

# Contract Pricing in Consumer Credit Markets\*

Liran Einav, Mark Jenkins and Jonathan Levin<sup>†</sup>

January 2008

**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 contracting terms, in particular car price and required down payment, resolve 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, so firms care about both the quantity and quality of demand.

---

\*We are grateful to many seminar participants for many useful comments. 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). The help of former Stanford students Will Adams and Ernie Garcia has been invaluable in advancing this research.

<sup>†</sup>Department of Economics, Stanford University, Stanford CA 94305-6072 and NBER (Einav and Levin); email: leinav@stanford.edu, mwj@stanford.edu, jdlevin@stanford.edu.

# 1 Introduction

In recent years, the industrial organization of credit and insurance markets has been of central importance to policy-makers and practitioners. Yet compared to the rapid advance of empirical methods for studying demand and pricing in product markets, the analysis of these contract markets has lagged behind. A primary reason for this has been the perceived difficulty of estimating demand and supply behavior in markets with adverse selection and/or moral hazard. In this paper, we illustrate how one can adapt standard empirical tools for demand and pricing analysis to contract markets that are characterized by both incentive and informational problems. We apply these methods to study the market for subprime auto loans.

Our development builds on models of insurance demand formulated by Cardon and Hendel (2001) and Cohen and Einav (2007).<sup>1</sup> In a similar spirit to those papers, we build a demand system for loan contracts in which choice behavior and transaction outcomes arise from a combination of consumer and contract characteristics. We then go a step further by marrying the demand system to supply-side decisions about contract design. This requires us to tackle several complications relative to a standard product market analysis. In credit markets, firms care about the identity of their customers as well as the quantity of sales. In addition, contract terms may affect transaction outcomes; for instance, charging a higher interest rate may increase the likelihood of default. Also, contracts may have several dimensions that are easy to adjust, undermining the usual assumption that non-price product characteristics are fixed, at least in the short run. Despite these complications, we provide a relatively simple framework that permits analysis of cost conditions and optimal contract design.<sup>2</sup>

While one of our goals is to illustrate a general approach to studying contract markets, our focus is on the specific market for used auto sales and subprime loans.<sup>3</sup> Here we make use of extraordinarily rich transaction-level data from a large auto sales company. The company specializes in selling to consumers with low incomes or poor credit histories — the so-called “subprime” market. This market is attractive for studying pricing and contract design in the presence of informational

---

<sup>1</sup>See also Einav, Finkelstein and Schrimpf (2007) for a related application to annuity demand, and Chiappori and Salanie (2000) and Finkelstein and McGarry (2006) as representative of a closely-related literature on “testing for asymmetric information.”

<sup>2</sup>Some of the issues that arise here are conceptually similar to those in Wolak (1994) and Perrigne and Vuong (2006), who focus on the behavior of a regulated utility.

<sup>3</sup>For additional work on informational problems in consumer lending, see Ausubel (1991, 1999), Edelberg (2004, 2006), and Karlan and Zinman (2007a, 2007b).

problems. There is substantial borrower heterogeneity and default risk, so the ability to originate profitable loans depends crucially on designing offers that attract lower risk borrowers and that facilitate repayment. In addition, the availability of credit bureau information allows firms to base financing options on customer risk profiles and allows us to study the value of risk-based financing. Finally, offers to customers can vary on multiple dimensions: car price, required down payment, interest rate and loan length, so we can investigate the screening and incentive roles played by different contract parameters.

We use a model of subprime auto transactions to quantify the trade-offs involved in contract design. In earlier work (Adams, Einav and Levin, 2007), we found evidence that subprime lenders face both repayment problems (with larger loans less likely to be repaid) and selection problems (substantial observed and unobserved borrower heterogeneity). We take the descriptive findings reported in that paper as a guide in formulating the present model. The advantage of the current approach is that it unifies the insights from the earlier paper in a single empirical model, and allows us to quantify in dollar terms the importance of incentive and selection effects, the value of information about consumers, the ability to vary prices over time, and possible departures from optimal pricing.

The demand system we develop captures a customer's decision of whether to purchase a car and how to finance the transaction jointly with the loan repayment process. In a nutshell, the model consists of four equations that link individual customer, car and financing characteristics to (i) a purchase decision, (ii) a financing decision, (iii) a repayment history, and (iv) a recovery in the event of default. The demand equations are linked structurally, so an individual's financing decision affects loan size, which in turn affects repayment. The model also permits individual customer and contract characteristics to affect each decision, so customers with a higher credit score may be less likely to default but also less inclined to take the largest possible loan. Importantly we allow consumer decisions to be correlated conditional on observables, so we account for the possibility that buyers who are inclined to borrow more for unobservable reasons are also more likely to default.

Our demand estimates reflect the importance of consumer liquidity, and highlight the significance of both moral hazard and adverse selection. Consistent with the results of Adams, Einav and Levin (2007), we find that purchasing decisions are highly sensitive to minimum down payment requirements and substantially less sensitive to car prices. Changes in car prices appear to translate primarily into larger loans. We also find a strong correlation between consumer liquidity

and subsequent default. An implication is that “marginal” buyers, who are just able to meet the required down payment, represent much worse risks than average buyers; roughly 60 percent of buyers default on their loans, but marginal buyers default at an even greater rate of 69 percent. Finally, default rates are quite sensitive to loan size: a \$1,000 larger loan increases the probability of default by 5 percent.

The estimated demand model provides a building block to study pricing and contract design. We use relatively weak assumptions about optimality, combined with observed pricing decisions, to estimate the firm’s indirect or shadow cost of capital adjustment. Because we start with excellent data on observed costs, we are able to explore the implications of different notions of pricing “optimality”, and also assess the profitability of alternative pricing policies.

The main idea we explore with the pricing model is that changes in offered terms have very different effects on the composition of purchasers and their borrowing and repayment behavior. We focus on two particular dimensions: car prices and required down payments. The required down payment plays perhaps the most distinct role. For many buyers, a small increase in the minimum down payment essentially has no effect; they intend to make a down payment above the minimum in any case. For buyers who intend to make the minimum down payment, however, an increase in the requirement either leads them to take a smaller loan or causes them to forego the purchase altogether. Because these buyers are relatively illiquid compared to an average buyer, and represent relatively high risks, there can be a benefit to reducing their loan size and potentially even a benefit to screening them out.

Changes in car prices play a dramatically different role. An increase in car prices has relatively little effect on the volume of sales or on the size of buyers’ down payments. Instead, the primary effect of an increase in car prices is to increase loan sizes. This raises monthly payments but also the probability of default. Optimal prices resolve this trade-off to balance the benefits of larger payments with the correspondingly faster and more likely default. The model suggests that optimal interest rate offers involve a similar balancing effect.

Having outlined the basic trade-offs in contract design, we use the pricing model to quantify the value of using individual information about consumers to make financing offers. 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 their current liquidity. Compared to the case of uniform minimum down requirement, we find that the observed pricing increases profits by 10%, that optimal pricing would increase profits by 18%, and that perfect information about current liquidity would increase expected profits by 90%. Our findings suggest that the company might benefit from lowering minimum down payments somewhat for the best risks, while raising them for the highest risks. Finally, we quantify the value of information as increasing the barrier for potential entrants.

The plan of the paper is as follows. Section 2 outlines a general model of demand and pricing in contract markets. Section 3 describes subprime lending and our auto sales data. We develop the demand and supply model, and discuss identification and estimation in Sections 4-6. Sections 7 and 8 present the results, and Section 9 concludes.

## 2 Demand and Pricing in Contract Markets

We consider a market consisting of a population of consumers, each described by a vector of characteristics  $\zeta$ , and a single seller. To keep things simple, in this section we imagine the seller offers a single contract described by a vector of terms  $\phi$ . In the case of a financed auto sale, the contract terms might include the car being offered, its price, a maximum loan size or down payment requirement, an interest rate, a schedule for payments, and so forth.

Each consumer chooses whether or not to accept the contract. We represent this decision by the function  $g(\phi, \zeta)$ ; that is, a consumer with characteristics  $\zeta$  accepts a contract  $\phi$  if and only if  $g(\phi, \zeta) \geq 0$ . With this notation, total sales are

$$Q(\phi) = \int \mathbf{1}\{g(\phi, \zeta) \geq 0\} dF(\zeta), \tag{1}$$

where  $F(\cdot)$  is the population distribution of individual characteristics.

Conditional on purchase, a transaction results in an outcome  $y(\phi, \zeta)$ , which in a loan market might be the fraction of loan payments that are made. The seller realizes a variable profit or net revenue  $r(\phi, y)$  that depends on the contract terms and the transaction outcome. Again in the loan setting, a lender’s return depends on the size of the loan, the interest rate, and the repayment history.

This minimal model already allows for both selection and incentive effects. Selection effects

arise if buyer characteristics affect both the decision to purchase and transaction outcomes. In adverse selection models of borrowing, for example, buyers at greater risk of default demand larger loans. Incentive effects arise if contract terms affect transaction outcomes. For instance, in moral hazard models of borrowing, a larger loan increases the probability of default.

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

$$\max_{\phi \in \Phi} \Pi(\phi) = Q(\phi) \cdot \mathbb{E}[r(\phi, y) | g(\phi, \zeta) \geq 0]. \quad (2)$$

It is typical in analyzing product markets to treat all product characteristics other than price as fixed, at least in the short run. In credit and insurance markets, firms may be able to adjust several dimensions of their offers fairly easily. In our auto sales context, we focus on two particular pricing terms: car price and the minimum down payment.

For now, however, let's focus on a single contract dimension and assume that  $g(\cdot)$  is continuous and strictly decreasing in  $\phi$ . In this case, the effect of a small change in the offered contract is:

$$\frac{d\Pi(\phi)}{d\phi} = \frac{dQ(\phi)}{d\phi} \cdot \mathbb{E}[r(\phi, y(\phi, \zeta)) | g(\phi, \zeta) = 0] + Q(\phi) \cdot \mathbb{E}\left[\frac{dr(\phi, y(\phi, \zeta))}{d\phi} | g(\phi, \zeta) \geq 0\right]. \quad (3)$$

The first term reflects the loss of customers who are just on the margin. The cost of losing these customers depends on their profitability — in principle, a marginal buyer could be more or less profitable than an average buyer. The second term reflects the change in the return on inframarginal buyers. In traditional product markets a one dollar increase in price translates directly to a one dollar increase in revenue; here a change in the signed contract may have a more complex effect. For instance, an increase in the interest rate on a loan raises the monthly payment but might lower the fraction of payments that are made.

It is possible to connect optimal pricing to the standard Lerner condition. To do this, let  $R(\phi) = \mathbb{E}[r(\phi, y) | g(\phi, \zeta) \geq 0]$  denote the seller's expected revenue conditional on sale. With this notation, expected profit is  $\Pi(\phi) = Q(\phi)R(\phi)$  and the first order condition for optimal pricing is  $(R/\phi)/R_\phi = (-Q/\phi)/Q_\phi$ . The firm equates the inverse elasticity of net revenue with the inverse elasticity of demand. In the standard product market case, the relevant contract dimension is price ( $\phi = p$ ) and the revenue from a transaction is  $r(p, \zeta) = p - c$ , so the inverse elasticity of net revenue is simply the markup  $(p - c)/p$ . Viewed in this light, pricing in contract markets obeys a Lerner equation, only the impact of a price change on the average revenue from each sale might include

selection and incentive effects.

Now consider an empirical perspective. The economic fundamentals of the model are the distribution of customer characteristics  $F(\cdot)$ , the choice function  $g(\phi, \zeta)$ , the outcome function  $y(\phi, \zeta)$ , the net revenue function  $r(\phi, y)$ , and the set of possible contracts  $\Phi$ . In our analysis below, individual data on choices and outcomes is available. For each individual  $i$ , we observe a subset of individual characteristics, the contract she faces  $\phi_i$ , her purchase decision  $q_i \in \{0, 1\}$ , and if she purchases, an outcome  $y_i$ .

A natural approach to demand estimation therefore combines a latent variable “selection” equation,  $q_i = 1$  if and only if  $g(\phi_i, \zeta_i) \geq 0$ , and a “treatment” equation  $y_i = y(\phi_i, \zeta_i)$ . These equations permit estimates of  $F(\cdot)$ ,  $g(\cdot)$  and  $y(\cdot)$ . The second step is to use observed pricing behavior to recover unobserved components of the cost structure, i.e.  $r(\cdot)$ , from the first-order conditions for optimal pricing. Moreover, to the extent that observed pricing behavior generates more restrictions on the data than there are unknowns, one can test if pricing decisions are indeed optimal.

The differences between this approach and the standard analysis of product markets are small. On the demand side, the only difference is the existence of the outcome function. In estimating demand for something like cereal, the idea is to use observed market shares or individual choice data to estimate the distribution of customer characteristics  $F(\cdot)$  and the choice function  $g(\cdot)$ . Here we just add an additional outcome equation and use data on outcomes to also recover  $y(\cdot)$ . On the supply side, the econometric differences are even less important — we still use first-order conditions for profit maximization to recover marginal cost parameters — the only difference is that the first-order conditions must be modified to account for selection and incentive effects.

### 3 Used Cars and Subprime Auto Loans

Our study makes use of data from a company that operates used car dealerships across the United States. The company specializes in selling to individuals with low incomes or poor credit histories. Customers who arrive at a dealership fill out a loan application, identify a car they might purchase and are quoted a price for it, and are given financing options that reflect their credit-worthiness. Virtually all buyers finance a large fraction of their purchase, so the company originates a substantial number of subprime loans. Defaults are common, and recoveries typically constitute only a small fraction of car cost. For this reason, both customer selection and the structure of financing are

of central importance, making this an attractive setting to study optimal pricing and design of consumer credit contracts.

For the present study, we obtained data on all loan applications and sales from June 2001 through December 2004. We observe well over 50,000 loan applications (we do not report the exact number which is proprietary), about a third of which result in a purchase.<sup>4</sup> We also obtained data on the loan terms being offered at any given time and the cost and list price of each car on the lot. In addition to this data, we are able to track loan repayments and recoveries up to April 2006. Compared to most industry studies, the data are extraordinarily detailed and complete.

Table 1 reports summary statistics on the applicant population and the terms and outcomes of observed transactions. The typical applicant has a household income just under \$29,000 a year, and appears to have relatively little access to savings or credit. Only 17 percent of the applicants have a FICO score above 600, a typical cut-off for obtaining a standard bank loan, and 18 percent of applicants have no FICO score at all. Although a small fraction, fifteen percent, are homeowners, almost a third have no bank account.

The company's inventory consists primarily of used cars between three and five years old. The company sets a list price for each car, but actual sale prices are negotiated at the dealership and can depart somewhat from the list price. The average sale price is just under \$11,000. Buyers must make a minimum down payment of up to two thousand dollars but may finance any fraction of the remainder of the purchase. Most of the loans originated by the company have three to four year terms and annual interest rates of 25-30%.

Both the minimum down payment and the offered interest rate depend on an applicant's credit category. The credit category is a discretized version of a proprietary credit score the company assigns based on the applicant's characteristics and credit history. Although interest rates are based on credit category, around half of the loans we observe are at state-mandated maximum rates, and much of the interest rate variation in the data arises from cross-state differences in rate caps.

In our empirical analysis, we focus primarily on minimum down payments and car prices, rather than interest rates or loan lengths. This focus is dictated partly by the available variation in the data. During the sample period, we observe over twenty company-wide changes in down payment

---

<sup>4</sup>We report summary statistics based on the full sample of applicants and loans, but to reduce computational time, we use a random subsample of 45,000 applicants to estimate the model. For the baseline model, the results that are based on the full sample are very similar. However, because it takes more than a month to estimate the model using the full sample, all the estimation results we report are based on this random subsample of 45,000 applicants.

requirements and two company-wide changes in car pricing. These changes, and additional discontinuities in the down payment requirements and the pricing schedule, allow consistent estimates of demand and revenue elasticities. Our analysis controls for loan length and the offered interest rate, but we are somewhat less confident in our ability to identify how changes in these financing terms affect the quality and quantity of demand.

We also emphasize the decision of whether or not to purchase and how much to finance, rather than the choice among cars. Again the motivation is two-fold. First, we are more interested in the borrowing decision than in whether customers choose a Ford Escort that is three rather than four years old. Second, adding a car choice dimension to the model adds complexity that appears to have little effect on the insights we derive about selection, liquidity and optimal pricing. Adams, Einav, and Levin (2007) provide additional discussion and evidence on the latter point.

Just over one-third of applicants purchase a car, and these individuals tend to have somewhat higher income and credit-worthiness than the average applicant. Virtually all buyers finance a large fraction of the purchase price. Forty-three percent make exactly the minimum down payment, and 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 loans end in default. Our data ends before the last payments are due on some loans, but of the loans with uncensored payment periods, only 39% are repaid in full. Moreover, when defaults occur, they 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 occur within the first half of the loan term. The recovered value in the event of default is typically a fairly small fraction of the car cost. For 22 percent of defaults we observe, no recovery is made at all, sometimes because the car has been in an accident or stolen. But 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. Taken together, these facts lead to a highly bimodal distribution of per-sale profits (Figure 1).

Three economic features of the market deserve particular attention. First, as emphasized by Adams, Einav and Levin (2007) purchasing decisions are highly sensitive to customer liquidity. Relatively small increases in the required down payment appear to have a disproportionately large effect on purchasing, and transitory income shocks appear to have a similarly dramatic effect. For instance, Adams, Einav and Levin (2007) document a nearly 50 percent increase in sales in early

February, and connect this spike to the arrival of tax rebates, particularly for consumers who are eligible for the Earned Income Tax Credit.

A second feature of the market is that the probability of loan repayment decreases fairly dramatically with loan size. Figure 2(a) provides some rough evidence of this by plotting loan sizes in the data against repayment probabilities. Here we group buyers into “high,” “medium,” and “low” risk using the company’s credit categories, and smooth the raw data using local linear regression. For each group of buyers, the probability of repayment falls steadily with loan size.

Finally, there is substantial heterogeneity in the likelihood of default, which is strongly correlated both with buyers’ observed characteristics and with their initial financing decisions. The former is already suggested by Figure 2(a), where the likelihood of default is substantially higher for buyers with worse credit scores. To investigate further, we divide each risk group into individuals that made minimum down payments and those whose down payments exceed the minimum, and plot repayment probabilities for each of the subgroups separately. These are presented in Figure 2(b), restricting attention to the sample of uncensored loans. The default rate is 71 percent for high risk buyers, compared to 44 percent for the low risk buyers. Moreover, buyers who make a down payment of exactly the required minimum have an average default rate of 67 percent compared to a rate of 56 percent for buyers who make a down payment above the minimum. As Figure 2(b) suggests, this pattern is fairly uniform across different risk groups. The strong correlation between financing decisions and default survives the addition of controls for buyer and car characteristics as well as fixed effects for dealership and time periods.

The correlation between financing decisions and default rates has two natural explanations. One is selection: buyers who choose to finance more heavily are those buyers who are more likely to default. The alternative is a repayment or moral hazard effect: a buyer who takes a larger loan is less likely to repay, either because she cannot come up with the loan payments or because the incentive to prioritize payments is reduced. Our analysis below shows that both effects are operative and quantifies them in dollar terms.

## 4 The Empirical Model: Demand

In this section we develop a model of consumer demand. We incorporate the institutional details described in the previous section to model the decision of whether to accept a contract and the

resulting outcome. These two ingredients,  $g(\cdot)$  and  $y(\cdot)$  in our earlier notation, feed into the pricing decisions, which we model later.

Consumers in our data make several decisions: whether to purchase, which car to purchase, how much of the purchase price to pay upfront, and how long to continue making loan payments. The data provide information about each decision. So in principle, we could formulate a demand model that takes them all into account. In practice, however, we make two simplifications. First, we focus on each consumer’s decisions about whether to purchase, how much to borrow and whether and when to default, but, as already mentioned, we leave aside the issue of car choice. Second, we reduce the dimensionality of the problem by assuming that unobservable heterogeneity is captured by two variables, one that affects decisions at the time of purchase, and one that affects decisions during repayment. This assumption simplifies computation and estimation, but we believe it doesn’t affect much our results.

## 4.1 Preliminaries

We describe each applicant in the data by a vector of characteristics  $\zeta = (x^a, x^d, \varepsilon, \eta)$ . Here  $x^a$  is a vector of observed individual characteristics including age, income, credit category, and proxies for wealth, and  $x^d$  includes dealership and time dummies. The scalar characteristics  $\varepsilon, \eta$  are not observed in the data. We assume  $\varepsilon$  is known to the applicant at the time of purchase, and hence affects purchasing and borrowing, while  $\eta$  is determined later and affects repayment. Naturally they may be correlated. One interpretation is that  $\varepsilon$  and  $\eta$  reflect unobserved aspects of liquidity at the time of purchase and later, and are mechanically related. Another possibility is that buyers have private information about the likelihood of future repayment (i.e. about  $\eta$ ) and because they are forward-looking, this information is reflected in their purchasing and financing decisions (i.e. in  $\varepsilon$ ). We discuss this further in the end of this section.

As discussed earlier, we view car price and the minimum down payment as key components of the firm’s offer. Given this, we summarize contract terms by  $\phi = (x^c, p, d)$ , where  $x^c$  includes the characteristics of the applicant’s preferred car on the lot, the offered interest rate and loan length,  $p$  is the price of the applicant’s preferred car, and  $d$  the required down payment. It is useful to let  $x = (x^a, x^d, x^c)$  denote the complete vector of observed characteristics other than price and minimum down payment.

A potential transaction is described by  $(\phi_i; \zeta_i) = (p_i, d_i; x_i, \varepsilon_i, \eta_i)$ . We let  $q_i \in \{0, 1\}$  denote

the decision of whether to purchase,  $D_i$  the choice of down payment, and  $s_i$  the fraction of loan payments that get made. The purchase decision is observed for all applicants, while the down payment and repayment decisions are observed only for buyers. The goal of the model is to map the characteristics  $(\phi, \zeta)$  into observed outcomes  $(q, D, s)$ .

## 4.2 Price Negotiation

As mentioned earlier, 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  $p_i$  and the list price  $l_i$ :

$$p_i = l_i + x_i' \lambda + \nu_i. \quad (4)$$

Here  $\nu_i$  is an unobservable aspect of negotiation that we allow to be correlated with  $\varepsilon_i$ , the buyer's unobserved information at the time of purchase. A rough way to view the pricing equation is that in the context of our nonlinear demand model, it plays the same role that using list price as instrument for negotiated price would play in a linear demand model.

Note that we restrict the coefficient on list price to be one. The advantage of forcing the coefficient to be exactly one is that when we consider optimal price-setting, a one-dollar increase in list price translates exactly into a one dollar increase in negotiated price, so we can think interchangeably of the firm controlling list or (expected) negotiated prices.

## 4.3 Purchase and Financing Decisions

Faced with an offer, the consumer decides whether or not to purchase and how large a down payment to make. We model the purchase decision in standard discrete choice fashion as

$$q_i = 1 \quad \Leftrightarrow \quad g(x_i, p_i, d_i, \varepsilon_i) = x_i' \alpha_x + p_i \alpha_i + d_i \alpha_d + \varepsilon_i \geq 0. \quad (5)$$

By modeling the purchase decision as a binary choice, we are thinking about the customer's decision of whether or not to purchase her preferred car on the lot.<sup>5</sup>

To model the down payment decision, we need to account for the minimum down payment

---

<sup>5</sup>That is, the preferred car from among a small set of cars assigned to the applicant by the company. See Adams, Einav, and Levin (2007) for more details.

requirement. We specify the down payment decision as

$$D_i = \begin{cases} D_i^* = x_i' \beta_x + p_i \beta_p + \varepsilon_i & \text{if } D_i^* \geq d_i \\ d_i & \text{if } D_i^* < d_i \end{cases} \quad (6)$$

A buyer will never put down more than the purchase price  $p_i$ , but this constraint is never binding in the data, so we omit it in presenting the model.

A simplifying feature of the model is that we allow only one dimension of unobserved heterogeneity at the time of purchase. That is, it is the same unobservable  $\varepsilon_i$  that enters both equations (5) and (6). This allows us to re-write the purchasing equation as

$$g(x_i, p_i, d_i, \varepsilon_i) = D_i^* + x_i' \delta_x + p_i \delta_p + d_i \alpha_d, \quad (7)$$

where  $\delta_x = \alpha_x - \beta_x$  and  $\delta_p = \alpha_p - \beta_p$ . In our preferred specification, we only allow a subset of the observed buyer characteristics (in particular, credit category and time dummies) to enter  $g(\cdot)$  directly, rather than through  $D_i^*$ . It is also useful to define  $Z_i = d_i + x_i' \delta_x + p_i \delta_p + d_i \alpha_d$ .

With this reformulation, we can view the purchase and down payment decisions as being determined by a single latent variable,  $D_i^*$ , interpreted as  $i$ 's “ideal” down payment, and a purchase “threshold”  $d_i - Z_i$ . Provided that  $Z_i \geq 0$  (which we do not impose but obtain in our estimated model), we have the following characterization depicted in Figure 3(a). If  $D_i^* \geq d_i$ , customer  $i$  purchases and makes her ideal down payment. If  $D_i^* \in [d_i - Z_i, d_i)$ ,  $i$  purchases but is forced to pay more than her desired down payment. Finally, if  $D_i^* < d_i - Z_i$ ,  $i$  chooses not to purchase. A desirable feature of the model is that it can rationalize the large fraction of buyers who make exactly the minimum down payment.

Figure 3 is useful for thinking about the effect of changes in car prices and minimum down payments. An increase in the required down payment has no effect on a buyer's ideal down payment  $D_i^*$ , but by reducing a buyer's options, it makes purchasing less desirable; i.e. it raises the purchase threshold  $d_i - Z_i$ . It also may change the actual down payment for buyers paying the minimum. In contrast, an increase in car price may well change the ideal down payment, while at the same time raising the purchase threshold. The anticipated effects — in both cases, fewer buyers but larger down payments for at least some of the applicants that do buy — are depicted in Figures 3(b) and 3(c).

#### 4.4 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). Rather than build the model specifically around the scheduled payments, we specify a continuous-time model of repayment. Specifically, we posit that consumer  $i$  will make a fraction of payments  $s_i \in [0, 1]$ , where:

$$s_i = \begin{cases} s_i^* = \exp(x_i' \gamma_x + (p_i - D_i) \gamma_L + \eta_i) & \text{if } s_i^* \leq 1 \\ 1 & \text{if } s_i^* > 1 \end{cases}. \quad (8)$$

For loans that occur later in our sample, we do not observe the full repayment period. This creates additional censoring that we account for in estimating the model, but we defer a complete discussion of this detail to Appendix C.

We expect a key determinant of default to be the loan size  $p_i - D_i$ , which depends on the earlier financing decision. The fact that repayment depends on loan size, and the potential for correlation between  $\varepsilon$  and  $\eta$ , creates two links between choices at the time of purchase and loan performance.

#### 4.5 Stochastic Assumptions

To close the model, we specify a stochastic structure for the unobservables  $(\nu_i, \varepsilon_i, \eta_i)$ . We assume that they are normally distributed, as follows:

$$\begin{pmatrix} \nu_i \\ \varepsilon_i \\ \eta_i \end{pmatrix} \sim N(0, V) \text{ with } V = \begin{pmatrix} \sigma_\nu^2 & \rho_{\nu\varepsilon} \sigma_\nu \sigma_\varepsilon & \rho_{\nu\eta} \sigma_\nu \sigma_\eta \\ \rho_{\nu\varepsilon} \sigma_\nu \sigma_\varepsilon & \sigma_\varepsilon^2 & \rho_{\varepsilon\eta} \sigma_\varepsilon \sigma_\eta \\ \rho_{\nu\eta} \sigma_\nu \sigma_\eta & \rho_{\varepsilon\eta} \sigma_\varepsilon \sigma_\eta & \sigma_\eta^2 \end{pmatrix} \quad (9)$$

The correlation parameter  $\rho_{\varepsilon\eta}$  plays a central role. If  $\rho_{\varepsilon\eta} = 0$ , an individual's purchasing and financing decisions reveal no new information about later default. If  $\rho_{\varepsilon\eta} > 0$  then buyers who are more liquid are also better risks. In this case, conditional on the information available to the firm, the group of buyers who demand the largest loans is *adversely selected*, and in addition a marginal buyer represents a worse risk than the average buyer.

The other correlation parameters,  $\rho_{\nu\varepsilon}$  and  $\rho_{\nu\eta}$ , play a role in identification. If  $\rho_{\nu\varepsilon} = \rho_{\nu\eta} = 0$ , the negotiated price is exogenous from the standpoint of an individual customer and we can estimate demand without worrying about the negotiation process. In estimating the model, we assume that

consumers internalize relevant information revealed in price negotiation, so that  $\varepsilon$  is sufficient for  $(\varepsilon, \nu)$  in predicting  $\eta$ . It is not hard to show that this assumption is equivalent to a parameter restriction of  $\rho_{v\eta} = \rho_{\nu\varepsilon}\rho_{\varepsilon\eta}$ .

Finally, the variance parameters  $\sigma_\nu, \sigma_\varepsilon, \sigma_\eta$  capture the importance of unobserved characteristics relative to observed characteristics in negotiation and customer decisions.

## 4.6 Economic Interpretation of the Demand Model

The model we have described is a statistical representation of observed choice behavior. It is designed to be consistent with a variety of underlying behavioral assumptions, but our intent is to remain somewhat agnostic about the precise behavioral patterns underlying consumer choices. For example, we allow for correlation between desired borrowing and propensity to default. This correlation could reflect a causal link — buyers who anticipate a high chance of default know they should not make a large down payment — or simply the fact that buyers who are illiquid today and cannot make a large down payment are likely to be illiquid tomorrow and unable to make their loan payments. Similarly, we do not attempt to distinguish whether buyers default for discretionary reasons (as in moral hazard models of consumer lending) or because of changes in their employment or health status that leave them simply unable to make payments. Finally, we do not attempt to estimate behavioral parameters such as individual discount factors or the accuracy of individual expectations that might be important for welfare analysis.

There are several reasons for this. First, our main focus is on firm behavior and pricing decisions. As should be clear from Section 2, for this particular problem, what matters is what consumers do rather than why they do it. Second, building estimation around a full-blown behavioral model likely would require strong and difficult to test assumptions about consumer rationality, far-sightedness and so forth. That being said, we want some assurance that our statistical model is consistent with plausible economic behavior. In Appendix A, we provide one possible behavioral foundation for the demand model, based on rational, forward-looking utility maximization by consumers.<sup>6</sup>

---

<sup>6</sup>Although we do not pursue it, in principle it would be possible to parameterize and estimate that model using our current demand estimates as a starting point, along the lines of the two-stage estimation procedure in Bajari, Benkard and Levin (2007).

## 5 The Empirical Model: Pricing

We now turn from the demand side to consider contract pricing from the perspective of the firm. We first derive conditions for optimal pricing and then explain how we can use the conditions to infer unobserved cost parameters and assess the optimality of observed pricing decisions.

To analyze optimal pricing, we follow the framework presented in Section 2. We focus on the seller’s choice of car prices, or more precisely list prices, and minimum down payments, and treat the interest rate and loan length as fixed, although it would be possible to extend the analysis in these directions. Conceptually, the extension from the earlier set-up is straightforward. The operational difficulty lies in specifying what pricing strategies are available to the firm, and in determining how strong an assumption of optimality to impose in estimation.

### 5.1 Optimal Pricing

Let’s start by thinking about a single applicant (and an associated preferred car) with characteristics given by  $(x, \omega)$ . Here  $x = (x^a, x^d, x^c)$  are the observable characteristics and  $\omega = (\nu, \varepsilon, \eta)$  are the unobservables.<sup>7</sup> Suppose this individual is faced with a minimum down payment  $d$ , and the list price on her preferred car is  $l$ . Drawing on our modeling above, the applicant will negotiate a price  $p = l + x'\lambda + \nu$  and purchase the car if  $g(p, d, x, \varepsilon) \geq 0$ . Finally, let  $r(p, d, x, \varepsilon, \eta)$  denote the seller’s net revenue from such a transaction — we will derive a detailed expression for this term in the next section.

Putting these ingredients together, the seller’s profit from applicant  $(x, \omega)$  given a minimum down payment  $d$  and car list price  $l$  is

$$\pi(l, d; x, \omega) = \mathbf{1}\{g(l + x'\lambda + \nu, d, x, \varepsilon) \geq 0\} \cdot r(l + x'\lambda + \nu, d, x, \varepsilon, \eta). \quad (10)$$

Now consider the set of possible policies for setting the minimum down payment and list price. The information available to the firm consists of the observable characteristics  $x$  so in theory any functions  $d(x)$  and  $l(x)$  could be a feasible policy for minimum down payments and list prices. We assume that in setting offer terms the distribution of applicant characteristics is known, and

---

<sup>7</sup>In developing the pricing model, we continue to abstract from car choice. This involves a more substantive restriction than when we consider the demand model alone, because a dramatic change in the pricing policy might cause an applicant to substitute to a different preferred car. What we assume in our actual estimation is that a *small* and *uniform* increase in *all car prices* will not change an applicant’s preferred car on the lot. This is true, for example, if the indirect utility of consumers is separable and linear in price.

denote this distribution by  $F(x, \omega)$ . If the company adopts a pricing policy  $l(\cdot)$  and minimum down payment schedule  $d(\cdot)$ , total profits are

$$\Pi(l, d) = \int \pi(l, d; x, \omega) dF(x, \omega). \quad (11)$$

Therefore if  $\Phi$  is the set of feasible pricing policies, the policy  $(l, d)$  is optimal if and only if  $\Pi(l, d) \geq \Pi(l', d')$  for all  $(l', d') \in \Phi$ .

From the perspective of the manager choosing list prices and a minimum down payment schedule, a critical decision is how finely to tailor offered contract terms to the individual characteristics of applicants and the characteristics of the cars on the lot, and also how often to make adjustments. Given the wealth of available information, this problem is non-trivial. For instance, car prices could be contingent on the precise description of the car — make, model, color, cost at auction, and the price could be discounted if the car does not sell for some period of time. Similarly, financing terms such as the minimum down payment can be made contingent on an individual’s credit history, her verified income, or on the vehicle she is purchasing.<sup>8</sup>

The minimum down payment and list price schedules we observe in the data, while sophisticated, are significantly coarser than what is feasible. At any point in time at a given dealership, the minimum down payment depends only on an applicant’s credit category, and the list price depends only on car cost. A textbook analysis might also suggest that these schedules should be changed in response to any new information about the distribution of applicant characteristics. In addition, changes in the list price schedule occur relatively infrequently, only twice during the sample period. We take this coarseness into account in our estimation strategy, and then return to it in our analysis of alternative pricing policies in Section 8.

## 5.2 Revenue Accounting

In this section we derive an expression for the firm’s net revenue from a given sale. Net revenue is the sum of four components: the initial down payment, the discounted value of the stream of loan payments, the discounted recovery in the event of default, and finally the total costs of the sale. Let  $D$  denote the initial down payment,  $p - D$  the amount that is borrowed,  $z$  the interest rate on

---

<sup>8</sup>As a matter of policy the company is committed to treating applicants equally with respect to the list prices on its cars. The company does this so that differences in the financing arrangements offered to buyers are transparent and depend only on standard loan features: interest rate, length of loan, and minimum down payment (or maximum loan size).

the loan,  $T$  the length of the loan,  $S$  the length of time for which loan payments are made,  $k$  the nominal time- $S$  recovery value, and  $C$  the costs incurred in selling the car. Finally, let  $\kappa$  denote the firm's internal discount rate.

With this notation in place, the present value of net revenue from the sale is

$$r = D + \frac{\frac{1}{\kappa}(1 - e^{-\kappa S})}{\frac{1}{z}(1 - e^{-zT})}(p - D) + e^{-\kappa S}k(S) - C. \quad (12)$$

The first and last terms, the down payment and cost of the car, are realized immediately. The second term is the present value of loan payments, where the fraction in the expression represents the present value return on each dollar of loan principal. The third term is the discounted value of recovery. Clearly if the loan is paid in full so  $S = T$ , there is no associated recovery and  $k(T) = 0$ .

To relate this accounting exercise to our statistical model, consider an individual applicant with characteristics  $(x, \omega)$ , who faces a car price  $p$  and minimum down payment  $d$ . The down payment  $D(p, d, x, \omega)$ , and resulting repayment length  $S = s(p, d, x, \omega)T$  are given by the demand model of Section 4. The other loan terms  $z$  and  $T$  are taken as given (i.e. they are elements of  $x$ ). The firm's internal discount rate  $\kappa$  is a new parameter. Industry knowledge suggests that this is likely to be somewhere in the 8-12% range.

Recovery value is not a component of the demand model of Section 4. We simplify computation by modeling and estimating this quantity separately. 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. The details are described in Appendix B. In estimating recoveries separately from the rest of the demand system we assume that unobserved heterogeneity in the recovery value is independent of other unobservables. We view this as relatively unproblematic, particularly as net recovery value is a fairly small fraction of the total revenue for most consumers and is realized only for those who default.

The final component of profitability is the marginal cost incurred from a sale. We observe detailed information on the cost of acquiring each car and transporting it to the lot, so it seems reasonable to assume that we observe the direct financial costs associated with each sale. Discussions with the firm, however, indicate that for various reasons limiting deal flow is a significant concern, and enters their thinking in setting list prices and particularly minimum down payments. Because of this, we assume that in addition to the direct dollar cost  $c$  of a given car, there is an additional

indirect or shadow cost  $\psi$  associated with making an extra sale, so that total costs are

$$C = c + \psi. \tag{13}$$

Our baseline model assumes  $\psi$  is constant across our data sample, though we report other specifications that relax this assumption.

### 5.3 Empirical Specification

The goal of this section is to derive empirical restrictions arising from assumptions about optimal pricing. Our basic idea is to require that observed list prices and minimum down payments result in higher expected profit than viable alternatives. The key modeling issue is how large a set of alternative policies to consider. Because the environment is complicated, we are hesitant to impose too strong an assumption of optimality. Instead we consider two alternatives.

The first restriction we consider takes the observed pricing structure as essentially fixed in terms of changes over time, across cars and across applicants. We require only that on average the general level of prices and minimum down payments was correct from a profit-maximization standpoint. More specifically, we assume that the firm would not benefit by *uniformly* raising or lowering its list prices or minimum down payments. Letting  $l(x), d(x)$  denote the observed policies for list prices and minimum down payments, we require that:

$$\int \pi(l(x), d(x); x, \omega) dF(x, \omega) \geq \int \pi(l(x) + a, d(x) + b; x, \omega) dF(x, \omega) \quad \text{for all } a, b \in \mathbb{R}. \tag{14}$$

The second restriction we consider is motivated by the idea that the company may be “satisficing” or looking for marginal improvements in its pricing structure. For this strategy, we assume that each observed change in the minimum down payment or list price schedule improves over the prior schedule for the price period that it is in effect. To this end, we break the data into pricing periods indexed by  $\tau$ , and let  $l_\tau(x), d_\tau(x)$  denote the observed pricing policies in period  $\tau$ . We then assume that for each  $\tau$ , the observed policies  $l_\tau, d_\tau$  generate more profit in expectation than  $l_{\tau-1}, d_{\tau-1}$ :

$$\int \pi(l_\tau(x), d_\tau(x); x, \omega) dF(x, \omega) \geq \int \pi(l_{\tau-1}(x), d_{\tau-1}(x); x, \omega) dF(x, \omega) \quad \text{for all } \tau > 1. \tag{15}$$

Note that this “satisficing” approach is neither more nor less restrictive than the first approach.

## 6 Identification and Estimation

In this section we discuss estimation and the variation in the data that identifies the unknown demand and supply parameters. The variation in the data allows us to take an empirical approach that separates the estimation of demand and supply parameters. Under this approach, we start by estimating the demand system, making no assumptions about the optimality of observed contract terms. We then combine the estimated demand system with the restrictions derived from the pricing model to estimate the remaining supply parameters. Although one could, in principle, estimate demand and supply jointly, we view it as preferable to use credible identifying variation to recover demand and avoid imposing specific pricing structure except where it is required. Throughout this section, we keep the discussion at a verbal level and defer specific formulas and details of implementation to Appendix C.

### 6.1 Exploiting the Individual-Level Data

To fully exploit the rich individual-level nature of the data, we estimate the model using the choices and outcomes of loan applicants. The use of applicant data raises two issues that merit discussion.

The first issue concerns the process by which applicants arrive in the sample. By focusing on the pool of applicants, we in effect take the arrival of customers at the lot to be independent of the company’s pricing decisions, at least conditional on year and month dummies. We think this assumption is reasonable for at least two reasons. First, for many of the dealerships in our sample, pricing information was not publicly posted. Second, most of the customers who arrive at the lot are referred by standard car dealers who cannot offer financing to individuals with poor credit history. There is sufficient market segmentation that these dealers have little reason to be aware of or care about pricing changes at the firm we are studying.<sup>9</sup>

The second issue concerns the completeness of the data for non-purchasers. For applicants who do not purchase a car, we can compute their potential financing terms, but we don’t observe their preferred car on the lot or the price they might have negotiated. The obvious remedy and the

---

<sup>9</sup>An alternative approach would be to estimate demand at a more aggregated level, implicitly or explicitly constructing a pool of potential applicants, or alternatively to develop a more formal model of the applicant arrival process. In our view, the former makes sub-optimal use of the data, while the latter adds extra complication with little benefit.

one we adopt is to impute the missing data. For each applicant who does not purchase, 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-purchasing applicant the same car and negotiated price.<sup>10</sup>

## 6.2 Identification

Our analysis emphasizes car prices and minimum down payments as the two key contract terms that are determined endogenously in the demand and supply model. By focusing on these contract terms, we treat the structure of interest rates and loan lengths, the pool of applicants arriving on the lot, and their car choices as exogenous. How then do we identify the effect of car prices and minimum down payments on consumer choices?

First consider minimum down payments. Because these are set at the company level, and the company has available to it precisely the information in our data, we are not much concerned with a correlation between the minimum down payment faced by an individual in the data and her individual-level unobservable characteristics. That is, traditional endogeneity seems unlikely to be a problem. What then creates identifying variation? The pricing model suggests that the company should adjust its minimum down payment schedule in response to changes in the distribution of applicant and car characteristics. It is also possible that rougher forms of experimentation than are suggested by the model of optimal pricing would create identifying variation.

In fact, we observe more than twenty changes in the minimum down payment schedule. Even with year and month controls, these changes provide time-series variation in the minimum payments for each credit category. And because the changes are rarely uniform across credit categories, we also have a source of difference-in-differences identification. 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 (2007) we provide a more detailed discussion of this variation.

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

---

<sup>10</sup>We also experimented with iterating the price imputation based on our estimated correlation structure between  $\varepsilon_i$  and  $\nu_i$ . A short summary of those experiments is that it was a lot of work with little impact on the results.

lose some of the true variation in the data. A second difficulty, accounted for in our demand model, is that negotiated prices may be correlated with unobserved buyer 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. We observe two major changes in the margin schedule used to generate list prices, creating identifying time-series and differences-in-differences variation. 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 (2007) provide more detail.

Especially in contrast to the demand side, identification of the supply parameters is straightforward. Given consistent estimates of the demand system and the distribution of applicant characteristics, each of our assumptions about the optimality of observed pricing generates at least enough restrictions on the data to identify the main supply-side parameter: the firm's indirect cost of a sale  $\psi$ .

### 6.3 Estimating the Model

Our approach to estimation separates demand estimation from inference about the supply parameters. To do this, we first use maximum likelihood to estimate the demand system, which consists of the pricing equation (4), the purchasing and borrowing equations (5) and (6), and the repayment equation (8), combined with the stochastic assumption on the unobservables in equation (9). The likelihood function is written out in Appendix C. We then condition on the demand parameters and use the supply-side restrictions (14) and (15) to estimate the supply parameters. To do this, we focus on small uniform changes in list prices and minimum down payments (i.e. small discrete changes in  $a$  and  $b$  in the supply inequalities) and use a grid search to find values of  $\psi$  that make the observed prices optimal against these very limited alternatives. This exercise is a special case of Pakes et al. (2006), who provide a more general approach for estimation in the presence of inequality constraints. Complete details are in the Appendix.

The advantage of the two-stage approach is that it allows us to estimate the demand system using minimal assumptions about the process driving prices. We require only that decisions made at the company level, conditional on year and month dummies, are based on available company data. The disadvantage is that if one strongly believes pricing is optimal, or at least preferable to some limited and identifiable set of alternative policies, this information is not used to refine the

demand estimates.

## 7 Results

In this section, we discuss our results from the estimated demand and supply model. We break the discussion into four parts: the ability of the model to fit that data, the determinants of purchasing, borrowing and repayment behavior, the correlation between individual decisions, and finally the estimated supply parameters.

### 7.1 Comparing the Estimated Model and the Data

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 4(a), 4(b) and 4(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.

### 7.2 Purchasing, Borrowing and Repayment Behavior

The first two columns of Table 3(a) report our estimates of purchasing and borrowing behavior. The first column reports the marginal effects of the variables on the probability of sale. The second column reports the determinants of individual's down payments. Consistent with the findings of Adams, Einav and Levin (2007), customer liquidity appears to be an important factor in explaining purchase decisions. The close rate, or probability that an applicant purchases, is very sensitive to the required down payment and much less sensitive to changes in the car price. A 100 dollar increase in the required down payment lowers the probability of sale by 2.2 percentage points, which is equivalent to about 6.4% reduction in volume. In contrast, a 100 dollar increase in car prices doesn't have any economically meaningful effect on the probability of sale. We estimate that an increase in car prices also has a relatively small effect on a buyer's desired down payment; in particular, we estimate that a 100 dollar increase in car prices raises a buyer's desired down

payment by about \$9. Therefore, it appears that the primary effect of a change in car prices is to increase the size of loans that buyers take. We return to this point below.

Further evidence of the importance of liquidity can be seen in the coefficient on the month dummy variable for February. All else fixed, the close rate in February is 17 percentage points higher, or 50%, and desired down payments are 500 dollars higher. Applicant characteristics are also consistent with liquidity effects. For example, applicants with higher income and applicants with a bank account are more likely to purchase and to make larger down payments while applicants who own a house are less likely to purchase and conditional on purchase make smaller down payments.

We control for individual credit scores by including dummy variables for the company's discrete credit categories. Table 3(a) reports the coefficients for three representative categories — a high-risk category, a medium-risk category and a low-risk category. The estimates imply that medium-risk applicants are the most likely to purchase, while the high credit risk have the highest desire to borrow. One interpretation for the non-monotonicity in purchase probability is that low-risk buyers have better outside opportunities.

Our estimates of repayment behavior are reported in the third column of Table 3(a). As one might expect, 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 19 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. The company's credit score varies strongly with the expected repayment. A representative low risk buyer is expected to make 88 percent more payments than a representative high risk buyer and is 22 percent less likely to default.

### 7.3 Selection Effects

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 \$100 larger down payment is expected to make 4.2 percent more of her payments.

Now consider two buyers who are identical on all observables, only one chooses to make a \$500

larger down payment. This decision has a direct effect on loan repayment because it reduces the size of the loan. This direct effect increases the expected fraction of payments made by 9.3 percent. Moreover, there is a signalling value. The large down payment reveals greater liquidity at the time of sale, which is correlated with making payments later on. The model suggests that facing equal loan principal, the buyer who pays \$500 more down will make 11.5 percent more payments. Thus the signalling value of a larger down payment is significant, and similar in magnitude to the direct effect of lowering loan size.

An alternative way to view the selection is to compare an average buyer with an average non-buyer. The estimated model suggests that the average buyer, given an average loan size, will make 61 percent of her payments on average, and has a 60 percent chance of default. An average non-buyer, given the same loan size, would be expected to make 49 percent of her payments and would have a 73 percent chance of default. So we can view selection on the purchase decision as advantageous — the average buyer represents a substantially better risk than the average non-buyer — but selection on the borrowing decision as adverse, in the sense that buyers who demand larger loans are worse risks.

Interestingly, we estimate that the correlation between the negotiated price and the (unobserved) credit risk is small but negative. The small correlation may suggest that credit worthiness that is unobserved to headquarters is also unlikely to be observed to the sales person, so that price negotiation outcome is mainly driven by aspects that are orthogonal to liquidity and repayment risk. The negative correlation reflects the better risk applicants are able to negotiate slightly better prices. For example, an applicant who signals his better type by paying down \$100 more, faces on average a 0.15 percent lower price for the car.

## 7.4 Supply Estimates

Table 3(b) reports estimates of the key supply parameter obtained in the second stage estimation. Based on our conversation with the company, we assume that the firm’s internal discount rate is 10%, and estimate the shadow cost of adjusting inventory. We find that for offered prices and financing to be profit-maximizing, even against a very limited class of alternatives that consists of uniform shifts in the pricing or minimum down payment schedules, the unobserved component of costs must be fairly large, on the order of 2,400 dollars for a marginal sale. This is our baseline estimate, which is used for the counterfactual exercises in the next section. Supply side estimation

that relies on various specifications of learning or “satisficing” models – as reported in the top panel of Table 3(b) – produce remarkably similar estimates.

This high inferred cost reflects hesitancy to scale up operations, which may be driven by several explanations. One natural explanation is that the company cannot replicate its business instantaneously, and scaling up takes time. Indeed, the company opened a number of dealerships, including some in areas served by older dealerships, during the observation period. 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, our estimates are driven by the realized risks in the data, while the company may have optimized against a wider distribution of events.

The data are less favorable to two other hypotheses we have considered. One is that the high inferred cost is driven primarily by a hesitancy to expand sales to higher-risk borrowers. As Table 3(b) suggests, when we allow the shadow cost parameter to vary by the observed credit risk, we estimate the shadow cost to be significantly higher for good risks rather than for bad risks. The second possibility is that it is driven by high costs of short-run scale adjustments, as are required when demand spikes in tax season. As Table 3(b) shows, however, when we allow the shadow cost parameter to vary in tax season, the estimates do not change much, suggesting that costly short-run scale adjustments are not driving the overall estimates.

## 8 Implications for Contract Pricing

In this section we use the estimated model to illuminate the trade-offs involved in optimizing repayment revenue and customer selection, and to calculate the value generated by credit scoring, and also seasonal pricing.

### 8.1 Differential Effect of Contract Terms

A useful starting point for thinking about the trade-offs involved in setting car prices and minimum down payments is to trace graphically how changes in contract terms affect the probability of purchase, the down payments made by buyers and the expected repayment.

Figure 5 illustrates the effect of a uniform 500 dollar increase in the minimum down payment. To compute this, we take as a starting point the minimum down payments observed in the data. We use the model to find the average close rate, down payment and repayment probability for each

credit category. These relationships are traced out by the solid lines in the first three panels of Figure 5. We then use the model to compute counterfactual outcomes, facing each applicant in the data with a 500 dollar higher required down payment. Again we compute the average close rate, down payment and repayment probability for each credit category and trace these relationships in the same panels using dotted lines.

As Figure 5 shows, an increase in the required down payment substantially reduces close rates across credit categories, substantially increases down payments, and substantially raises the probability of repayments. To understand these effects, it is useful to break the set of applicants into three categories, and consider how the increase in the required down payment affects each group. The first set of applicants are “marginal buyers”: this group was making minimum or near-minimum down payments at the lower requirement, and decides not to purchase at the higher down payment. The second set of applicants are “inframarginal buyers” who nevertheless would ideally make a down payment below the higher minimum. These applicants continue to purchase when the required down payment is raised, but are forced to raise their down payments accordingly. The final set of applicants are those who are unaffected by the increase in the required down payment, either because they were unwilling to purchase even at the lower minimum, or because their ideal down payment is above the higher minimum, so neither requirement constrains their behavior.

Aggregating these three groups of applicants delivers the effects illustrated in Figure 5. The substantial drop in the close rate shows that the set of marginal applicants is substantial relative to the overall applicant population. The increase in average down payments and in the probability of repayment results from a combination of selection and incentive effects. The group of marginal buyers is no longer represented once the down payment requirements is raised and this group both makes lower immediate payments and defaults at a higher rate than other buyers. Also, the fact that some inframarginal buyers are forced to raise their down payments means they now take smaller loans and consequently their probability of payment increases.

The final panel – panel (d) – of Figure 5 shows the effect on profits of an increase in the required down payment. The panel suggests that increasing the required down payment would increase the expected profit from high risk applicants, by screening out marginal buyers who generate losses for the firm. For low risk applicants, however, the model suggests that an increase in the required down payment would decrease profits. This reflects a finding we discuss in the next subsection, that even marginal low risk applicants appear to represent positive profit opportunities for the firm.

An important point that is reflected only indirectly in this picture is the effect of applicants with different credit scores being faced with different required down payments. Applicants who are observably higher risk face substantially higher required down payments. This has two effects. First, the set of these applicants who select into purchasing is much more selected — the close rate for high risk applicants is 25 percent compared to 45 percent for low risk applicants. Second, conditional on purchase, high risk buyers are constrained in their ability to borrow. A much higher proportion of high risk buyers make the minimum down payment relative to low risk buyers (57 percent versus 23 percent). The model suggests that both effects should serve to strongly lower the default likelihood of the high risk buyers in the data.

While changes in the required down payment operate primarily by screening marginal buyers, changes in car prices have a very different effect, illustrated in Figure 6. We generate this figure in the same way as the previous figure, only with the counterfactual outcomes incorporating a \$1,000 increase in car prices. The results (in Table 3) imply that such an increase has a very small change on both the probability of purchase and down payment behavior, as illustrated in the first two panels of Figure 6. That is, the primary effect of an increase in car prices is to increase the loan size of buyers by (almost) a corresponding amount. This increase in loan size, however, has a substantial effect on repayment, as illustrated in the third panel of Figure 6. The probability of repayment falls substantially across credit categories. The final panel shows the effect on profits, showing the increase in payment size and the decrease in payments made just about cancel out, so that an increase in car prices has very little effect on the expected profits per applicant.

The discussion of Figures 5 and 6 illustrates the very different roles that required down payment and car prices play in this setting. Required down payment is primarily a screening device, which affects volume (the probability of sale) and has only a minimal direct effect on repayment (that is, not through selection of better risks). In contrast, car prices are primarily used to control repayment, and their level reflects a trade-off between the dollar amount of monthly payments and the probability of default.

## 8.2 Optimal Pricing

In this Section we translate the effects displayed in Figures 5 and 6 into numbers and quantify the trade-offs in contract design more precisely. To this end, it is again useful to think about a uniform increase in minimum down payments or in car prices across the sample. To understand

the trade-off, analogously to the first order condition presented in equation (3), the effect on profits from a discrete change in pricing can be approximated by

$$\Delta\Pi \approx \Delta Q \cdot \mathbb{E}[r \mid \textit{marginal buyer}] + Q \cdot \mathbb{E}[\Delta r \mid \textit{average buyer}]. \quad (16)$$

For example, an increase in the required minimum can be decomposed of two parts: the lost profits (which could be positive or negative) from marginal buyers who decide not to purchase after the change, and the average effect on revenues from inframarginal buyers.

In light of this decomposition, Table 4 uses our parameter estimates to report these quantities separately. The top rows in Table 4 present the strong selection effects we already mentioned: the average buyer is much more profitable to the firm than the marginal buyer, who is in turn much more profitable than the average non-buyer. While this pattern holds for all risk categories, it is interesting to note that for low risk applicants, the average non-buyer would still be more profitable to the firm than the average buyer in worse risk categories, reflecting the importance of credit scoring.

The rest of Table 4 calculates the effect of a 100 dollar increase in the minimum down and a 1,000 dollar increase in car prices. Using the equation above, consider first the effect of increasing the minimum down. The first effect is the loss of marginal consumers: overall close rates decrease by 2.2 percentage points. Since the marginal buyer is worth, in expectation, about \$600 to the firm, it makes the increase undesirable. However, the same change implies an increase in profits of almost \$30 from each inframarginal buyer, balancing this effect. The effect of a change in price is quite different. Here, as already suggested by Figure 6, the close rate is not affected by much, so the first part of the equation above is small. The second part is high, reflecting the fact that, given our results and the internal discount rate we impose, profitability can be higher by charging higher prices. We return to this below.

Figure 7 illustrates the implications of these results by graphing the expected per-applicant profits for the firm as a function of the level of the required down payment. The figure shows several things. First, note that it verifies that, given the other parameters, satisfying the first order conditions indeed leads to profit maximization. Second, it shows how profitability can be improved. Recall that the preferred specification constrains the firm to only choose the level of required down payment, and doesn't allow the firm to change the spread of required down payments across credit

categories. Figure 7 shows that widening the spread, by reducing the required minimum for good risks and increasing the required minimum for bad risks, could substantially increase profits. Of course, there may be other reason why such a spread may be difficult to achieve in practice.

One could repeat the same exercise for the choice of price (or margin). As already mentioned, given our current estimates and the level we impose for the internal discount rate, optimal car prices are higher than what we observe them to be. Given that company has experimented frequently with changes in the required down payment, and much less often with changes in car prices, it is not implausible that deviations from optimality are greater in the price dimension. One can also ask how optimal prices would vary across credit risks. For a given car, the company always sets a single list price for all consumers, but interestingly, if it was possible to offer different prices to different applicants, the firm would have had an incentive to reduce relative prices for high risk applicants in order to increase their likelihood of repayment.

### **8.3 Risk-Based Financing and the Value of Information**

In this section we calculate the value of making financing offers contingent on applicant information. Tables 5(a), 5(b), and 5(c) summarize our findings, by reporting the profitability of the company under different scenarios. Starting with Table 5(a), we report the current profitability of the company, as well as the counterfactual profitability, if the company could price more efficiently, or if it had no access to its credit scoring technology. Overall, our estimates imply that, on average, each applicant arriving at the lot is worth to the company \$402, given its current pricing policy. We estimate that optimal setting of required down payment – that is, the increased spread presented in Figure 7 – would increase these profits by 26 dollars or 6.5%. In contrast, without the ability to categorize applicants by credit risk, optimal (uniform) pricing implies a per-applicant profits of \$364, which is 9.4% less than current profits and 15% less than optimal risk-contingent profits.

Table 5(a) also reports profitability for different risk groups. It is illustrative of the mechanism through which risk-based financing is valuable. Compared to risk-based pricing, uniform pricing requires higher down payments from low risks and lower down payment from high risks. This significantly changes the selection of customers, reducing the close rate of low risks by more than a third and more than doubles the close rate of high risks. As the marginal good risk is much more profitable to the company than the average bad risk, this worse selection results in a significant loss of profits.

Finally, Table 5(a) also computes how much “potential” there is in risk-based financing, by allowing the firm the ability to price on individual characteristics it doesn’t observe. In particular, we assume the firm could price on the individual’s  $\varepsilon_i$  (but not on  $\eta_i$ ) and recompute profits. This should provide an upper bound for the value of information unless the firm can get access to predictors of  $\eta_i$  which are better than current  $x_i$  and  $\varepsilon_i$ . Our results suggest that “perfect” information would increase profits substantially. Per-applicant profits would increase by 90% compared to the case of optimal uniform pricing, by 62% compared to the case of optimal pricing based on current risk categorization, and by 72% compared to current observed pricing. These results may indicate that despite the recent expansion of credit markets, partly due to improvement in risk classification and credit scoring, consumers still have significant private information that remain unpriced.

In Table 5(b) we report a similar exercise, where we ask how important seasonal pricing is. In particular, as we mentioned earlier and confirm by our large estimates of the February effects, the tax rebate season introduces a larger pool of applicants who face less liquidity constraints at the time of purchase. Indeed, the company raises the required down payments during tax season. The results are summarized at the bottom panel of Table 5(b), by computing profitability on a per-tax-season-applicant basis. Optimal seasonal pricing improves profitability from tax-season applicants by \$32 or about 7% compared to pricing schedule which is constant over the year. Both optimal pricing schedules are more profitable than the observed pricing.

Tables 5(a) and 5(b) quantify the value of contingent financing offers. This value is created by a combination of the better ability of contingent offers to screen bad risks, as well as the ability to better tailor financing offers to specific customers and to offer terms that are more customized to individual applicants. More contingent financing terms also create value through their effect on the likelihood of potential entry. Table 5(c) summarizes our results from exercises that consider such potential entry. While these exercises clearly take us much more out of sample and rely on many stylized assumptions, they are illustrative as to the extent to which information may operate as an entry barrier.

In Table 5(c) we compute equilibrium per-applicant profits for a set of monopolistic and duopolistic scenarios. Each cell in the matrix presents the profits for the incumbent company first, and the profits for the potential entrant second. When an entrant is not present (top row), the profits of the incumbent simply replicate the figures already presented in Table 5(a). When an entrant is present, profits depend on whether the incumbent can use credit grades for pricing,

and whether this technology is also available to the entrant. In each scenario, we find the Nash Equilibrium of the duopoly game, where each of the firms simultaneously set uniform or grade-based required down payments, depending on the situation. 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 car prices remain the same through this analysis.<sup>11</sup>

The top row of Table 5(c) imply that the value of information for a monopolist is \$64 per applicant. Information also has an entry-detering benefit: better selection of applicants for an 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 \$45 per applicant. That is, the break-even sunk entry cost that would justify entry need to be 27% lower when an incumbent can offer contingent financing terms. Interestingly, Table 5(c) also shows that even if entry is accommodated, the incentive to offer grade-based pricing is higher in the presence of competition, with per-applicant profits increasing by \$81 or 48% compared to \$64 or 17% when the company is a monopolist.

## 9 Conclusion

Economists increasingly have access to detailed transaction data from insurance, credit and other contract markets. These data offer the promise of radically advancing our understanding of markets with asymmetric information, and providing a laboratory to test and apply the large theoretical literature on pricing and contract design. In this paper, we have tried to take a small step toward realizing this agenda by analyzing optimal pricing and contract design in the market for subprime auto sales and loans. We hope the approach taken here will encourage future empirical work on pricing and contract design in settings of asymmetric information.

Our principal economic findings concern the design of contracts in the face of informational imperfections in credit markets. In subprime lending, correlation between default risk and current liquidity make up-front payment requirements a powerful screening instrument. Our estimates show that even modestly relaxing these requirements can greatly expand and increase the riskiness of the

---

<sup>11</sup>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.

pool of borrowers. 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 this paper on contract design and firm behavior. To do this, we formulated a structural model of firm behavior, but a fairly “reduced-form” model of consumer behavior. A benefit of this approach is that our conclusions do not rely on assumptions about financial sophistication, the accuracy of consumer expectations, discounting and so forth that would likely be controversial and hard-to-assess. There are, however, two costs. One is that assumptions along these lines are necessary to draw quantitative welfare conclusions. The other is that our more reduced form model limits the range of “out-of-sample” counterfactual exercises we can perform. For example, absent such variation in the data, we cannot use our approach to evaluate a change in the firm’s collection policy or the introduction of a front-loaded or back-loaded repayment schedule. In future work we plan to bring to bear additional repayment data that will allow us to go deeper into these issues.

## References

- Adams, William, Liran Einav and Jonathan Levin**, “Liquidity Constraints and Imperfect Information in Subprime Lending,” NBER Working Paper No. 13067, 2007.
- Ausubel, Larry**, “The Failure of Competition in the Credit Card Market,” *American Economic Review*, 91, 1991, 50-81.
- Ausubel, Larry**, “Adverse Selection in the Credit Card Market,” University of Maryland Working Paper, 1999.
- Bajari, Patrick, C. Lanier Benkard and Jonathan Levin**, “Estimating Dynamic Models of Imperfect Competition,” *Econometrica*, 75(5), July 2007, 1331–1370.
- Cardon, James H., and Igal Hendel**, “Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey,” *Rand Journal of Economics*, 32(3), Autumn 2001, 408-427.
- Chiappori, Pierre-Andre and Bernard Salanie**, “Testing for Asymmetric Information in Automobile Insurance,” *Journal of Political Economy*, 108(1), February 2000, 56-78.
- Cohen, Alma and Liran Einav**, “Estimating Risk Aversion from Deductible Choice,” *American Economic Review*, 97(3), June 2007, 745-788.
- Edelberg, Wendy**, “Testing for Adverse Selection and Moral Hazard in Consumer Loan Markets,” Finance and Economics Discussion Series 2004-9, Washington: Board of Governors of the Federal Reserve System, 2004.

- Edelberg, Wendy**, “Risk-based Pricing of Interest Rates for Consumer Loans,” *Journal of Monetary Economics*, 53, November 2006, 2283-2296.
- Einav, Liran, Amy Finkelstein and Paul Schrimpf**, “The Welfare Cost of Asymmetric Information: Evidence from the U.K. Annuity Market,” NBER Working Paper No. 13228, 2007.
- Finkelstein, Amy, and Kathleen McGarry**, “Multiple dimensions of private information: evidence from the long-term care insurance market,” *American Economic Review*, 96(4), September 2006, 938-958.
- Karlan, Dean and Jonathan Zinman**, “Credit Elasticities in Less Developed Countries: Implications for Microfinance,” forthcoming, *American Economic Review*, 2007a.
- Karlan, Dean and Jonathan Zinman**, “Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment,” Yale University Working Paper, 2007b.
- Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii**, “Moment Inequalities and Their Applications,” Harvard University Working Paper, 2006.
- Perrigne, Isabelle M., and Quang H. Vuong**, “Econometrics of Incentive Regulation: Non-parametric 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.

## Appendix A: An economic model of consumer behavior

In this appendix we present an economic model of consumer behavior that maps to the econometric model we estimate. We consider a  $T+1$  period setting. In period 0, a consumer with initial income  $y_0$  has the option to purchase a car with financing. Let  $p$  denote the car price, and  $d$  the required down payment. The consumer can borrow  $L$ , up to  $p-d$ , to purchase the car, with a required payments of  $m(L) = (rL) / (1 - (1+r)^{-T})$  in each of periods  $t = 1, \dots, T$ . The consumer cannot do any additional borrowing, nor can he save. In each period  $t = 1, \dots, T$ , the consumer receives a stochastic income  $y_t$ ; if she has not yet defaulted on a payment, she decides whether to make her payment and keep the car for another period.

Let  $F_0$  denote the distribution of initial income and suppose that over time income follows a Markov process, so that  $y_t$  has a distribution  $F_t(y_t|y_{t-1})$ . Assume that the income process is persistent so that  $F_t$  is increasing in  $y_{t-1}$  in the first-order stochastic dominance sense.

Consumer utility comes from consuming a car and/or money. Let  $v_t$  be the consumption value of the car and  $u_t(c)$  the utility derived from consuming  $c$  dollars in period  $t$ . We assume that each  $u_t$  is increasing and concave, and satisfies the boundary condition that  $u'_t(0) = \infty$ , and that there is no discounting.<sup>12</sup>

We solve the consumer's problem by backward induction. At time  $T$ , if the consumer does not have the car, she consumes her income  $y_T$ , realizing utility  $u_T(y_T)$ . If she has the car, she makes her final payment if and only if  $v_T + u_T(y_T - m(L)) \geq u_T(y_T)$ , which obtains if  $y_T$  is above some threshold  $y_T^*(L)$ . Note that the income threshold will be higher for a larger loan, i.e.  $y_T^*(L)$  is increasing in  $L$ .

Let  $U_T(y_{T-1}, L)$  denote the consumer's expected utility just prior to the realization of  $y_T$ , so:

$$U_T(y_{T-1}, L) = \mathbb{E}[\max\{v_T + u_T(y_T - m(L)), u_T(y_T)\} | y_{T-1}]. \quad (17)$$

Let  $\bar{U}_T(y_{T-1}) = E[u_T(y_T) | y_{T-1}]$  denote her expected utility if she does not have the car. Note that the value of having the car,  $U_T(y_{T-1}, L) - \bar{U}_T(y_{T-1})$  is increasing in  $y_{T-1}$  and decreasing in  $L$ .

Proceeding recursively, suppose we have identified  $U_{t+1}(y_t, L)$  and  $\bar{U}_{t+1}(y_t)$  and that the value of having the car entering time  $t+1$ ,  $U_{t+1}(y_t, L) - \bar{U}_{t+1}(y_t)$  is increasing in  $y_t$  and decreasing in  $L$ . Consider the consumer at time  $t \in \{1, \dots, T\}$ . If she does not own the car, she consumes her income  $y_t$ , realizing utility  $u_t(y_t)$ , so her expected present value just prior to the realization of  $y_t$  is  $\bar{U}_t(y_{t-1}) = E[u_t(y_t) + \bar{U}_{t+1}(y_t) | y_{t-1}]$ . If she owns the car, she optimally repays if

$$v_t + u_t(y_t - m(L)) + U_{t+1}(y_t, L) \geq u_t(y_t) + \bar{U}_{t+1}(y_t). \quad (18)$$

Under our induction hypothesis, she optimally repays if  $y_t$  is above some threshold  $y_t^*(L)$ , with the threshold increasing in  $L$ . Let  $U_t(y_{t-1}, L)$  denote the consumer's present value just prior to the realization of  $y_t$ , assuming she still has the car. It follows that  $U_t(y_t, L) - \bar{U}_t(y_t)$  is increasing in  $y_t$  and decreasing in  $L$ .

This argument yields a complete characterization of repayment behavior in terms of the re-payment thresholds  $y_1^*(L), \dots, y_T^*(L)$ . A consumer who owns the car at date  $t$  makes payments if and only if  $y_t \geq y_t^*(L)$ . The length of repayment  $S$  is a random variable equal to  $\max\{t : y_k \geq y_k^*(L) \text{ for all } k \leq t\}$ . The theoretical model does not pin down a functional form for its distribution, which depends on the income process, the consumption value of the car over time, and consumer preferences. In our empirical model, we assume that the distribution of  $s = S/T$  is truncated log-normal. As we showed earlier, this assumption appears to fit the empirical distribution of repayment lengths well. The model does imply that an increase in loan size, by increasing every payment threshold, shortens the time to default. This feature is borne out in our estimates.

---

<sup>12</sup>Throughout this example we assume that the only uncertainty pertains to an individual's future income. It would be fairly straightforward to allow for uncertainty about future preferences or the consumption value of the car.

Now consider the consumer’s down payment decision at  $t = 0$  assuming she decides to purchase the car. If her current income is  $y_0$ , her optimal down payment solves:

$$\max_D v_0 + u_0 (y_0 - D) + U_1 (y_0, p - D). \quad (19)$$

Let  $D^*(y_0)$  denote the optimal down payment. If the consumer chooses to purchase, and faces a minimum down payment  $d$ , her constrained optimal down payment will be  $D(y_0) = \min\{d, D^*(y_0)\}$ . Note that in choosing the down payment, an increase in  $y_0$  has two effects: it reduces the need for cash but also increases expected future income. The former effect makes a higher down payment desirable; the latter effect has an ambiguous effect on the desired down payment. We make the assumption that  $D^*(y_0)$  (and hence  $D(y_0)$ ) is increasing in  $y_0$ .

Finally, we turn to the consumer’s decision of whether or not to purchase the car. Purchasing is optimal if and only if:

$$\max_{D \geq d} v_0 + u_0 (y_0 - D) + U_1 (y_0, p - D) \geq u_0 (y_0) + \bar{U}_1 (y_0). \quad (20)$$

Not surprisingly, a lower car price  $p$  or minimum down payment  $d$  makes purchasing more attractive. Holding the offer terms fixed, purchasing is optimal if and only if initial income  $y_0$  is above some threshold level  $y_0^*$ .

Recall that our empirical model collapsed the borrowing and purchasing decisions. To do this, it is useful to re-state the optimal purchasing rule to say that the consumer purchases if and only if  $D^*(y_0)$  is above some threshold  $d^*$ , where  $d^* = D^*(y_0^*)$ . This involves no loss of generality so long as  $D^*(y_0)$  is increasing in  $y_0$ . We can then define  $Z(d) = d - d^*$  to match the notation of the theoretical model to that of the empirical model.

Just as our theoretical model does not imply a functional form for the distribution of repayment times, it also does not imply a functional form for distribution of the ideal down payment  $D^*$ , which depends on the distribution of  $y_0$  and other parameters of the model. In our empirical estimation, we assume that  $D^*$  has a normal distribution, so that the actual distribution of down payments has a truncated normal distribution. As we showed earlier, this assumption seems to align quite well with the empirical distribution of down payments we observe in the data.

## Appendix B: Specification and estimation of the recovery value

This appendix describes the recovery model used to estimate expected recovery amounts conditional on default. The recovery model is estimated separately from the purchase and repayment model, and the recovery parameter estimates are taken as given when computing expected net revenues in supply side estimation and counterfactuals. The recovery model consists of two separate equations. The first equation is a probit regression with a positive recovery indicator as the dependent variable, estimated using all loans that end in default.

The second equation is a linear regression with recovery amount as the dependent variable, estimated using all observations with positive recoveries. Both equations use the same set of explanatory variables, which include car characteristics, applicant characteristics, time and city fixed effects, and the number of months that loan payments were made before default. The last variable is of particular importance, since it provides a link between recovery amount and the endogenous loan repayment variable  $s_i$ . Credit category fixed effects are not included in either recovery equation since they are found to have very little explanatory power. Estimating the recovery equations separately from the main model essentially assumes that the residuals in the recovery model are independent of the other unobservables in the model.

Table A1 shows the parameter estimates from both recovery equations. The results are fairly intuitive. An increase in the number of months before default decreases the probability of a non-zero recovery and also decreases the expected recovery amount, conditional on a non-zero recovery, by about \$45 per month.

Individual characteristics, such as higher monthly income, possession of a bank account, and home ownership, do not have a significant effect on the probability of recovery, but do significantly increase the expected recovery amount, conditional on recovery occurring.

## Appendix C: Derivations and estimation details

### C.1 Derivation of Likelihood Function for Demand

In this section we derive the likelihood function used to estimate the parameters of the price negotiation equation, the purchase and financing equation, and the loan repayment equation. We begin with the three equations:

$$p_i = l_i + x_i' \lambda + \nu_i \quad (21)$$

$$D_i^* = x_i' \beta_x + p_i \beta_p + \varepsilon_i \quad (22)$$

$$\ln(s_i^*) = x_i' \gamma_x + (p_i - D_i) \gamma_L + \eta_i. \quad (23)$$

The variables  $p_i$  (negotiated price),  $D_i^*$  (desired down payment),  $\ln(s_i^*)$  (log of fraction of payments made), and  $D_i$  (observed down payment - a nonlinear function of  $D_i^*$ ) are the system's endogenous variables, and  $l_i$  (list price) and  $x_i$  (a vector of offer, car, applicant, location, and time characteristics common to all equations) are considered exogenous. The variables  $D_i^*$  and  $\ln(s_i^*)$  are latent variables that are not observed for all applicants; specifically,  $D_i^*$  is not observed if applicants either make the required minimum down payment or do not purchase at all, and  $\ln(s_i^*)$  is not observed if loan repayment is censored either due to full payment or the end of our sample. We discuss the relationship between these latent variables and their observable counterparts,  $D_i$  and  $\ln(s_i)$ , in more detail below.

The unobservables  $\nu_i$ ,  $\varepsilon_i$ , and  $\eta_i$  are distributed jointly normal as:

$$\begin{pmatrix} \nu_i \\ \varepsilon_i \\ \eta_i \end{pmatrix} \sim f(\nu_i, \varepsilon_i, \eta_i) = N(0, V) \text{ with } V = \begin{pmatrix} \sigma_\nu^2 & \rho_{\nu\varepsilon} \sigma_\nu \sigma_\varepsilon & \rho_{\nu\eta} \sigma_\nu \sigma_\eta \\ \rho_{\nu\varepsilon} \sigma_\nu \sigma_\varepsilon & \sigma_\varepsilon^2 & \rho_{\varepsilon\eta} \sigma_\varepsilon \sigma_\eta \\ \rho_{\nu\eta} \sigma_\nu \sigma_\eta & \rho_{\varepsilon\eta} \sigma_\varepsilon \sigma_\eta & \sigma_\eta^2 \end{pmatrix}. \quad (24)$$

This joint density provides the foundation for deriving the likelihood function of the data  $\mathcal{L}(p_i, D_i, \ln(s_i) | d_i, l_i, x_i)$ . We begin by rewriting the joint density

$$f(\nu_i, \varepsilon_i, \eta_i) = f(\eta_i | \nu_i, \varepsilon_i) f(\varepsilon_i | \nu_i) f(\nu_i), \quad (25)$$

where  $f(\nu_i) = N(0, \sigma_\nu^2)$ ,  $f(\varepsilon_i | \nu_i) = N(\mu_{\varepsilon|\nu}, \sigma_{\varepsilon|\nu}^2)$  with  $\mu_{\varepsilon|\nu} = \frac{\rho_{\varepsilon\nu} \sigma_\varepsilon}{\sigma_\nu} \nu_i$  and  $\sigma_{\varepsilon|\nu}^2 = \sigma_\varepsilon^2 (1 - \rho_{\varepsilon\nu}^2)$ , and  $f(\eta_i | \varepsilon_i, \nu_i) = N(\mu_{\eta|\nu, \varepsilon}, \sigma_{\eta|\nu, \varepsilon}^2)$  with

$$\mu_{\eta|\nu, \varepsilon} = \begin{pmatrix} \rho_{\nu\eta} \sigma_\nu \sigma_\eta & \rho_{\varepsilon\eta} \sigma_\varepsilon \sigma_\eta \end{pmatrix} \begin{pmatrix} \sigma_\nu^2 & \rho_{\nu\varepsilon} \sigma_\nu \sigma_\varepsilon \\ \rho_{\nu\varepsilon} \sigma_\nu \sigma_\varepsilon & \sigma_\varepsilon^2 \end{pmatrix}^{-1} \begin{pmatrix} \nu_i \\ \varepsilon_i \end{pmatrix}, \quad (26)$$

and

$$\sigma_{\eta|\nu, \varepsilon}^2 = \sigma_\eta^2 - \begin{pmatrix} \rho_{\nu\eta} \sigma_\nu \sigma_\eta & \rho_{\varepsilon\eta} \sigma_\varepsilon \sigma_\eta \end{pmatrix} \begin{pmatrix} \sigma_\nu^2 & \rho_{\nu\varepsilon} \sigma_\nu \sigma_\varepsilon \\ \rho_{\nu\varepsilon} \sigma_\nu \sigma_\varepsilon & \sigma_\varepsilon^2 \end{pmatrix}^{-1} \begin{pmatrix} \rho_{\nu\eta} \sigma_\nu \sigma_\eta \\ \rho_{\varepsilon\eta} \sigma_\varepsilon \sigma_\eta \end{pmatrix}. \quad (27)$$

Since the Jacobian of the transformation of  $(\nu_i, \varepsilon_i, \eta_i)'$  to  $(p_i, D_i^*, \ln(s_i^*))'$  is 1, we can write the joint density of  $(p_i, D_i^*, \ln(s_i^*))'$  as:

$$\begin{aligned}
f(p_i, D_i^*, \ln(s_i^*) | l_i, x_i) &= f(\ln(s_i^*) - x_i' \gamma_x - (p_i - D_i) \gamma_L | p_i - l_i - x_i' \lambda, D_i^* - x_i' \beta_x - p_i \beta_p) \\
&\times f(D_i^* - x_i' \beta_x - p_i \beta_p | p_i - l_i - x_i' \lambda) \\
&\times f(p_i - l_i - x_i' \lambda).
\end{aligned} \tag{28}$$

If  $D_i^*$  and  $\ln(s_i^*)$  were observed for all applicants, this expression would provide the likelihood function for the data; however, since  $D_i^*$  and  $\ln(s_i^*)$  are not always observed, we must rewrite this expression in terms of the observable  $D_i$  and  $\ln(s_i)$ . We proceed in four steps. First, we derive the likelihood of observing a given negotiated price. Since we assume price is observed for all applicants (we impute prices for non-buyers), this step is straightforward. The probability of observing a negotiated price  $p_i$  is given simply by

$$p_{p_i} = \phi \left[ \frac{p_i - l_i - x_i' \lambda}{\sigma_\nu} \right] \tag{29}$$

where  $\phi$  denotes the standard normal pdf. We do not account for censoring at list price, since this would require integration over  $\nu_i$  for some observations, thus complicating the derivation of the likelihood function with limited added benefit.

Second, conditional on a negotiated price, we derive the likelihood of observing three possible purchase and financing outcomes: sale with a down payment above the minimum, sale with a minimum down payment, and no sale. We define the observed down payment  $D_i$  as:

$$D_i = \begin{cases} D_i^* = x_i' \beta_x + p_i \beta_p + \varepsilon_i & \text{if } D_i^* \geq d_i \\ d_i & \text{if } D_i^* \in [d_i - Z_i, d_i) \\ \emptyset & \text{if } D_i^* < d_i - Z_i \end{cases} \tag{30}$$

where  $Z_i = d_i + x_i' \delta_x + p_i \delta_p + d_i \alpha_d$ , and  $\varepsilon_i$  is correlated with  $\nu_i$ . The first case applies applicants who purchase a car and make a down payment above the minimum, the second to applicants who purchase a car and make a minimum down payment, and the third to applicants who do not purchase. In the first case,  $D_i^*$  is observed, and the likelihood is defined by the pdf of  $\varepsilon_i$ . In the latter two cases,  $D_i^*$  is unobserved, and the likelihood of each outcome is defined by the corresponding cdf. To incorporate the correlation between  $\varepsilon_i$  and a known  $\nu_i$ , we calculate the probability of each outcome conditional on  $\nu_i$ .

The probability of observing a sale with a given down payment above the minimum is:

$$p_{D_i = D_i^* | \nu_i} = \Pr(D_i^* = x_i' \beta_x + p_i \beta_p + \varepsilon_i | \nu_i) = \phi \left[ \frac{D_i^* - x_i' \beta_x - p_i \beta_p - \mu_{\varepsilon | \nu}}{\sigma_{\varepsilon | \nu}} \right] \tag{31}$$

We can write this expression in terms of the standard normal pdf since  $\varepsilon_i | \nu_i \sim N(\mu_{\varepsilon | \nu}, \sigma_{\varepsilon | \nu}^2)$  implies that conditional on  $\nu_i$ ,  $(\varepsilon_i - \mu_{\varepsilon | \nu}) / \sigma_{\varepsilon | \nu} \sim N(0, 1)$ .

The probability of observing a sale with a minimum down payments is:

$$\begin{aligned}
p_{D_i = d_i | \nu_i} &= \Pr(D_i^* \in [d_i - Z_i, d_i) | \nu_i) \\
&= \Pr(d_i - Z_i - x_i' \beta_x - p_i \beta_p < \varepsilon_i < d_i - x_i' \beta_x - p_i \beta_p | \nu_i) \\
&= \Phi \left[ \frac{d_i - x_i' \beta_x - p_i \beta_p - \mu_{\varepsilon | \nu}}{\sigma_{\varepsilon | \nu}} \right] - \Phi \left[ \frac{d_i - Z_i - x_i' \beta_x - p_i \beta_p - \mu_{\varepsilon | \nu}}{\sigma_{\varepsilon | \nu}} \right]
\end{aligned} \tag{32}$$

where  $\Phi$  denotes the standard normal cdf. For certain parameter values,  $Z_i$ , may be negative, meaning  $p_{D_i = d_i} = 0$ , though in practice, this occurs for less than 0.1 percent of observations.

The probability of observing no sale is:

$$\begin{aligned} p_{D_i=0|\nu_i} &= \Pr(D_i^* < d_i - Z_i | \nu_i) = \Pr(\varepsilon_i < d_i - Z_i - x_i'\beta_x - p_i\beta_p | \nu_i) \\ &= \Phi \left[ \frac{d_i - Z_i - x_i'\beta - p_i\beta_p - \mu_{\varepsilon|\nu}}{\sigma_{\varepsilon|\nu}} \right] \end{aligned} \quad (33)$$

The third step is to derive the likelihood of observing three possible loan repayment outcomes: no payments, default after at least one payment, and payments censored due to full payment or the end of our sample, conditional on a negotiated price and a purchase and financing decision. We begin with the model in equation (8), but adapt it to account for loans that are censored by the end of our sample period. We define the censoring point,  $c_i \in (0, 1]$  as the fraction of the loan observed in our data. The observed fraction of payments made,  $s_i$ , is then

$$s_i = \begin{cases} s_i^* = \exp(x_i'\gamma_x + (p_i - D_i)\gamma_L + \eta_i) & \text{if } s_i^* < c_i \\ c_i & \text{if } s_i^* \geq c_i \end{cases}, \quad (34)$$

where  $\eta_i$  is correlated with  $\nu_i$  and  $\varepsilon_i$ . The first case applies to buyers with observed default (including default after zero payments), and the second applies to buyers with censored repayment. For loans that have been repaid in full,  $c_i = 1$ . To account for the correlation between  $\eta_i$  and  $(\nu_i, \varepsilon_i)$ , we calculate the probability of each repayment outcome conditional on  $(\nu_i, \varepsilon_i)$ . As shown below, when  $\varepsilon_i$  is unobserved (that is, when a minimum down payment is made), this requires integration over  $\varepsilon_i$ .

With  $\varepsilon_i$  known, the likelihood of observing censored payments is:

$$\begin{aligned} p_{s_i=c_i|\nu_i,\varepsilon_i} &= \Pr(s_i = c_i|\nu_i,\varepsilon_i) = \Pr(s_i^* \geq c_i|\nu_i,\varepsilon_i) = \Pr(\exp(x_i'\gamma_x + (p_i - D_i)\gamma_L + \eta_i) \geq c_i|\nu_i,\varepsilon_i) \\ &= \Pr(\eta_i < -\ln(c_i) + x_i'\gamma_x + (p_i - D_i)\gamma_L|\nu_i,\varepsilon_i) = \Phi \left[ \frac{-\ln(c_i) + x_i'\gamma_x + (p_i - D_i)\gamma_L + \mu_{\eta|\nu,\varepsilon}}{\sigma_{\eta|\nu,\varepsilon}} \right] \end{aligned} \quad (35)$$

The likelihood of observing no payments is:

$$\begin{aligned} p_{s_i=0|\nu_i,\varepsilon_i} &= \Pr(s_i = 0|\nu_i,\varepsilon_i) = \Pr(s_i^* < g|\nu_i,\varepsilon_i) = \Pr(\exp(x_i'\gamma_x + (p_i - D_i)\gamma_L + \eta_i) < g|\nu_i,\varepsilon_i) \\ &= \Pr(\eta_i < \ln(g) - x_i'\gamma_x - (p_i - D_i)\gamma_L|\nu_i,\varepsilon_i) = \Phi \left[ \frac{\ln(g) - x_i'\gamma_x - (p_i - D_i)\gamma_L - \mu_{\eta|\nu,\varepsilon}}{\sigma_{\eta|\nu,\varepsilon}} \right], \end{aligned} \quad (36)$$

where  $g$  represents the fraction of the loan paid in each installment. The probability of zero payments thus equals the probability of default before the first payment is made.

The likelihood of observing payments through  $s_i^*$  prior to the censoring point is:

$$\begin{aligned} p_{s_i=s_i^*|\nu_i,\varepsilon_i} &= \Pr(s_i = s_i^*|\nu_i,\varepsilon_i) = \\ &= \Phi \left[ \frac{\ln(s_i^* + g) - x_i'\gamma_x - (p_i - D_i)\gamma_L - \mu_{\eta|\nu,\varepsilon}}{\sigma_{\eta|\nu,\varepsilon}} \right] - \Phi \left[ \frac{\ln(s_i^*) - x_i'\gamma_x - (p_i - D_i)\gamma_L - \mu_{\eta|\nu,\varepsilon}}{\sigma_{\eta|\nu,\varepsilon}} \right] \end{aligned} \quad (37)$$

The conditional moments in these expressions,  $\mu_{\eta|\nu,\varepsilon}$  and  $\sigma_{\eta|\nu,\varepsilon}$ , are functions of  $\nu_i$  and  $\varepsilon_i$ . We can calculate  $\nu_i$  as  $p_i - l_i - x_i'\lambda$  for all buyers. When a down payment above the minimum is observed, we can calculate  $\varepsilon_i$  as  $D_i^* - x_i'\beta_x - p_i\beta_p$ , which yields a likelihood for each repayment outcome written in terms of the observed  $D_i^*$ . However, when a minimum down payment is observed, or when  $\varepsilon_i$  is unknown, we must integrate over all  $\varepsilon_i$  in the minimum down payment region. After replacing  $\varepsilon_i = D_i^* - x_i'\beta_x - p_i\beta_p$ , this yields

the following likelihood for each repayment outcome:

$$p_{s_i=0|p_i, D_i=d_i} = \int_{d_i-Z_i}^{d_i} \Phi \left[ \frac{\ln(g) - x'_i \gamma_x - (p_i - D_i) \gamma_L - \mu_{\eta|\nu, \varepsilon}(D_i^*)}{\sigma_{\eta|\nu, \varepsilon}} \right] \phi \left[ \frac{D_i^* - x'_i \beta_x - p_i \beta_p}{\sigma_{\varepsilon|\nu}} \right] dD_i^* \quad (38)$$

$$p_{s_i=c_i|p_i, D_i=d_i} = \int_{d_i-Z_i}^{d_i} \Phi \left[ \frac{-\ln(c_i) + x'_i \gamma_x + (p_i - D_i) \gamma_L + \mu_{\eta|\nu, \varepsilon}(D_i^*)}{\sigma_{\eta|\nu, \varepsilon}} \right] \phi \left[ \frac{D_i^* - x'_i \beta_x - p_i \beta_p}{\sigma_{\varepsilon|\nu}} \right] dD_i^* \quad (39)$$

$$p_{s_i=s_i^*|p_i, D_i=d_i} = \int_{d_i-Z_i}^{d_i} \left( \Phi \left[ \frac{\ln(s_i^* + g) - x'_i \gamma_x - (p_i - D_i) \gamma_L - \mu_{\eta|\nu, \varepsilon}}{\sigma_{\eta|\nu, \varepsilon}} \right] - \Phi \left[ \frac{\ln(s_i^*) - x'_i \gamma_x - (p_i - D_i) \gamma_L - \mu_{\eta|\nu, \varepsilon}}{\sigma_{\eta|\nu, \varepsilon}} \right] \right) \phi \left[ \frac{D_i^* - x'_i \beta_x - p_i \beta_p}{\sigma_{\varepsilon|\nu}} \right] dD_i^* \quad (40)$$

The final step is to combine the set of negotiate price probabilities  $p_{p_i}$ , purchase and down payment probabilities ( $p_{D_i=\emptyset|p_i}$ ,  $p_{D_i=d_i|p_i}$ , and  $p_{D_i=D_i^*|p_i}$ ) and loan repayment probabilities ( $p_{s_i=0|p_i, D_i}$ ,  $p_{s_i=c_i|p_i, D_i}$ , and  $p_{s_i=s_i^*|p_i, D_i}$ ) into a full likelihood function for data,  $\mathcal{L}(p_i, D_i, \ln(s_i)|d_i, l_i, x_i)$ . Before writing the full likelihood function, it is useful to define the set of possible outcomes observed in the data:

- $I_0$ : no sale
- $I_1$ : sale, down payment above minimum, no payments
- $I_2$ : sale, down payment above minimum, censored payments
- $I_3$ : sale, down payment above minimum, observed default after at least one payment
- $I_4$ : sale, minimum down payment, no payments
- $I_5$ : sale, minimum down payment, censored payments
- $I_6$ : sale, minimum down payment, observed default after at least one payment

Using the notation  $i \in I$  to indicate that applicant  $i$  chose outcome  $I$ , we can write the full log-likelihood function for the data as:

$$\begin{aligned} \log \mathcal{L} &= \sum_i \log(p_{p_i}) + \sum_{i \in I_0} \log(p_{D_i=\emptyset|p_i}) + \\ &+ \sum_{i \in I_1} \{\log(p_{D_i=D_i^*|p_i}) + \log(p_{s_i=0|p_i, D_i=D_i^*})\} + \sum_{i \in I_4} \{\log(p_{D_i=d_i|p_i}) + \log(p_{s_i=0|p_i, D_i=d_i})\} + \\ &+ \sum_{i \in I_2} \{\log(p_{D_i=D_i^*|p_i}) + \log(p_{s_i=c_i|p_i, D_i=D_i^*})\} + \sum_{i \in I_5} \{\log(p_{D_i=d_i|p_i}) + \log(p_{s_i=c_i|p_i, D_i=d_i})\} + \\ &+ \sum_{i \in I_3} \{\log(p_{D_i=D_i^*|p_i}) + \log(p_{s_i=s_i^*|p_i, D_i=D_i^*})\} + \sum_{i \in I_6} \{\log(p_{D_i=d_i|p_i}) + \log(p_{s_i=s_i^*|p_i, D_i=d_i})\} \end{aligned} \quad (41)$$

Our estimates of the parameters  $\lambda, \beta_x, \beta_p, \delta_x, \delta_p, \alpha_d, \gamma_x, \gamma_L, \rho_{\nu\varepsilon}, \rho_{\nu\eta}, \rho_{\varepsilon\eta}, \sigma_{\nu}, \sigma_{\varepsilon}$ , and  $\sigma_{\eta}$  maximize this log-likelihood function.

## C.2 Supply Side Estimation

In this section we derive the moment conditions used to estimate the supply-side parameter  $\psi$ . We first derive detailed expressions for net revenue,  $r_i$ , expected net revenue conditional on sale,  $R_i$ , and expected profits,  $\Pi_i$ , and describe how these quantities are computed in practice. We then derive a set of moment

conditions from the firm's optimal pricing problem and describe the method used to estimate supply-side parameters.

When the number of payments made,  $S_i$ , is known, net revenue is given by equation (12):

$$r_i = D_i + (p_i - D_i) \frac{\frac{1}{\kappa} (1 - e^{-\kappa S_i})}{\frac{1}{z_i} (1 - e^{-z_i T_i})} + e^{-\kappa S_i} k_i - C_i. \quad (42)$$

When  $S_i$  is unknown but  $\varepsilon_i$  is known, we can integrate over  $S_i$  to calculate expected net revenues conditional on  $\varepsilon_i$ . Replacing  $S_i = s_i T_i$  in the above expression, and integrating over  $s_i$  instead of  $S_i$ , expected net revenue conditional on  $\varepsilon_i$  is:

$$\begin{aligned} \mathbb{E}[r_i | \varepsilon_i] &= p_{s_i=0|\varepsilon_i} \mathbb{E}[r_i | s_i = 0] + p_{s_i=1|\varepsilon_i} \mathbb{E}[r_i | s_i = 1] + \int_0^1 p_{s_i=s_i^*|\varepsilon_i} \mathbb{E}[r_i | s_i = s_i^*] ds_i^* \\ &= p_{s_i=0|\varepsilon_i} (D_i + k_i - C_i) + p_{s_i=1|\varepsilon_i} (D_i + (p_i - D_i) \frac{\frac{1}{\kappa} (1 - e^{-\kappa T_i})}{\frac{1}{z_i} (1 - e^{-z_i T_i})} - C_i) + \\ &\quad + \int_0^1 p_{s_i=s_i^*|\varepsilon_i} (D_i + (p_i - D_i) \frac{\frac{1}{\kappa} (1 - e^{-\kappa s_i^* T_i})}{\frac{1}{z_i} (1 - e^{-z_i T_i})} + e^{-\kappa s_i^* T_i} k_i - C_i) ds_i^* \end{aligned} \quad (43)$$

where  $p_{s_i=0|\varepsilon_i}$ ,  $p_{s_i=s_i^*|\varepsilon_i}$ , and  $p_{s_i=1|\varepsilon_i}$  are defined above. The first term on the right hand side of the equation is equal to the probability of zero payments times the net revenue from zero payments, the second term is equal to the probability of full payment times the net revenue from full payment, and the third term is equal to the expected net revenue from between 1 and  $T_i - 1$  payments.

This expression for expected revenue can be used when  $\varepsilon_i$  is known, or alternatively when  $D_i^*$  has been observed by the firm. To compute the expected net revenue conditional on sale prior to observing  $D_i^*$ , we integrate over the region of sale, namely the region where  $\varepsilon_i > d_i - Z_i - x'_i \beta_x - p_i \beta_p$ :

$$R_i = \mathbb{E}[r_i | g_i \geq 0] = \int_{d_i - Z_i - x'_i \beta_x - p_i \beta_p}^{\infty} \mathbb{E}[r_i | \varepsilon_i] \phi \left[ \frac{\varepsilon_i}{\sigma_\varepsilon} \right] d\varepsilon_i = \int_{d_i - Z_i}^{\infty} \mathbb{E}[r_i | D_i^*] \phi \left[ \frac{D_i^* - x'_i \beta_x - p_i \beta_p}{\sigma_\varepsilon} \right] dD_i^* \quad (44)$$

where the second equality follows by replacing  $\varepsilon_i = D_i^* - x'_i \beta_x - p_i \beta_p$ .

Expected profits are then equal to the probability of sale times the expected revenue conditional on sale, or:

$$\Pi_i = Q_i R_i = (1 - p_{D_i=\emptyset}) \mathbb{E}[r_i | g_i \geq 0] = \left( 1 - \Phi \left[ \frac{d_i - Z_i - x'_i \beta - p_i \beta_p}{\sigma_\varepsilon} \right] \right) \mathbb{E}[r_i | g_i \geq 0]. \quad (45)$$

For our current estimation, we assume that

$$\int \pi(l(x), d(x); x, \omega) dF(x, \omega) \geq \int \pi(l(x) + a, d(x) + b; x, \omega) dF(x, \omega) \quad \text{for all } a, b \in \mathbb{R}. \quad (46)$$

We approximate the integral by using its empirical analog, summing over the observed applicants. To operationalize the optimality assumption, and because the firm prefers to offer prices on discrete intervals (e.g. minimum down payments and margins are typically in units of \$100), we do not compute the first order conditions, but rather we compute the expected profits that could be gained from increasing or decreasing contract terms by \$100. If the firm is pricing optimally, these potential gains should be non-positive. Our current procedure computes the gains from these possible deviations and searches for the supply side parameter  $\psi$  that make the observed pricing optimal. Specifically, we search for  $\psi$  that minimizes the two

equations below:

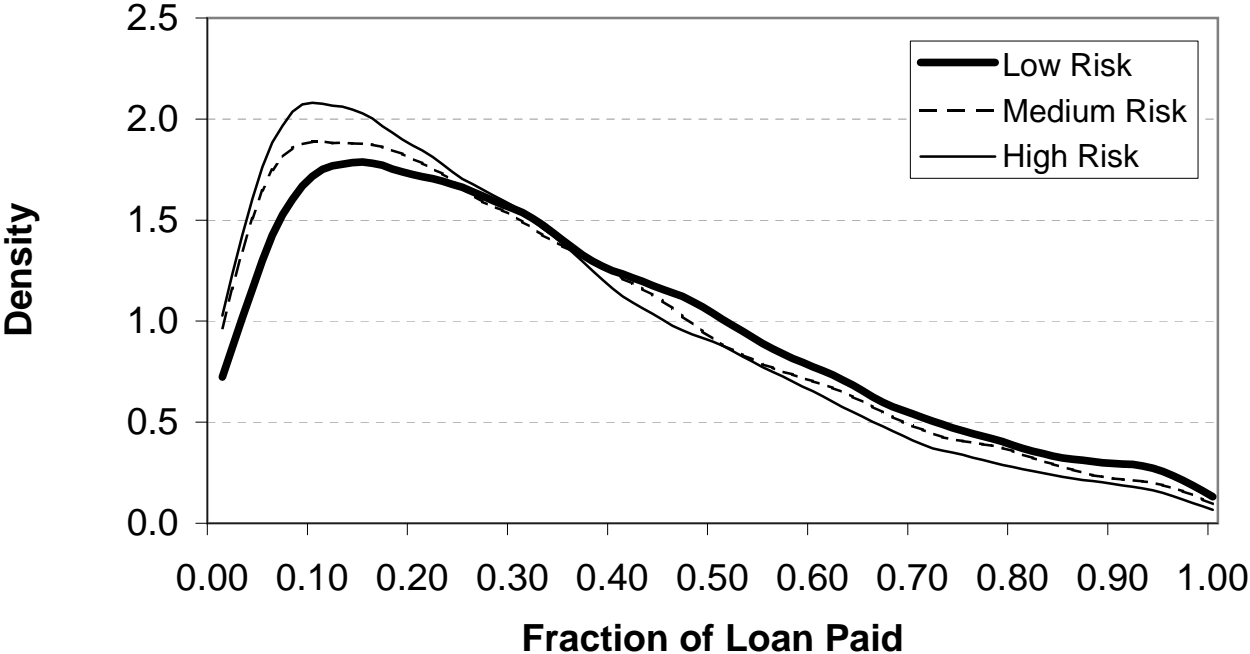
$$\Omega_d[d_i, p_i|\psi] = \left[ \min \left\{ 0, \sum_i (\Pi_i(d_i, p_i) - \Pi_i(d_i + 100, p_i)) \right\} \right]^2 + \left[ \min \left\{ 0, \sum_i (\Pi_i(d_i, p_i) - \Pi_i(d_i - 100, p_i)) \right\} \right]^2 \quad (47)$$

and

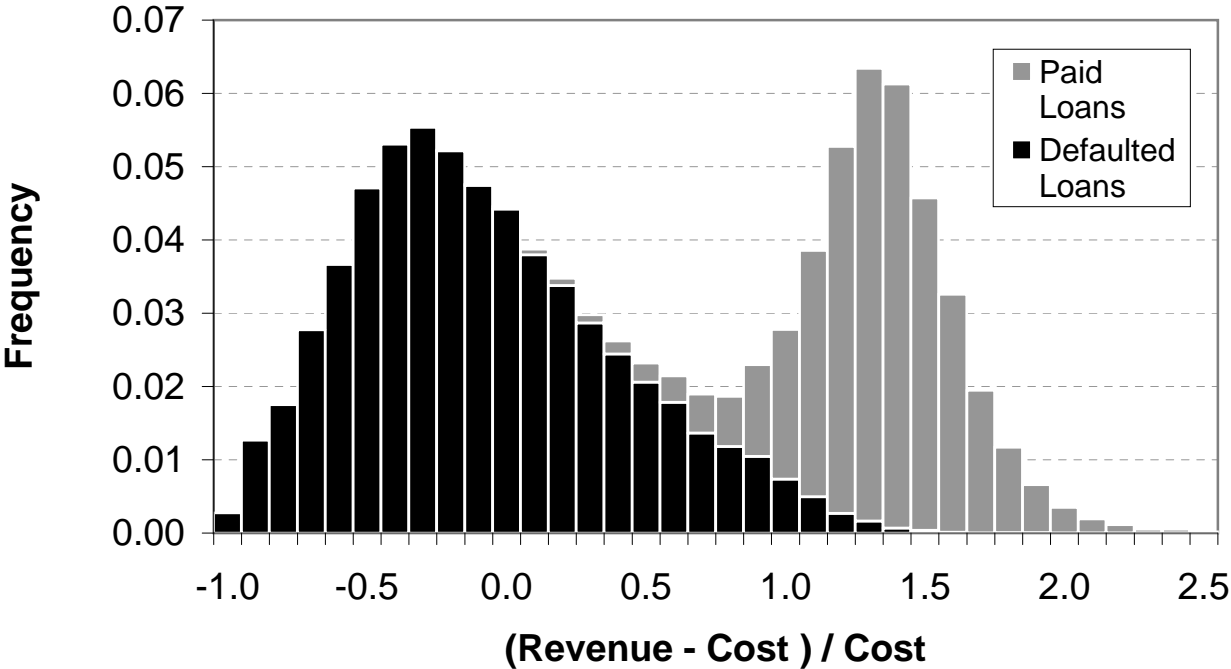
$$\Omega_p[d_i, p_i|\psi] = \left[ \min \left\{ 0, \sum_i (\Pi_i(d_i, p_i) - \Pi_i(d_i, p_i + 100)) \right\} \right]^2 + \left[ \min \left\{ 0, \sum_i (\Pi_i(d_i, p_i) - \Pi_i(d_i, p_i - 100)) \right\} \right]^2 \quad (48)$$

The parameters are found by grid search over a grid with increments of \$100 for the firm's shadow cost of capital adjustment  $\psi$ .

**Figure 1(a): Kernel Density of Fraction of Loan Paid Conditional on Default**

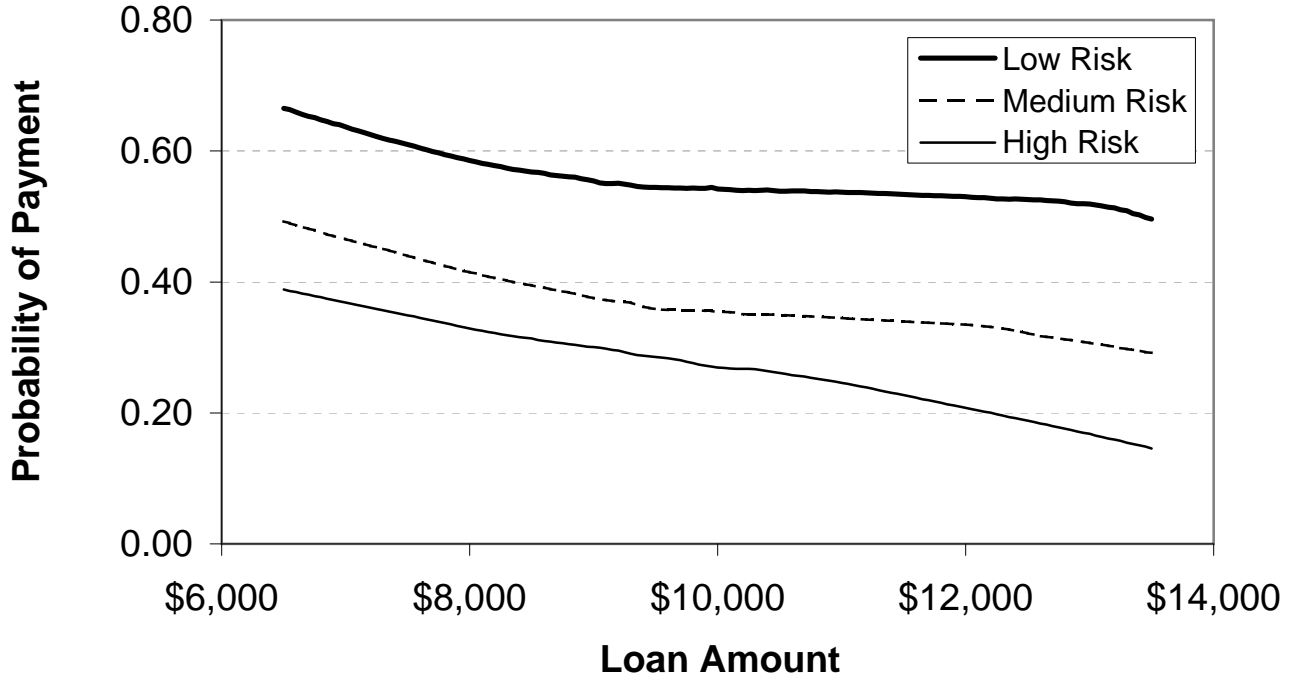


**Figure 1(b): Rate of Return Histogram**



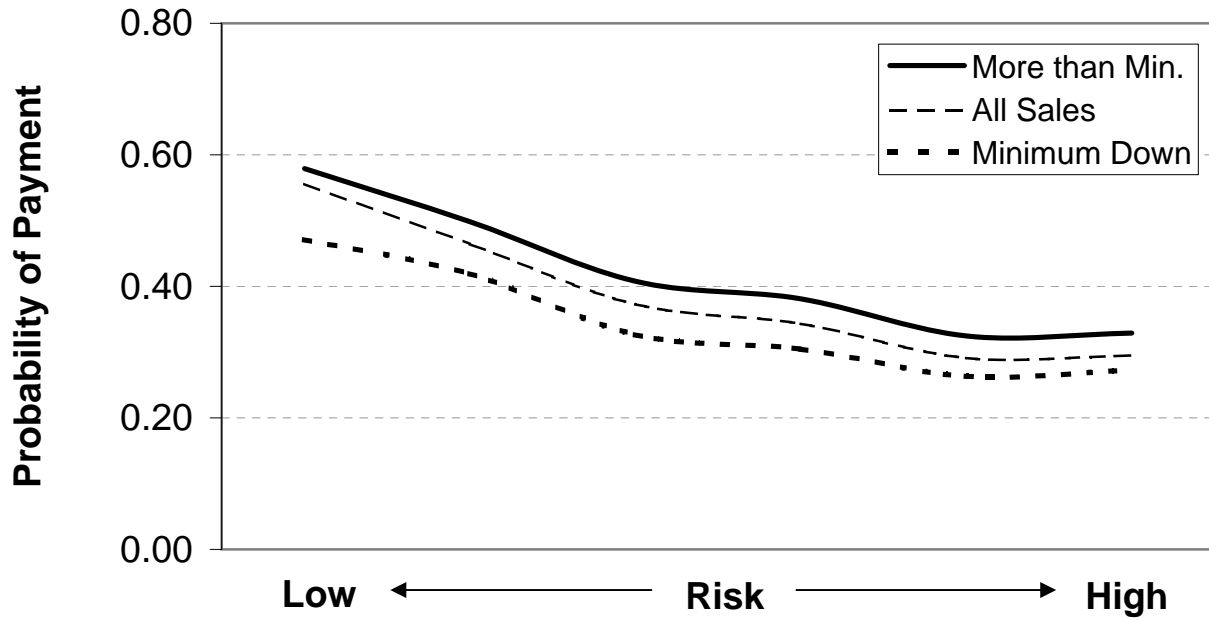
Notes: Figure 1(a): Based on data from uncensored loans that ended in default. Figure 1(b): Based on data from uncensored loans. Revenue is calculated as down payment + PV of loan payments + PV of recovery, assuming an internal firm discount rate of 10 percent.

**Figure 2(a): Probability of Payment vs. Loan Amount**



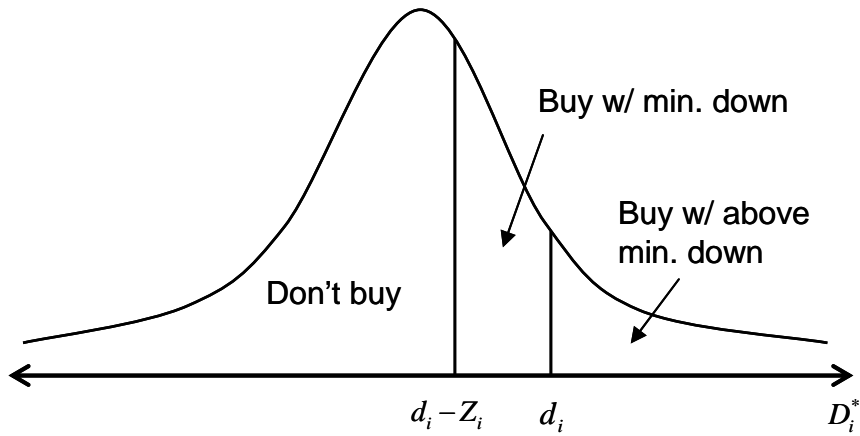
Notes: Based on data from uncensored loans. The x-axis represents the size of the loan at the time of origination, not including finance charges. The y-axis represents the probability that the loan is repaid in full. The "Low Risk" line (solid dark), "Medium Risk" line (dashed), and "High Risk" line (solid light) show the average relationship between loan amount and the probability of full payment for buyers in each risk group, where risk groups are defined by internal company credit scoring. All lines are constructed by local linear regression of a payment dummy on loan amount for buyers in each risk group.

**Figure 2(b): Probability of Payment by Risk Type and Down Payment**

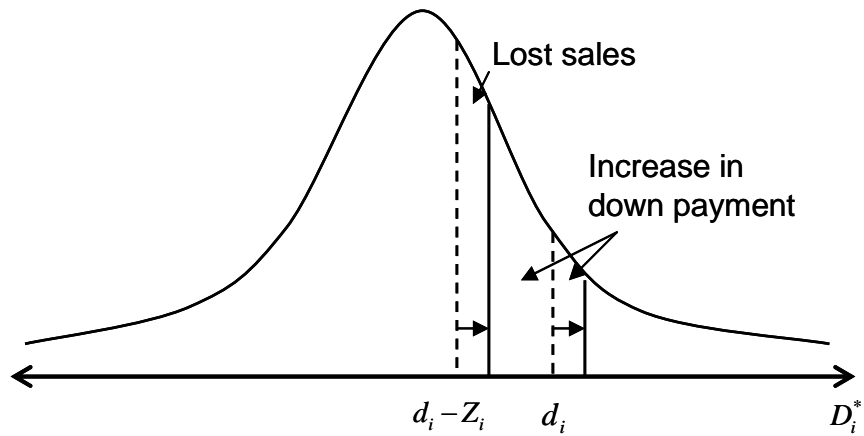


Notes: Based on data from uncensored loans. The x-axis represents a measure of buyer riskiness based on internal company credit scoring. The y-axis represents the probability that buyers repay the loan in full. The "More than Min." line (solid) shows the relationship between credit score and the probability of payment for buyers who put down more than the required minimum down payment. The "Minimum Down" line (short-dashed) shows this relationship for buyers who put down exactly the required minimum down payment. The "All Sales" line (long-dashed) shows the relationship for all buyers. All lines are constructed by grouping buyers at similar risk levels, calculating probabilities, and smoothing.

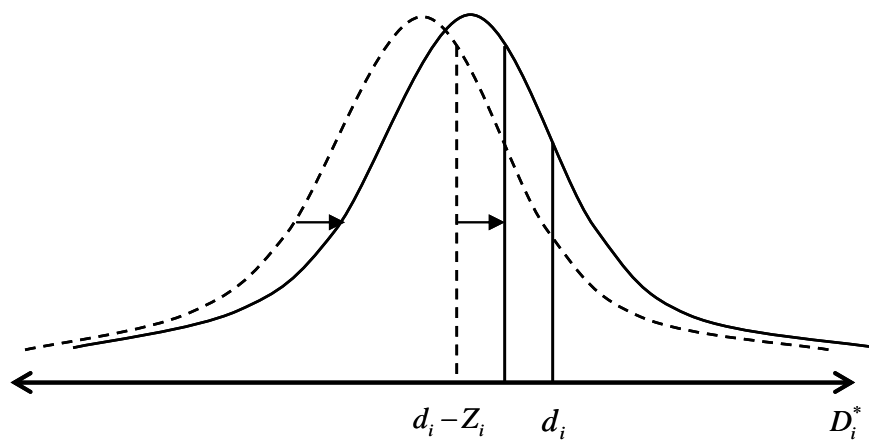
**Figure 3(a): Purchasing and Down Payment**



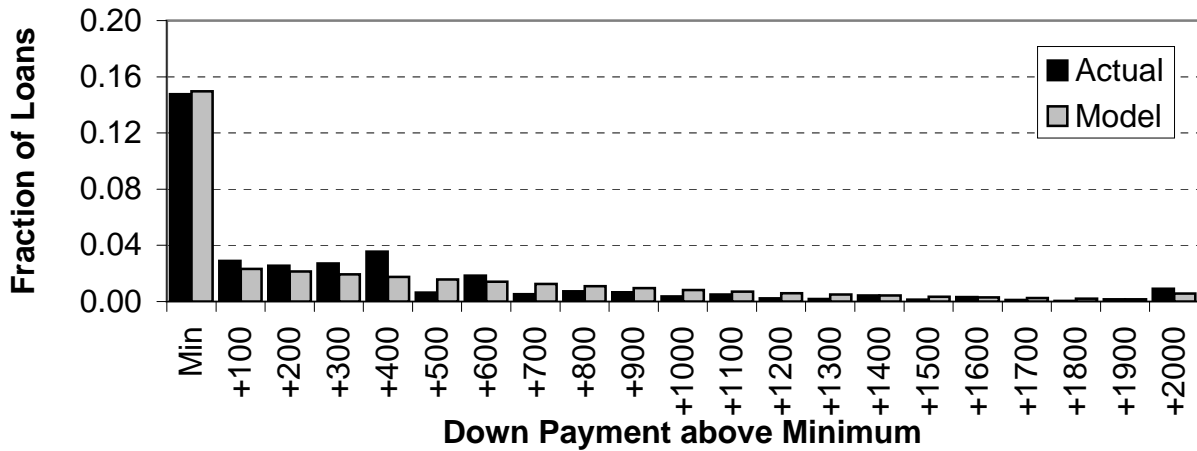
**Figure 3(b): An Increase in Minimum Down**



**Figure 3(c): An Increase in Car Price**

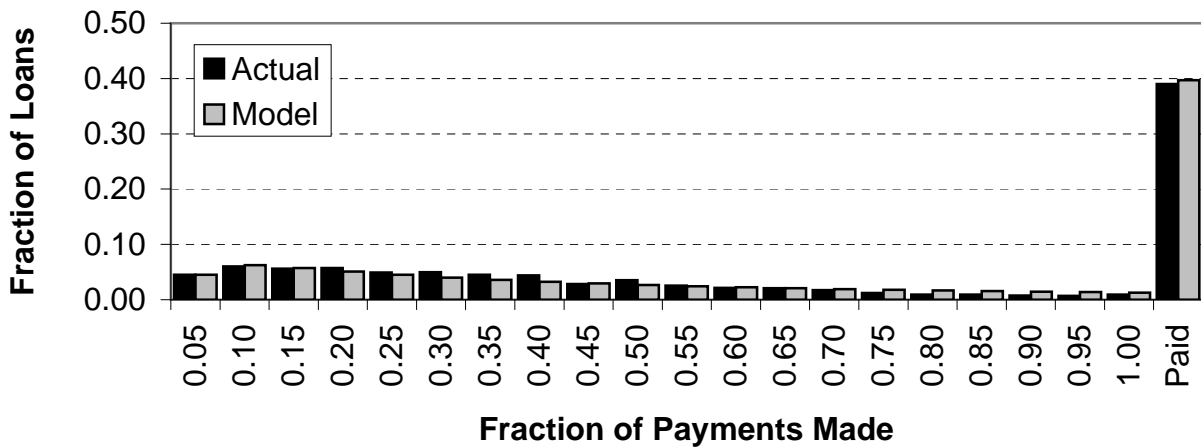


**Figure 4(a): Distribution of Down Payments**



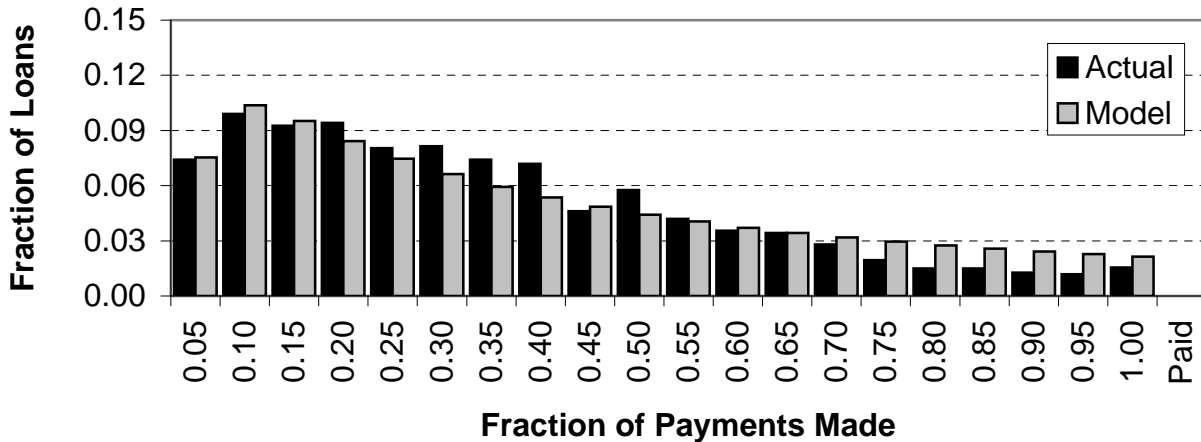
Notes: Based on raw data and model output for all sales. The bucket labeled "Min" includes all buyers who put down exactly the required minimum down payment. Each bucket labeled "+X" (except for the "+2000" bucket) includes all buyers who put down an amount between \$(X-100) and \$X more than the required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

**Figure 4(b): Distribution of Default Timing**



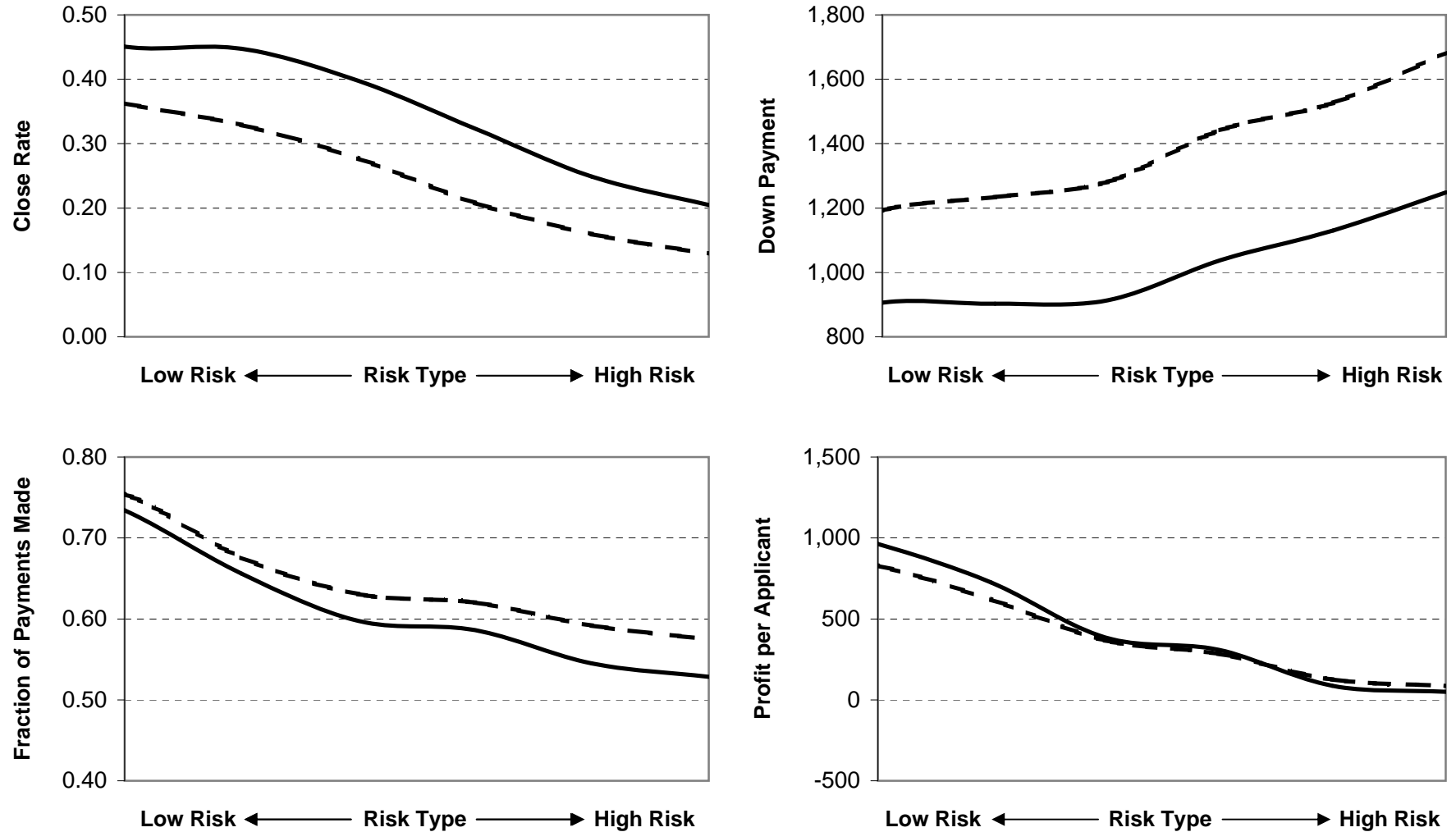
Notes: Based on raw data and model output for uncensored loans. Each bucket labeled "X" includes all buyers who defaulted after making a fraction of their payments between (X-0.05) and X.

**Figure 4(c): Distribution of Default Timing Conditional on Default**



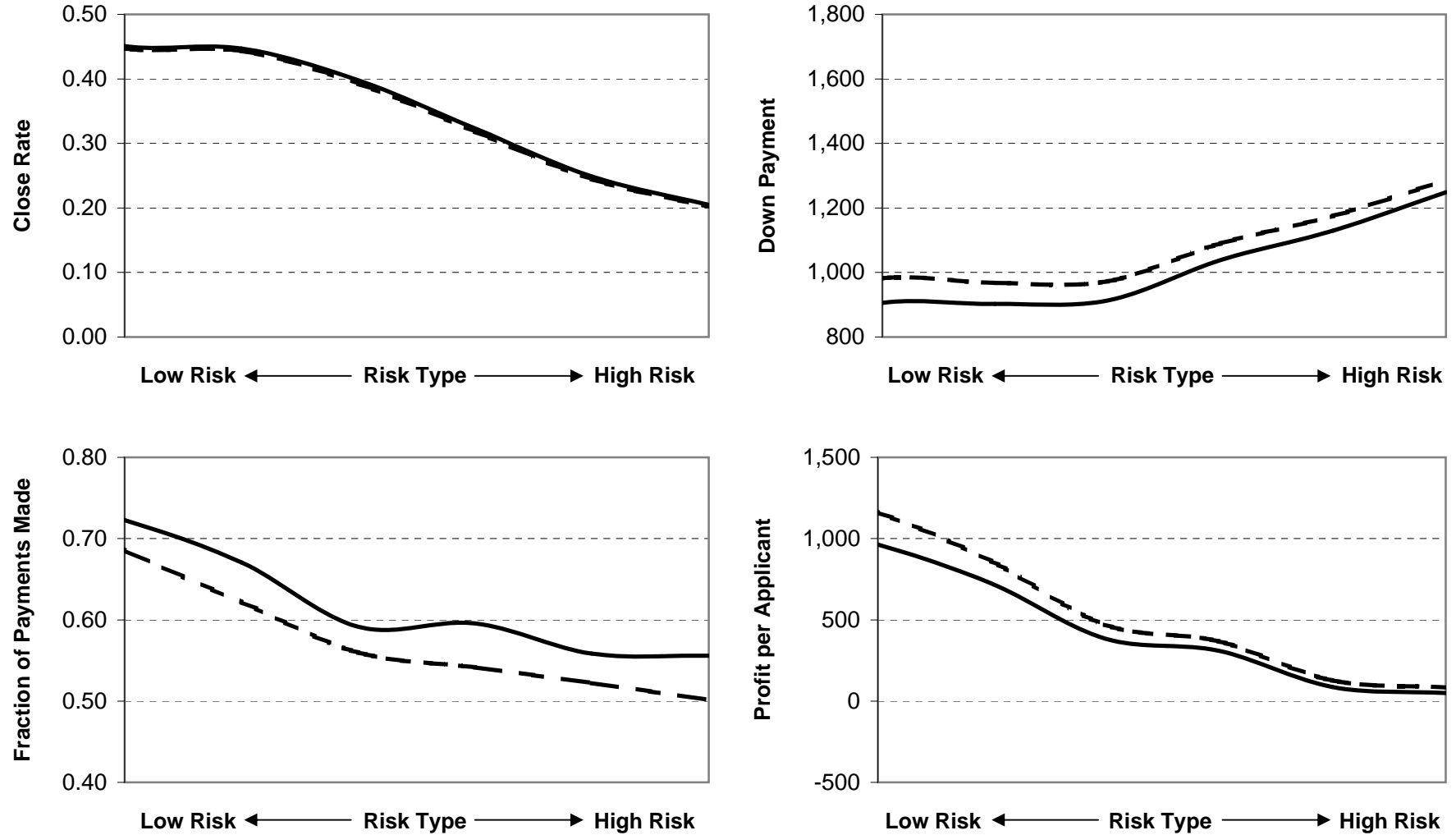
Notes: Based on raw data and model output for uncensored loans. Each bucket labeled "X" includes all buyers who defaulted after making a fraction of their payments between (X-0.05) and X.

**Figure 5: Effects of a \$500 Increase in Minimum Down Payment**



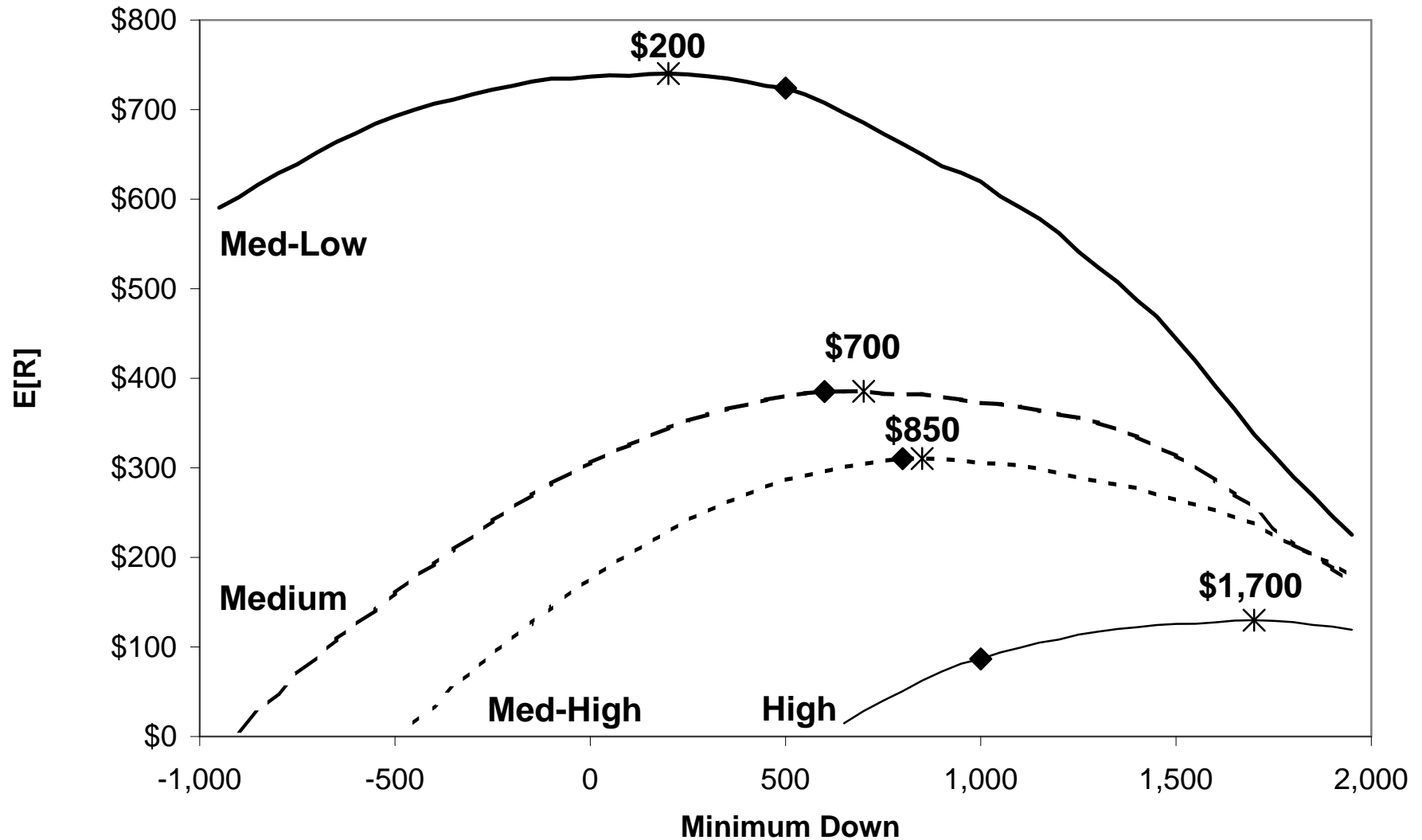
Notes: Solid lines represent model-estimated values at observed company prices. Dashed lines represent model-estimated values for a counterfactual in which the minimum down payments for all buyers are raised by \$500 from their observed levels, while margins are held fixed. The x-axis on each chart represents the internal company credit score assigned to buyers prior to loan origination. Profit per applicant is equal to 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 - shadow cost. The shadow cost is estimated using the supply-side moments described in Section 5. All series are constructed by grouping buyers with similar internal company credit scores, calculating the variable of interest for each group, and smoothing.

**Figure 6: Effects of a \$1,000 Increase in Car Price**



Notes: Solid lines represent model-estimated values at observed company prices. Dashed lines represent model-estimated values for a counterfactual in which the margins for all buyers are raised by \$1,000 from their observed levels, while minimum down payments are held fixed. The x-axis on each chart represents the internal company credit score assigned to buyers prior to loan origination. Profit per applicant is equal to 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 - shadow cost. The shadow cost is estimated using the supply-side moments described in Section 5. All series are constructed by grouping buyers with similar internal company credit scores, calculating the variable of interest for each group, and smoothing.

**Figure 7: Expected Profit vs. Minimum Down Payment**



*Notes:* Based on model estimates for all applicants. Med-Low, Medium, Med-High, and High labels indicate risk category, with High being most risky. The x-axis represents the required minimum down payment applied to applicants in each credit category. The y-axis represents expected profit per applicant, calculated as the probability of sale times net operating revenue per sale (see Figure 5 notes for details). Diamonds show observed average minimum down payments for each credit category. Stars and the associated labels show optimal minimum down payments based on model estimates.

**Table 1: Summary Statistics**

	Obs*	Mean	Std. Dev.	5%	95%
<i>Applicant Characteristics</i>					
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	-	-	-
<i>Buyer Characteristics</i>					
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	-	-	-
<i>Car Characteristics</i>					
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
<i>Transaction Characteristics</i>					
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
<i>Loan Outcomes (uncensored sales only)</i>					
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: \* To preserve the confidentiality of the company that provided the data, we do not report the exact number of applications. 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.

**Table 2: Model Fit**

---

---

	<u>Raw Data</u>	<u>Demand Model</u>
<i>Close Rate</i>		
All Applicants	0.343	0.343
Low Risk	0.451	0.451
Medium Risk	0.398	0.398
High Risk	0.249	0.249
<i>Probability of Making Minimum Down Payment</i>		
All Buyers	0.431	0.430
Low Risk	0.234	0.234
Medium Risk	0.428	0.428
High Risk	0.570	0.570
<i>Average Loan Size</i>		
All Buyers	\$10,709	\$10,690
Low Risk	\$11,047	\$11,034
Medium Risk	\$10,660	\$10,649
High Risk	\$9,992	\$9,967
<i>Probability of Payment</i>		
All Buyers	0.390	0.397
Low Risk	0.559	0.559
Medium Risk	0.363	0.384
High Risk	0.289	0.320
<i>Fraction of Payments Made</i>		
All Buyers	0.594	0.612
Low Risk	0.715	0.764
Medium Risk	0.576	0.596
High Risk	0.521	0.560
<i>Correlation between Loan Size and Fraction of Payments Made</i>		
All Buyers	-0.056	-0.055
Low Risk	-0.013	-0.078
Medium Risk	-0.064	-0.049
High Risk	-0.142	-0.111

---

Notes: Close rate, probability of making minimum down payment, and average loan amount based on all observations. Probability of making minimum down payment is conditional on sale. Probability of payment, fraction of payments made, and correlation between loan size and fraction of payments made based on uncensored sales only.



**Table 3(b): Supply Estimates**

	Level Changes <sup>a</sup>		Percent Changes <sup>b</sup>	
	Unobs.		Unobs.	
	Cost	Std. Err.	Cost	Std. Err.
<b>Uniform</b>				
Baseline	2,400	(84)	2,300	(90)
Learning #1	2,700	(527)	-	-
Learning #2	2,500	(493)	-	-
Learning #3	2,900	(580)	-	-
Learning #4	2,600	(542)	-	-
<b>By Credit Category - Broad</b>				
Low Risk	3,300	(130)	3,200	(152)
Medium Risk	2,200	(101)	2,300	(119)
High Risk	1,700	(107)	1,800	(127)
<b>By Credit Category - Fine</b>				
1 (lowest risk)	4,200	(272)	4,600	(2688)
2	3,800	(254)	3,800	(260)
3	2,800	(168)	2,700	(200)
4	2,300	(118)	2,300	(133)
5	2,200	(183)	2,300	(194)
6	1,800	(166)	1,800	(176)
7	1,500	(275)	1,700	(239)
8 (highest risk)	1,900	(241)	1,900	(220)
<b>Seasonal</b>				
Non Tax Season	2,400	(81)	2,300	(101)
Tax Season	2,200	(258)	2,100	(402)

Notes: All results are estimates of the firm's shadow cost of capital adjustment based on the supply-side moment inequalities described in Section 5 and Appendix C.2. Estimates assume a firm discount rate of 10 percent. Standard errors are bootstrap standard errors from 30 resamplings.

The two columns correspond to alternative methods of computing moment inequalities, specifically:  
a Moment inequalities are based on \$100 uniform changes to minimum down and price.  
b Moment inequalities are based on 10% changes in minimum down and 1% changes in price.

The four estimates based on "learning" moment inequalities reflect the following four conditions:  
#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)

**Table 4: First Order Condition Quantities by Risk Category**

Description	Representation in Paper	All Sales			
		Low Risk	Med Risk	High Risk	All Buyers
<i>Close Rate</i>	$Q(p_i, d_i)$	0.451	0.398	0.249	0.343
<i>Profit Conditional on Sale</i>					
Average Buyer	$E[r(p_i, d_i)   g_i > 0]$	\$2,137	\$969	\$348	\$1,174
Marginal Buyer	$E[r(p_i, d_i)   g_i = 0]$	\$1,719	\$481	-\$117	\$609
Average Non-Buyer	$E[r(p_i, d_i)   g_i < 0]$	\$1,036	-\$252	-\$745	-\$226
<i>Changes in Close Rates</i>					
\$100 increase in minimum down	$Q(p_i, d_i) - Q(p_i, d_i + 100)$	-0.023	-0.026	-0.020	-0.022
\$1,000 increase in list price	$Q(p_i, d_i) - Q(p_i + 1000, d_i)$	-0.004	-0.005	-0.004	-0.004
<i>Changes in Profit Conditional on Sale</i>					
\$100 increase in minimum down	$E[r(p_i, d_i) - r(p_i, d_i + 100)   g_i > 0]$	\$11	\$30	\$43	\$29
\$1,000 increase in list price	$E[r(p_i, d_i) - r(p_i + 1000, d_i)   g_i > 0]$	\$456	\$223	\$167	\$267

Notes: All results based on model estimates at observed prices and minimum down payments. Close Rate is the probability that an applicant purchases a car. Profit Conditional on Sale from the Average Buyer is expected profits from all applicants who purchase a car. Profit from the Marginal Buyer is the expected profit from all buyers whose desired down payments are between \$0 and \$100 above the purchase threshold. Profit from the Average Non-Buyer is the expected profit from applicants who do not purchase a car, assuming these applicants make the minimum down payment. Profit is defined as net operating revenue (down payment + PV of loan payments + PV of recovery - vehicle cost) minus a shadow cost. The shadow cost is estimated using the supply-side moments described in Section 5. Increases in price and minimum down payment describe uniform changes across all credit categories and time periods.

**Table 5(a): Value of Credit Scoring**

	Low Risk	Med Risk	High Risk	All Applicants
<i>Minimum Down Payment</i>				
Observed pricing	\$400	\$600	\$1,000	-
Optimal credit-based pricing	\$0	\$700	\$1,550	-
Optimal uniform pricing	\$800	\$800	\$800	\$800
Pricing with perfect knowledge of liquidity	-	-	-	-
<i>Close Rate</i>				
Observed pricing	0.451	0.398	0.249	0.343
Optimal credit-based pricing	0.568	0.377	0.150	0.340
Optimal uniform pricing	0.381	0.352	0.291	0.328
Pricing with perfect knowledge of liquidity	0.577	0.472	0.273	0.408
<i>Profit Conditional on Sale</i>				
Observed pricing	\$2,137	\$969	\$348	\$1,174
Optimal credit-based pricing	\$1,924	\$1,026	\$914	\$1,258
Optimal uniform pricing	\$2,254	\$1,092	\$218	\$1,112
Pricing with perfect knowledge of liquidity	\$2,499	\$1,543	\$1,154	\$1,695
<i>Expected Profit per Applicant</i>				
Observed pricing	\$963	\$385	\$87	\$402
Optimal credit-based pricing	\$1,093	\$387	\$137	\$428
Optimal uniform pricing	\$859	\$384	\$63	\$364
Pricing with perfect knowledge of liquidity	\$1,443	\$728	\$314	\$692

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 a shadow cost. The shadow cost is estimated using the supply-side moments described in Section 5. For observed pricing, profit conditional on sale corresponds to the term  $E[r(\pi_i, d_i) | g_i > 0]$  in 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-based pricing describes a counterfactual in which minimum down payments are constant over time, but vary by credit category in order to maximize expected profit per applicant. (3) Optimal uniform pricing describes a counterfactual in which a single minimum down payment, which is constant over time and across credit categories, is chosen to maximize expected profit per applicant. (4) Pricing with perfect knowledge of liquidity describes 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 who are profitable at this minimum down payment.

**Table 5(b): Value of Tax Season Pricing**

	<b>Low Risk</b>	<b>Med Risk</b>	<b>High Risk</b>	<b>All Applicants</b>
<i>Minimum Down Payment</i>				
Observed Tax-Season Pricing	\$400	\$600	\$1,000	-
Optimal Tax-Season Pricing	\$50	\$1,300	\$1,900	-
Optimal Full-Year Pricing	\$0	\$700	\$1,550	-
<i>Close Rate</i>				
Observed Tax-Season Pricing	0.506	0.463	0.288	0.402
Optimal Tax-Season Pricing	0.615	0.368	0.172	0.350
Optimal Full-Year Pricing	0.615	0.466	0.220	0.417
<i>Profit Conditional on Sale</i>				
Observed Tax-Season Pricing	\$2,040	\$929	\$396	\$1,085
Optimal Tax-Season Pricing	\$1,818	\$1,313	\$1,114	\$1,429
Optimal Full-Year Pricing	\$1,812	\$944	\$815	\$1,121
<i>Expected Profit per Applicant</i>				
Observed Tax-Season Pricing	\$1,033	\$431	\$114	\$436
Optimal Tax-Season Pricing	\$1,118	\$483	\$192	\$500
Optimal Full-Year Pricing	\$1,114	\$440	\$179	\$468

Notes: All results based on model estimates for tax season only. Tax season is defined as the month of February. 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 a shadow cost. The shadow cost is estimated using the supply-side moments described in Section 5. Expected profit per applicant is equal to the close rate times profit conditional on sale.

Each counterfactual represents a different minimum down payment policy, and each allows minimum down paymentes to vary by grade. List prices are held fixed at observed values in all counterfactuals. Observed pricing describes outcomes based on the company's observed minimum down payments during tax seasons. In this scenario, minimum downs may vary across different tax seasons. Optimal tax-season pricing describes a counterfactual in which minimum down payments are chosen to maximize expected profit per applicant during tax seasons. Optimal full-year pricing describes a counterfactual in which minimum down payments are chosen to maximize expected profit per applicant over the entire sample.

**Table 5(c): 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)	(\$364, \$0)	(\$428, \$0)
Entrant Prices Uniformly	(\$168, \$168)	(\$249, \$123)
Entrant Prices by Grade	(\$123, \$249)	(\$202, \$202)

*Equilibrium Minimum Downs*

	Uniform	By Risk Category (Low, Med, High)
Monopoly	\$800	\$0, \$700, \$1550
Vs. Uniform	\$350	\$0, \$450, \$750
Vs. Risk-Based	\$500	\$0, \$250, \$1000

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(a). 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 of the firms simultaneously either set 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(a), column 3) and estimated costs (Table 3(b)).

**Table A1: Recovery Parameter Estimates**

<i>Dependent Variable</i>	<i>Nonzero recovery ind.</i>		<i>Recovery Amt (\$1000s)</i>	
	Probit		OLS	
	dF/dx	Std. Err.	Coeff.	Std. Err.
Months Paid	-0.006	(0.0002)	-0.045	(0.001)
<i>Car Characteristics</i>				
Car Cost (\$1000s)	0.027	(0.002)	0.526	(0.006)
Premium (Cost > \$7,500)	-0.097	(0.008)	0.079	(0.024)
Car Age (years)	-0.008	(0.001)	-0.068	(0.004)
Odometer (10,000s)	0.010	(0.001)	-0.029	(0.003)
Lot Age (months)	-0.004	(0.001)	-0.131	(0.004)
<i>Individual Characteristics</i>				
Income (\$1,000s/month)	0.000	(0.002)	0.048	(0.006)
Bank Account	0.001	(0.003)	0.078	(0.012)
House Owner	0.001	(0.004)	0.118	(0.015)
Other Fixed Effects	Year, Month, City,		Year, Month, City,	

Notes: The sample for the probit equation is all defaults; sample size is  $\sim 0.18N$ , where  $N \gg 50,000$  (see Table 1). Reported coefficients show the marginal effect of a one unit change in the explanatory variable on the probability of making a positive recovery. The sample for the OLS equation is all recoveries; sample size is  $\sim 0.14N$ .