)) \\
\geq (\tau_i p_i^* - d_i)r_i - b_i[\tau_i p_i^* - d_i + \tau_i (\tau_i p_i^* - d_i)\alpha(r_i^* - b_i)] - \tau_i (\tau_i p_i^* - d_i)\alpha(r_i^* - b_i) \\
= (\tau_i p_i^* - d_i)(r_i - b_i)[1 - \tau_i b_i \alpha - \tau_i \alpha],
\]

which concludes the prof of the first part of the proposition. Part (ii) of the proposition follows from Proposition 2 and the assumed convexity of \( M \). In particular, convexity implies that

\[
M((\tau_i p_i^* - d_i)r_i) - M_i(0) \leq M_i'((\tau_i p_i^* - d_i)r_i)(\tau_i p_i^* - d_i)r_i \\
\leq \frac{\tau_i \alpha}{1 - \tau_i \alpha b_i}(\tau_i p_i^* - d_i)r_i
\]

Which implies that a lower Bound for the Mistake is given by:

\[
(\tau_i p_i^* - d_i)r_i - M((\tau_i p_i^* - d_i)r_i) \geq (\tau_i p_i^* - d_i)r_i - \frac{\tau_i \alpha}{1 - \tau_i \alpha b_i}(\tau_i p_i^* - d_i)r_i = \\
(\tau_i p_i^* - d_i)r_i - \left[1 - \frac{\tau_i \alpha}{1 - \tau_i \alpha b_i}\right]
\]

44
Proof of Corollary 1

With the same arguments developed in the proof of Proposition 2, we know that either only the first one or both constraints are binding at the optimum. If only the first constraint is binding, we get the following Lagrange function:

$$\max_{p_i,r_i} L = p_i + (\tau_i p_i - d_i)(r_i - b_i)\alpha \gamma + \beta - c + \mu (\bar{u}_i + \tau_i p_i + \gamma M_i(r_i(\tau_i p_i - d_i)))$$

Leading to the following first order conditions:

$$\frac{dL}{dp_i} = 1 + \tau_i \gamma \alpha (r_i - b_i) + \mu \left( \tau_i + \gamma M_i'(r_i(\tau_i p_i - d_i))r_i r_i \right) = 0$$

$$\frac{dL}{dr_i} = \gamma \alpha (\tau_i p_i - d_i) + \mu \gamma M_i'(r_i(\tau_i p_i - d_i))(\tau_i p_i - d_i) = 0$$

Since both of these hold in equilibrium we get:

$$\frac{1 + \tau_i \gamma \alpha (r_i^* - b_i)}{\tau_i + \gamma M'_i(r_i^*(\tau_i p_i^* - d_i))r_i r_i} = \frac{\alpha}{M'_i(r_i^*(\tau_i p_i^* - d_i))}$$

$$\Leftrightarrow M'_i(r_i^*(\tau_i p_i^* - d_i)) = \frac{\tau_i \alpha}{1 - \tau_i \gamma \alpha b_i}$$

If both constraints are binding, we can again use the same arguments as in the proof of Proposition 2 to show that:

$$M'_i(r_i^*(\tau_i p_i^* - d_i)) < \frac{\tau_i \alpha}{1 - \gamma \tau_i \alpha b_i}$$

A.1 Model with Exogenous Downpayment

In this subsection we derive the results in Proposition 2 and 1 taking the time structure of the decision problem into account. For this purpose, suppose the consumer discounts each month with $\delta$. If he signs a fully amortizing loan with a term of $x$ months, his overall payments on the loan are going to be

$$x \cdot P = x \cdot \frac{r(1 + r)^x}{(1 + r)^x - 1}(\tau_i p_i - d_i) \equiv (1 + \hat{r}_i)(\tau_i p_i - d_i)$$

Recall that we assume financial payments to be perceived differently than payments on the car. $\hat{r}_i(\tau_i p_i - d_i)$ is the part of the payments that are finance charges. Assume that the consumer has a stream of utility over money where he splits up payments every period into finance charges and non finance charges:

$$u = -d_i - \sum_{i=1}^{x} \frac{\delta^j}{x} \left[ (\tau_i p_i - d_i) + M((\tau_i p_i - d_i)\hat{r}_i) \right] =$$

$$-d_i - \frac{\delta - \delta^{x+1}}{(1 - \delta)x} \left[ (\tau_i p_i - d_i) + M((\tau_i p_i - d_i)\hat{r}_i) \right]$$
Hence, the utility is the same as before except for the multiplication by a constant and subtracting and adding exogenous variables. Therefore, we can redefine the utility of the outside option such that the maximization problem is identical to the one studied before. None of the calculations from above would change.

### A.2 Downpayment as a Fixed Fraction of the Price

In this subsection, we assume that the downpayment \(d_i\) is given by \(d \tau_i p_i\), where \(d < 1\) is an exogenous fraction. Moreover, we assume that the downpayment is due immediately while all other payments are financed by a fully amortizing loan. Denote by \(x\) the monthly discount rate, and by \(\hat{r}_i = \frac{(1 + \frac{r}{12})^x}{(1 + \frac{r}{12})^x - 1}\) the fraction of the payment that is due to finance charges if the yearly interest rate is \(r\). Taking the arguments from before the utility of a consumer from buying the car is then given by:

\[
u = \sum_{j=1}^{x} \frac{\delta^j}{\delta^j} [(1 - d) \tau_i p_i + M((1 - d) \tau_i p_i h)] =
(d \tau_i p_i - \frac{\delta - \delta^{x+1}}{(1 - \delta) x} [(1 - d) \tau_i p_i + M((1 - d) \tau_i p_i h)])
\]

The car dealers optimization problem then becomes:

\[
\max_{\hat{r_i}, p_i} \quad p_i + (1 - d) \tau_i p_i (\hat{r}_i - \hat{b}_i) \alpha + \beta - c_i
\]

\[
s.t. \quad - d \tau_i p_i - \frac{\delta - \delta^{x+1}}{(1 - \delta) x} [(1 - d) \tau_i p_i + M((1 - d) \tau_i p_i h)] \geq \bar{\alpha}
\]

\[
-M_i((1 - d) \tau_i p_i h) \geq - \int M_i((1 - d) \tau_i p_i h g(r^f) - s_i)
\]

\[
r_i, p_i \geq 0
\]

With the same arguments as above, we can concentrate on the case where the first constraint is binding, but the second not. The Lagrangian is then given by:

\[
\mathcal{L} = p_i + (1 - d) \tau_i p_i (\hat{r}_i - \hat{b}_i) \alpha + \beta - c_i + \mu \left[u + d \tau_i p_i + \frac{\delta - \delta^{x+1}}{(1 - \delta) x} [(1 - d) \tau_i p_i + M((1 - d) \tau_i p_i h)]\right]
\]

The derivatives are given by:

\[
\frac{d \mathcal{L}}{d \hat{r}_i} = (1 - d) \tau_i p_i \alpha + \mu \frac{\delta - \delta^{x+1}}{(1 - \delta) x} M'((1 - d) \tau_i p_i h) \tau_i p_i (1 - d)
\]

\[
\frac{d \mathcal{L}}{d p_i} = 1 + (1 - d)(\hat{r}_i - \hat{b}_i) \tau_i \alpha + \mu \left[d \tau_i + \frac{\delta - \delta^{x+1}}{(1 - \delta) x} \tau_i (1 - d) + M'((1 - d) \tau_i p_i h) \tau_i (1 - d)\right]
\]
Since both have to equal to zero simultaneously, we get:

\[
\frac{\alpha}{\frac{\partial - \delta^{x+1}}{(1-\delta)^x} M'(1-d)p_i \tau_i \hat{r}_i} = 1 + (1-d)(\hat{r}_i - b_i) \tau_i \alpha
\]

\[
\Leftrightarrow \frac{\delta - \delta^{x+1}}{(1-\delta)^x} M'(1-d) \tau_i \hat{r}_i [1 + (1-d) \tau_i \alpha (\hat{r}_i - b_i) - (1-d) \tau_i \hat{r}_i \alpha] = \tau_i \alpha \left[ d + (1-d) \frac{\delta - \delta^{x+1}}{(1-\delta)^x} \right]
\]

\[
\Leftrightarrow M'(1-d) \tau_i \hat{r}_i = \tau_i \alpha \left[ d + (1-d) \frac{\delta - \delta^{x+1}}{(1-\delta)^x} \right]
\]

**Proof of Proposition 4**

We construct \( M(x, \rho_i) \) such that the ordering of \( B_i^O = [1-M'(x, \rho_i)]x \) induces an ordering on \( \rho_i \). Hence, \( \rho_i > \rho_j \Leftrightarrow B_i^O > B_j^O \) for all \( i, j \). As a direct consequence of continuity it them must hold that \( \rho_i = \rho_j \) if and only if \( B_i^O = B_j^O \). Consider the set of observations such that this is true. Hence,

\[
[1 - M'(x, \rho)]x = C
\]

Then with changing \( x, \rho \) needs to stay constant:

\[
\frac{dp}{dx} = \frac{1 - M'(x, \rho) - M''(x, \rho)x}{x M'(x, \rho)} = 0
\]

\[
\Leftrightarrow 1 - M'(x, \rho) - M''(x, \rho)x = 0
\]

This is a separable ODE such that we can solve it the following way:

\[
\frac{M''(x, \rho)}{1 - M'(x, \rho)} = \frac{1}{x}
\]

\[
\Leftrightarrow \int \frac{1 - M'}{d - M'} = \int \frac{1}{x} \, dx
\]

\[
\Rightarrow -\ln(1 - M') = \ln(x) + g(\rho)
\]

\[
\Leftrightarrow \frac{1}{1 - M'} = xg(\rho)
\]

\[
\Rightarrow M' = 1 - \frac{1}{xg(\rho)}
\]

\[
\Rightarrow M = x - h(\rho)\ln(x),
\]

where \( h(\rho) \) is some arbitrary function of \( \rho \). Hence, every \( M \) function that satisfies the above induces the ordering on the bounds to be also an ordering on the rhos. Note that these functions are convex as long as \( h(\rho) > 0 \) and that higher \( \rho \) corresponds to a larger misperception if \( h'(\rho) > 0 \).
Proof of Proposition 5

The lenders maximization problem is given by:

$$\max_{b_k} s_{j,a}(b) \cdot \left(p_{j,a}(b_k) - d\right) \cdot \left(1 - F(\beta_{j,a}^{-1}(b_k))\right)^{N-1} \cdot \left[(1 - \alpha) \cdot m_{j,a}(b) + b_k - r'_k\right]. \quad (18)$$

Taking the first order condition, we obtain:

$$\left[(1 - \alpha) \cdot m'_{j,a}(b_k) + 1\right] \cdot \psi(b_k) \cdot \left(1 - F(\beta_{j,a}^{-1}(b_k))\right)^{N-1} + \left[(1 - \alpha) \cdot m_{j,a}(b_k) + b_k - r\right] \cdot \left[\psi'(b_k) \cdot \left(1 - F(\beta_{j,a}^{-1}(b_k))\right)^{N-1} - \psi(b_k) \cdot (N - 1) \cdot f(r) \cdot \frac{1}{\beta'(b_k)} \cdot \left(1 - F(\beta_{j,a}^{-1}(b_k))\right)^{N-2}\right] = 0 \quad (19)$$

Which can be simplified to the following equation, using the insight that $G(b) = F(\beta_{j,a}^{-1}(b_k)) \ \forall b$

and therefore $g(b) = \frac{1}{\beta'(b)} \cdot f(r)$.

$$\left[(1 - \alpha) \cdot m'_{j,a}(b_k) + 1\right] \cdot \psi(b_k) \cdot \left(1 - F(\beta_{j,a}^{-1}(b_k))\right)^{N-1} + \left[(1 - \alpha) \cdot m_{j,a}(b_k) + b_k - r'_k\right] \cdot \left[\psi'(b_k) \cdot \left(1 - F(\beta_{j,a}^{-1}(b_k))\right)^{N-1} - \psi(b_k) \cdot (N - 1) \cdot g(b_k)\right] = 0$$

Solving this equation for $r$ and defining $\tilde{\psi}(b_k) = \frac{\psi'(b_k)}{\psi(b_k)}$ to simplify the exposition we get the following expression for the bank’s refinancing costs.

$$r_k = (1 - \alpha) \cdot m_{j,a}(b_k) + b_k + \frac{1 + (1 - \alpha) \cdot m'_{j,a}(b_k)}{\psi'(b_k) - (N - 1) \cdot g(b_k) \cdot (1 - G(b_k))^{-1}} \quad (20)$$

B Additional Facts, Graphs and Figures

B.1 Dealer Issued Loans

This section explains our procedure to estimate the fraction of auto loans originated “indirectly”, i.e. through auto dealers. The CCP includes data on hard credit inquiries, and so is one of very few datasets that covers search behavior outside of online markets. When an auto dealer intermediates a loan, the process should always begin with a hard credit inquiry so that the dealer can determine the kind of loan a borrower can qualify for. We take the percent of loans originated within a short window of at least one hard credit inquiry from a dealer as our proxy for the fraction of indirect loans. If the CCP included data on all hard credit inquiries, this would be straightforward to estimate. The main difficulty is that we only observe hard credit inquiries reported to the credit bureau our data is from; we do not see hard credit inquiries reported to the other two major credit bureaus.
We can deal with this difficulty if we assume a constant probability $P_o$ that a hard credit inquiry is observed in the CCP. Let $P_d(i)$ denote the probability that a loan is originated within a short time window of $i$ hard credit inquiries from dealers, and let $P_{do}(i)$ denote the probability that a loan is originated within a short time window of $i$ hard credit inquiries from dealers \textit{that we observe}. Finally, assume that no more than $N$ dealers perform a hard credit pull on a consumer within the time window. Then we need to estimate $P_d(0)$, and have the following equation:

$$P_{do}(i) = \sum_{n=i}^{N} P_d(n) \binom{n}{i} P_o^i (1 - P_o)^{n-i}$$

(21)

First, we set $N = 3$. The next step is to estimate $P_{do}(i)$ and $P_o$. We do that by matching new auto loans to auto loan inquiries from the company the loan is from (for $P_o$) and to inquiries from auto dealers (for $P_{do}$). We match auto loans to auto loan inquiries if they are for the same consumer and if the inquiry date is no more than 14 days before or 7 days after the origination date of the auto loan. We restrict the sample during this step to consumers with credit scores above 680, to minimize the possibility of one dealer pulling credit records from multiple credit bureaus.

For a given guess of the vector $P_d(n)$, Equation 21 yields implied values of the vector $P_{do}(i)$. We take as our estimate of $P_d(n)$ the vector that minimizes the sum of squared deviations between implied and estimated values of $P_{do}(i)$. Using data for the U.S. as a whole, this yields an estimate of $P_d(0) = .17$. This implies that an estimated 83% of auto loans are opened a short time before or after a hard credit inquiry from an auto dealer, which we interpret to mean that roughly 83% of auto loans are indirect. A number of assumptions were required to obtain this estimate, and we do not view it as precise. However, we do interpret it as strong evidence that the vast majority of auto loans in the United States are indirect.

### B.2 Other regressions etc

A regression of markup on log applicant income, log price, credit score bins (of width 20), mileage, log loan amount, and state, bank, model, and new car fixed effects yields an $R^2$ of .07. This goes up to

---

40 In particular, we assume that the probability of observing a given hard credit inquiry from a lender (which we can estimate) is the same as observing a given hard credit inquiry from a dealer (which we need). This is equivalent to assuming that the credit bureau’s market share for dealer inquiries is the same as its market share for lender inquiries. Unfortunately we cannot test this assumption, but our conversations with market experts lead us to believe it is reasonable. Dealers have an incentive to pull credit from the same credit bureau as the lenders they work with, so that they are operating with the same information.

41 Lenders pay credit bureaus for every inquiry they make, so when deciding on the number of bureaus to pull from, they face a tradeoff between the cost of an additional pull and the benefit of obtaining more information. Because of the very large sums of money involved, mortgage lenders nearly always pull from all three major credit bureaus. Auto lenders typically only pull information from one credit bureau for borrowers who do not appear to be a credit risk, which is why we focus on consumers with good credit scores. Auto dealers have even less incentive to pull from multiple bureaus than auto lenders do, because auto dealers do not bear default risk.
.11 when using zip fixed effects instead of state fixed effects, or .10 if restricting the regression to non-negative markups no higher than 2.5 (the 99th percentile in the data). We interpret this as evidence that banks cannot predict markups.

Table 9: Regressions of $M'$ on Finance Charges, instrumented by Buy Rate

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finance Charges, in thousands, instrumented by Buy Rate</td>
<td>0.00512***</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Credit score score, in 100s</td>
<td>-0.00801***</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Log Monthly Income</td>
<td>-0.00261***</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

Figure 7: Distribution of Down Payment as a Fraction of Price
Figure 8: This map shows the residual of bound $B^O$ (the overall bound) after taking out bank fixed effects.
Figure 9: This map shows the residual of bound $B^M$ (the bound due to markups) after taking out bank fixed effects.
Table 10: *Estimates for $M'_i(\cdot)$ for different down-payments*

<table>
<thead>
<tr>
<th>Down Payment</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Downpayment=0</td>
<td>0.65</td>
<td>0.73</td>
<td>0.87</td>
<td>0.90</td>
<td>0.93</td>
</tr>
<tr>
<td>Below 10th</td>
<td>0.67</td>
<td>0.71</td>
<td>0.82</td>
<td>0.86</td>
<td>0.92</td>
</tr>
<tr>
<td>10th-25th</td>
<td>0.66</td>
<td>0.68</td>
<td>0.74</td>
<td>0.85</td>
<td>0.92</td>
</tr>
<tr>
<td>25th-50th</td>
<td>0.65</td>
<td>0.68</td>
<td>0.75</td>
<td>0.84</td>
<td>0.91</td>
</tr>
<tr>
<td>50th-75th</td>
<td>0.43</td>
<td>0.67</td>
<td>0.80</td>
<td>0.84</td>
<td>0.89</td>
</tr>
<tr>
<td>75th-90th</td>
<td>0.41</td>
<td>0.67</td>
<td>0.81</td>
<td>0.83</td>
<td>0.88</td>
</tr>
<tr>
<td>Above 90th</td>
<td>0.41</td>
<td>0.69</td>
<td>0.81</td>
<td>0.83</td>
<td>0.87</td>
</tr>
</tbody>
</table>

*Note:* The table shows the estimated values for $M'_i(\cdot)$ for different percentiles of the population by the size of their down payment.
Table 11: *Estimates for $B^O_i$ for different downpayments*

<table>
<thead>
<tr>
<th>Down Payment Percentile</th>
<th>$B^O_i$ Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p10</td>
</tr>
<tr>
<td>Downpayment=0</td>
<td>161.01</td>
</tr>
<tr>
<td>Below 10th</td>
<td>217.41</td>
</tr>
<tr>
<td>10th-25th</td>
<td>220.59</td>
</tr>
<tr>
<td>25th-50th</td>
<td>223.79</td>
</tr>
<tr>
<td>50th-75th</td>
<td>219.43</td>
</tr>
<tr>
<td>75th-90th</td>
<td>196.04</td>
</tr>
<tr>
<td>Above 90th</td>
<td>172.61</td>
</tr>
</tbody>
</table>

*Note:* The table shows the estimated values for $B^O_i$ for different percentiles of the population by the size their down payment.

Table 12: *Estimates for $M'(-\cdot)_i$ for different prepayment probabilities*

<table>
<thead>
<tr>
<th>Prepayment Percentile</th>
<th>$M'(-\cdot)_i$ Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p10</td>
</tr>
<tr>
<td>Below 10th</td>
<td>0.78</td>
</tr>
<tr>
<td>10th-25th</td>
<td>0.78</td>
</tr>
<tr>
<td>25th-50th</td>
<td>0.77</td>
</tr>
<tr>
<td>50th-75th</td>
<td>0.77</td>
</tr>
<tr>
<td>75th-90th</td>
<td>0.77</td>
</tr>
<tr>
<td>Above 90th</td>
<td>0.77</td>
</tr>
</tbody>
</table>

*Note:* The table shows the estimated values for $M'(-\cdot)_i$ for different percentiles of the population by the probability that they prepay early. Prepayment rates are estimated using the CCP.
Table 13: *Mark ups for different prepayment probabilities*

<table>
<thead>
<tr>
<th>PREPAYMENT PERCENTILE</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 10th</td>
<td>0</td>
<td>0.22</td>
<td>1.01</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>10th-25th</td>
<td>0</td>
<td>0.21</td>
<td>1.25</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>25th-50th</td>
<td>0</td>
<td>0.25</td>
<td>1.23</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>50th-75th</td>
<td>0</td>
<td>0.36</td>
<td>1.2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>75th-90th</td>
<td>0</td>
<td>0.48</td>
<td>1.19</td>
<td>1.98</td>
<td>2</td>
</tr>
<tr>
<td>Above 90th</td>
<td>0</td>
<td>0.5</td>
<td>1.16</td>
<td>1.91</td>
<td>2.1</td>
</tr>
</tbody>
</table>

*Note:* The table shows mark ups for different percentiles of the population by the probability that they prepay early. Prepayment rates are estimated using the CCP.

Table 14: *Estimates for $B^M_i$ for different prepayment probabilities*

<table>
<thead>
<tr>
<th>PREPAYMENT PERCENTILE</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 10th</td>
<td>0</td>
<td>14.6</td>
<td>73.68</td>
<td>159.61</td>
<td>247.61</td>
</tr>
<tr>
<td>10th-25th</td>
<td>0</td>
<td>9.29</td>
<td>71.45</td>
<td>147.04</td>
<td>225.02</td>
</tr>
<tr>
<td>25th-50th</td>
<td>0</td>
<td>12.38</td>
<td>72.21</td>
<td>143.88</td>
<td>220.72</td>
</tr>
<tr>
<td>50th-75th</td>
<td>0</td>
<td>18.80</td>
<td>72.27</td>
<td>142.47</td>
<td>223.99</td>
</tr>
<tr>
<td>75th-90th</td>
<td>0</td>
<td>21.86</td>
<td>70.97</td>
<td>137.76</td>
<td>220.82</td>
</tr>
<tr>
<td>Above 90th</td>
<td>0</td>
<td>24.51</td>
<td>71.93</td>
<td>141.61</td>
<td>233.89</td>
</tr>
</tbody>
</table>

*Note:* The table shows the estimated values for $B^M_i$ for different percentiles of the population by the probability that they prepay early. Prepayment rates are estimated using the CCP.
<table>
<thead>
<tr>
<th>Markups</th>
<th>( M'(\cdot) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lender finances 5 – 20% of sales</td>
<td>0.0435***, 0.0405***, -0.00123***, -0.00126***</td>
</tr>
<tr>
<td></td>
<td>(0.00331), (0.00332), (0.000131), (0.000131)</td>
</tr>
<tr>
<td>Lender finances &gt; 20% of sales</td>
<td>0.0126***, 0.00773*, 0.000122, 0.0000675</td>
</tr>
<tr>
<td></td>
<td>(0.00411), (0.00411), (0.000162), (0.000163)</td>
</tr>
<tr>
<td>Credit Score</td>
<td>-0.000843***, -0.000840***, -0.0000805***, -0.0000803***</td>
</tr>
<tr>
<td></td>
<td>(0.0000298), (0.0000297), (0.000000883), (0.000000883)</td>
</tr>
<tr>
<td>Log Monthly Income</td>
<td>-0.0260***, -0.0261***, -0.00186***, -0.00187***</td>
</tr>
<tr>
<td></td>
<td>(0.00205), (0.00205), (0.0000694), (0.0000694)</td>
</tr>
<tr>
<td>Loan Term</td>
<td>0.00174***, 0.00177***, 0.000761***, 0.000762***</td>
</tr>
<tr>
<td></td>
<td>(0.000133), (0.000133), (0.00000410), (0.00000410)</td>
</tr>
<tr>
<td>New Car</td>
<td>-0.217***, -0.218***, -0.00727***, -0.00727***</td>
</tr>
<tr>
<td></td>
<td>(0.00278), (0.00278), (0.0000790), (0.0000789)</td>
</tr>
<tr>
<td>Dealer market share 2.5 – 10%</td>
<td>-0.0155***, -0.00895***</td>
</tr>
<tr>
<td></td>
<td>(0.00382), (0.00111)</td>
</tr>
<tr>
<td>Dealer market share &gt; 10%</td>
<td>-0.0453***, -0.00520***</td>
</tr>
<tr>
<td></td>
<td>(0.00499), (0.00161)</td>
</tr>
<tr>
<td>HH index 500 – 1500</td>
<td>-0.00813, -0.000300*</td>
</tr>
<tr>
<td></td>
<td>(0.00495), (0.000157)</td>
</tr>
<tr>
<td>HH index &gt; 1500</td>
<td>0.00644, -0.000897***</td>
</tr>
<tr>
<td></td>
<td>(0.00610), (0.000197)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.608***, 1.625***, 0.777***, 0.778***</td>
</tr>
<tr>
<td></td>
<td>(0.0663), (0.0664), (0.00269), (0.00270)</td>
</tr>
</tbody>
</table>

Note: Bank, month, model, and zip code fixed effects also included, but not shown. Standard errors clustered at zip code level.
Table 16: *Estimates if down payments are a fixed fraction of the car price*

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M'$ with $\delta = 1$</td>
<td>0.80</td>
<td>0.64</td>
<td>0.69</td>
<td>0.84</td>
<td>0.89</td>
<td>0.93</td>
</tr>
<tr>
<td>$M'$ with $\delta = 0.98$</td>
<td>0.80</td>
<td>0.64</td>
<td>0.69</td>
<td>0.84</td>
<td>0.89</td>
<td>0.92</td>
</tr>
<tr>
<td>$M'$ with $\delta = 0.95$</td>
<td>0.81</td>
<td>0.66</td>
<td>0.70</td>
<td>0.85</td>
<td>0.91</td>
<td>0.94</td>
</tr>
<tr>
<td>$M'$ with $\delta = 0.90$</td>
<td>0.84</td>
<td>0.67</td>
<td>0.73</td>
<td>0.86</td>
<td>0.94</td>
<td>0.98</td>
</tr>
<tr>
<td>$M'$ with $\delta = 0.6$</td>
<td>1.08</td>
<td>0.73</td>
<td>0.87</td>
<td>1.03</td>
<td>1.27</td>
<td>1.51</td>
</tr>
</tbody>
</table>

Note: Table shows percentiles of the distribution of estimated $M'$, assuming loan term length varies as in the data but all consumers have an annual discount factor equal to the given value of $\delta$. 
Table 17: Logit Regression of Prepayment Risk

<table>
<thead>
<tr>
<th>Credit Score Range</th>
<th>Prepayment Early Prepayment</th>
<th>Late Prepayment</th>
</tr>
</thead>
<tbody>
<tr>
<td>710 ≤ Credit score ≤ 729</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>730 ≤ Credit score ≤ 749</td>
<td>-0.0338***</td>
<td>0.0427</td>
</tr>
<tr>
<td>750 ≤ Credit score ≤ 769</td>
<td>-0.0367***</td>
<td>0.0963***</td>
</tr>
<tr>
<td>770 ≤ Credit score ≤ 789</td>
<td>-0.0117</td>
<td>0.252***</td>
</tr>
<tr>
<td>790 ≤ Credit score ≤ 809</td>
<td>-0.0307**</td>
<td>0.378***</td>
</tr>
<tr>
<td>810 ≤ Credit score ≤ 829</td>
<td>-0.0140</td>
<td>0.537***</td>
</tr>
<tr>
<td>830 ≤ Credit score ≤ 849</td>
<td>0.0350</td>
<td>0.545***</td>
</tr>
<tr>
<td>Log loan size</td>
<td>-0.381***</td>
<td>-0.316***</td>
</tr>
<tr>
<td>25 ≤ Loan term ≤ 29</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30 ≤ Loan term ≤ 41</td>
<td>-0.962***</td>
<td>0.00358</td>
</tr>
<tr>
<td>42 ≤ Loan term ≤ 53</td>
<td>-1.306***</td>
<td>-0.429***</td>
</tr>
<tr>
<td>54 ≤ Loan term ≤ 65</td>
<td>-1.472***</td>
<td>-0.149</td>
</tr>
<tr>
<td>66 ≤ Loan term ≤ 77</td>
<td>-1.403***</td>
<td>-0.134</td>
</tr>
<tr>
<td>78 ≤ Loan term ≤ 89</td>
<td>-1.306***</td>
<td>-0.257**</td>
</tr>
<tr>
<td>90 ≤ Loan term ≤ 101</td>
<td>-1.603***</td>
<td>-0.229</td>
</tr>
<tr>
<td>102 ≤ Loan term ≤ 113</td>
<td>-1.935***</td>
<td>0</td>
</tr>
<tr>
<td>114 ≤ Loan term ≤ 125</td>
<td>-1.491***</td>
<td>-0.966*</td>
</tr>
</tbody>
</table>

Note: The table shows the results from a logit regression. The dependent variable for column (1) is one if the consumer prepaid the loan in the first two years. The dependent variable for column (2) is one if the consumer prepaid the loan in the first 120 days, which is our proxy for prepayment during the clawback period. The dependent variable for column (3) is one if the consumer prepaid the loan after the first 120 days but within the first two years. All regressions condition on loans of length greater than two years.
Table 18: Regressions of Prepayment Risk on Estimated Bias

<table>
<thead>
<tr>
<th></th>
<th>$B_{ij}^{O}$</th>
<th>$B_{ij}^{M}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.244)</td>
<td>(0.244)</td>
</tr>
<tr>
<td>Credit score, 100 points</td>
<td>-30.58***</td>
<td>-32.82***</td>
</tr>
<tr>
<td></td>
<td>(0.377)</td>
<td>(0.481)</td>
</tr>
<tr>
<td>Mileage, Tens of Thousands</td>
<td>4.524***</td>
<td>4.543***</td>
</tr>
<tr>
<td></td>
<td>(0.077)</td>
<td>(0.078)</td>
</tr>
<tr>
<td></td>
<td>(0.390)</td>
<td>(0.390)</td>
</tr>
<tr>
<td>Log Loan Amount</td>
<td>381.4***</td>
<td>384.5***</td>
</tr>
<tr>
<td></td>
<td>(1.253)</td>
<td>(1.272)</td>
</tr>
<tr>
<td>Average years of education</td>
<td>-3.807***</td>
<td>-3.791***</td>
</tr>
<tr>
<td></td>
<td>(0.335)</td>
<td>(0.335)</td>
</tr>
<tr>
<td></td>
<td>(1.482)</td>
<td>(1.482)</td>
</tr>
<tr>
<td>Estimated prepayment probability</td>
<td>-24.16***</td>
<td>-69.71***</td>
</tr>
<tr>
<td></td>
<td>(6.432)</td>
<td>(1.722)</td>
</tr>
<tr>
<td>Estimated early prepayment probability</td>
<td>96.93***</td>
<td>19.18***</td>
</tr>
<tr>
<td></td>
<td>(14.727)</td>
<td>(6.114)</td>
</tr>
<tr>
<td>Estimated late prepayment probability</td>
<td>-12.77**</td>
<td>-69.75***</td>
</tr>
<tr>
<td></td>
<td>(6.453)</td>
<td>(1.790)</td>
</tr>
</tbody>
</table>

**Note:** Bank, model, and state fixed effects also included, but not shown. Standard errors clustered at zip code level. Regressions drop observations in the top percentile of loan size. Early prepayment is defined as prepayment in the first 120 days. Late prepayment is defined as prepayment after 120 days but within the first 2 years. Estimated prepayment probabilities in the supervisory data are imputed using coefficient estimates from the CCP.
Table 19: $M'(\cdot)$, $B^O$, and $B^M$ by Car Model

<table>
<thead>
<tr>
<th>Model</th>
<th>$M'(\cdot)$</th>
<th>$B^O$</th>
<th>$B^M$</th>
<th>Model</th>
<th>$M'(\cdot)$</th>
<th>$B^O$</th>
<th>$B^M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ram</td>
<td>0.76</td>
<td>795.89</td>
<td>239.31</td>
<td>Chevrolet Equinox</td>
<td>0.81</td>
<td>499.41</td>
<td>169.74</td>
</tr>
<tr>
<td>Chevrolet Silverado</td>
<td>0.77</td>
<td>693.59</td>
<td>219.19</td>
<td>Ford Escape</td>
<td>0.79</td>
<td>493.89</td>
<td>170.68</td>
</tr>
<tr>
<td>Ford F-150</td>
<td>0.79</td>
<td>660.66</td>
<td>220.28</td>
<td>Ford Fusion</td>
<td>0.79</td>
<td>480.00</td>
<td>161.90</td>
</tr>
<tr>
<td>Kia Sorento</td>
<td>0.78</td>
<td>614.92</td>
<td>189.59</td>
<td>Toyota Camry</td>
<td>0.77</td>
<td>475.85</td>
<td>143.56</td>
</tr>
<tr>
<td>Ford Explorer</td>
<td>0.80</td>
<td>602.00</td>
<td>212.74</td>
<td>Hyundai Elantra</td>
<td>0.77</td>
<td>474.41</td>
<td>156.04</td>
</tr>
<tr>
<td>Jeep Gr. Cherokee</td>
<td>0.80</td>
<td>585.28</td>
<td>200.07</td>
<td>Chevrolet Malibu</td>
<td>0.80</td>
<td>450.21</td>
<td>154.90</td>
</tr>
<tr>
<td>Jeep Wrangler</td>
<td>0.81</td>
<td>576.16</td>
<td>201.04</td>
<td>Honda Civic</td>
<td>0.78</td>
<td>442.62</td>
<td>139.26</td>
</tr>
<tr>
<td>Hyundai Sonata</td>
<td>0.77</td>
<td>529.34</td>
<td>157.96</td>
<td>Chevrolet Cruze</td>
<td>0.81</td>
<td>437.70</td>
<td>144.71</td>
</tr>
<tr>
<td>Honda Odyssey</td>
<td>0.80</td>
<td>519.13</td>
<td>169.13</td>
<td>Honda Accord</td>
<td>0.78</td>
<td>437.70</td>
<td>150.26</td>
</tr>
<tr>
<td>Nissan Altima</td>
<td>0.79</td>
<td>507.10</td>
<td>159.52</td>
<td>Ford Focus</td>
<td>0.80</td>
<td>424.82</td>
<td>147.52</td>
</tr>
</tbody>
</table>

Note: The table shows for the 20 most common car models in the data the average estimates for the lower bound $B^O_i$, $B^M_i$ and $M(\cdot)_i$.

Table 20: Selected summary statistics of consumers’ misperception of financial charges.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\rho}$</td>
<td>0.77</td>
<td>0.67</td>
<td>0.74</td>
<td>0.78</td>
<td>0.82</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Note: Table gives summary statistics on the distribution of estimated values of $\rho$. 
Table 21: Demand Model Coefficients

<table>
<thead>
<tr>
<th>Make</th>
<th>Mean Utility Shift</th>
<th>Standard Error</th>
<th>Own Price Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chevrolet</td>
<td>-4.1</td>
<td>0.047</td>
<td>-4.2</td>
</tr>
<tr>
<td>Chrysler</td>
<td>-4.4</td>
<td>0.047</td>
<td>-4.2</td>
</tr>
<tr>
<td>Dodge</td>
<td>-4.8</td>
<td>0.044</td>
<td>-3.7</td>
</tr>
<tr>
<td>Ford</td>
<td>-4.0</td>
<td>0.047</td>
<td>-4.1</td>
</tr>
<tr>
<td>GMC</td>
<td>-2.7</td>
<td>0.073</td>
<td>-5.4</td>
</tr>
<tr>
<td>Honda</td>
<td>-3.5</td>
<td>0.047</td>
<td>-3.2</td>
</tr>
<tr>
<td>Hyundai</td>
<td>-4.0</td>
<td>0.047</td>
<td>-2.5</td>
</tr>
<tr>
<td>Jeep</td>
<td>-3.7</td>
<td>0.04</td>
<td>-4.1</td>
</tr>
<tr>
<td>Kia</td>
<td>-3.9</td>
<td>0.045</td>
<td>-2.7</td>
</tr>
<tr>
<td>Mazda</td>
<td>-4.1</td>
<td>0.048</td>
<td>-2.7</td>
</tr>
<tr>
<td>Nissan</td>
<td>-4.3</td>
<td>0.045</td>
<td>-2.9</td>
</tr>
<tr>
<td>RAM</td>
<td>-3.7</td>
<td>0.051</td>
<td>-4.9</td>
</tr>
<tr>
<td>Subaru</td>
<td>-3.8</td>
<td>0.044</td>
<td>-2.7</td>
</tr>
<tr>
<td>Toyota</td>
<td>-3.8</td>
<td>0.047</td>
<td>-3.4</td>
</tr>
<tr>
<td>Volkswagen</td>
<td>-4.2</td>
<td>0.052</td>
<td>-2.8</td>
</tr>
</tbody>
</table>

Note: The table shows how mean utility is shifted for different types of brands, the standard errors of those mean shifts, as well as own-price elasticities for different makes.
Table 22: Regressions of observables on markup (new cars only)

<table>
<thead>
<tr>
<th></th>
<th>Loan Markup</th>
<th>Loan Markup</th>
<th>Loan Markup</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Monthly Income</td>
<td>-0.0158***</td>
<td>-0.0242***</td>
<td>-0.0176***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Log Car Price</td>
<td>0.166***</td>
<td>0.397***</td>
<td>0.357***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Credit score, 100 points</td>
<td>-0.0989***</td>
<td>-0.0840***</td>
<td>-0.0860***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Mileage, Tens of Thousands</td>
<td>0.0228***</td>
<td>0.0286***</td>
<td>0.0217***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Log Loan Amount</td>
<td>-0.133***</td>
<td>-0.241***</td>
<td>-0.223***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Buy Rate</td>
<td>-0.0980***</td>
<td>0.000952</td>
<td>-0.000826</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
</tbody>
</table>

Model Fixed Effects:  Yes   Yes   Yes
Term Fixed Effects:    Yes   Yes   Yes
Month-Year Fixed Effects: Yes   Yes   Yes
Bank Fixed Effects: No   Yes   Yes
Zip Fixed Effects: No   No   Yes
Adjusted $R^2$: .049   .08   .11

Note: The table shows three different regressions of dealer markup on customer observables that are available to the lender at the time of the buy-rate auction. Even though these regressions are very saturated, only a small percent of the variation in markups can be explained by the independent variables.