# Contract Pricing in Consumer Credit Markets*

Liran Einav, Mark Jenkins and Jonathan Levin†

March 2010

**Abstract.** We study pricing and contract design in the subprime auto sales market. We develop a model of the demand for financed purchases that incorporates both adverse selection and moral hazard effects, and estimate the model using detailed transaction-level data. We use the model to quantify selection and repayment problems and show that different elements of the financial contract, in particular car price and required down payment, generate very different pricing trade-offs. We also evaluate the returns to credit scoring that allows sellers to customize financing terms to individual applicants. Our empirical approach shows how standard tools for analyzing demand and supply in traditional product markets extend to contract markets, where agreement and performance are separated in time and firms care about both the quantity and quality of demand.

*We have benefited from the suggestions of the co-editor and three referees, and from the comments of many seminar participants. We acknowledge the support of the Stanford Institute for Economic Policy Research, the National Science Foundation (Einav and Levin), and the Alfred P. Sloan Foundation (Levin).

†Einav and Levin: Department of Economics, Stanford University, and NBER; Jenkins: Wharton School, University of Pennsylvania. E-mail addresses: leinav@stanford.edu; mjenk@wharton.upenn.edu; jdlevin@stanford.edu.

# 1 Introduction

This paper is an empirical study of pricing in a consumer credit market. In credit markets, lenders offering credit terms have imperfect information about future events. The terms they offer can affect the subsequent behavior of borrowers and the payoffs of both parties. These features of credit markets give rise to selection and incentive problems for which effective contract design must account. Our analysis attempts to illuminate how different contract dimensions have differential effects on the quantity and quality of demand, leading to different pricing trade-offs. We also show, and quantify, how improved information about potential customers can lead to better screening, better matching of contracts to customers, and higher profits. To do this, we adapt standard empirical tools for demand and pricing analysis to contract markets and apply them to consumer lending.

Our empirical application makes use of rich transaction-level data from a large auto sales company. The company specializes in financing consumers with low incomes or poor credit histories, the so-called "subprime" market. This is an ideal setting to study pricing and contract design in the presence of informational problems. There is substantial default risk and borrower heterogeneity, so the ability to originate profitable loans depends crucially on designing offers that attract lower risk borrowers and encourage repayment. As offers to customers can vary on multiple dimensions – car price, required down payment, interest rate, and loan length – we can investigate the screening and incentive roles played by different contract terms. Pricing possibilities are further enriched by the ability of lenders to price based on a borrower's credit history, allowing us to study the value of credit scoring and risk-based financing. The market is also of substantial economic interest given the expansion and subsequent turmoil in subprime lending in recent years. As we discuss below, our results offer a number of useful insights into these events.

We develop our analysis in several steps. In Section 2, we describe the simple economics of contract pricing, relating it to the traditional pricing problem of a seller with market power. One of the key points is that pricing incentives depend on a small number of demand and revenue elasticities, and on the relationship between marginal and inframarginal buyers. Estimating these quantities from data requires a joint model of consumer demand and transaction outcomes. After describing the empirical setting in Section 3, we develop such a model. We first describe in Section 4 a theoretical model of consumers that provides insight into how standard optimizing behavior can

rationalize certain features observed in the data. This helps guide the empirical model we develop in Section 5. Because our main focus is on pricing, the empirical model focuses on capturing the market-level elasticities that matter for pricing, without trying to directly map these elasticities to the underlying consumer primitives.

Our estimates reveal several striking features of the market, starting with the observation that consumer purchasing decisions are highly sensitive to down payment requirements. A natural interpretation of this finding, and one that is consistent with our theoretical model, is that consumers are constrained by cash-on-hand. This makes them sensitive to current expenditures, especially compared to price increases that can be financed. Indeed, we find that price increases have relatively little effect on the decision to purchase a car or on how much to put down; instead they translate mostly into larger loans. The estimates also allow us to identify separately the selection effects of pricing changes, and their effect on borrower repayment.

The central part of our analysis uses the demand estimates to analyze contract design: how different aspects of the financial contract affect the quantity and quality of loan originations. Down payment requirements generate a trade-off between loan volume and loan quality. A down payment requirement places a limit on borrower leverage. For some buyers, the limit is irrelevant because they intend to make a substantial down payment in any case. A buyer who does not intend to pay more than the minimum down, however, either has to come up with extra cash or forego the purchase if the down payment requirement is raised. Our estimates suggest that these "marginal" buyers tend to have trouble with repayment later on, so screening them out improves the composition of borrowers. Once the firm has more information about individual borrowers, the optimal down payment schedule exhibits the intuitive property that safer borrowers are allowed to borrow more, a feature shared, to a lesser extent, by the firm's observed contracts.

In contrast to down payment requirements, changes in car prices have a very different effect. In our data, increases in car prices have little effect on sales volume or down payments. Their primary effect is to increase loan size and consequently the monthly payments. This means more revenue while borrowers are making payments but also a higher rate of default. We describe how optimal prices resolve this trade-off. Perhaps surprisingly, the optimal policy calls for reduced prices for high-risk borrowers whose likelihood of repayment is more sensitive to loan size.

We also use our model and data on the firm's costs to assess whether the firm's loan offerings were profit-maximizing, or at least consistent with a weak set of revealed preference assumptions.

Here we find that relative to directly measured costs, the firm appears to have been relatively conservative in making loans. We back out an implied cost of lending that rationalizes the firm's decisions. One interpretation of this cost, which is roughly $1,500 per loan, is that the firm's lending practices exhibited a concern for aggregate risk, a justifiable concern given the events that have occurred subsequent to our data (which ends in spring 2006).

The last part of the paper uses the pricing model to quantify the value of using individual information. The advent of sophisticated credit scoring has revolutionized consumer credit markets over the last quarter century. Because the firm sets car prices independent of the characteristics of individual customers, we focus on the value of credit scoring to set minimum down payment requirements. As a comparison, we compute the firm's per-customer profits if it could not distinguish at all between customers, and if it had perfect information about them at the time of sale. Compared to the case of uniform minimum down requirement, we find that risk-based financing can increase profits by 22%, and perfect information about consumer characteristics would increase expected profits by 96%.[1] We also make a rough attempt to quantify how proprietary information on the part of an incumbent lender can create a barrier for potential entrants by leaving them with an adversely selected pool of borrowers.

The time period we study was notable for the dramatic increase and subsequent fall in subprime lending. Subprime mortgage originations rose from $65 billion in 1995 to over $650 billion in 2005, before dropping sharply. During the same period, loan originations by AmeriCredit, the largest subprime auto lender in the United States, grew from $230 million to over $5 billion. Losses on some of these subprime loans helped to precipitate the recent financial crisis. This episode has raised questions about the operation of subprime credit markets, the design of loan contracts in these markets, and the potential for informational problems to create market failures. Although we are cautious about extrapolating too far from our specific setting, the mechanisms we identify are natural candidates to help explain these broader events. In particular, our findings indicate how sharp declines in down payment requirements – an important shift in mortgage lending – can attract latent demand from consumers eager to take on additional leverage. This leads to an increase in loan originations, but increases default risk and generates a worsening borrower pool. It is also possible that the incentives to reduce down payments and boost originations were higher in

[1] In a related paper (Einav, Jenkins, and Levin, 2009) we assess the value of credit scoring by applying a more descriptive "before and after" comparison using data from before the advent of credit socring. The results are qualitatively similar, although the current study is much more explicit about the mechanism.

mortgage lending than in our setting because lenders were not forced to retain a substantial stake when they sold loans into the secondary market.

In emphasizing the importance of borrower leverage, our study is related to some recent analyses of mortgage lending (e.g. Geanakoplos, 2009; Mayer, Pence, and Sherlund, 2009). In emphasizing the importance of adverse selection and borrower incentives, our study relates to work by Ausubel (1991, 1999), Edelberg (2004, 2006), and Karlan and Zinman (2008, 2009), who provide evidence on asymmetric information in a variety of consumer credit markets. Our specific analysis relates most closely to Adams, Einav, and Levin (2009) and Einav, Jenkins, and Levin (2009), who use data from the same lender. The former study provides descriptive evidence on subprime consumers and documents informational frictions, but without examining the pricing implications. The latter compares lending before and after the advent of credit scoring, an analysis that complements the last subsection of this paper.

From a conceptual standpoint, our empirical approach is related to recent models of demand in insurance markets (see Einav, Finkelstein and Levin, 2010, for a recent review). Similar to that research, we build a model of consumer demand that integrates choices between contracts (e.g. how large a loan to take, how much insurance to purchase) with subsequent contract outcomes. In this paper, however, we take the additional step of marrying the demand system to supply-side decisions about contract design, and showing how concerns about asymmetric information and incentive provision guide pricing decisions.[2]

## 2 Simple Economics of Contract Pricing

We begin by describing the simple economics of contract pricing. Our goal here is to highlight how optimal pricing incorporates concerns about selection and incentive effects, and what empirical objects are most relevant for the seller's pricing problem.

Consider a unit mass of potential buyers, each described by a vector of characteristics . Let $\mathsf{F}$ denote the distribution of consumer characteristics in the population. In our empirical application, these characteristics will include consumer demographics, as well as controls for geographic location

down payment, an interest rate, payment schedule for the corresponding loan, and so forth. Let $\mathsf{u}(\ ;\ )$ denote the expected utility a type- consumer assigns to a contract , and let $\overline{\mathsf{u}}(\ )$ denote the expected utility the consumer assigns to not accepting the contract. In Section 4, we describe an explicit dynamic model of consumer behavior that gives rise to an expected utility of making or not making a financed car purchase. A type- consumer will select the contract if $\mathsf{u}(\ ;\ ) \geq \overline{\mathsf{u}}(\ )$.

If a consumer enters into the contract, subsequent events will lead to an outcome $\mathsf{y}(\ ;\ )$, and to a seller's profit, $(\ ;\mathsf{y})$, that depends on the outcome and on the contract itself.[3] In our empirical application, the outcome is the duration of time over which the consumer makes loan payments, and the seller's profit depends on the repayment duration, as well as on the cost of the car, the amount of financing, the interest rate, and the loan length.

We now take up the seller's pricing problem. Suppose the seller offers a single contract to the population of potential buyers. To the extent that the seller might vary pricing based on observable consumer characteristics, we can think of an analogous problem applying for each observed category of potential buyers. If the firm offers a contract , its total sales will be

$$\mathsf{Q}(\ ) = \int \mathbf{1}\{\mathsf{u}(\ ;\ ) \geq \overline{\mathsf{u}}(\ )\}\mathsf{dF}(\ ). \tag{1}$$

The firm's problem is to choose contract terms to maximize expected profit:

$$\max_{\phi \in \Phi} \Pi(\ ) = \mathsf{Q}(\ ) \cdot \mathbb{E}[\ (\ ;\mathsf{y}) \mid \mathsf{u}(\ ;\ ) \geq \overline{\mathsf{u}}(\ )]: \tag{2}$$

The model incorporates two important elements of pricing under asymmetric information. First, there can be selection effects: buyer characteristics that affect the purchasing decision may also be relevant for later outcomes. Second, there can be incentive effects: the terms of the contract may affect the outcome of the transaction. Permitting borrowers to take more leveraged loans, for instance, might increase the number of low-quality borrowers (adverse selection), and reduce the incentive for borrowers to repay (moral hazard).

To see how optimal pricing must address these issues, suppose that the contract has a single dimension, so that is a scalar variable. Our empirical application focuses on both car price and

[3] Note that the model allows for the consumer to be initially uncertain about some of her own characteristics, say her future income, that subsequently will affect outcomes. If $\zeta$ includes characteristics that are initially unknown to the buyer, then $u$ and $\overline{u}$ will not vary with these components of $\zeta$, although they may affect the contract outcome $y$, and hence the seller's profit $\pi$.

required down payment, but a single dimension suffices to illustrate the main ideas. Suppose that $u$ is strictly decreasing in $\quad$, and the functions defined above are differentiable. The effect of a small increase in $\quad$ is then

$$\begin{aligned} \frac{d\Pi(\ )}{d} &= \frac{dQ(\ )}{d} \cdot \mathbb{E}[\ (\ ;y(\ ;\ ))\mid u(\ ;\ ) = \bar{u}(\ )] \\ &\quad + Q(\ ) \cdot \mathbb{E}\left[\frac{d\ (\ ;y(\ ;\ ))}{d} \mid u(\ ;\ ) \geq \bar{u}(\ )\right]: \end{aligned} \tag{3}$$

The first term reflects the loss of "marginal" customers. The cost of losing these customers depends on their profitability, e.g. a marginal buyer could be more or less profitable than an average buyer. The second term reflects the change in profit from inframarginal buyers. In a traditional product market, a one dollar increase in price means a one dollar increase in revenue, and no change in cost. Here, the effect can be more complex because a change in the contract terms may alter the incentives for performance, for instance making default more or less likely.

Another way to express the first order condition for optimal pricing is in terms of elasticities. To do this, let $R(\ ) = \mathbb{E}[\ (\ ;y)\mid u(\ ;\ ) \geq \bar{u}(\ )]$ denote the seller's expected profit (or net revenue) conditional on sale. Expected profit is then $\Pi(\ ) = Q(\ )R(\ )$, and we can write the first order condition for optimal pricing as

$$\frac{R(\ )=}{dR(\ )=d} = \frac{-Q(\ )=}{dQ(\ )=d}: \tag{4}$$

If pricing is optimal, the firm equates the inverse elasticity of net revenue with the inverse elasticity of demand. In a standard product market, the relevant contract characteristic is price ($\ = p$), and $(p;\ ) = p - c$, so equation (4) reduces to the standard formula equating the markup $(p - c)=p$ with the inverse demand elasticity. Contract pricing follows the same principles, only the relevant revenue elasticity must account for selection and incentive effects.

From an empirical perspective, it is clear that the objects of interest for pricing are the sensitivities of demand and outcomes to the relevant contract terms, and the relationship between marginal and inframarginal buyers. Because changes in pricing have both a selection effect (altering the composition of buyers) and a direct effect (altering the incentives of purchasers), it is natural to consider an empirical strategy based around traditional selection methods. This is the approach we pursue in Section 5, combining a model of contract selection (a loan applicant chooses whether or not to purchase and how much of the purchase to finance) with a model of loan repayment, and

dollars. Credit categories are assigned using a proprietary credit scoring algorithm developed by the firm. Buyers are offered financing on the remainder of the car price, although by making a larger down payment they can reduce their financing. The offered loans typically have annual interest rates of 25-30%, and the rate is often set at the state-mandated maximum. Loans are expected to be repaid over three to four years.

Just over one-third of the loan applicants in our data purchase a car. Forty-three percent of these buyers make exactly the minimum down payment. The remainder make somewhat larger down payments, but frequently not much larger. Fewer than ten percent make a down payment that exceeds the minimum by a thousand dollars. The average down payment is around $1,000, so that after taxes and fees the average loan size is a bit under $11,000. This translates into monthly payments on the order of $400.

A large portion of borrowers subsequently default. Our data ends before the last payments are due on some loans, but of the loans with uncensored payment periods, 61% end in default. Defaults tend to come early in the loan period. Nearly half the defaults occur before a quarter of the payments have been made, and nearly 80 percent in the first half of the loan length. A notable feature of the data is that the value recovered following a default is rather low. For 22 percent of defaults we observe, no recovery is made at all, sometimes because the car has been in an accident or stolen. Even when the recovery value is positive, the average present value of the recovery is less than $1,600, compared to an average car cost of around $6,000. Jenkins (2009) explores this aspect of subprime lending in more detail.

From the lender's perspective there is a large difference between loans that pay off and loans that end in default. Figure 1 plots the distribution of per-sale profits, computed by adding up the firm's revenue from the down payment, loan payments and recoveries, and subtracting off the cost of the car and collections. The distribution is highly bimodal. Paid loans are quite profitable. Defaulted loans are not. Note that these profit calculations, and ones we will make subsequently in the paper, are made as if loans are held to maturity. In fact, loans during the sample period were securitized and sold into the secondary market. We return to this point when it becomes relevant in our discussion of optimal pricing.

Figure 1 raises the obvious question of whether it is possible ex ante to predict likely defaulters. To provide a rough assessment, we group buyers into "high," "medium," and "low" risk using the company's proprietary credit categories. In the sample of uncensored loans, the default rate for

high risk buyers is 71 percent, compared to 44 percent for the low risk buyers. This suggests that the credit score is highly predictive, a conclusion borne out by our later estimates that include more detailed controls. Figure 2 presents a graphical version of this point, plotting the likelihood of default against the underlying credit score, again using the sample of uncensored loans.

Figure 2 also addresses a more nuanced question of whether, conditional on observed credit risk, a buyer's actions at the dealership reveal information about the subsequent probability of default. The two other lines in Figure 2 plot default rates for two groups of buyers: those who make a minimum down payment and those who voluntarily made a larger down payment. The default rate for the former group is 67 percent compared to 56 percent for the latter group. There are two explanations for the correlation between financing decisions and default rates (which as we will see survives the addition of detailed controls for buyer and car characteristics, as well as fixed effects for dealership and time periods). The first is selection: high risk buyers choose to make low down payments and leverage aggressively. The second is causal: buyers who leverage more aggressively end up with larger debt burdens, and this makes default more likely. Our empirical model below disentangles these channels and verifies the importance of both.

### 3.2 Making Use of the Data

In modeling consumer demand and repayment, and analyzing pricing decisions, we confront a number of decisions about how to best use the rich data. Many of these decisions are informed by the institutional details of the setting, visits to the firm, and discussions with their employees and executives. This subsection provides some background.

The first issue concerns the process by which customers arrive at a dealership and enter our data. The company attracts customers primarily through street presence and referrals. Referrals often come from car dealerships that are unwilling to offer financing to borrowers with very poor credit histories. The market segmentation means there is little reason for these dealers to be aware of, or respond to, pricing changes at the firm we are studying. Also, the company's financing terms are not publicly posted and are tailored to a borrower's customized credit score. To the extent that there are dealerships or lenders that might compete to offer financing to these consumers, it would be extremely difficult for them to assess the company's current financing terms. As a result, our later analysis will treat both the customer arrival process and the outside options of customers as independent of the company's pricing decisions, conditional on location, year and month dummies.

When customers do arrive at a dealership, they face decisions about whether to purchase, what car to purchase, and how much of the transaction they should finance. In this paper, we focus on the purchase and financing decision, rather than on car selection. An important reason for this is that we are interested mainly in the financial aspects of the market rather than consumer preferences over used vehicles. A second reason is that many customers in fact do not have a very rich set of alternatives. The company has an algorithm to determine whether a buyer is eligible to receive financing for a given car, and this often leads to tight restrictions. So while one could enrich the demand analysis to incorporate car choice, our view is that the additional complexity would not be matched by sharper insights into selection, consumer behavior, and optimal pricing. Adams, Einav, and Levin (2009) provide evidence and further discussion on this point.

In modeling and estimating consumer demand, we face a standard problem related to non-purchasers. Although we have remarkably detailed demographic and credit history data on applicants who do not purchase a car, and can compute their potential financing terms, we do not observe their preferred car on the lot or the discount from the car's list price they might have negotiated. The obvious remedy, and the one we adopt, is to impute the missing data. For each non-purchasing applicant, we select at random an applicant in the same credit and income category who purchased a car in the same week at the same dealership, and assign the non-purchaser the same car and negotiated price.

Finally, an important feature of the data is the observed variation in the firm's pricing policy. Our data includes a number of sharp changes in the company's schedules for list prices and required down payments. We describe these changes in Section 5, along with additional discontinuities in the pricing schedules that allow us to estimate demand and revenue sensitivities. We also observe some variation in interest rates and loan lengths. For instance, the length of loans has expanded somewhat over time, and interest rate caps vary by state, creating variation in the company's rates. We control for this variation, but feel somewhat less confident in our ability to identify convincingly how changes in these financing terms affect the quality and quantity of demand. As a result, we focus our pricing analysis mainly on down payment and car price.

We should note that while the data and modeling decisions we have described were informed by institutional specifics, we have subjected many of them to extensive robustness checks. For instance, we have tried estimating the quantity of car purchases in an aggregated fashion that allows the number of applicants to respond to price changes. We have looked at whether changes

in pricing affected the distribution of cars being purchased. We have considered many variations on the imputation process for non-purchasers. We have considered demand estimates that focus on narrow windows around price changes, or rely only on discontinuities in the pricing schedules. Many of these variations and checks are detailed in Adams, Einav, and Levin (2009).

# 4 A Model of Consumer Behavior

In this section, we develop a stylized model of consumer behavior. The model incorporates some of the main features of the market and the central consumer decisions: whether to purchase a car, how large a down payment to make, and whether to continue making payments over time or default on the loan. We model these decisions in textbook fashion, assuming that consumers have standard preferences, are forward-looking, and form rational expectations about the likelihood of future default. These assumptions are perhaps debatable in a setting where individuals are entering into high-cost loans with substantial chance of default. This is one reason why we do not impose strong rationality assumptions in the statistical model we estimate in the next section.

Nevertheless, the model we develop here provides useful guidance in choosing our eventual empirical specification, and is a way to gain insight into how a standard model of consumer behavior can match the main facts in the data. In particular, the model of this section highlights how liquidity constraints can lead to consumer demand being much more sensitive to a higher down payment requirement than to an increase in car price that can be borrowed. It also illustrates why buyers paying exactly the minimum down can be appropriately viewed as marginal buyers, and also as worse default risks than inframarginal buyers. These ideas are useful later in interpreting our empirical results.

## 4.1 Modeling Consumer Demand

Our model begins at the point when a consumer enters a dealership and is offered a car and possible financing. Let $p$ denote the car price and $d$ the required down payment. If the consumer chooses to purchase, she can choose any down payment $D \in [d; p]$ and borrow the remaining $L = p - D$. Loans carry a monthly interest rate $z$ and a loan length of $T$ months. For a loan of size $L$, the monthly payment is $m(L) = (zL) = \left(1 - (1+z)^{-T}\right)$. In each subsequent month, the individual chooses whether to make the loan payment. Failure to pay means default, and the car is repossessed.

An individual incurs a non-monetary cost $'$ from default, which we think of as the effect on the individual's credit record. In practice, default and repossession may not be immediate, but we view this as a reasonable modeling simplifications.

We characterize individuals at the time of purchase along two dimensions. Let $y_0 > 0$ be the individual's "liquidity" or available cash, and $v_0 > 0$ be the monthly use value the individual assigns to the car. We assume that the car value declines deterministically, so that $v_t = (1 - \ )\, v_{t-1}$. This is a simple way to capture asset depreciation. Consumer liquidity evolves stochastically, so v

ing the car, she will choose a down payment so that

$$\mathsf{U}(\mathsf{y}_0;\mathsf{v}_0) = \max_{D \geq d} (\mathsf{v}_0 + \ln(\mathsf{y}_0 - \mathsf{D}) + \quad \mathbb{E}\left[\mathsf{U}_1(\mathsf{y}_1;\mathsf{v}_1;\mathsf{p} - \mathsf{D}) \mid \mathsf{y}_0\right]); \tag{7}$$

and will optimally purchase if $\mathsf{U}(\mathsf{y}_0;\mathsf{v}_0) \geq \overline{\mathsf{U}}(\mathsf{y}_0)$. For later purposes, it is useful to let $\mathsf{D}^*(\mathsf{y}_0; \quad _0)$ denote the unconstrained solution to the maximization problem in equation (7), i.e. the solution ignoring the constraint that $\mathsf{D} \geq \mathsf{d}$. Note that if $\mathsf{y}_0 < \mathsf{d}$, the consumer simply doesn't have enough cash to purchase the car. In this case, we can set $\mathsf{U}(\mathsf{y}_0;\mathsf{v}_0) = -\infty$.

The model, which we analyze in more detail in Appendix A (which is available online), has a number of intuitive properties. At month $\mathsf{t}$, a consumer with an active loan will make the required payment if $\mathsf{y}_t$ is above a threshold $\mathsf{y}_t^*(\mathsf{v}_t;\mathsf{L})$. A higher car value and smaller loan size both make payment more likely. We can define the length of repayment as a random variable $\mathsf{S}$ equal to $\max\{\mathsf{t} : \mathsf{y}_\tau \geq \mathsf{y}_\tau^*(\mathsf{v}_\tau;\mathsf{L})$ for all $\quad \leq \mathsf{t}\}$. Conditional on loan size, the repayment period will be longer (stochastically) for a consumer with higher car value or greater initial liquidity. Conditional on initial car value and liquidity, a larger loan reduces the repayment period. That is, larger loans make default more likely. The model also generates intuitive predictions about purchasing and borrowing. Consumers with higher car values are more likely to purchase, as are consumer with greater initial liquidity. And as we will see momentarily, consumers with low liquidity can prefer to make minimal down payments, leading to larger loans and further increasing their odds of default.

$\mathsf{p}$ equal to \$11;000, a required down payment $\mathsf{d}$ equal to \$1;000, an annual interest rate of 29:9%, and a loan duration of 42 months. We assume that loan payments are due monthly (here to reduce computation we depart slightly from the data where bi-weekly payment is more common). We assume car value depreciates at an annual rate of 0:88 (so $= 0{:}989$), and that individuals have an annual discount rate of 0:75 (so $= 0{:}976$).

Finally, we calibrate the other parameters of the model – $_v$, $_y$, $_v$, $_y$, $_{vy}$, $_\varepsilon$, and ' – to match some key moments in the data. Specifically, we choose values for these seven parameters so that the model predictions match the following seven moments in the data: the probability of sale (0:345), the probability of paying exactly the minimum down payment conditional on sale (0:427), the average additional down payment paid by those who make more than the required minimum (\$577), the default probability conditional on purchase (0:608), the average fraction of payments made conditional on default (0:344), the semi-elasticity of demand with respect to a \$100 increase in the minimum down payment ($-0{:}022$), and the semi-elasticity of demand with respect to a \$100 increase in the price of the car ($-0{:}002$). The calibrated model is able to fit these moments relatively well, and does a reasonable job matching some other features in the data such as the distribution of default times. The appendix describes the calibrated parameters, the associated computational details, and the model fit in more detail.

*Graphical Intuition.* Figure 3(a) provides intuition for the model implications by characterizing the purchase and down payment decisions in the space of initial liquidity $\mathsf{y}_0$ and car utility.$\mathsf{v}_0$. Each point in this space represents a possible consumer at the time of purchase. The black solid lines divide individuals into three groups: non-purchasers, purchasers who make exactly the minimum down payment, and purchasers who make a larger down payment. The grey lines show predicted default rates with each line corresponding to an iso-default curve.

The figure highlights some basic properties of the model. Individuals in the bottom-left region are non-purchasers. These individuals either assign relatively low value to the car (low $\mathsf{v}_0$) or have little available cash (low $\mathsf{y}_0$). Some of these individuals have such low liquidity, $\mathsf{y}_0 < \mathsf{d}$, that the required down payment is infeasible. Others have enough liquidity to purchase, but find it costly due to their high marginal utility from other consumption, and realize that they are unlikely to be able to make payments later. The second group of individuals, in the middle region, purchase and make the minimum down payment. Many of these individuals have a relatively high car value, but are sufficiently illiquid that they would find it costly to come up with additional cash, and also risky

given the high likelihood of default. The final group of individuals are those with high liquidity and high car utility. These individuals make additional down payments above the minimum, and have the lowest rate of subsequent default. So a negative correlation between additional down payment and default, which as noted above is an important feature of the data, arises naturally in the model.

Figures 3(b) and 3(c) illustrate how changes in the minimum down payment and the price of the car affect individual behavior, with the scatter plots in both figures representing draws of individuals based on the calibrated parameters. In Figure 3(b) we consider a substantial \$500 increase in the required down payment. This shifts both solid curves up and to the right. Fewer individuals purchase, and the individuals who cease to purchase (the "marginal" purchasers) are almost entirely individuals with low initial liquidity who would make a \$1,000 minimum down payment but are unwilling to make the higher down payment. These individuals represent high default risks relative to an average buyer. So the model predicts than an increase in the minimum down payment will reduce the number of loans being made, but improve loan performance.

Figure 3(c) shows the effect of a substantial $\$2;000$ increase in car price. While the solid curves shift up in the same direction, they shift much less, indicating that the increase in car price has only a small effect on which consumers choose to purchase, and on which consumers pay more than the minimum down. An important point here is that in order to rationalize the car price and minimum down payment demand elasticities (two of the moments used to calibrate the model), the model needs a relatively high distribution of car values and a relatively low distribution of consumer liquidity. An implication is that in response to a price increase, consumers primarily respond by taking extra financing because of their tight liquidity position. The larger loans that result increase the rate of subsequent default. So the model predicts that raising car price generates a trade-off: higher monthly payments from borrowers but higher likelihood of default.

*Summary of the Insights.* The theoretical model we have described provides several useful insights. It shows how a lack of consumer liquidity helps to explain the observed purchasing and borrowing patterns. It shows why larger loans are less likely to be repaid, both due to a direct causal effect (increasing the debt of any given consumer makes it less likely she will repay), and a selection effect (consumers with less liquidity take larger loans, and these consumers are worse default risks). It also shows why "marginal" purchasers may be virtually all consumers who, conditional on purchase, would borrow the maximum amount. While in principle there could be marginal consumers with higher liquidity and lower value for the car, the low elasticity of demand

with respect to the price of the car implies that such individuals are not common. Finally, the model provides a way to understand how direct financial constraints and strategic decision-making can both affect purchasing and subsequent repayment.

### 4.3 Moving from Theory to Empirics

Our next step is to move from this theoretical example to an empirical model of demand and repayment that will allow us to study issues of optimal loan pricing. The observed decisions in the data are the choice of whether or not to purchase, the down payment decision, and repayment. The model above places some structure on these decisions. Given consumer and offer characteristics, the consumer will purchase if $\mathsf{U} \geq \overline{\mathsf{U}}$, will make a down payment $\mathsf{D} = \min\{\mathsf{D}^*; \mathsf{d}\}$, and subsequently will make payments for some period $\mathsf{S}$ that depends (stochastically) on characteristics and the loan size $\mathsf{L} = \mathsf{p} - \mathsf{D}$. An important point which we will emphasize below is that the joint distribution of $\mathsf{U}; \overline{\mathsf{U}}; \mathsf{D}^*; \mathsf{S}$ is precisely the information needed to analyze pricing, making this the central empirical object in our analysis.

In modeling the empirical quantities we have just described, we face a choice. One possibility is to work directly from the theoretical model above, expanding the model to allow observed and unobserved characteristics to affect consumer liquidity and car value, and hence purchasing, down payment, and default. This approach would have the virtue of letting us draw detailed inferences about "deep" primitives such as individual discount rates and cash flows, perform welfare calculations, and analyze a potentially rich set of counterfactuals, perhaps even counterfactuals well outside our data. At the same time, it would be computationally demanding, and would require us to maintain assumptions about consumer foresight and rationality that many economists would find strained in our specific empirical setting. In addition, it would yield estimates of underlying parameters that might be interesting, but are not necessarily needed for our primary task of studying pricing decisions.

The alternative is to directly specify a model for the empirical objects $\mathsf{U}; \overline{\mathsf{U}}; \mathsf{D}^*; \mathsf{S}$. Such specification is informed by the theoretical model but is not directly derived from underlying assumptions about intertemporal utility maximization. This is the approach we follow in the next section. An advantage of this approach, apart from the ability to bypass difficult-to-assess rationality and foresight assumptions,[5] is that it lets us focus directly on the empirical objects that are relevant for

[5] For example, we allow for correlation between desired borrowing and propensity to default. This correlation

$= (\mathsf{x}^c; \mathsf{p}; \mathsf{d})$. The vector $\mathsf{x}^c$ includes the characteristics of the applicant's preferred car on the lot, including its cost, as well as the offered interest rate and loan length. The remaining terms are the car price $\mathsf{p}$ and the required down payment $\mathsf{d}$. It is useful to let $\mathsf{x} = \left(\mathsf{x}^a; \mathsf{x}^d; \mathsf{x}^c\right)$ denote the complete vector of observed characteristics other than price and minimum down payment. Using our notation from Section 2, a potential transaction is described by $(\ _i;\ _i) = (\mathsf{p}_i; \mathsf{d}_i; \mathsf{x}_i;\ _i; \mathsf{"}_i;\ _i)$.

*Purchase and Financing Decisions.* We write the purchasing decision as

$$\mathsf{Q}_i = 1 \qquad \Leftrightarrow \qquad \mathsf{U}(\ _i;\ _i) \geq \overline{\mathsf{U}}(\ _i); \tag{8}$$

where $\mathsf{U}(\ _i;\ _i)$ is the utility from purchasing and $\overline{\mathsf{U}}(\ _i)$ the utility from not purchasing. Because the data only can identify the difference $\mathsf{U} - \overline{\ }$

presenting the model.

*Loan Repayment.* Once a purchase is made and a loan is extended, buyers make payments on a regularly scheduled basis, most often bi-weekly but sometimes more or less often. Because there is some variation in this schedule, and because we sometimes observe deviations such as off-schedule or partial payments, it is convenient to work with a continuous model of repayment. Specifically, we posit that consumer $\mathsf{i}$ will make a fraction of payments $\mathsf{S}_i \in [0; 1]$, where:

$$\mathsf{S}_i = \begin{cases} \mathsf{S}_i^* = \exp\left(\mathsf{x}_i' \quad_x + (\mathsf{p}_i - \mathsf{D}_i) \quad_L + \quad_i\right) & \text{if } \mathsf{S}_i^* \leq 1 \\ 1 & \text{if } \mathsf{S}_i^* > 1 \end{cases} : \tag{12}$$

The parameter $\quad_L$ is an important one in this specification, as it defines the sensitivity of the repayment duration to changes in loan size. Increases in car price and changes in the down payment will both affect the loan size $\mathsf{p}_i - \mathsf{D}_i$. The larger the value of $\quad_L$, the greater the causal link between financing decisions and subsequent repayment. Of course, correlation between , ", and , will also connect choices at the time of purchase and loan performance.

Note that for loans that occur later in our sample, there is some additional censoring because we do not observe the full repayment period. We account for it in estimating the model, but we defer a complete discussion of this detail to Appendix B (available online), where we also provide additional details about estimation.

*Price Negotiation.* As mentioned in Section 3, the company sets a list price for each car, but customers have some ability to negotiate at the dealership. To model price determination at the dealership level, we specify a simple linear relationship between the negotiated price $\mathsf{p}_i$ and the list price $\mathsf{l}_i$:

$$\mathsf{p}_i = \mathsf{x}_i' \quad_x + \quad_l \mathsf{l}_i + \quad_i : \tag{13}$$

Although our nonlinear demand and repayment model requires the list price to satisfy stronger assumptions for consistent estimation, a rough way to view the pricing equation is as a "first stage" regression, where list price $\mathsf{l}_i$ is used as an instrumental variable for the possibly endogenous covariate $\mathsf{p}_i$. Here the possible endogeneity is captured by allowing $\quad_i$, the unobservable aspect of negotiation, to be correlated with $\quad_i$ and $"_i$, the buyer's unobserved information at the time of purchase, or even with $\quad_i$, which is unknown by the buyer but may be inferred by the dealer selling

the car. In estimating the pricing equation, the quantity $_{l}$ is of particular interest. When we consider optimal price-setting, we consider the company influencing prices through its choice of list price, and $_{l}$ defines the rate at which list price changes pass through to actual transaction prices.

*Stochastic Assumptions.* To close the model, we specify a stochastic structure for the unobservables ( $_{i}$; $''_{i}$; $_{i}$; $_{i}$). We assume that they are normally distributed, as follows:

$$\begin{pmatrix} {}_{i} \\ {}''_{i} \\ {}_{i} \\ {}_{i} \end{pmatrix} \sim \mathsf{N}\,(0;\Sigma) \text{ with } \Sigma = \begin{pmatrix} {}^{2}_{\xi} & {}_{\xi\varepsilon} & {}_{\xi} & {}_{\varepsilon} & {}_{\xi\eta} & {}_{\xi} & {}_{\eta} & {}_{\xi\nu} & {}_{\xi} & {}_{\nu} \\ {}_{\xi\varepsilon} & {}_{\xi} & {}_{\varepsilon} & {}^{2}_{\varepsilon} & {}_{\varepsilon\eta} & {}_{\varepsilon} & {}_{\eta} & {}_{\varepsilon\nu} & {}_{\varepsilon} & {}_{\nu} \\ {}_{\xi\eta} & {}_{\xi} & {}_{\eta} & {}_{\varepsilon\eta} & {}_{\varepsilon} & {}_{\eta} & {}^{2}_{\eta} & {}_{\eta\nu} & {}_{\eta} & {}_{\nu} \\ {}_{\xi\nu} & {}_{\xi} & {}_{\nu} & {}_{\varepsilon\nu} & {}_{\varepsilon} & {}_{\nu} & {}_{\eta\nu} & {}_{\eta} & {}_{\nu} & {}^{2}_{\nu} \end{pmatrix} : \tag{14}$$

The correlation parameters $_{\xi\eta}$ and $_{\varepsilon\eta}$ have important economic meaning. They characterize the relationship between the applicant's unobserved motives for purchasing and down payment, and her subsequent repayment behavior. If both are zero, the purchase and financing decisions would reveal no new information about later default. Based on the descriptive evidence in Section 3, one expects that all else equal, purchasers and especially purchasers making larger down payments are better risks. This would imply that $_{\xi\eta}$; $_{\varepsilon\eta} > 0$.

The correlation parameters, $_{\xi\nu}$, $_{\varepsilon\nu}$, and $_{\eta\nu}$ play a role in identification. If all three were zero, the negotiated price could be treated as exogenous in modeling the purchase, financing and repayment decisions. Finally, the variance parameters $_{\xi}$; $_{\varepsilon}$; $_{\eta}$; $_{\nu}$ capture the importance of unobserved characteristics relative to observed characteristics in negotiation and customer decisions.

## 5.2 Estimation and Identification

The loan demand and repayment model consists of four equations: the purchasing and borrowing equations (8)-(9) and (10)-(11), the repayment equation (12), and the pricing equation (13). These equations map the observed and unobserved characteristics of each applicant $\mathsf{i}$ into purchase, borrowing, and repayment decisions, $\mathsf{Q}_i$; $\mathsf{D}_i$; and $\mathsf{S}_i$. Combining these equations with the stochastic assumption on unobservables in equation (14), therefore, leads to a likelihood function for the observed decisions. In Appendix B (available online) we write out the full likelihood function and provide computational details on the parameter estimation, which we perform using maximum likelihood. Maximizing the likelihood is fairly standard, although somewhat cumbersome due to the

need to integrate over multiple unobservables. The reported standard errors are computed using a bootstrap method.

The pricing variation in our data plays an important role in identifying the key parameters of the consumer model. We focus here on the variation we observe in the two main pricing terms, the minimum down payment requirement and the markup of car price over cost.

As described above, the company sets a down payment requirement for each credit category of applicants. The company has adjusted the schedule over time, and the variation is substantial. Over the observation period, we observe more than twenty changes in the minimum down payment schedule. Figure 4 shows the time pattern of required down payments for three representative credit categories. Some of these changes can be explained by seasonal shifts in demand. Demand peaks notably in February when potential customer become eligible for tax rebates, and the company typically increases down payment requirements. Some of these changes arose from gradual shifts in corporate thinking about whether to focus on higher or lower risk borrowers. Finally, the company reports that over time it became more confident in its credit scores and relied on them more heavily leading to the increased spread in down payments as seen in Figure 4. We use year and month controls to capture seasonal shifts and slow-moving trends at the firm, and rely on the abrupt variation from the shifts to identify the effect of minimum payment changes. Note that the identification is both in the time-series within a credit category and, because the changes are rarely uniform across credit categories, from differences in the changes across categories.[6] An important feature of the data is that we have available to us precisely the data available at headquarters. This greatly mitigates the typical concern in demand estimation that prices are being adjusted in response to shifts in demand that the researcher does not see, but would pick up with access to all of the firm's information.

Identifying the effect of changes in car price on customer purchasing and borrowing decisions is more challenging. One difficulty is that because we impute prices for non-purchasers, we potentially lose some of the true variation in the data. A second difficulty, accounted for in our demand model – by introducing the pricing equation (13) and allowing  to be correlated with the other unobservables in the model – is that negotiated prices may be correlated with unobserved buyer

[6] It is also possible to exploit additional variation by controlling continuously for the underlying credit score, but not credit category per se, and using regression discontinuity to compare buyers with credit scores just above and below the threshold for different credit categories. In Adams, Einav, and Levin (2009) we provide a more detailed discussion of this variation. We have also estimated the model restricting the sample to narrow windows around changes in the schedule, obtaining similar results.

characteristics. To the extent that such a correlation exists, identification in the model comes from changes in the list prices set at the company level, which are analogous to an "instrumental variable." We observe two major changes in the margin formula used to generate list prices,[7] creating identifying time-series and differences-in-differences variation.[8]

The other key parameters in the model are the correlations between the unobservables that affect repayment, $_i$, and the unobservables that affect decisions at the time of purchase, $_i$ and $"_i$. These correlations affect our inference about the extent of adverse selection, which translates to the implied difference between the average and marginal customer, which are key for optimal pricing as described in Section 2. Here the variation described above is also useful – as minimum down requirement change we observe the change in average per-loan profits and back out the implied profitability of the marginal applicant – and is aided by two additional modeling assumptions. The first is the parameter restriction implied by our assumption that car price and down payment affect repayment in the same way, as only their difference (that is, the loan amount) enters into the repayment equation (12). The second is the structure imposed by the joint normality assumption which helps us back out the profitability of the marginal borrowers from changes in the average profitability of inframarginal ones. We have experimented with other distributional assumptions and are reassured that the key parameters are not overly sensitive to our choice of distribution, which is clearly not implied by economic theory.

One aspect of our general empirical strategy that deserves mention is that we do not impose assumptions about the optimality of observed prices to identify demand elasticities. Instead, we use the observed variation in prices to estimate the loan demand and repayment system, without assuming that observed contracts are profit-maximizing. We then use the demand estimates to analyze pricing. The advantage of this approach is that it relies on relatively weak assumptions about price-setting. Essentially, we require only that decisions made at the company level, conditional on year and month dummies, are orthogonal to individual applicant unobservables. The downside of our approach is that if one strongly believed that prices were profit-maximizing, or at least preferable to some identifiable alternatives, this information is not used to refine the demand estimates. In weighing these approaches, our view has been that it makes more sense (in our set-

[7] We focus on variation in the margin formula rather than in the list price of the car because the car cost (to the company) is used as an important right-hand-side variable, which we view as an excellent proxy for the car quality.

[8] In addition, list prices jump discontinuously at certain cost thresholds. By controlling continuously for car cost, this provides a source of regression discontinuity identification. Again, Adams, Einav and Levin (2009) provide more detail.

ting) to use the relatively convincing identifying variation in prices to recover demand and avoid imposing specific pricing structure except where it is required.

### 5.3 Results

Table 2 reports summary statistics for the data next to averages predicted by the model, showing that we are able to fit the key moments in the data fairly well. Figures 5(a), 5(b), and 5(c) show the distributions of down payments and repayment lengths observed in the data and predicted by the model. The model does well in matching the distribution of down payments and repayment length. This suggests that the distributional assumptions imposed by the model – truncated normal in the case of down payments and truncated lognormal in the case of repayment length – are not particularly restrictive. In fact, we tried to estimate versions of the model with an additional parameter that tilted the repayment distribution, and couldn't reject the baseline specification.

The first two column of Table 3 report our estimates of purchasing and borrowing behavior. The first column reports the marginal effects of the variables on the probability of sale. The probability that an applicant purchases, the "close rate," is very sensitive to the required down payment and much less sensitive to changes in the car price. A $100 increase in the required down payment lowers the probability of sale by 2.3 percentage points, which is equivalent to about 6.7% reduction in volume. In contrast, a $100 increase in car prices has essentially no economically meaningful effect on the probability of sale. An increase in car prices also has a relatively small effect on a buyer's desired down payment. As reported in the second column of Table 3, we estimate that a $100 increase in car prices raises a buyer's desired down payment by about $11. Therefore, it appears that the primary effect of higher car prices is to increase the size of loans that buyers take.

These findings — the substantial effect of down payment requirements, and the tendency of buyers to finance price increases — are consistent with the idea that customer liquidity is an important factor in explaining purchasing and borrowing decisions. Applicant characteristics are consistent with this hypothesis. Applicants with higher income and applicants with a bank account are more likely to purchase, and those with bank accounts tend to make larger down payments.[9] Adams, Einav and Levin (2009) provide additional evidence for consumer liquidity effects by analyzing an annual spike in applications that occurs each February when consumers become eligible for tax re-

[9] Applicants who own a house are less likely to purchase, a finding which may reflect their potential access to better credit terms at other lenders.

bates. We account for this in our demand specification by including month dummies, reporting the coefficient on the February tax season dummy in Table 3. All else equal, the close rate in February is 14 percentage points higher, or 40%, and desired down payments are 545 dollars higher.

Our estimates of repayment behavior are reported in the third column of Table 3. Loan size is a primary determinant of payment duration and hence the likelihood of default. All else equal a buyer who takes a $1,000 larger loan (which translates into monthly payments being about 35 dollars higher) makes about 32 percent fewer payments. Payment duration varies with individual and car characteristics in ways that are largely predictable. All else equal, buyers are less likely to default on higher quality cars, and buyers with greater income or with a bank account are more likely to make payments.

As discussed earlier, the company's credit score is a strong predictor of expected repayment. Table 3 reports the coefficients for two representative categories — a low-risk category and a medium-risk category — relative to the omitted category, which is a representative high-risk category. A representative low risk buyer is expected to make 87 percent more payments than a representative high risk buyer and is 24 percent less likely to default. Individual credit scores are also predictive of decisions at the time of purchasing. High credit risks appear to have the highest desire to borrow, while medium-risk applicants are the most likely to purchase. One interpretation for the non-monotonicity in purchase probability is that low-risk buyers may have better outside opportunities.

The bottom of Table 3 reports the estimated variances and covariances of the unobserved individual characteristics. Consistent with our earlier discussion of Figure 2, unobserved drivers of purchasing and down payment are positively correlated with the fraction of loan payments made. All else equal, a buyer who is inclined to make a $500 larger down payment is expected to make 21.6 percent more of her payments.

To put this in context, a thought experiment is useful. Suppose two buyers who are identical on observables arrive at the lot and are offered the same car and the same financing terms. Both choose to purchase, but the first applicant chooses to make a down payment that is $500 more than the second applicant. From this decision, the company can infer that the first applicant is in fact a better credit risk. If the applicants had identical loans, the first would be expected to make 2.3 percent more payments. But the additional down payment also has a second, larger, effect, which is to reduce the first applicant's loan principal. The 500 dollar reduction in loan size increases

the expected fraction of payments by 19.3 percent. So the fact that the first applicant volunteers a larger down payment has both a signalling aspect and a direct repayment effect and both are important.

An alternative way to think about buyers selecting into larger and smaller loans, and one that is useful for pricing calculation, is to compare the predicted repayment of the average buyer with the predicted repayment of a marginal buyer who is just indifferent between purchasing and not purchasing, and of an average non-buyer. The estimates imply that the average buyer, given an average loan size, will make 62 percent of her payments on average, and has a 59 percent chance of default. A marginal buyer, who puts down exactly the required minimum given the same loan size, would be expected to make 55 percent of her payments and would have a 66 percent chance of default. An average non-buyer, given the same loan size, would be in between and would be expected to make 58 percent of her payments and would have a 62 percent chance of default. So we can view both selection into purchasing and into larger loans as advantageous — the average buyer represents a substantially better risk than either the average non-buyer or the buyers who demand larger loans. However, relaxing the down payment requirement selects relatively bad risks even when compared to the general pool of non-buyers.

A final point about the estimates pertains to the price negotiation. We find a relatively small correlation between the negotiated price and the (unobserved) credit risk. One interpretation of the low correlation is that dealership managers cannot infer much about credit risk beyond what is contained in the credit score, so that conditional on observables, the price negotiation outcome is driven mainly by factors that are unrelated to applicant creditworthiness. The small correlation that we do find is positive, which could be driven either by the higher risk borrowers performing somewhat better in negotiation, or because managers anticipate that they might benefit more from a slightly lower repayment burden (which is consistent with the results we report in the next section).

# 6 Empirical Analysis of Contract Pricing

We now turn from the demand side to consider contract pricing from the perspective of the firm. We start by defining profits and deriving conditions for optimal pricing. We then use these conditions to assess the optimality of pricing decisions, and the implied costs of lending that rationalize the

firm's observed policies. We explore the implications of the demand estimates for pricing decisions, and then use the pricing model to estimate the value that can be derived from risk-based pricing, and the extent to which better information about borrower credit risk can serve as a barrier to entry.

### 6.1 Revenue Accounting

The firm's profit from a given sale is the difference between the total revenue it receives and its costs. There are three sources of revenue: the initial down payment, the discounted value of the stream of loan payments, and the discounted recovery in the event of default. Let $\mathsf{D}$ denote the initial down payment, $\mathsf{p}-\mathsf{D}$ the amount that is borrowed, $\mathsf{T}$ the calendar length of the loan, $\mathsf{z}$ the interest rate on the loan, and  the firm's internal discount rate. As in the previous section, let $\mathsf{S}$ denote the fraction of loan payments that are made, so that $\mathsf{ST}$ is the calendar time during which payments are made. Finally, let $\mathsf{k}(\mathsf{ST})$ be the expected recovery if loan payments cease after a length of time $\mathsf{ST}$, and let $\mathsf{C}$ denote the costs incurred in selling the car.

With this notation in place, the present value of a given sale is

$$= \mathsf{D} + \left( \frac{\frac{1}{\kappa}\left(1-\mathsf{e}^{-\kappa ST}\right)}{\frac{1}{z}\left(1-\mathsf{e}^{-zT}\right)} \right) (\mathsf{p}-\mathsf{D}) + \mathsf{e}^{-\kappa ST}\mathsf{k}(\mathsf{ST}) - \mathsf{C}: \tag{15}$$

The down payment $\mathsf{D}$ is realized immediately. The second term is the present value of loan payments. The fraction in the expression represents the present value return on each dollar of loan principal. The third term is the value of the expected recovery, also discounted back to the time of purchase. If the loan defaults, i.e. $\mathsf{S} < 1$, the expected value of the recovery is positive. If the loan is paid in full, i.e. $\mathsf{S} = 1$, there is no recovery, so $\mathsf{k} = 0$. The final term in the profit expression is the company's cost of sale.

The empirical model from the previous section allows us to fill in the variables in the profit expression. To see how this works, consider an applicant with observed characteristics $\mathsf{x}$ and unobservables  $= (\ ;\ ";\ ;\ )$. The interest rate $\mathsf{z}$ and length of the loan $\mathsf{T}$ are part of the characteristics vector $\mathsf{x}$. Suppose the applicant faces a car price $\mathsf{p}$ and minimum down payment $\mathsf{d}$. If she purchases a car, then according to the model, she will make a down payment $\mathsf{D}\,(\mathsf{p};\mathsf{d};\mathsf{x};\ )$ and subsequently make a fraction $\mathsf{S}\,(\mathsf{p};\mathsf{d};\mathsf{x};\ )$ of her scheduled repayments.

The recovery value is not a component of the model of Section 5. We simplify computation by

modeling and estimating this object separately. We model expected recovery value $\mathsf{k}$ as a function of the transaction characteristics and the time elapsed between the start of the loan and default. The model we consider assumes there is a discrete probability of no recovery. Conditional on a recovery being made, we specify a linear model for the dollar value. In the end of Appendix B we describe the details of the estimation and report the estimates. In estimating recoveries separately from the rest of the model we ignore the possibility that unobserved heterogeneity in the recovery value might be related to other unobservables. We view this as relatively unproblematic for our analysis, where we focus more on the variation in default. Jenkins (2009) considers a more detailed, and unified, analysis of default and recovery.

The final components of profitability are the firm's internal discount rate and the cost of the sale. We assume an annual discount rate of $= 10\%$ based on discussions with firm executives.[10] For the cost of sale, our data includes detailed information on the cost of acquiring each car and transporting it to the lot. We treat these as the direct financial cost associated with a given sale. Naturally we expect that sales may involve additional indirect costs, for instance from the additional collections effort. We address this in more detail in the next section.

## 6.2 Optimal Pricing and Implied Costs of Lending

We now examine the extent to which the firm's observed pricing can be viewed as profit-maximizing, and estimate the implied indirect costs of lending necessary to rationalize observed policies. At first glance, this problem appears straightforward. Given our estimates of the profit components, we can simply ask whether there are plausible cost assumptions that rationalize observed prices as expected profit-maximizing. The difficulty, however, is that the firm in principle could use an almost unlimited set of pricing policies, varying down payment requirements, prices, interest rates, and loan lengths with a host of underlying applicant and car characteristics. It seems unreasonable to expect the firm's pricing decisions to be optimally designed in every instance.

We therefore focus on a single dimension of pricing, the down payment requirement, and on two relatively weak assumptions about optimality. The first is that over the sample period, the firm's down payment requirements were on average correct, in the sense that a uniform increase

[10] Discussions with industry participants suggest that at the time of our study lenders may have been using internal discount rates in the range of 8 to 12 percent. These relatively high rates, particularly compared to the low rates on treasury bills at the time, reflect the risk involved in subprime lending. We also experimented by using values in the range of 5 to 15 percent, and found that our results were not much affected.

or decrease in the down payment required of every applicant in the data would not have improved expected profit. Here we explicitly do not assume that the structure of down payments was optimal for each risk category, allowing us to examine later whether the firm could have profitability shifted its lending by making down payment requirements more or less sensitive to assessed risk.

We also start with a very simple specification of the indirect costs of lending. We write the potential cost of a sale to a given applicant $\mathsf{i}$ as the sum of the financial cost of the offered car $\mathsf{c}_i$ and a constant per-loan indirect cost . So total costs are

$$\mathsf{C}_i = \mathsf{c}_i + \quad : \tag{16}$$

We discuss the interpretation of the indirect cost below and also consider more flexible specifications in which the per-loan cost is not constant but varies over time or according to borrower risk. For now, we treat it as a single parameter to be estimated.

To proceed with our strategy, we construct the expected profit from the observed pricing and from alternative policies. Consider an applicant in the data with observed characteristics $(\mathsf{x}_i; \mathsf{p}_i; \mathsf{d}_i)$. For any such applicant, we can use the demand and repayment estimates to compute the probability of purchase $\Pr_\zeta [\mathsf{U}(\mathsf{x}_i; \mathsf{p}_i; \mathsf{d}_i; \quad) \geq 0]$, and expected profit conditional on purchase $\mathbb{E}_\zeta [\quad (\mathsf{x}_i; \mathsf{p}_i; \mathsf{d}_i; \quad ; \quad) \mid \mathsf{U}(\mathsf{x}_i; \mathsf{p}_i; \mathsf{d}_i; \quad) \geq 0]$

Our second approach to estimating implied lending costs also relies on a weak revealed preference assumption, but has a different motivation. As described above, at any point in time applicants face a required down payment that depends on their credit category. We observe twenty-two changes in the underlying schedule, and hence twenty-three pricing periods. Index these periods by $= 1; :::; 23$, and let $\mathsf{I}_\tau$ denote the set of applicants in period . For any applicant $\mathsf{i}$, $\mathsf{d}_i = (\mathsf{x}_i; \ _i)$, where $_i$ is the pricing period and $\mathsf{x}_i$ includes $\mathsf{i}$'s credit category. Our alternative assumption about pricing is that while changes in the pricing schedule over time may not have been optimal, each change represents an improvement over the current status quo.

To capture this in our notation, let $\mathsf{d}'_i = (\mathsf{x}_i; \ _i - 1)$ denote the down payment requirement that applicant $\mathsf{i}$ would have faced under the prior pricing schedule. We assume that the observed prices should satisfy:

$$\sum_{i \in I_\tau} \Pi(\mathsf{x}_i; \mathsf{p}_i; \mathsf{d}_i; \ ) \geq \sum_{i \in I_\tau} \Pi(\mathsf{x}_i; \mathsf{p}_i; \mathsf{d}'_i; \ ) \qquad \text{for all} \quad = 2; :::; 23. \tag{19}$$

In contrast to equation (18), there is not necessarily a single indirect cost parameter that satisfies (19). In fact, given our demand estimates, there is no such value. We therefore use as our estimate of in this exercise, the value that minimizes the sum of squared violations of the 22 inequalities in equation (19).

We report our estimates of the implied lending costs in Table 4. We find that for offered financing to be preferable to any uniform shift in the down payment schedule, the unobserved component of costs must be fairly large, on the order of 2,500 dollars. Our alternative approach, based on the idea that changes in down payments are revealed preferred to the status quo, produces similar estimates, ranging from 1,800 to 2,800 dollars. One way to understand these results is to observe that relative to the direct financial costs, the observed down payments are too high, and the number of loan originations too low to be optimal. Instead the down payment requirements are set "as if" there is a fairly high indirect cost of extending credit.

It is natural to ask why exactly the observed down payment requirements are high relative to the direct financial costs, or what the implied indirect costs might represent. Loan servicing is an obvious explanation. While we have no direct way to measure the cost of servicing each additional

on the sale. Write the latter term as $R(x_i, p_i, d_i + a) - \psi$ where $R$ is the expected revenue net of the direct financial cost of the car. Because $Q$ is decreasing in $a$, $\Pi(x_i, p_i, d_i + a; \psi)$ has increasing differences in $(a, \psi)$, and $a^*(\psi) = \arg\max_a \sum_i \Pi(x_i, p_i, d_i + a; \psi)$ is an increasing function. Condition (18) is equivalent to the requirement that $a^*(\psi) = 0$ which holds for a unique value of $\psi$.

loan, the firm places this cost at around $1,000. This leaves another $1,500 or so to explain. One possibility is that expanding loan originations is costly due to various adjustment costs of increasing the scale of operations. An alternative explanation is that the company was concerned about the overall risk of its originated loans – in particular the possibility of macroeconomic shocks. That is, while our default estimates are driven by realized loan outcomes up to April 2006, the company may have realized that there was a risk of an aggregate downturn, as occurred subsequently. Finally, a third possibility is that there are frictions between the lender and the secondary market that arise when loans are securitized.

To explore the adjustment cost story, we re-estimated indirect costs allowing them to differ in tax season, and assuming that equation (18) holds separately for tax season and non-tax season applicants. To the extent that scale or capacity constraints are important, we should expect higher costs in tax season when demand is at a peak. As Table 4 shows, however, we do not observe this. The implied costs are similar and perhaps even slightly lower in tax season. We used a similar approach to estimate costs separately for each credit category and to explore if the indirect costs estimate is driven primarily by a hesitancy to lend to the highest-risk borrowers. As reported in Table 4, however, we find little evidence for the hypothesis: estimated costs are higher for better risks. These estimates have an interesting alternative explanation, which is that if the indirect costs are constant across risk categories, then the observed down payments should have been somewhat lower for low risks and higher for high risks. The general trend in the company's pricing, as shown in Figure 4, is consistent with the company figuring this out slowly over time.

This leaves us with possible costs from securitization and aggregate risk exposure, which are connected if the secondary market pricing was distorted or sheltered the firm from risk exposure. We have no data on secondary market prices, but given the time period, one suspects that prices were, if anything, quite optimistic. Indeed a concern with mortgage lending during this period is that securitization gave lenders an excessive incentive to make marginal loans. We see little evidence of this in our estimates, and one explanation is that the securitization contracts in our setting called for substantial penalties in the event of excessive defaults. Because of this, we view aggregate risk exposure as a fairly plausible cost of lending and rationale for our estimates, although fees or other transaction costs of securitization might also explain some fraction of the residual implied cost.

### 6.3 Differential Effect of Contract Terms

Our informal discussion of pricing, and the theoretical model of Section 4, highlighted the possibility that different contract dimensions might give rise to very different pricing trade-offs. In this section, we use our demand estimates to illustrate these trade-offs graphically. We do this for both the down payment requirement and the mark-up of car prices over cost in order to emphasize the very different nature of these contract terms.

Although substantively quite different, the mechanics of this exercise are similar to the previous section. For each applicant in the data, we compute the probability of purchase $\Pr_{\zeta}\left[\mathsf{U}(\mathsf{x}_i;\mathsf{p}_i;\mathsf{d};\ \ ) \geq 0\right]$, the probability of default conditional on purchase, $\Pr_{\zeta}\left[\mathsf{S}(\mathsf{x}_i;\mathsf{p}_i;\mathsf{d};\ \ ) < 1 | \mathsf{U}(\mathsf{x}_i;\mathsf{p}_i;\mathsf{d};\ \ ) \geq 0\right]$, and the expected profit $\Pi(\mathsf{x}_i;\mathsf{p}_i;\mathsf{d})$ (specified in equation (17) above) for varying levels of the down payment requirement $\mathsf{d}$. We include in the profit calculation our indirect cost estimate $\ \ = 2;481$ (see Table 4). We then average over applicants to find the expected outcomes that would result from specifying a down payment requirement $\mathsf{d}$ for a given credit category. In this exercise, as above, we hold fixed the car assignment, offered price and other contract characteristics for each applicant. We repeat this same exercise to look at car pricing. In this case, we vary the mark-up of price over cost for the applicants, holding fixed the other contract elements including the down payment requirement.[12]

Figure 6(a) illustrates the effect of changes in minimum down payment requirement. The top panel does so for low risk applicants and the bottom panel for high risk applicants. In both panels, the probability of sale line reflects the large estimated elasticity of purchase with respect to the down payment requirement. Our estimate, extrapolated a bit out of sample, suggests that absent a down payment requirement, the probability of purchase would reach 50%, but it sharply declines and comes close to zero as the requirement reaches \$2,000 or \$2,500. The second line in Figure 6(a) presents the probability of default, and shows how the default probability decreases with required down payments. As discussed earlier, this happens for two reasons. First, an increased down payment requirement (holding fixed car prices) leads to smaller loans, and hence greater ability and incentive to repay. Second, an increased down payment requirement screens out the relatively

[12] In analyzing price-cost margins, we continue to abstract from car choice. This involves a more substantive restriction than when we estimated the demand model alone, because a large out-of-sample change in the pricing policy might cause an applicant to substitute to a different preferred car. More precisely, the way to think about the current exercise is that we are considering a change in the mark-up for all cars on the lot, and assuming that such a change does not affect the identity of a buyer's preferred car. This seems like a natural assumption for small price changes, but perhaps more questionable for very large ones.

high risk marginal borrowers, so the remaining pool of purchasers is of better quality.

The qualitative pattern is similar for low and high risk borrowers. In both cases increasing minimum down requirement trades off a reduction in loan originations with an improvement in loan repayments. However, the default probability is significantly higher for high risk borrowers, implying that the resolution to this trade off is quite different. Because high risk applicants are less profitable, the company has a greater incentive to screen them out, leading to a relatively high optimal down payment requirement. In contrast, even marginal low risk applicants are quite profitable, implying that expected profits are maximized at a much lower level of down payment requirement, which screens out only a small fraction.

Figure 6(b) repeats the exercise, focusing on changes in the mark-up of prices over cost, and holding the down payment requirements fixed. Compared to Figure 6(a), the patterns are dramatically different. Here the probability of sale line is almost flat, reflecting our finding that purchase probability is hardly sensitive to the mark-up. What then is the downside to higher mark-ups? The answer becomes clear by looking at the probability of default. Increased car prices substantially raise the probability of default. So in looking at how expected profits vary with price, the curve in Figure 6(b) reflects the trade-off between the size of the payments and the probability they will be made, with customer selection playing essentially no important role. Comparing the two panels suggests that the optimal price is actually higher for low risk applicants, who are less likely to default when margins are increased.

To summarize, the empirical findings in Figures 6(a) and 6(b) highlight the discussion in the earlier sections. Different elements of the loan contract introduce very different pricing trade-offs. Setting down payment requirements leads in large part to a trade-off between origination volume and borrower selection. Setting price-cost margins leads to a trade-off between the specified payments and the probability that payments will actually be received. Note that we might suspect interest rates to play a similar role, with higher rates leading to larger specified payments but more chance of default. The idea, therefore, that one could track a credit market by focusing on a single pricing dimension, as is commonly done when, say, regulators or firms emphasize an average interest rate, is clearly not the right approach to thinking about how markets may be functioning or evolving.

### 6.4 Risk-Based Financing and the Value of Information

The importance of credit scores in consumer lending is widely appreciated. Most consumers these days understand that their credit history may disqualify them from certain loan offers or qualify them for better terms. Our data, and our estimated model, provides an opportunity to assess quantitatively the value that this sort of risk-based pricing generates for a lender. The setting is an interesting one for this calculation both because subprime lending is so risky and because we have already seen that the company's credit scores have substantial predictive power and are used heavily in setting down payment requirements.

To assess the value generated by risk-based pricing, we use our model to calculate the expected profitability of the company in different informational scenarios. The estimates are reported in Table 5. The table shows the company's per-loan profitability under the observed prices, the expected profitability if the company had set down payment requirements optimally using its existing credit categories, and its expected profit if it had not been able to use credit scores and instead set an optimal uniform down payment in each pricing period. Overall, our estimates imply that under the observed pricing, an applicant arriving at the lot was worth on average $341. Setting down payments optimally given the credit categories, which as noted above would have meant somewhat lower down payments for low risk applicants and somewhat higher ones for high risk applicants, increases the expected profit per-applicant by $25 or 7%. In contrast, without the ability to categorize applicants by credit risk, optimal uniform pricing implies a per-applicant profits of $299, which is 18% less than optimal risk-contingent profits.

It is instructive to compare these estimates to the analysis in Einav, Jenkins and Levin (2009). In that paper, we use different data from the same company to compare profits from before and after the company adopted credit scoring technology and began to use risk-based pricing. Because we observe the full history of each loan in that paper, the statistical analysis is straightforward. We simply compute observed profits (revenues minus direct financial costs) and compare before and after the advent of credit scoring. The numbers we obtain there are quite comparable to what we get from the current modeling exercise, if we compare profits per borrower with the observed prices to profits per borrower with a uniform down payment set at the level the company used prior to credit scoring. This provides something of an out-of-sample reality check on our estimates.

An interesting question from a conceptual perspective is whether existing credit scoring technology extracts most of the value that could be extracted with available information. A rough way

to get at this is to ask how well the firm could have done with additional information, in particular if it was able to set down payment requirement's based not just on the credit score, but also on each applicant's private information at the time of purchase (i.e. $_i$ and $"_i$). Our results suggest that using this additional information would increase profits substantially: by 96% compared to the case of optimal uniform pricing, by 60% compared to the case of optimal pricing based on current risk categorization, and by 72% compared to observed pricing. These numbers suggest that despite dramatic improvements in risk classification and credit scoring, consumers still have significant private information that remain unpriced.

To understand exactly how credit risk information benefits the lender, rather than just the size of the benefit, we can use a graphical treatment to decompose the effect of risk-based down payment requirements. The top panel of Figure 7 illustrates a situation where the firm has no minimum down payment requirement.[13] Each bubble in the plot represents a credit category. The size of the bubble represents the number of loan originations that would have resulted according to the estimated model. The location on the horizontal axis represents the default rate and the location on the vertical axis the average down payment. Two features are notable. The first is the large number of high-risk loan originations. The second is the correlation across credit categories between down payment and default. Observably high risk borrowers demand larger loans and subsequently default at much higher rates.

The middle panel of Figure 7 illustrates how imposing the optimal risk-based down payment requirements affects the number of loan originations and the resulting loan sizes. The required down payment is highest for the high risk borrowers, and loan originations for these applicants fall dramatically. Moreover, the high risk applicants who do borrow are forced to make substantial down payments. So conditional on purchase there is a better matching between credit risk and loan size. Rather than taking the smallest down payments, the highest risk borrowers are forced to make the largest. The final panel in Figure 7 shows the effect on subsequent default rates. Default rates fall for all credit categories, but most dramatically for the highest risk category. As explained above, there are really two forces at work. One is selection: the down payment requirement screens out the risky borrowers within a credit category. The second channel is through repayment: forcing borrowers to make larger down payments results in smaller loans, which in turn decrease default

[13] The analysis would be qualitatively similar if we started with a non-zero, but still uniform, down payment requirement. This version is useful in also highlighting the general importance of the down payment.

rates.

### 6.5 Proprietary Information as an Entry Barrier

The analysis above quantifies the value of contingent financing offers, holding fixed the competitive environment. Contingent financing increases profitability both by allowing the lender to screen bad risks, and by allowing it to more effectively match loan sizes with default risk. Another potential effect, however, is strategic. In consumer credit markets, it is possible that having a better credit scoring technology or an informational advantage, imposes a cost on competitors who may end up facing an adversely selected pool of borrowers. Our model provides a simple way to explore this idea, albeit one that takes us *much* more out of sample and relies on many stylized assumptions. Despite these limitations, however, we view the exercise as useful to illustrate the extent to which information may operate as an entry barrier.

For this exercise, we use the estimated model to compute per-applicant profits for a set of monopolistic and duopolistic scenarios. We examine cases where neither the incumbent firm nor the entrant has access to credit scores, where both have access to credit scores, and where only the incumbent has access to credit scoring. The results are reported in Table 6. Each cell in the table presents the profits for the incumbent first, and for the potential entrant second. In each scenario, we take our estimated demand system to be the market demand (as opposed to our earlier interpretation as residual demand), so the monopoly profits mirror those reported in Table 5. When an entrant is present, we find the Nash Equilibrium of the duopoly game, where the firms simultaneously set uniform or category-based required down payments, depending on the information they have available. Note that in our model a fraction of the applicants are unaffected by the required down payment, and these applicants represent better-than-average risks. We assume that such applicants are randomly split between the two firms. We also make the strong assumption that in all other respects, including the car offered to each applicant and the car price, the competing firms are identical.[14]

The top row of Table 6 implies that the value of information for a monopolist is $67 per applicant. Information also has an entry-deterring benefit: better selection of applicants for an

[14] We calculate a pure strategy Nash equilibrium (PSNE) for all cases assuming firms must choose from a discretized set of prices. As a general theoretical matter, the existence of a PSNE is not assured when one firm chooses a grade-based minimum and the other firm a uniform minimum. Fortunately, we do not encounter a non-existence problem for our particular calibrated model.

incumbent implies also that a competitor who doesn't have that information (and therefore has to price uniformly) would face worse selection of applicants. We compute that information reduces the potential competitor's profits by $39 per applicant. That is, the break-even sunk entry cost that would justify entry need to be 30% lower when an incumbent can offer contingent financing terms. Interestingly, Table 6 also shows that the unilateral incentive to offer risk-based pricing is quite similar in the presence of competition, with per-applicant profits increasing by $64 compared to $67 when the company is a monopolist.

# 7 Conclusion

This paper has analyzed contracting in the market for subprime loans. Our principal findings highlight how informational imperfections, and consumer liquidity effects, lead to sharp pricing trade-offs. Such trade-offs need to be considered separately in thinking about different elements of a loan contract. For instance, our results illustrate the central role that down payment requirements play in limiting loan originations and constraining borrower leverage. Our estimates show that even modestly relaxing these requirements can greatly expand and increase the riskiness of the pool of borrowers. As noted in the introduction, these insights seem particularly relevant in light of the disastrous events in the subprime mortgage market, in which lax down payment requirements allowed borrowers to become highly leveraged and therefore vulnerable in the face of declining house prices and underlying income or liquidity risk. Our estimates also reveal a high value, both direct and strategic, to innovations in credit scoring that allow offers to be based on the observed riskiness of loan applicants.

We have focused the empirical work in this paper primarily on contract design and firm behavior. To do this, we formulated a structural model of firm behavior, but a relatively "reduced-form" model of consumer behavior, albeit one that can be motivated with a standard model of intertemporal utility maximization. A benefit of our modeling approach is that our conclusions do not rely on specific assumptions about the financial sophistication of consumers, the accuracy of their expectations, or the form of intertemporal preferences. Such assumptions, particularly in our setting, likely would be controversial and certainly hard to assess.

Having said that, our approach also has at least two limitations. One is that we are limited in the range of "out-of-sample" counterfactual exercises we can perform. For example, we have no easy

**Jenkins, Mark.** "Subprime Defaults: Measuring the Costs of Borrower Bad Behavior," Wharton School Working Paper, 2009.

**Karlan, Dean, and Jonathan Zinman**. "Credit Elasticities in Less Developed Countries: Implications for Microfinance," *American Economic Review*, 98(3), June 2008, 1040-1068.

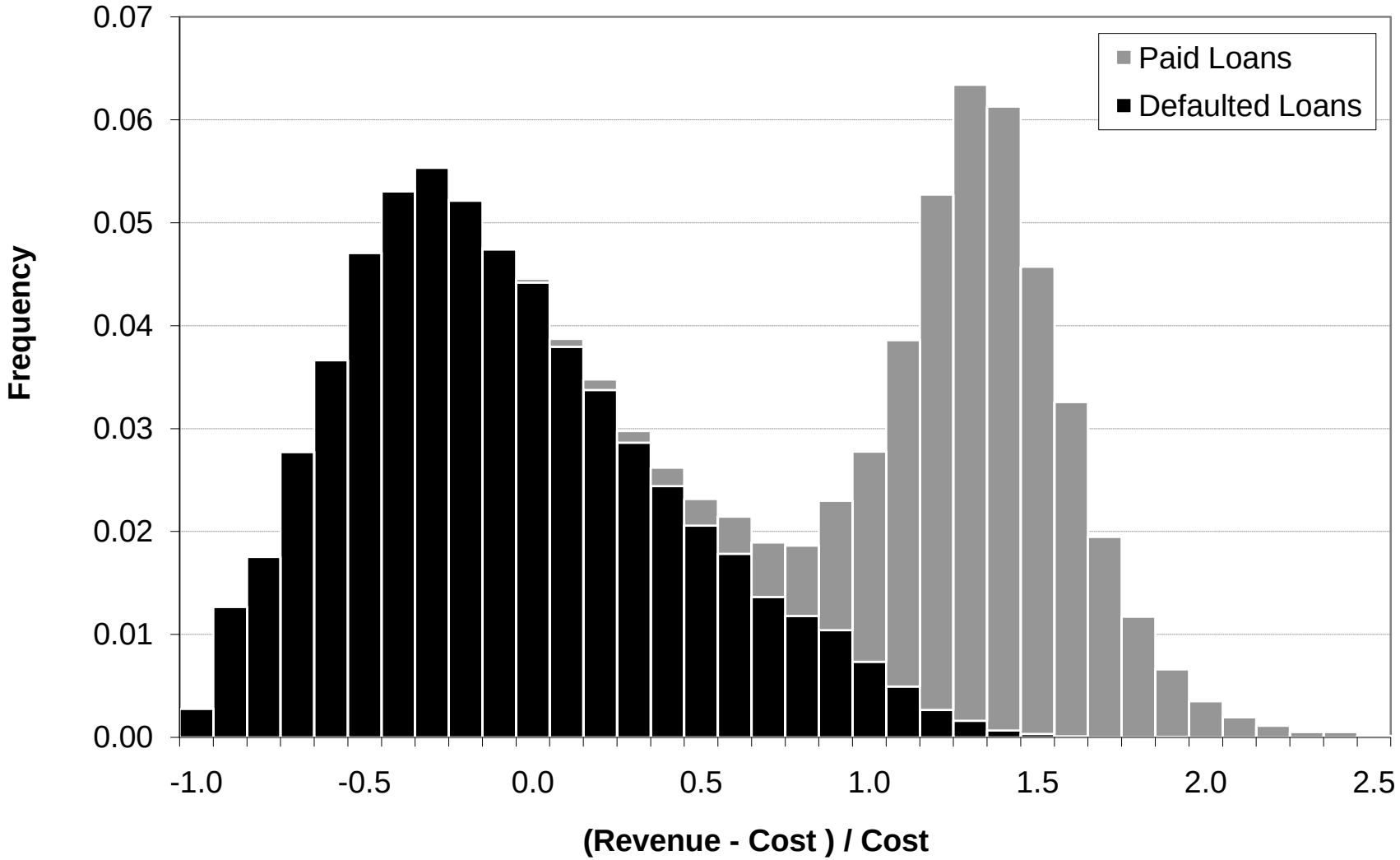
**Karlan, Dean, and Jonathan Zinman**. "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment," *Econometrica*, 77(6), November 2009, 1993-2008.

**Mayer, Christopher, Karen Pence, and Shane M. Sherlund.** "The Rise in Mortgage Defaults," *Journal of Economics Perspectives*, 23(1), Winter 2009, 27-50.

**Perrigne, Isabelle M., and Quang H. Vuong**. "Econometrics of Incentive Regulation: Nonparametric Identification," mimeo, Penn State University, 2006.

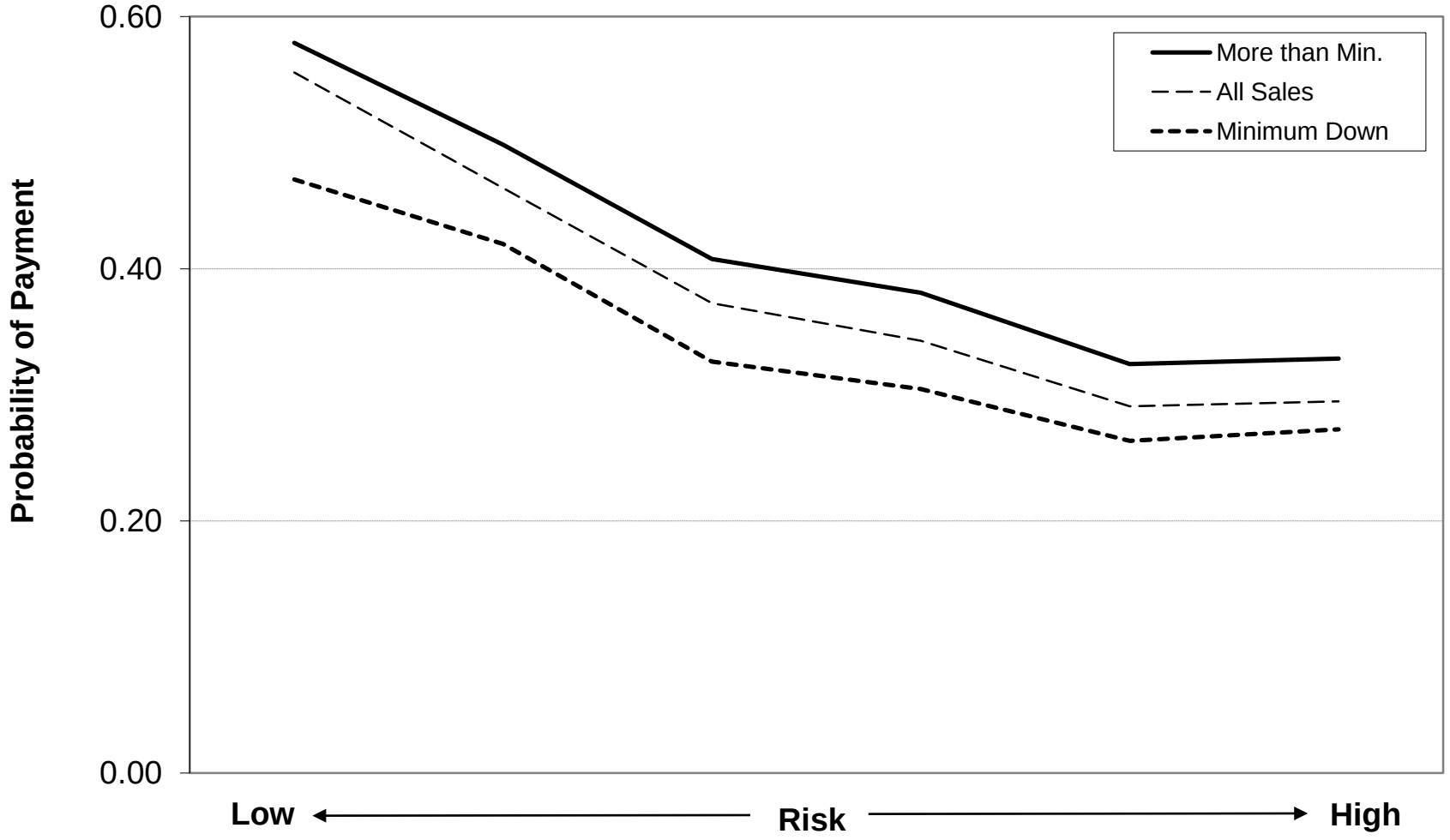
**Wolak, Frank A.** "An Econometric Analysis of the Asymmetric Information, Regulator-Utility Interaction," *Annales D'Economie et de Statistique*, 34, 1994, 13-69.

## Figure 1: Rate of Return Histogram

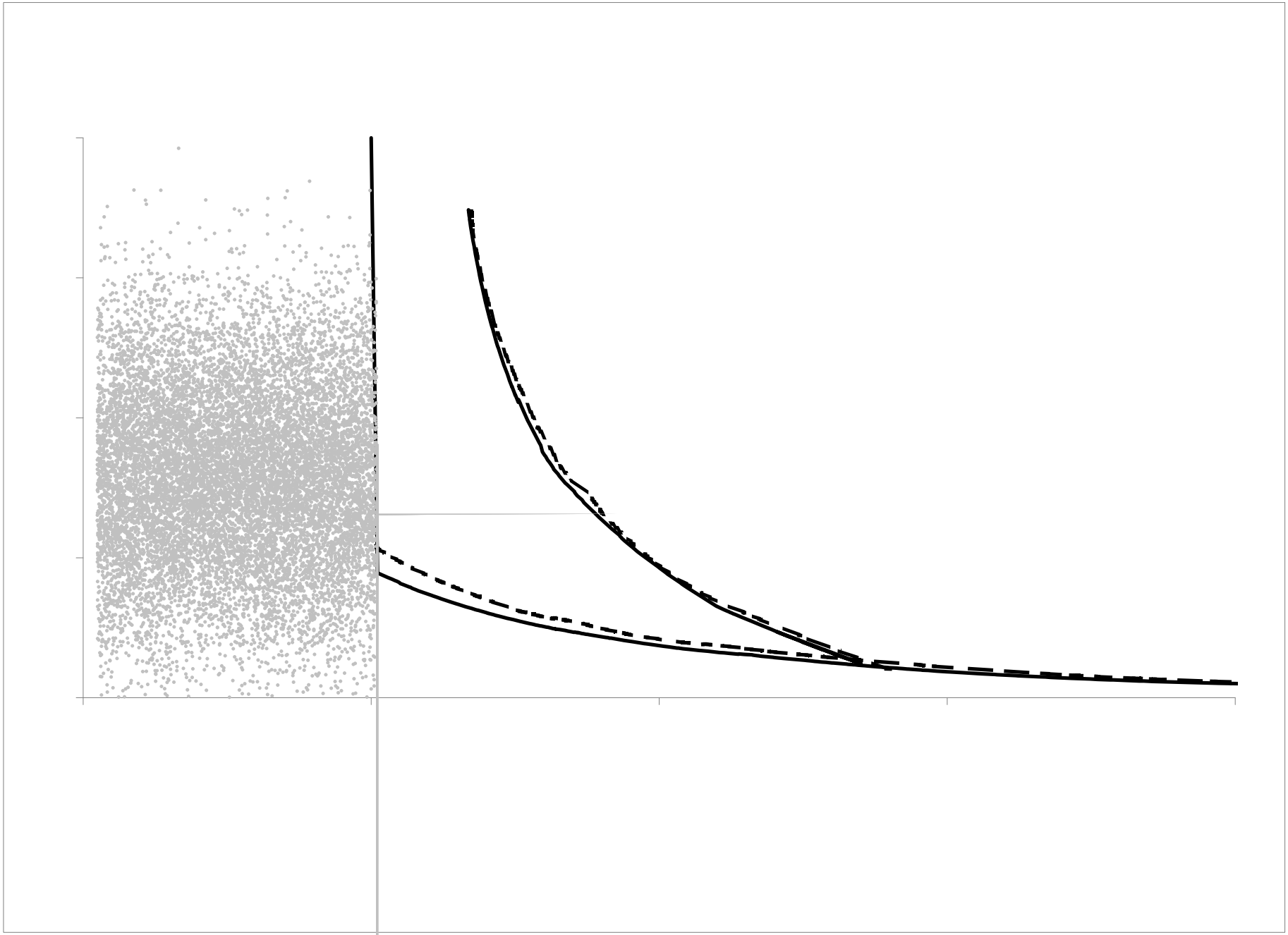



*Notes:* The figure is based on data from uncensored loans. Revenue is calculated as down payment + PV of loan payments + PV of recovery, assuming an internal firm discount rate of 10 percent. Cost includes the purchase cost at the auction and reconditioning cost.

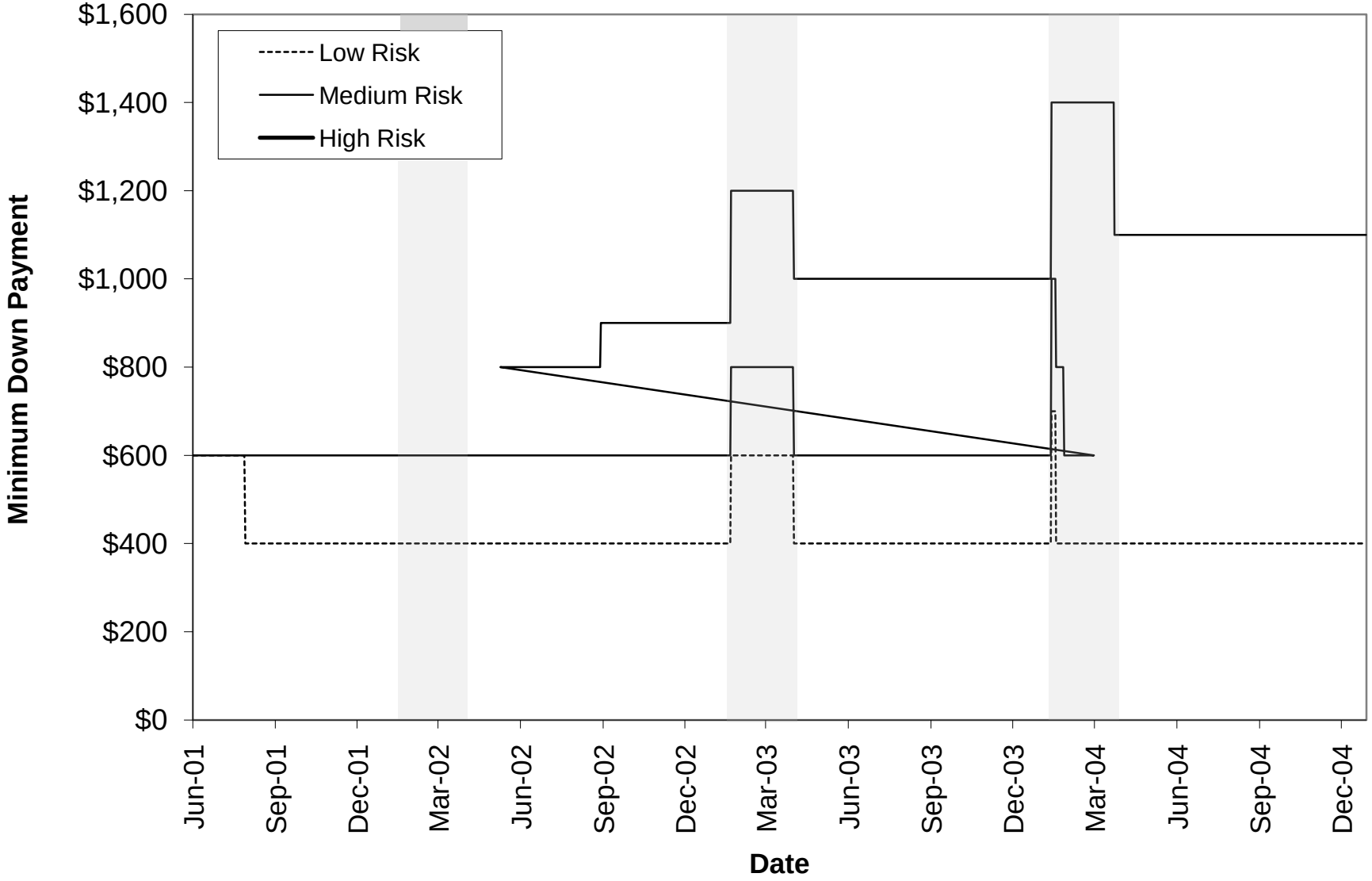
## Figure 2: Probability of Payment by Risk Type and Down Payment




Notes: Based on data from uncensored loans. The horizontal axis represents [illegible]

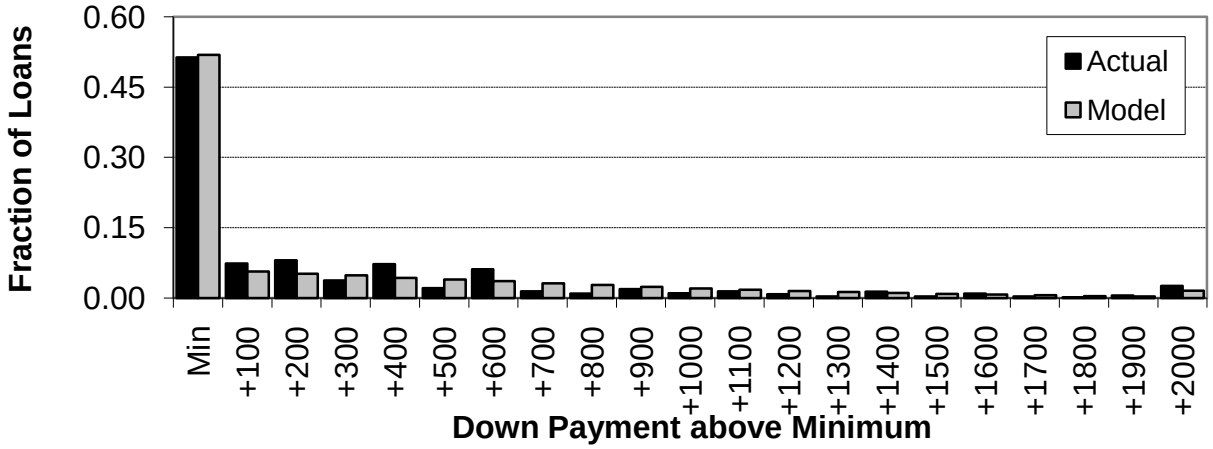


## Figure 4: Minimum Down Payments Over Time



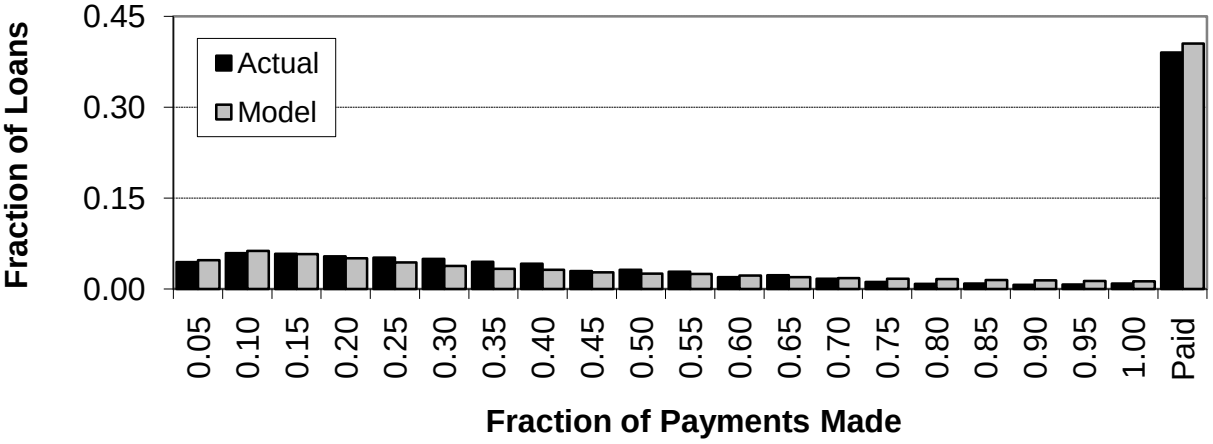

*Notes*: Shaded areas represent "tax season" (mid-January through end of March) during which volume is higher and down payment requirements increase. Each line represents the time series of required down payment for a representative high-risk, medium-risk, and low-ris k applicant, based on the firm's credit scoring.

## Figure 5(a): Distribution of Down Payments



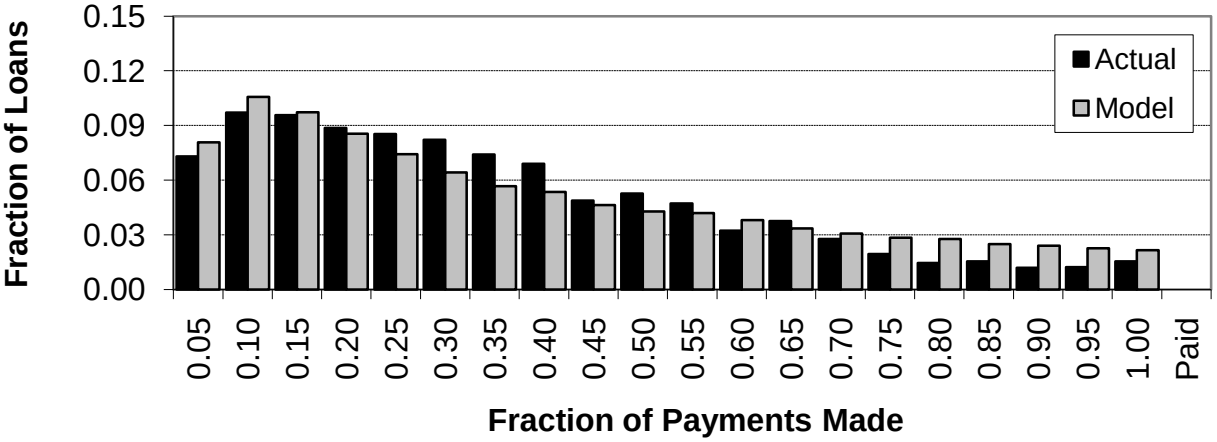

*Notes*: Based on raw data and model output for all sales. The bin labeled "Min" includes all buyers who put down exactly the required minimum down payment. Each bin labeled "+X" (except for the "+2000" bin) includes all buyers who put down an amount between $(X-100) and $X more than the required minimum. The bin labeled "+2000" includes all buyers whose down payments exceeded the minimum by more than $2,000.

## Figure 5(b): Distribution of Default Timing



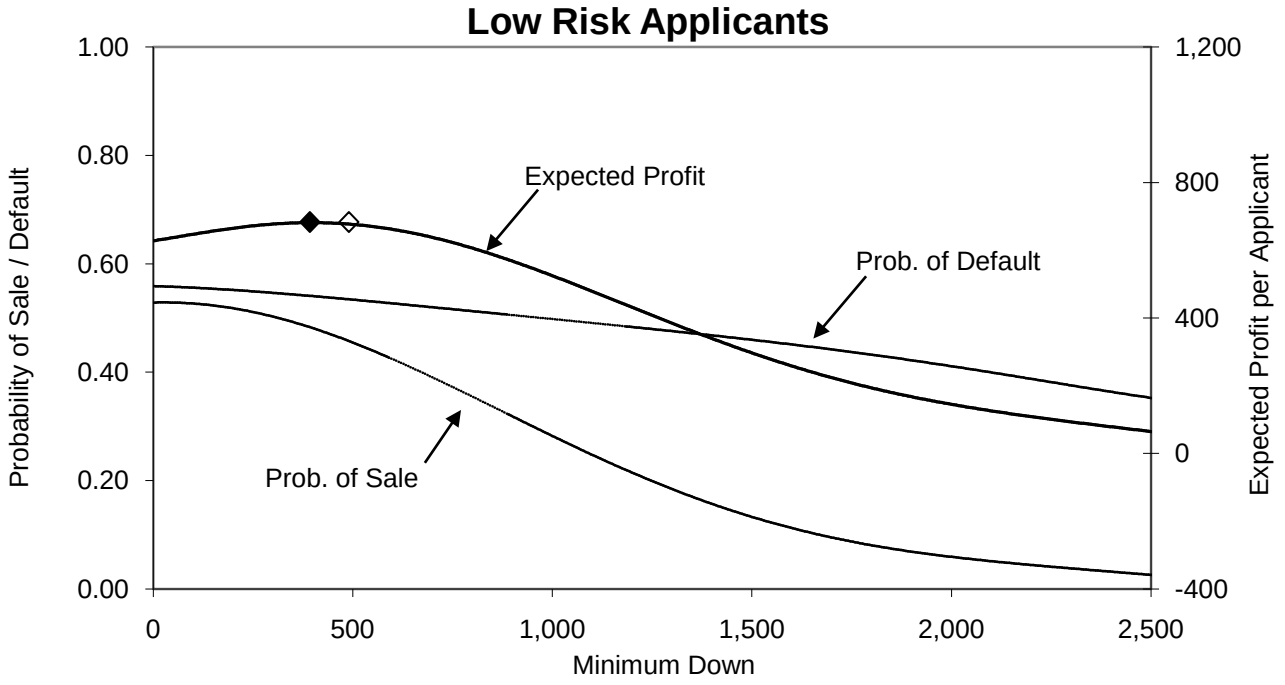

*Notes*: Based on raw data and model output for uncensored loans. Each bin labeled "X" includes all buyers who defaulted after making a fraction of their payments between (X-0.05)

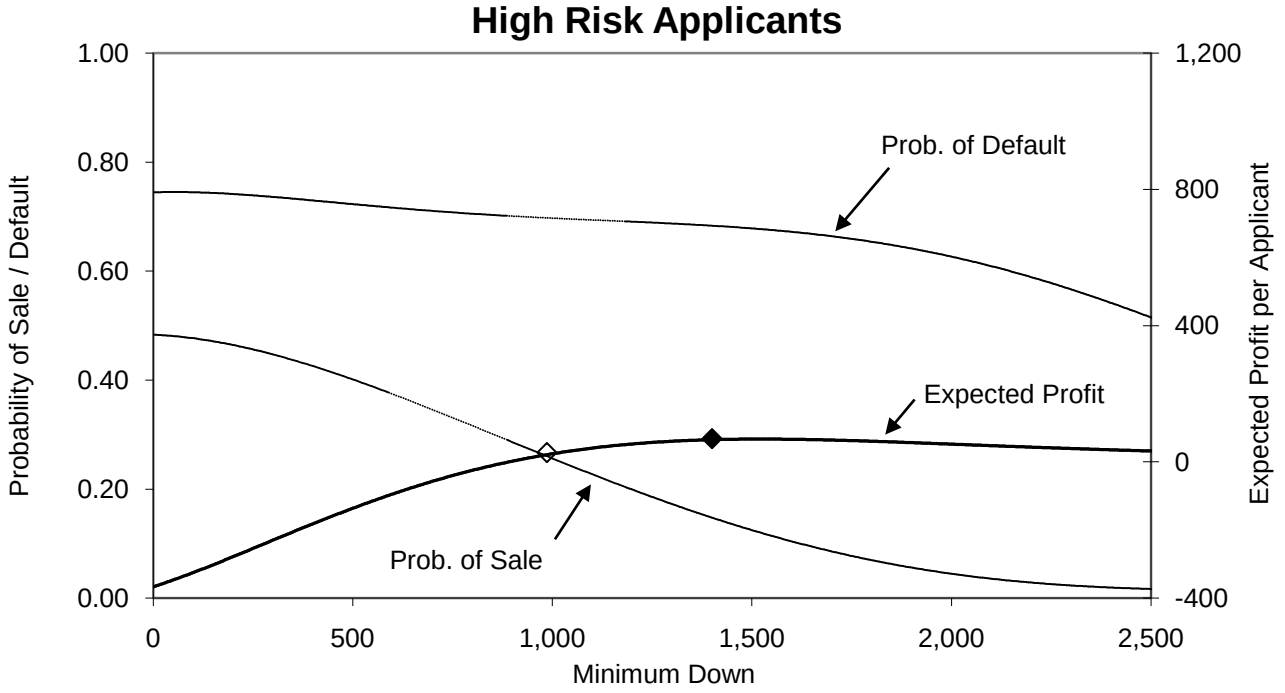
## Figure 5(c) : Distribution of Default Timing Conditional on Default




*Notes*: Based on raw data and model output for uncensored loans. Each bin labeled "X" includes all buyers who defaulted after making a fraction of their payments between (X-0.05) and X.

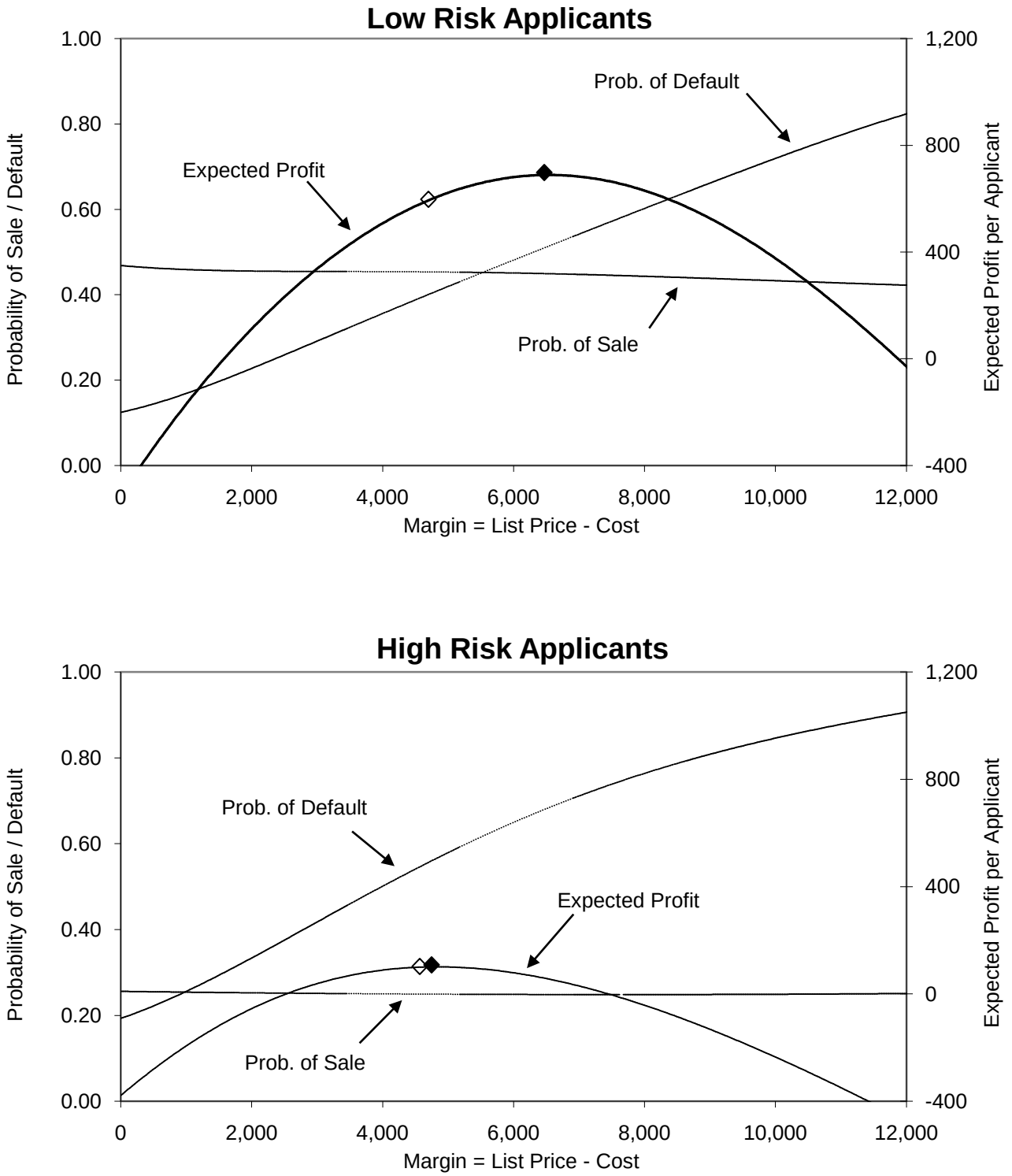
## Figure 6(a): Effect of Minimum Down Changes




*Notes*: Based on model estimates for all applicants. The horizontal axis represents the required minimum down payment applied to applicants in each risk category. The left-hand y-axis represents the probabability of sale (for applicants) and probability of default (for buyers). The right-hand y-axis represents expected profit per applicant, calculated as the probability of sale times net operating revenue per sale, where net operating revenue = down payment + PV of loan payments + PV of recovery - vehicle cost - unobserved cost. The unobserved cost is estimated as described in Section 6. White diamonds show observed average minimum down payments for each credit category. Black diamonds show optimal minimum down payments based on model estimates.

## Figure 6(b): Effect of Price-Cost Margin Changes




*Notes*: Based on model estimates for all applicants. The horizontal axis represents the target margin (list price minus cost) for applicants in each risk category. The left-hand y-axis represents the probabability of sale (for applicants) and probability of default (for buyers). The right-hand y-axis represents expected profit per applicant, calculated as the probability of sale times net operating revenue per sale, where net operating revenue = down payment + PV of loan payments + PV of recovery - vehicle cost - unobserved cost. The unobserved cost is estimated as described in Section 6. White diamonds show observed average targeted price-cost margins for each credit category (which vary across consumers only slightly as a function of the assigned car they are offered). Black diamonds show optimal price-cost margins based on model estimates.

**Figure 7(a): Down Payments vs. Default Rates without Credit Scoring**

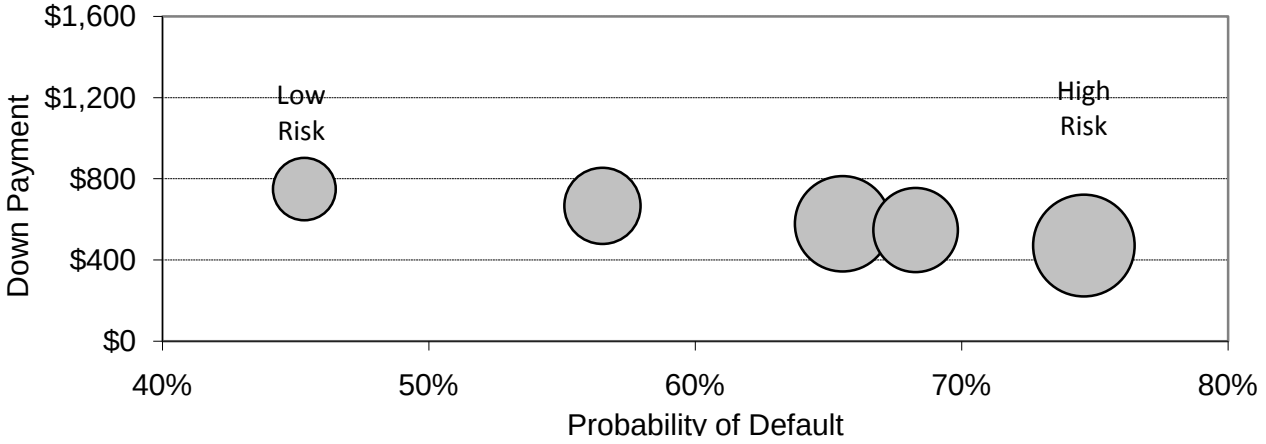



**Figure 7(b): Down Payments with Scoring vs. Non-Scoring Default Rates**

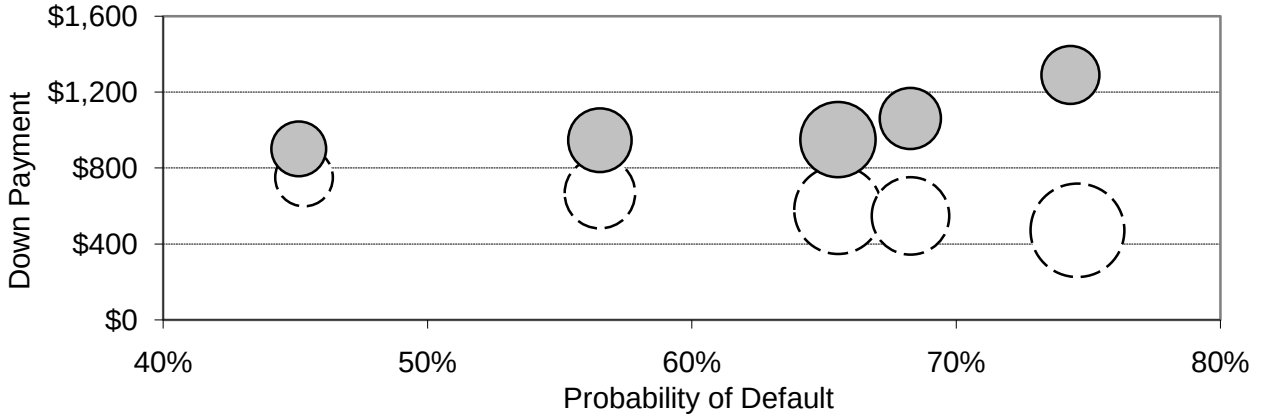
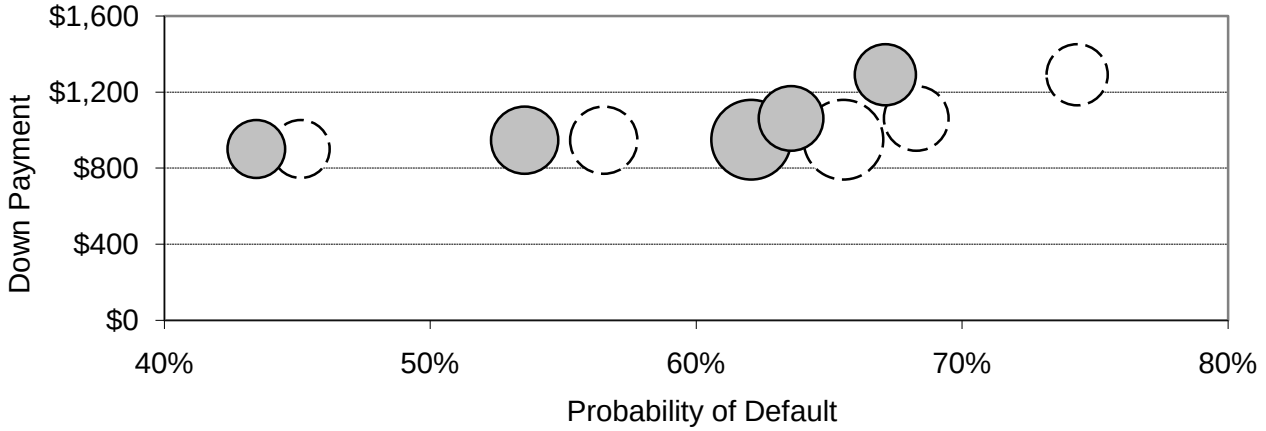



**Figure 7(c): Down Payments vs. Default Rates with Credit Scoring**




Notes: This figure shows model-estimated average down payments and default rates for borrowers in different risk categories under two different pricing regimes, one with no minimum down payment and one with observed minimum down payments that vary by risk category. Each bubble in the figure's panels represents an observable risk category defined by the lender's proprietary scoring system. Risk categories are monotonically increasing from the left (low risk) to the right (high risk). The size of each bubble represents the estimated loan volume for a given risk category. Estimated average down payments, default rates, and loan volumes are calculated for all applicants based on the demand model estimated in the paper. Panel 7(a) shows estimated average down payments, default rates, and loan volumes for each risk category assuming all applicants face observed car prices and no minimum down payment. In the absence of minimum down payments, higher risk borrowers put less down and represent a large fraction of loans. Panel 7(b) shows estimated average down payments assuming applicants in each risk category face their observed average minimum down payment, vs. default rates from Panel 7(a). This illustrates two effects of the lender's ability to offer different contract terms to borrowers with different observable risk types: riskier borrowers put more down on average and riskier borrowers are less likely to buy. Panel 7(c) adjusts Panel 7(b) to include estimated default rates based on observed average minimum down payments. The average down payments and bubble sizes are the same as in Panel 7(b). This illustrates a third benefit of tailoring contract terms to different observable risk types: higher minimum down payments lead to lower default rates due smaller loan sizes and improved selection.

## Table 1: Summary Statistics

| | Obs* | Mean | Std. Dev. | 5% | 95% |
|---|---|---|---|---|---|
| *Applicant Characteristics* | | | | | |
| Age | N | 32.8 | 10.7 | 19 | 53 |
| Monthly Income | N | 2,414 | 1,074 | 1,299 | 4,500 |
| Home Owner | N | 0.15 | - | - | - |
| Live With Parents | N | 0.18 | - | - | - |
| Bank Account | N | 0.72 | - | - | - |
| Risk Category | | 0.34 | - | - | - |
| Low | N | 0.27 | - | - | - |
| Medium | N | 0.45 | - | - | - |
| High | N | 0.29 | - | - | - |
| Car Purchased | N | 0.34 | - | - | - |
| *Buyer Characteristics* | | | | | |
| Age | 0.34N | 34.7 | 10.8 | 20 | 55 |
| Monthly Income | 0.34N | 2,557 | 1,089 | 1,385 | 4,677 |
| Home Owner | 0.34N | 0.17 | - | - | - |
| Live With Parents | 0.34N | 0.16 | - | - | - |
| Bank Account | 0.34N | 0.76 | - | - | - |
| Risk Category | | | | | |
| Low | 0.34N | 0.35 | - | - | - |
| Medium | 0.34N | 0.47 | - | - | - |
| High | 0.34N | 0.17 | - | - | - |
| *Car Characteristics* | | | | | |
| Acquisition Cost | 0.34N | 5,090 | 1,329 | 3,140 | 7,075 |
| Total Cost | 0.34N | 6,096 | 1,372 | 4,096 | 8,212 |
| Car Age (years) | 0.34N | 4.3 | 1.9 | 2.0 | 8.0 |
| Odometer | 0.34N | 68,775 | 22,091 | 31,179 | 102,299 |
| Lot Age (days) | 0.34N | 33 | 44 | 1 | 122 |
| Car Price | 0.34N | 10,777 | 1,797 | 8,095 | 13,595 |
| *Transaction Characteristics* | | | | | |
| Min. Down Payment (applicants) | N | 648 | 276 | 400 | 1,200 |
| Min. Down Payment (buyers) | 0.34N | 750 | 335 | 400 | 1,400 |
| Interest Rate (APR) | 0.34N | 26.2 | 4.4 | 17.7 | 29.9 |
| Loan Term (months) | 0.34N | 40.5 | 3.7 | 35.0 | 45.0 |
| Down Payment | 0.34N | 942 | 599 | 400 | 2,000 |
| Loan Amount | 0.34N | 10,740 | 1,801 | 7,982 | 13,559 |
| Monthly Payment | 0.34N | 395 | 49 | 314 | 471 |
| *Loan Outcomes (uncensored sales only)* | | | | | |
| Default | 0.13N | 0.61 | - | - | - |
| Fraction of Payments Made | 0.13N | 0.58 | 0.38 | 0.04 | 1.00 |
| Loan Payments | 0.13N | 7,972 | 5,635 | 491 | 16,587 |
| Nonzero Recovery (all defaults) | 0.08N | 0.78 | - | - | - |
| PV of Recovery (all recoveries) | 0.06N | 1,579 | 1,328 | 231 | 4,075 |
| Gross Operating Revenue | 0.13N | 9,614 | 5,192 | 2,169 | 17,501 |
| Net Operating Revenue | 0.13N | 3,333 | 5,020 | -3,906 | 10,284 |

Notes: Loan Payments, Gross Operating Revenues, and Net Operating Revenues are in present value terms. Gross Operating Revenue = Down Payment + PV of Loan Payments + PV of Recovery. Net Operating Revenue = Gross Operating Revenue - Total Car Cost including sales tax.

* To preserve the confidentiality of the company that provided the data, we do not report the exact number of applications.

| | | |
|---|---|---|
| *Probability of Minimum Down Payment* | | |
| All Buyers | 0.431 | 0.461 |
| Low Risk | 0.234 | 0.335 |
| Medium Risk | 0.428 | 0.463 |
| High Risk | 0.570 | 0.566 |
| *Average Loan Size* | | |
| All Buyers | $10,709 | $10,674 |
| Low Risk | $11,047 | $10,837 |
| Medium Risk | $10,660 | $10,636 |
| High Risk | $9,992 | $10,093 |
| *Probability of Payment* | | |
| All Buyers | 0.390 | 0.405 |
| Low Risk | 0.559 | 0.551 |
| Medium Risk | 0.363 | 0.378 |
| High Risk | 0.289 | 0.330 |
| *Fraction of Payments Made* | | |
| All Buyers | 0.594 | 0.620 |
| Low Risk | 0.715 | 0.740 |
| Medium Risk | 0.576 | 0.599 |
| High Risk | 0.521 | 0.554 |

Notes: Raw data moments are computed directly from the estimation sample (see notes to Table 3 for details). Demand model moments are computed based on the econometric model described in Section 5 and the parameter estimates presented in Table 3. Close rate, or probability of sale, is computed using data on all applicants. Probability of making minimum down payment and average loan size are conditional on sale. Probability of payment and fraction of payments made are computed using data on uncensored loans only.

## Table 3: Demand Estimates

| ***Dependent Variable*** | ***Probability of Sale*** | | ***Down Payment*** | | ***Payments Made*** | |
|---|---|---|---|---|---|---|
| | Marg. Eff. | Std. Err. | Coeff. | Std. Err. | Coeff. | Std. Err. |
| *Offer Variables* | | | | | | |
| Minimum Down ($1000s)* | -0.231 | (0.011) | - | - | - | - |
| Negotiated Price ($1000s) | -0.025 | (0.006) | 0.106 | (0.086) | - | - |
| Maximum Interest Rate (%) | 0.002 | (0.001) | 0.024 | (0.004) | -0.030 | (0.006) |
| Term (years) | -0.060 | -(0.019) | -0.375 | (0.100) | 0.070 | (0.101) |
| Loan Amount ($1000s) | - | - | - | - | -0.386 | (0.041) |
| *Car Characteristics* | | | | | | |
| Car Cost ($1000s) | 0.011 | (0.003) | 0.162 | (0.087) | 0.389 | (0.043) |
| Premium (Cost > $7,500) | 0.013 | (0.004) | 0.458 | (0.036) | 0.059 | (0.070) |
| Car Age (years) | 0.001 | (0.000) | 0.000 | (0.005) | -0.049 | (0.009) |
| Odometer (10,000s) | -0.003 | -(0.001) | 0.000 | (0.004) | 0.005 | (0.008) |
| Lot Age (months) | 0.001 | (0.000) | -0.041 | (0.017) | -0.079 | (0.011) |
| *Individual Characteristics* | | | | | | |
| Income ($1,000s/month) | 0.021 | (0.007) | -0.002 | (0.009) | 0.078 | (0.014) |
| Age | 0.007 | (0.002) | -0.015 | (0.004) | 0.014 | (0.007) |
| Age squared | -6E-05 | -(2E-05) | 2E-04 | (5E-05) | -1E-04 | (8E-05) |
| Bank Account | 0.034 | (0.011) | 0.036 | (0.017) | 0.214 | (0.031) |
| House Owner | -0.032 | -(0.010) | 0.006 | (0.019) | 0.008 | (0.042) |
| Lives with Parents | 0.003 | (0.001) | 0.043 | (0.019) | -0.111 | (0.036) |
| *Credit Grade Fixed Effects* | | | | | | |
| Representative Low Risk Grade | -0.012 | -(0.004) | 0.349 | (0.060) | 0.869 | (0.101) |
| Representative Medium Risk Grade | 0.004 | (0.001) | 0.260 | (0.056) | 0.316 | (0.081) |
| Representative High Risk Grade | omitted | omitted | omitted | omitted | omitted | omitted |
| *Seasonal Effects* | | | | | | |
| Tax Season | 0.137 | (0.043) | 0.545 | (0.027) | 0.023 | (0.065) |
| Other Fixed Effects | | | Year, Month, City, Credit Grade | | | |

| *Covariance Matrix* | Xi | | Eps | | Eta | | Nu | |
|---|---|---|---|---|---|---|---|---|
| Xi (purchase equation) | 1.000 | - | 0.062 | (0.007) | 0.003 | (0.014) | -0.018 | (0.006) |
| Epsilon (down equation) | 0.062 | (0.007) | 0.557 | (0.027) | 0.022 | (0.014) | -0.024 | (0.058) |
| Eta (repayment equation) | 0.003 | (0.014) | 0.022 | (0.014) | 2.334 | (0.047) | 0.078 | (0.015) |
| Nu (price equation) | -0.018 | (0.006) | -0.024 | (0.058) | 0.078 | (0.015) | 0.283 | (0.020) |

Notes: All estimates based on demand model described in Section 5. The sample for the purchase and down payment equations is a random sample of all applicants; the sample size is 0.1N, where N >> 50,000 (see Table 1). The sample for the fraction of payments made equation is all sales; sample size is ~0.034N. Reported estimates in the first colunm show the marginal effects of a one unit change in each of the explanatory variables on the probability of sale. For dummy variables, this is computed by taking the difference between the probability of sale when the variable is equal to 1 and the probability when the variable is equal to 0 (holding other variables fixed). For continuous variables, this is computed by taking a numerical derivative of the probability of sale w.r.t. the continuous variable. Estimates in the second column show the effects of a one unit change in each explanatory variable on desired down payment (in $1,000s). For instance, a $1,000 increase in price raises the desired down payment of the average applicant by $106. Estimates in the third column show the effects of a one unit change in each explantory variable on the log of fraction of payments made. For example, a $1,000 increase in loan amount decreases the fraction of payments made by 1-exp(-0.386), or 32 percent. Standard errors are bootstrap standard errors from 60 resamplings.

## Table 4: Estimates of Indirect Cost

| | Cost Estimate | Standard Error |
|---|---|---|
| Uniform | 2,481 | (68) |
| Credit Category | | |
| Low Risk | 3,265 | (111) |
| Medium Risk | 2,373 | (74) |
| High Risk | 1,932 | (88) |
| Seasonal | | |
| Non Tax Season | 2,498 | (70) |
| Tax Season | 2,285 | (151) |
| Learning | | |
| Learning #1 | 1,800 | (332) |
| Learning #2 | 2,200 | (240) |
| Learning #3 | 2,800 | (222) |
| Learning #4 | 2,600 | (207) |

Notes: All results are estimates of the firm's unobserved indirect cost of sale. The top six rows are based on the first-order condition for pricing described in Section 6. Estimates assume a firm discount rate of 10 percent. Standard errors are bootstrap standard errors from 60 resamplings.

The four estimates which are based on "learning" inequalities reflect the four conditions below. In each case, the estimates are computed by minimizing the sum of squared violations of the inequality dictated by each condition. For indirect cost estimates (2) and (4), violations are weighted by the number of applicants in a given period.

(1) Period t+1 pricing dominates period t pricing in period t+1
(2) Period t+1 pricing dominates period t pricing in period t+1 (weighted by apps)
(3) Period t+1 pricing dominates period t pricing in period t
(4) Period t+1 pricing dominates period t pricing in period t (weighted by apps)

For example, to

## Table 5: Value of Credit Scoring

| | Low Risk | Med Risk | High Risk | All Applicants |
|---|---|---|---|---|
| *Minimum Down Payment* | | | | |
| Observed pricing | $400 | $600 | $1,000 | - |
| Optimal credit-based pricing | $0 | $700 | $1,550 | - |
| Optimal uniform pricing | $800 | $800 | $800 | $800 |
| Pricing with perfect information | - | - | - | - |
| *Close Rate* | | | | |
| Observed pricing | 0.451 | 0.398 | 0.249 | 0.343 |
| Optimal credit-based pricing | 0.505 | 0.355 | 0.119 | 0.305 |
| Optimal uniform pricing | 0.345 | 0.355 | 0.312 | 0.339 |
| Pricing with perfect information | 0.504 | 0.443 | 0.266 | 0.394 |
| *Profit Conditional on Sale* | | | | |
| Observed pricing | $1,967 | $797 | $106 | $996 |
| Optimal credit-based pricing | $1,840 | $944 | $679 | $1,201 |
| Optimal uniform pricing | $2,137 | $944 | -$34 | $883 |
| Pricing with perfect information | $2,292 | $1,333 | $918 | $1,490 |
| *Expected Profit per Applicant* | | | | |
| Observed pricing | $886 | $317 | $26 | $341 |
| Optimal credit-based pricing | $930 | $335 | $81 | $366 |
| Optimal uniform pricing | $736 | $335 | -$11 | $299 |
| Pricing with perfect information | $1,156 | $591 | $244 | $587 |

Notes: All results based on model estimates. Close Rate is the probability that an applicant purchases a car. Profit Conditional on Sale is defined as net operating revenue (down payment + PV of loan payments + PV of recovery - vehicle cost) minus our estimate of indirect cost in the top row of Table 4. Expected profit per applicant is equal to the close rate times profit conditional on sale. Each counterfactual represents a different minimum down payment policy. List prices are held fixed at observed values in all counterfactuals. (1) "Observed pricing" describes outcomes based on the company's observed minimum down payments, which vary both over time and across credit categories. (2) "Optimal credit-basd pricing" describes a counterfactual in which minimum down payments vary by credit category in order to maximize expected profit per applicant in each observed pricing period. (3) "Optimal uniform pricing" describes a counterfactual in which a single minimum down payment, which is constant across credit categories, is chosen to maximize expected profit per applicant in each pricing period. (4) "Pricing with perfect information" escribes a counterfactual in which the firm sets a minimum down payment for each applicant equal to the maximum amount that the applicant is able to put down, and only sells to applicants with positive expected hprofits at this minimum down payment., and given their unobservables at the time of purchase (xi and epsilon).

## Table 6: Credit Scoring as a Barrier to Entry

*(Incumbent Profit per Applicant, Entrant Profit per Applicant)*

| | Incumbent Prices Uniformly | Incumbent Prices by Risk Category |
|---|---|---|
| No Entrant (Monopoly) | ($299, $0) | ($366, $0) |
| Entrant Prices Uniformly | ($132, $132) | ($196, $93) |
| Entrant Prices by Grade | ($93, $196) | ($166, $166) |

*Equilibrium Minimum Downs*

| | Uniform | By Risk Category (Low, Med, High) |
|---|---|---|
| Monopoly | $750 | $50, $750, $1450 |
| Vs. Uniform | $450 | $50, $400, $1000 |
| Vs. Risk-Based | $750 | $0, $400, $1050 |

Notes: All results based on model estimates. Each cell in the first panel of the table presents the expected profits per applicant for an incumbent lender (first) and the profits per applicant for a potential entrant (second), calculated at the equilibrium minimum down payments shown in the corresponding cell of the second panel.
The top row presents the case of no entrant, which is also presented in Table 5. The second row presents the case where an entrant prices uniformly (i.e. sets one minimum down for all risk categories), and the third row represents the case where an entrant prices by risk category. Each column represents the pricing strategy of the incumbent firm.
In each scenario, we find the Nash Equilibrium of the duopoly game in which each firm simultaneously either sets uniform or risk-based minimum down payments. We assume that applicants who choose to put down more than either firm's minimum down payment are randomly split between the two firms, and other applicants choose the lender with the lowest minimum down payment. We also assume that car prices remain the same in all scenarios. Expected profits conditional on sale are then calculated using the estimated repayment equation (Table 3, column 3) and estimated indirect costs (Table 4).