Liquidity Constraints and Imperfect Information in Subprime Lending*

William Adams, Liran Einav and Jonathan Levin†

September 2007

Abstract. We present new evidence on consumer liquidity constraints and the credit market conditions that might give rise to them. Our analysis is based on unique data from a large auto sales company that serves the subprime market. We first document the role of short-term liquidity in driving purchasing behavior, including sharp increases in demand during tax rebate season and a high sensitivity to minimum down payment requirements. We then explore the informational problems facing subprime lenders. We find that default rates rise significantly with loan size, providing a rationale for lenders to impose loan caps because of moral hazard. We also find that borrowers at the highest risk of default demand the largest loans, but the degree of adverse selection is mitigated substantially by effective risk-based pricing.

*We thank two anonymous referees and Judy Chevalier, the Editor, as well as Raj Chetty, Amy Finkelstein, Robert Hall, Richard Levin, and many seminar participants for suggestions and encouragement. Mark Jenkins provided stellar research assistance and Ricky Townsend greatly assisted our early data analysis. Einav and Levin acknowledge the support of the National Science Foundation and the Stanford Institute for Economic Policy Research, and Levin acknowledges the support of the Alfred P. Sloan Foundation.

†Adams is at Citigroup; email: william.adams@citigroup.com; Einav and Levin are at the Department of Economics, Stanford University, Stanford CA 94305-6072; emails: leinav@stanford.edu and jdlevin@stanford.edu.

1 Introduction

Access to credit markets is generally considered a hallmark of developed economies. In the United States, most households appear to have substantial ability to borrow; indeed, the average household in the United States has over 23,000 dollars in non-mortgage debt alone. Nevertheless, economists often point to limited borrowing opportunities, or liquidity constraints, to explain anomalous findings about consumption behavior, labor supply, and the demand for credit. Despite a sizeable theoretical literature that explains why some borrowers might have trouble obtaining credit even in competitive markets (e.g., Jaffee and Stiglitz, 1990), there has been relatively little work relating the consumer behavior indicative of liquidity constraints to the actual functioning of the credit market.

In this paper, we use unique data from a large auto sales company to study liquidity constraints and credit market conditions for precisely the population that is most likely to have a difficult time obtaining credit, those with low incomes and poor credit histories. These consumers, who typically cannot qualify for regular bank loans, comprise the so-called subprime market. Because the company we study originates subprime loans, its loan applications and transaction records provide an unusual window into the consumption and borrowing behavior of households that are rationed in primary credit markets. Moreover, we track loan repayments allowing us to analyze the difficulties in lending to the subprime population and explain why their supply of credit may be limited.

We begin by documenting the importance of short-term liquidity constraints for individuals in our sample. We present two pieces of evidence. Both are based on purchasing behavior and indicate a high sensitivity to cash-on-hand. First, we document a dramatic degree of demand seasonality associated with tax rebates. Overall demand is almost 50 percent higher during tax rebate season than during other parts of the year. This seasonal effect varies with household income and with the number of dependents, closely mirroring the federal earned income tax credit schedule. Second, we find that demand is highly responsive to changes in minimum down payment requirements. A 100 dollar increase in the required down payment, holding car prices fixed, reduces demand by 9 percent. In contrast, generating the same reduction in demand requires an increase in car prices of almost 3,000 dollars. We calculate that in the absence of liquidity constraints these effects would imply an annual discount rate of 1,415 percent.[1]

[1] If s denotes the discount rate, the discount factor is $1/(1+s)$. So an annual discount rate of 1,415 percent implies

Taken together, these findings point to the conclusion that this population does not have ready access to credit that allows them to shift wealth across time. This raises the question of whether consumer liquidity constraints can be tied to underlying credit market conditions. One possibility is that high default rates, coupled with legal caps on interest rates, simply rule out some forms of lending. A second possibility is that fundamental features of the consumer credit market are responsible for credit constraints. We focus on the latter, turning to the information economics view of credit markets as developed by Jaffee and Russell (1976) and Stiglitz and Weiss (1981).

Modern information economics emphasizes that credit constraints can arise in equilibrium even if financing terms can adjust freely and lenders are fully competitive. Its explanation lies in the twin problems of moral hazard and adverse selection. In the moral hazard version of the story, individual borrowers are more likely to default on larger loans. This leads to problems in the loan market because borrowers do not internalize the full increase in default costs that come with larger loan sizes. As a result, lenders may need to cap loan sizes to prevent over-borrowing. In contrast, adverse selection problems arise if borrowers at high risk of default also desire large loans, as might be expected given that they view repayment as less likely. As the theoretical literature has pointed out, adverse selection can give rise not only to loan caps, but to some worthy borrowers being denied credit because they cannot distinguish themselves from the less worthy.[2]

The second half of the paper explores these ideas, first from the standpoint of theory and then empirically. In Section 4 we present a simple model of consumer demand for credit and competitive lending, along the lines of Jaffee and Russell (1976). We show that such a model can explain many of the institutional features we observe on the lender side of the market, such as the adoption of credit scoring and risk-based pricing, and the use of interest rates that increase with loan size. We also explain why informational problems, compounded by interest rate caps, create a rationale for lenders to limit access to credit. The model therefore provides a simple credit market based explanation for why purchasing behavior might reflect liquidity constraints.

Having outlined the theoretical framework, we investigate the empirical importance of moral hazard and adverse selection for subprime lending. Separately identifying these two forces is often a challenge because they have similar implications: both moral hazard and adverse selection imply

an annual subjective discount factor of less than 0.07. Such an individual is indiferent between paying 1,000 dollars today and 15,150 dollars in a year.

[2] The fact that imperfect information in the credit market leads to limits on lending is analogous to Rothschild and Stiglitz's (1976) famous observation that imperfect information in an insurance market may lead to under-insurance relative to the full-information optimum.

a positive correlation between loan size and default. A useful feature of our data is that we can exploit exogenous (to the individual) variation in car price and minimum down payment to isolate the moral hazard effect of increased loan size on default. This in turn allows us to back out a quantitative estimate of self-selection from the cross-sectional correlation between loan size and default. We explain the econometric strategy in detail in Section 5.2.

We find compelling evidence for both moral hazard and adverse selection. We estimate that for a given borrower, a 1,000 dollar increase in loan size increases the rate of default by over 17 percent. This alone provides a rationale for limiting loan sizes because the expected revenue from a loan is not monotonically increasing in the size of the loan. Regarding adverse selection, we find that borrowers who are observably at high risk of default are precisely the borrowers who desire the largest loans. The company we study assigns buyers to a small number of credit categories. We estimate that all else equal, a buyer in the worst category wants to borrow around 180 dollars more than a buyer in the best category, and is more than twice as likely to default given equally-sized loans.

This strong force toward adverse selection is mitigated substantially by the use of risk-based pricing. In practice, observably risky buyers end up with smaller rather than larger loans because they face higher down payment requirements. The finding is notable because the development of sophisticated credit scoring is widely perceived to have had a major impact on consumer credit markets, but there is relatively little empirical evidence on exactly what it accomplishes. Here we document its marked effect in matching high-risk borrowers with smaller loans. Of course, risk-based pricing only mitigates selection across observably different risk groups. We also look for, and find, evidence of adverse selection within risk groups, driven by unobservable characteristics. Specifically, we estimate that a buyer who pays an extra 1,000 dollars down for unobservable reasons will be 17 percent less likely to default than one who does not given identical cars and equivalent loan liabilities. This adverse selection on unobservables is both statistically and economically significant, but smaller in magnitude than our estimates of moral hazard.

We view these findings as broadly supportive of the information economics view of consumer lending and its explanation for the presence of credit constraints. Overall our evidence supports: (1) the underlying forces of informational models of lending, namely moral hazard and adverse selection; (2) the supply-side responses these models predict, specifically loan caps, variable interest rates, and risk-based pricing; and (3) the predicted consequences, specifically liquidity effects in purchasing behavior. So while there are limits to what we can conclude with data from a single lender, we think that our results highlight the empirical relevance of informational models of consumer credit markets.

Our paper ties into a large empirical literature documenting liquidity-constrained consumer behavior and a much smaller literature on its causes. Much of the accumulated evidence on the former comes from consumption studies that document relatively high propensities to consume out of transitory income, particularly for households with low wealth.[3] Some of the sharpest evidence in this regard comes from analyzing consumption following predictable tax rebates. For instance, Johnson, Parker and Souleles (2006) find that households immediately consumed 20-40 percent of the 2001 tax rebate, with the effect biggest for low-wealth households (see also Souleles, 1999, and Parker, 1999). A common explanation for these findings is that households with low wealth are unable to effectively access credit (Deaton, 1991; Zeldes, 1989).[4]

Further evidence on credit constraints comes from Gross and Souleles (2001), who use detailed data from a credit card company to look at what happens when credit limits are raised. They find that a hundred dollar increase in a card holder's limit raises spending by ten to fourteen dollars. Based on this, they argue that a substantial fraction of borrowers in their sample appear to be credit constrained. As will be apparent below, the population in our data is most likely in a substantially worse position to access credit than the typical credit card holder.

A distinct set of evidence on credit constraints comes from studying household preferences over different types of loan contracts. An early survey by Juster and Shay (1964) found striking differences between households in their willingness to pay higher interest rates for a longer loan with lower monthly payments. In particular, households likely to be credit constrained, e.g. those with lower incomes, were much more willing to pay higher interest rates to reduce their monthly payment. More recently, Attanasio, Goldberg and Kyriazidou (2006) use Survey of Consumer Finances data on auto loans to show that for most households, and particularly for low-income ones, the demand for loans is much more sensitive to loan maturity than to interest rate.[5] Their interpretation is

[3] Studies of the effects of unemployment insurance also provide evidence for credit constraints (e.g., Chetty, 2007; Card, Chetty and Weber, 2007).

[4] There is no clear consensus, however, on the exact story. For instance, Carroll (2001) argues that much of the evidence on consumption behavior can be explained by a buffer stock model where all agents can borrow freely at relatively low interest rates. Jappelli (1990) provides some limited evidence supporting credit rationing, based on the fact that nineteen percent of the households in the 1983 Survey of Consumer Finances report having had a credit application rejected or not applying for a loan for fear of being rejected.

[5] Karlan and Zinman (forthcoming) report a similar finding, that loan demand is more sensitive to maturity than to interest rate, based on a pricing experiment carried out by a South African lender. Their experiment also provides

that because of their limited access to credit, many consumers will pay a substantial premium to smooth payments over a longer period.

The purpose of the above studies is to document that a significant set of households has a limited ability to borrow at desirable rates. There is much less empirical work that addresses the causes of credit constraints. Ausubel (1991, 1999) argues that the high interest rates charged by credit card issuers are a market failure caused by adverse selection, a view that is supported by direct marketing experiments. Edelberg (2003, 2004) also finds evidence for adverse selection in both mortgage lending and automobile loans, and documents an increasing trend toward risk-based interest rates. We view it as a virtue of our data that we can tie together demand-side evidence for credit constrained behavior with evidence on the informational problems that might give rise to these constraints. Some of our ongoing work (Einav, Jenkins, and Levin, 2007) explores more deeply how lenders respond to informational problems by looking at the introduction of credit scoring and the problem of optimal loan pricing in the presence of moral hazard and adverse selection.

2 Data and Environment

Our data come from an auto sales company that operates used car dealerships in the United States. Each potential customer fills out a loan application and is assigned a credit category that determines the possible financing terms. Almost all buyers finance a large fraction of their purchase with a loan that extends over a period of several years. What makes the company an unusual window into consumer borrowing is its customer population. Customers are primarily low-income workers and a great majority are subprime borrowers. In the U.S., Fair Isaac (FICO) scores are the most-used measure of creditworthiness. They range from 350 to 800, with the national median between 700 and 750. Less than half of the company's applicants have a FICO score above 500, the second percentile of the national FICO score distribution. This kind of low credit score indicates either a sparse or, more often, checkered credit record.

The principal characteristics of subprime lending are high interest rates and high default rates. A typical loan in our data has an annual interest rate on the order of 25-30 percent. The flip side of high interest rates is high default rates. Over half of the company's loans end in default. With such a high probability of default, screening the good risks from the bad, and monitoring loan

some evidence for moral hazard and adverse selection (Karlan and Zinman, 2007).

payments, is extremely important. The company has invested significantly in proprietary credit scoring technology.

Our specific data consists of all loan applications and sales from June 2001 through December 2004. We combined this with records of loan payments, defaults and recoveries through April 2006. This gives us information on the characteristics of potential customers, the terms of the consummated transactions, and the resulting loan outcomes. We have additional data on the loan terms being offered at any given time as a function of credit score, and inventory data that allows us to observe the acquisition cost of each car, the amount spent to recondition it, and its list price on the lot.

The top panel of Table 1 contains summary statistics on the applicant population. There are well over 50,000 applications in our sample period (to preserve confidentiality, we do not report the exact number of applications). The median applicant is 31 years old and has a monthly household income of 2,101 dollars. We do not have a direct measure of household assets or debt, but we observe a variety of indirect measures. A small fraction of applicants are homeowners, but the majority are renters and more live with their parents than own their own home. Nearly a third report having neither a savings nor a checking account. The typical credit history is spotty: more than half of the applicants have had a delinquent balance within six months prior to their loan application. In short, these applicants represent a segment of the population for whom access to credit is potentially problematic.

Just over one third of the applicants purchase a car. As shown in the second panel of Table 1, the average buyer has a somewhat higher income and somewhat better credit characteristics than the average applicant. In particular, the company assigns each applicant a credit category, which we partition into "high", "medium" and "low" risk. The applicant pool is 27 percent low risk and 29 percent high risk, while the corresponding percentages for the pool of buyers are 35 and 17.

The sales terms, summarized in the third panel of Table 1, reflect the presumably limited options of this population. A typical car, and most are around 3-5 years old, costs around 6,000 dollars to bring to the lot. The average sale price is just under 11,000 dollars.[6] The average down payment is a bit less than 1,000 dollars, so after taxes and fees, the average loan size is similar to the sales price.

[6] Car prices are subject to some degree of negotiation, which we discuss in Section 3. The price we report here is the negotiated transaction price rather than the "list" price, which is slightly higher.

Despite the large loans and small down payments, it appears that many buyers would prefer to put down even less money. Forty-three percent make exactly the minimum down payment, which varies with the buyer's credit category but is typically between 400 and 1,000 dollars. Some buyers do make down payments that are substantially above the required minimum, but the number is small. Less than ten percent of buyers make down payments that exceed the required down payment by a thousand dollars.

In a financed purchase, the monthly payment depends on the loan size, the loan term and the interest rate.[7] Much of the relevant variation in our data is due to the former rather than the latter. Over eighty-five percent of the loans have an annual interest rate over 20 percent, and around half the loans appear to be at the state-mandated maximum annual interest rate.[8] Most states in our data have a uniform 30 percent cap.[9] These rates mean that finance charges are significant. For instance, a borrower who takes an 11,000 dollar loan at a 30 percent APR and repays it over 42 months will make interest payments totalling 6,000 dollars.

The main reason for the high finance charges is evident in the fourth panel of Table 1. Most 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 percent are repaid in full.[10] Moreover, loans that do default tend to default quickly. Figure 1(a) plots a kernel density of the fraction of payments made by borrowers who defaulted. Nearly half the defaults occur before a quarter of the payments have been made, that is, within ten months. This leads to a highly bimodal distribution of per-sale profits. To capture this, we calculated the present value of payments received for each uncensored loan in our data, including both the down payment and the amount recovered in the event of default, using an annual interest rate of 10 percent to value the payment stream. We then divide this by the firm's reported costs of purchasing and reconditioning the car to obtain a rate of return on capital for each transaction. Figure 1(b) plots the distribution of returns, showing the clear

[7] Letting p denote the car price, D the down payment, T the loan term in months and $R = 1 + r$ the monthly interest rate, the monthly payment is given by $m = (p - D) \cdot (R - 1)/(1 - R^{-T})$.

[8] The company offers lower rates to some buyers who have either particularly good credit records or make down payments above the minimum. Although we do not have direct data on the offers of competing lenders, it seems unlikely that this population has access to better rates. Fair Isaac's web page indicates that borrowers with FICO scores in the 500-600 range (that is, better than the majority of the applicants in our sample) should expect to pay close to 20 percent annual interest for standard used car loans in most states, and in some states will not qualify at all for "standard" loans.

[9] A few states have lower caps that depend on characteristics of the car.

[10] We have qualitative data on the causes of default for about 60 percent of the defaults in our sample. More than half of these are reported as driven by personal finances (job loss, overextended debt, and personal bunkraptcy). A non-trivial portion (10-25 percent) seems to be related to problems associated with the car (accident, mechanical breakdown, car theft).

bimodal pattern.

It is also interesting to isolate the value of each stream of loan repayments and compare it to the size of each loan. When we do this for each uncensored loan in our data (and use annual discount rates of 0 to 10 percent), we find an average repayment to loan ratio of 0.79-0.88. Moreover, a substantial majority of loans in the data, 54-57 percent, have a repayment to loan ratio below one. This calculation helps to explain why buyers who are going to finance heavily in any event might maximize their loan size. In the majority of cases, the present value of payments on an extra dollar borrowed is significantly less than a dollar paid up front.[11]

3 Evidence of Liquidity Constraints: Purchasing Behavior

A consumer is liquidity constrained if he cannot finance present purchases using resources that will accrue to him in the future. Subprime borrowers are obvious candidates to find themselves in this position. While we cannot observe directly individual household balance sheets and credit options, our data does permit us to investigate the behavioral implications of liquidity constraints. We consider two such implications in this section.

The first concerns purchasing sensitivity with respect to current and predictable future cash flow. For an individual who can borrow freely against future resources, the response should be equal. In contrast, a high purchase response to a predictable temporary spike in cash flow, such as a tax rebate, suggests an inability to shift resources over time. The first piece of evidence we present is a striking seasonal increase in applications and sales at precisely tax rebate time. Moreover, we show that there is a remarkably clear correlation between the seasonal effects we observe and the amount of the earned income tax credit, which is likely to be a significant portion of the tax rebate for many households in our data.

The second empirical implication is the mirror image of the first. An individual who is not liquidity constrained should evaluate the cost of a given payment schedule based on its present value. In contrast, a liquidity constrained individual values the opportunity to defer payments to the future, and therefore views current payments as more costly than the present value of future

[11] The point applies most clearly for small changes in loan size. As we show below, smaller loans decrease the probability of default, which generates a non-convexity in loan demand. This effect is not reflected in our calculation, which takes the default process as fixed. It is also worth noting that the incentive to borrow on the margin increases with buyers' subjective discount rates. Some researchers (e.g., Laibson, Repetto and Tobacman, 2003) have argued that borrowing behavior reflects a much higher degree of impatience than we assume here.

payments. This is consistent with the second piece of evidence we present: individual purchase elasticity with respect to current payment (down payment) is an order of magnitude higher than with respect to future payments.

Can we rationalize these findings in the absence of borrowing constraints? Explaining our seasonality finding is difficult. It seems unlikely that members of the population we study have a particular need for cars in the month of February. An alternative is that consumers view their purchase as a form of savings rather than consumption. But given the price margins and very low down payments, the immediate post-purchase equity share is negligible.[12] Moreover, given the high default rate, viewing the transaction as a form of saving seems implausible unless consumers are greatly over-optimistic about their likelihood of making payments, which in turn would make it even harder to rationalize our second finding.

If consumers are realistic about the possibility of default, our second finding can be explained without reference to borrowing constraints if individuals highly discount the future. In particular, we calculate that our estimated purchase sensitivity with respect to present and future payments is consistent with consumers equating a 1,000 dollar cost today with a 15,150 dollar cost in one year. This number would be higher if consumers were over-optimistic about repayment and viewed their car purchase as a form of saving. For this reason, we view the combination of our two findings as particularly convincing evidence that liquidity plays a key role in driving consumer purchasing behavior.

3.1 The Effect of Tax Rebate Season

We start by examining seasonal patterns in demand. Figure 2 displays the average number of applications and sales, by calendar week, over the 2002-2005 period. Both are markedly higher from late January to early March. Applications are 23 percent higher in February than in the other months, and the close rate (sales to applications ratio) is 40 percent compared to 33 percent over the rest of the year. These seasonal patterns cannot be attributed to sales or other changes in the firm's offers. In fact, required down payments are almost 150 dollars higher in February, averaging across applicants in our data, than in the other months of the year. Indeed we initially thought these patterns indicated a data problem until the company pointed out that prospective buyers

[12] This would not be the case for a non-financed car purchase, which is a reason that studies of the marginal propensity to consume out of tax rebates focus on expenditure on non-durables.

receive their tax rebates at precisely this time of year.

But can tax rebates be large enough to explain such a dramatic spike in demand? All loan applicants must hold a job to be eligible for a loan, and most are relatively low earners, making them eligible for the earned income tax credit (EITC). The associated rebate, which varies with income and the number of dependents, can be as high as 4,500 dollars. To assess whether purchasing patterns might reflect EITC rebates, we classified applicants into twelve groups depending on their monthly household income and their number of dependents. For each group, we calculated the earned income tax credit for the average household in the group,[13] and also the percent increase in applications, close rate and sales in February relative to the other months. Figure 2(b) plots the relationship between the calculated EITC rebate and the seasonal spike in demand for each group. There is a sharp correlation. For households with monthly incomes below 1,500 dollars and at least two dependents, for whom the EITC rebate could be around 4,000 dollars, the number of applications doubles in February and the number of purchases more than triples. In contrast, for households with monthly incomes above 3,500 dollars and no dependents, for whom the EITC rebate is likely zero, the number of applications and purchases exhibits virtually no increase in tax rebate season.

Because minimum down payment requirements are raised during tax season, it is interesting to isolate the seasonal effect in demand holding all else constant. Our demand estimates in the next section, which control for the relevant offer terms as well as individual characteristics such as credit score and household income, indicate that the demand of applicants who arrive on the lot is 30 percent higher in the month of February than in other months. There are also positive but less pronounced demand effects for January and March. Consistent with the liquidity story, we also find that the seasonal pattern reported above is mainly driven by cash transactions, while purchases that involve trade-ins, which are less likely to be affected by tax rebates, do not exhibit noticeable seasonal variation.

Our estimates of loan demand, discussed in Section 5, are also consistent with the hypothesis that tax rebates represent a substantial liquidity shock that significantly affects behavior. In particular, down payments are substantially higher in tax season even after factoring in the higher minimum requirements. About 65 percent of February purchasers make a down payment above the

[13] The details of the EITC schedule did not change much over our observation period (2001-2005). The particular numbers we report are based on the the 2003 schedule.

required (higher) minimum compared to 54 percent in the rest of the year. Moreover, we estimate that after controlling for transaction characteristics the desired down payment of a February buyer is about 300 dollar higher than that of the average buyer. This is a nontrivial effect given that the average down payment is under 1,000 dollars.[14]

3.2 Estimating Purchasing Demand

3.2.1 Empirical Strategy

Additional evidence on the role of short-term liquidity in purchasing comes from looking at the responsiveness of demand to changes in different components of the car/financing package. To study this, we use our data on applications and purchases to estimate a model of consumer demand. Specifically, we consider a probit model for the purchase decision, estimated at the level of the individual applicant. Let q_i denote a dummy variable equal to one if applicant i purchases a car. We assume that

$$q_i = 1 \qquad \Leftrightarrow \qquad q_i^* = x_i'\alpha + \varepsilon_i \geq 0, \tag{1}$$

where $x_i = (x_i^o, x_i^c, x_i^a)$ is a vector of transaction characteristics and ε_i is an i.i.d normally-distributed error term. Here, x_i^o denotes the offer characteristics: car price, baseline interest rate, loan term and minimum down payment. The vector x_i^c denotes car characteristics including the cost of acquiring and reconditioning the car, the mileage, car age and, as a useful proxy for any unobserved quality, the time the car has spent on the lot. Finally, x_i^a denotes applicant characteristics including the applicant's credit category and monthly income, as well as city, month and year dummies.

Before discussing our estimates, several points deserve attention. The first is our use of individual-level data. The use of individual level data to estimate demand, particularly for unique goods such as used cars, is preferable to the use of aggregate data. To take advantage of this, however, we have to address a missing data problem. We observe applicant characteristics for non-purchasers, but not the car and offer they considered. Our solution is imputation. We assign each applicant to one of three credit categories and one of three income categories (the income categories correspond to household annual income of less than \$24,000, \$24,000-30,000, and greater than \$30,000). Then for each non-purchaser, we randomly select a purchaser in the same credit and income category, and in the same city and week, and assume the non-purchaser faced the same

[14] Defaults on existing loans are also lower in tax season than in other months of the year, providing further evidence of a liquidity effect.

car and price.[15] At the end of this section, we verify that our conclusions are not sensitive to the specific imputation strategy. We also report similar results based on aggregated data, where we lose efficiency but do not have to rely explicitly on imputation.

The second point is our decision to model purchasing as an up or down decision made after the consumer arrives on the lot. By considering only the pool of applicants, we neglect the possibility that pricing might affect applications. This concern is mitigated by the fact that financing terms are discussed, and in the case of price negotiated, at the individual level. Financing terms aren't publicly posted and often no price is listed on the car. Therefore, it seems reasonable to model consumers as learning what specific terms apply only after they arrive on the lot and fill in the loan application (and, by that, enter our data).

By modeling the purchase decision as a binary choice, we also downplay the possibility that consumers might choose among multiple cars taking into account all of their prices. Incorporating this into the econometric model might improve the efficiency of our estimates. In our context, however, car choice is much less of an issue than at the car dealerships with which most professional economists are familiar, and the match between applicants and cars is driven substantially by company policy. We return to this point later, and also provide some evidence that price changes do not appear to induce much cross-car substitution.

3.2.2 Identification

From both our perspective and the company's perspective, the central decision variables are car prices and required down payments. To identify their effect on purchasing, we need to understand how they are set and why they vary in the data. The typical concern here is endogeneity — the firm's pricing choices may reflect information about demand that is not available in the data. In our case, we observe the same information as company headquarters so we feel comfortable making the assumption that with sufficient controls decisions made at the company level are exogenous to individual applicants, i.e. uncorrelated with unobservable individual characteristics (the ε_i's).

Minimum down payments indeed are set at the company level. There are separate requirements for each credit category, with some regional adjustment, and these requirements are adjusted pe-

[15] An obvious concern with this imputation is identifying the effect of price changes. Because prices are individually negotiated, it seems plausible that non-purchasers might have faced somewhat higher prices. Even if the difference in offers arises for random exogenous reasons, a straight demand regression would underestimate the effect of price changes. We address this problem, as well as the concern that negotiated prices may incorporate information not available to us as analysts, with the instrumental variables strategy described below.

riodically. Moreover, because minimum payment requirements are set for groups, two identical (or near-identical) applicants can face different down payment requirements due to variation in the characteristics of other applicants in their pool. Our data, therefore, contain three sources of identifying variation in minimum down payments: variation over time, variation across credit categories, and regional variation. In our baseline specification, we include city and category dummies, meaning that we focus on changes over time and on differential changes across categories and cities. We have also run a wide range of alternative specifications, summarized later, where we separately isolate each source of variation in the data.

Identifying the demand response to changes in car prices is more difficult because the actual transaction price is negotiated individually. Individual salespeople start with a "list" price for each car that is set centrally, but may incorporate further information into the negotiation. This additional information creates a possible endogeneity problem. Our solution is to use the centrally set list price as an instrument for the negotiated price.

To do this, we specify an additional equation for the negotiated price:

$$p_i = l_i \lambda_l + x_i' \lambda_x + \nu_i, \tag{2}$$

where l_i is the company list price, x_i includes all of the relevant car, buyer and offer characteristics apart from price, and ν_i is a normally distributed error, potentially correlated with ε_i (thus accounting for the possible endogeneity of price in the demand equation). Our instrumental variable estimates are based on joint maximum likelihood estimation of equations (1) and (2). The t-statistic on λ_l is over 100 indicating that a significant portion of list price changes are passed into negotiated prices.

In the instrumental variables specification, variation in list price identifies the price coefficient in the demand equation. The list price derives from a mechanical formula used to mark prices up over cost. As we control for cost, the relevant variation is in the margin formula. We have three separate sources of identifying variation in this formula. The first is variation over time: we observe one large and one small change in the margin schedule. The second is regional variation, which is substantial. The third arises from the fact that margins are different for different priced cars and the formula is highly discontinuous. Our baseline specification contains city dummies, so it combines the time variation, the non-linearity of the mark-up formula and differential changes across region. As with minimum down payment, we summarize later a wide range of alternative specifications that separately isolate each source of variation.

3.2.3 Demand Estimates

Table 2 reports our demand estimates. The first four specifications vary only in their treatment of car price. The first column contain ordinary probit estimate of the effect of negotiated price on the purchase decision, as described above. Columns (2)-(4) contain estimates from the two-equation model, which accounts for the endogeneity of the negotiated price. Column (2) uses the list price as the "instrument," as described above. Columns (3) and (4) use state dummy variables and dummies for ranges of car cost ("cost buckets"), respectively, in place of list price. As described above, variation in list prices is partly driven by variation across states and across car cost buckets, so one way to think about these columns is that they use only a specific part of the identifying variation in price. We focus on the second column as our preferred specification.

In addition to the offer terms, our demand specification includes detailed buyer and car characteristics, including city, month and year dummies. Because the realized interest rate can depend on the size of the down payment, we do not include it as part of the offer. Instead, we include the interest rate that the buyer would have paid if they made the minimum down payment. As an empirical matter, the differences are relatively small and using the realized interest rate has no effect on the other coefficients.

The final two columns of Table 2 report results using an aggregated demand specification that does not require us to impute the preferred car and price faced by non-purchasers. For these estimates, we aggregate the data to city-week-credit category-income category "cells," and estimate the number of purchases in each cell as a function of the number of applicants in the cell, the average applicant characteristics, the average car characteristics, price paid and loan term of the purchasers, and the baseline interest rate and minimum down payment. Letting Q_j denote the number of purchases in cell j, we specify a log-linear model:

$$\ln Q_j = \beta_A \ln A_j + \overline{x}_j \beta_x + u_j, \tag{3}$$

where A_j is the number of applicants in cell j, and $\overline{x}_j$ are the average applicant, car and offer characteristics.

We report ordinary least squares estimates in column (5) and instrumental variables estimates in column (6), where the we use the average list price in the cell as an instrument for the average transaction price. Again we focus on the instrumental variable estimates. To compare the individual and cell-level estimates, observe that the individual-level numbers in Table 2 are marginal effects on the probability of sale, while the cell-level numbers correspond to percent changes in the probability of sale. As thirty-four percent of applicants purchase, one can make a rough comparison between the two specifications by dividing the cell-level numbers by three.

Our main interest is the effect of car price and minimum down payment on purchasing decisions. Changes in these offer terms are not identical from a buyer's perspective. The down payment is made immediately as a lump-sum, while changes in car price can be spread over time (as we will see below, changes in car price translate almost one-for-one into larger loans rather than larger down payments). As a result, an applicant who is relatively impatient or liquidity constrained should be more sensitive to changes in the down payment, holding the loan amount fixed. Moreover, the high probability of default means that a purchaser often will not bear the full cost of a price increase. This should also reduce the sensitivity of demand to car price relative to the down payment requirement. At the same time, a higher down payment holding car price fixed implies a smaller loan, weakening the effect of a change in the required down payment.

Despite these various forces, it is still straightforward to look for evidence of liquidity constraints. If applicants are not liquidity constrained, they care about the present value of future payments. For a purchaser who agrees to a price p, makes a down payment D, and borrows the balance at a monthly interest rate r over a T-month term, the expected payment is

$$\mathbb{E}[Payment] = D + (p - D) \cdot \frac{\sum_{t=1}^{T} (1+s)^{-t} S_t}{\sum_{t=1}^{T} (1+r)^{-t}} = D + (p - D)\,\phi, \tag{4}$$

where s is the purchaser's subjective monthly discount rate and S_t is the probability that the loan will not be in default before the end of month t.

The value ϕ represents the expected present value of payment that will be made for each dollar that is borrowed. It is exactly analogous to the repayment to loan ratio introduced in the end of Section 2, the difference being that the relevant rate of discount is the customer's rather than the firm's. To construct a plausible estimate of ϕ, therefore, we again calculate the average repayment to loan ratio for uncensored loans in the data using a broader range of discount rates.[16] Using this approach, an applicant who is not liquidity constrained, plans to make the minimum down payment and has an annual discount rate of 5 percent should view a 100 dollar increase in the required down payment as equivalent to a 30 dollar increase in the car price. If the agent is more impatient, a down payment increase matters more. For annual discount rates of 10, 20 and 50 percent, the agent views a 100 dollar increase in the down payment as equivalent to 38, 55 and 108 dollar increases in the car price.

Our demand estimates, however, imply that applicants are far more sensitive to minimum down payment requirements than these calculations would suggest. We estimate that a 100 dollar increase in the minimum down payment reduces the probability that an applicant will purchase by 0.0301, while a 100 dollar increase in the car price reduces the purchase probability by only 0.0010. That is, a 100 dollar increase in the minimum down payment has the same effect as a 2,884 dollar increase in car price. This can still be explained in the absence of liquidity constraints, but it requires a much higher annual discount rate 1,415 percent. The conclusions are similar using the specification – in column (4) – that generates the highest effect of price implying an effect of 0.0032, which translates to an equivalent price increase of 956 dollars and an annual discount rate of 474 percent.

These calculations focus on the relative sensitivity of demand to car price and minimum down payment. The absolute sensitivity to minimum down payment is itself large. Our estimate implies that a 100 dollar increase in the minimum down payment reduces sales by nine percent. This number appears to be consistent with the company's own view of pricing responsiveness, but it is still notable given that subsequent monthly payments are on average 400 dollars. This, too, suggests that applicants face a high cost of coming up with extra cash.

Table 2 also reports estimates of how buyer and car characteristics affect demand. As might be expected, conditional on price, cars that cost more, have lower mileage and have spent less time on the lot are more desirable. Similarly, applicants with higher incomes and with bank accounts are more likely to purchase. Both effects make particular sense from a liquidity standpoint; these applicants are likely to have greater resources to make a down payment. A somewhat surprising result is that a buyer's credit category does not systematically influence the probability of purchase.

[16] One potentially could be more sophisticated here and account for changes in loan size affecting the default process, or differences between marginal applicants and the broader distribution of buyers. We think (and have checked with some simple back-of-the-envelope calculations) that our approach is a good enough approximation for the task at hand, however.

One possible explanation is that although lower risk buyers may have greater resources and access to immediate cash, they also have better alternatives. Our finding that applicants who own their own homes are less likely to buy than renters is consistent with this hypothesis.

3.2.4 Robustness of Demand Estimates

Having presented our main results, we now return to the specific aspects of the demand modeling discussed above, and describe a range of robustness checks and alternative specifications. The main point we want to convey is that our conclusions aren't very sensitive to the modeling specifics.

Imputation of Negotiated Prices

For our baseline individual-level estimates, we impute preferred cars and negotiated prices by matching each non-purchaser with a randomly selected purchaser in the same city, week, credit and income category. We have experimented with other approaches, for example by adjusting the imputed prices using additional buyer characteristics or by inferring these prices from our estimated equation (2). We have also tried a variety of cell constructions for our aggregate-level estimates, for example pooling income groups or using bi-weekly rather than weekly cells.

We report the results from a number of these specifications in an Appendix. The estimated effect of changes in the minimum down payment is remarkably stable. The estimated effect of pricing changes is a bit less stable, as might be expected given that we are experimenting with different ways to fill in missing prices, but the estimated effect is always very small. Our conclusion that the purchase decision is much more sensitive to minimum down payment than to price holds uniformly across all the specifications we have tried.

Car Choice

Our baseline specification views purchasing as a binary choice. This is not, by itself, a problem. Even if consumers faced a large set of cars and observed the price of each, the conditions we use for individual-level estimation – that purchasers preferred to buy their chosen car over not buying, and that non-purchasers preferred not to buy over their imputed preferred car – would still yield consistent estimates. Using the entire choice set could increase efficiency, but with a large cost in terms of complexity.

We want to emphasize, however, that our decision to model the purchasing decision as an up or down decision is also motivated by the particular context, and is consistent with the way the company thinks about its business. Applicants are matched with cars primarily through company headquarters policy. Once an application is filled and a credit category is assigned, an automated process combines the associated offer terms with the inventory in the lot to guide the salesperson as to which car to show the applicant.

These points aside, we can still explore whether price changes give rise to measurable substitution across cars. One specific hypothesis is that when margins increase, customers might substitute to cheaper cars. To check this, we focus on the single largest margin schedule change in our data, when the company moved from a uniform margin across all cars to a graduated schedule with higher margins for higher-cost cars. Figure 3 shows the margin schedules before and after the change. The smoothed histograms in the Figure show the distribution of costs of the purchased cars in the month before and after the change. If there was substantial substitution, we would expect the distribution after the change to be shifted toward cheaper cars. It is not. The distributions are right on top of one another, and we cannot reject the hypothesis that they are the same ($p - value$ of 0.364 using a Kolmogorov-Smirnov test). This, combined with the institutional details of the purchasing process, suggests to us that we are not missing much by abstracting from car choice.

Isolating Sources of Identification

Finally, we noted above that we have multiple sources of identifying variation in both the minimum down payments faced by applicants and the company's margin schedule. Our baseline estimates pool several sources of variation. We have also performed a wide range of robustness checks where we separately isolate each source of variation in the data, for instance by using only data from short time windows around changes in the minimum down payment or margin schedules, by including only applicants whose credit scores place them on the margin between two adjacent credit categories (leading them to be similar on observables yet have discretely different minimum down payments), or by including only applicants matched with cars that have costs close to jump points in the margin schedule.

The results from many of these alternative specifications are reported in the Appendix. Similar to the imputation checks discussed above, the estimated coefficient on minimum down payment is remarkably stable (ranging from -0.0238 to -0.0304). The estimated coefficient on price varies a bit more, in particular when we dramatically reduce the sample size, but again our conclusion about the relative demand sensitivities is highly robust across specifications. In particular, across the 21 specifications considered in the Appendix, the subjective discount rate implied by the relative sensitivities is never lower than 200 percent and is generally much higher.

4 Asymmetric Information and Liquidity Constraints: Theory

In this section, we develop a simple credit market model along the lines of Jaffee and Russell (1976), and show how moral hazard and adverse selection can lead to credit constraints being imposed in equilibrium. We also explain the effect of interest rate caps and how risk-based pricing mitigates adverse selection. The theory presented in this section will guide our empirical analysis in the next section. Because the basic ideas are familiar from the general theory of credit and insurance markets, we confine ourselves to a largely graphical analysis.

We consider a two-period model with a large number of firms and consumers. We assume that firms are integrated and the sales and finance market is perfectly competitive. Neither assumption is essential for the points we make. To begin, we also assume that customers are ex ante homogenous although we will relax this below.

In the first period, each consumer decides whether or not to buy a car, and if so, how large a loan to take. In the second period, the consumer decides whether or not to repay the loan. For expositional purposes, it is useful to think of a contract between a consumer and a firm as specifying a first period down payment D, equal to the price p minus the loan size L, and a second period payment M. The second period payment will equal to the loan size L times the contractual interest rate R. The borrower may or may not repay the loan. We make the natural assumption that the probability of repayment $\lambda(M)$ is decreasing in loan liability M.

We pause briefly to observe that there are many ways to motivate the assumption that larger loan liability increases the probability of default. One is that the customer's second period income or opportunity cost of paying back the loan is stochastic, so she may default on a larger payment. Another possibility is that the value of the car evolves stochastically and the customer opts to default if the value of paying falls below the loan liability. In our setting, there is little doubt that defaults arise from a combination of circumstances and individual decisions. We will use the term "moral hazard" to refer to the fact that default increases with loan size even though circumstance, rather than individual decisions alone, may be responsible for some of the relationship. As will be apparent, the equilibrium predictions of the model, and its ability to explain credit limits, simply do not depend on the specific breakdown.

Let $U(D, M)$ denote the expected utility of a consumer who agrees to a contract (D, M). We assume that U is decreasing in both arguments. Let $\Pi(D, M)$ denote the firm's expected profit from the same contract. We assume that Π is increasing in D because holding fixed the second period payment, a larger down payment is clearly advantageous for the firm. Firm profits, however, need not be increasing in M because a large loan size increases the probability of default. We assume instead that Π first increases and then decreases in M. We assess this below in our empirical work.

Figure 4(a) depicts the iso-profit line $\Pi(D, M) = 0$, where c denotes the firm's cost of acquiring the car so that $\Pi(c, 0)$ is on the zero-profit curve. An immediate observation is that moral hazard may imply loan limits. As we have illustrated the situation, no firm would write a contract that involves a down payment below d' regardless of the required second period payment. Therefore, given a car price $p \geq c$, loans will certainly be capped at $p - d'$.

The competitive outcome in this setting is the contract that maximizes customer utility subject to firms making non-negative profits. This contract is denoted by the point E in Figure 4(a). An interesting question is how this outcome might arise in practice. One possibility is that firms allow customers to choose any point on the zero-profit locus (i.e. the curve AEB), with customers choosing the optimal point E. The interest rate on small changes in loan size is described by the tangents of AEB, so this outcome involves firms pricing cars at cost, requiring a minimum down payment d', and charging lower interest rates to customers who make larger down payments and take smaller loans.

We noted earlier that interest rate caps appear to constrain subprime lenders. In the current setting, interest rate caps need not affect the competitive allocation, but they can have a dramatic effect on its implementation. If the seller sets $p = c$ and offers the competitive contract E, the contractual interest rate is given by the slope of the line between A and E. The seller can also offer E by charging $p > c$ and lowering the interest rate. Such an offer is depicted in Figure 4(b) by the line EF. This offer necessarily leads to a higher minimum payment requirement, equal to d. If customers were allowed to borrow more than $p - d$ at the capped rate, they would, and firms would lose money.

So far we have seen that even if consumers are homogeneous and lending is competitive, moral hazard can give rise to minimum down payments and interest rate caps can tighten these requirements. We now show that consumer heterogeneity can lead to still tighter restrictions. To introduce heterogeneity in the simplest way, suppose there are two types of customers, low and high risks.

Denote their utility functions by $U_L(D, M)$ and $U_H(D, M)$. We assume that high risk consumers are more likely to default for any given loan size, and because of this have a greater desire to backload payments.

This is depicted in Figure 4(c). Following the discussion, we have drawn the utility iso-quants so that the high-risk customers have a higher marginal rate of substitution between future and present payments — their iso-quants are steeper than those of low risk customers. The result is adverse selection: given a set of financing choices, high-risk customers select smaller down payments and larger loans. In the figure, we have drawn the offer curve as the set of contracts (D, M) that would yield zero profit to a firm if the contract were to attract a representative mix of high and low risks. These contracts are not, however, offered in equilibrium.

Figure 4(d) depicts a separating equilibrium with heterogenous customers.[17] The two iso-profit lines depict the locus of points that give firms zero profits assuming their customers are either all low-risk or all high-risk. The zero profit curve for high risk types lies to the right of that for low risk types because fixing a loan liability M, a high risk consumer will be less likely to repay. Just as no firm would contract with a low risk consumer without requiring at least a down payment d', no firm would transact with a high risk customer without requiring at least a down payment d''.

In the absence of credit scoring, firms must offer the same options to low and high risk types, who self-select into different contracts in equilibrium. High risk customers get their preferred allocation subject to the constraint that firms make zero profits. This point is denoted by H. Low risk customers get their preferred allocation subject to the constraint that firms break even and also that high risk customers prefer the allocation H to the contract intended for low-risk types. We denote this point by G. Note that in equilibrium, self-selection leads to a negative correlation between down payment and default rate. We examine this prediction below in our empirical analysis.

There are many financing offers that support the separating allocation, but a natural possibility in the presence of interest rate caps is that firms price cars at $p > c$ and require a down payment of at least d_H. The minimum down payment allows a customer to borrow at the capped interest rate; a customer who makes a larger down payment receives a lower interest rate, so that both G and H

[17] Depending on the parameters, a separating equilibrium may not always exist. One possibility is that the terms required by firms are simply too onerous for consumers. Another possibility is that the equilibrium can be upset by a firm that offers a profitable pooling contract. The intuition for the latter is the same as in the well-studied insurance framework of Rothschild and Stiglitz (1976).

are possible. Note that the possibility of adverse selection substantially constrains the equilibrium loan size for low-risk buyers. Rather than being able to purchase with only a down payment d, low-risk customers must make down payments of $d_L > d$ to distinguish themselves from high-risk customers.

The development of credit scoring has important consequences when customers are heterogenous in their underlying default risk. Suppose that firms can distinguish between risk types and price accordingly. The resulting allocation for high risks is the same, but low risks receive their optimal allocation E, i.e. they are allowed to take larger loans. As we have drawn it, low risk customers actually take larger loans than high risk customers once credit scoring is in effect, in direct contrast to self-selection that occurs given a common set of choices. More robustly, in the presence of credit scoring, the correlation between loan size and default rate will be lower if one doesn't condition on risk group than if one does. We return to this point in our empirical analysis.

4.1 Moral hazard and adverse selection: discussion of terminology

We conclude this section with a brief comment on our use of the terms adverse selection and moral hazard. As we have already noted, we use adverse selection to refer to a situation where high risk individuals self-select into larger loans. This could happen because forward-looking consumers who anticipate a high likelihood of default shy away from larger down payments, or because consumers who are illiquid today and unable to make sizeable down payments are also likely to be illiquid later and have trouble with repayment. Our goal in this paper is not to distinguish between these underlying behavioral stories; and indeed from the firm's profitability perspective the distinction is not important. Rather our goal is to explain how self-selection can lead to equilibrium liquidity constraints, to demonstrate that self-selection is operative in subprime lending, and to quantify its importance.

A similar point applies to the term moral hazard. We have used it to describe a situation where an increase in loan liability increases the likelihood of default. Again, many underlying behavioral models are consistent with such a relationship — ranging from models where individuals have a great deal of conscious control over their environment and default decisions have a sizeable strategic component, to models where individuals have relatively little agency and default is due largely to uncontrollable circumstances. From the standpoint of the market equilibrium model described above, the distinction is unimportant. What the model shows is that regardless of the exact channel, a causal relationship between loan size and default can lead to equilibrium liquidity constraints. Our goal in the next section, therefore, is to demonstrate that the relationship is operative in subprime lending and quantify its importance. That being said, we view parsing out the path to financial distress as an important avenue for future work, and return to this point at the end of the paper.

5 Asymmetric Information and Liquidity Constraints: Evidence

The preceding section developed a simple equilibrium model of the credit market we study. The model relies on certain assumptions regarding consumer behavior, specifically moral hazard and adverse selection, and generates several qualitative predictions. Our goal in this section is to assess the empirical validity of the assumptions and predictions of the theory and to quantify the relative magnitude of various forces. To this end, we document the following:

A. Moral hazard: For a given individual, the probability of default increases substantially with loan size. As a result, expected loan payments are not monotone in loan size.

B. Adverse selection: Individuals who are more likely to default demand larger loans. This effect operates through both predictable default risk and idiosyncratic default risk that is known to the individuals but cannot be predicted using characteristics observed by the firm.

C. Lender response: Loan sizes are capped using minimum down requirements, which are tighter for individuals with higher default risk. Interest rates are more favorable for individuals who are observably lower risks or who signal their lower risk by making a larger down payment.

D. Effect of credit scoring: Risk-based pricing mitigates some of the adverse selection problem described above, and reduces the positive correlation between loan size and default.

We address these points somewhat out of order. We start by estimating a simple down payment (equivalently, loan demand) model for a customer who has decided to purchase a car. Our estimates indicate a force toward adverse selection: high risk buyers systematically prefer to make smaller down payments. Having established this pattern, we specify a model of default behavior and explain how such a model, in conjunction with the loan demand model, can be used to separately quantify the effects of moral hazard and adverse selection. Such a separation is not immediate because the main empirical implication of moral hazard and adverse selection is the same: a positive relationship between loan size and default. Nevertheless, we are able to show that moral hazard is responsible for around sixty percent of the (large) within credit category correlation between loan size and default. Finally, we document the lender responses described above and show that adverse selection would be substantially worse if risk-based pricing did not force the highest risk buyers to make the largest down payments.

Before we begin the more formal analysis, we present two descriptive figures that motivate the direction we explore in this section. Figure 5(a) plots loan sizes in the data against repayment probabilities for different risk categories, smoothing the raw data using local linear regression. For each group of buyers, the probability of repayment falls steadily with loan size. The Figure also illustrates the strong correlation between default and buyers' credit scores; buyers assigned to high-risk credit categories are substantially more likely to default.

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 5(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 5(b) suggests, this pattern is fairly uniform across different risk groups. Indeed, this descriptive evidence is only suggestive, as it doesn't control for many confounding factors. These are addressed by the more formal analysis in the rest of this section.

5.1 Adverse Selection and the Demand for Loans

We begin by studying the financing decisions of car buyers. Consider a buyer who faces a minimum down payment d_i. By making a larger down payment, the buyer reduces her loan size and, as we discuss below, may receive a lower interest rate. So she faces a trade-off between a lower immediate payment and higher future payments. We want to understand how this trade-off is resolved depending on the buyer's risk characteristics, as well as her liquidity characteristics, the value of the car, and so on.

Because many buyers (43 percent) make exactly the minimum down payment, we specify a tobit model of the down payment decision, where

$$D_i = \begin{cases} D_i^* = x_i'\gamma + \varepsilon_i & \text{if } D_i^* \geq d_i \\ d_i & \text{if } D_i^* < d_i \end{cases} . \tag{5}$$

This model equivalently characterizes the choice of loan size $L_i = \min\{p_i - d_i, p_i - (x_i'\gamma + \varepsilon_i)\}$, so we will speak of loan demand and down payment choice interchangeably.

For obvious reasons, we observe only the down payments, and later loan repayments, of purchasers. For our present purpose, however, we don't think that selection is a major concern. First, we have extremely detailed individual data and sufficient observations to include controls for city, year and month of purchase, so we can control for most factors that affect both purchasing and borrowing behavior. Second, we focus primarily on the effect of price rather than of minimum down payment. In light of the extremely low sensitivity of purchasing to price, as documented in Section 3, we believe that the main effect of price is on the financing terms of purchasers rather than their composition.[18]

In estimating the down payment model, we include as controls all of the variables in our model of purchasing — car characteristics, individual characteristics including credit category and controls for city, year and month of purchase. We consider two specifications, the straight tobit model and an alternative where we estimate the tobit model (5) jointly with negotiated price, equation (2). In the latter specification, the coefficient on car price is identified through variation in the company margin schedule, so we account for the fact that unobserved determinants of the negotiated price may be correlated with the unobservable component of the down payment decision.

Table 3 presents our estimates, which are consistent with the importance of liquidity constraints. Notably, a price increase of 100 dollars has a relatively small effect on the desired down payment (between 5 to 18 dollars depending on the specification). The primary response to a higher price is to take a larger loan. A higher minimum down payment and a higher starting interest rate are also associated with larger desired down payments. The monthly dummy for tax season is also large and highly significant. As noted earlier, desired down payments are about 300 dollars higher in February than in the other eleven months.

The most striking results in Table 3 concern the relationship between desired down payments

[18] Further assurance about selection concerns comes from our ongoing work (Einav, Jenkins, and Levin, 2007), where in the process of studying optimal pricing we model purchasing, loan demand, and default behavior jointly, and obtain results similar to what we report here.

and observable risk characteristics. High risk buyers systematically prefer to make smaller down payments. Our estimates imply that, all else equal, the ideal down payment of a buyer in the worst credit category is 28 percent less than that of a buyer in the best credit category. The same relationship holds to some extent within credit categories. For instance, among the buyers with a given credit category, those with lower raw credit scores choose larger loans. There is also a similar pattern across cars. Buyers who have selected newer and more valuable cars, which are presumably less likely to break down, make larger down payments. These findings all indicate a tendency for adverse selection, our point B above. A key point we address below is that these estimates need not translate directly into realized adverse selection because, just as in the theory, observably high risk buyers face higher minimum down requirements.

5.2 Identifying Moral Hazard and Adverse Selection from Default Behavior

We now turn our attention to loan repayment, or default behavior. We start by specifying an empirical model of default. For this purpose, we use a Cox proportional hazard model. The model is convenient both because it allows for a flexible default pattern over time and because it allows us to work with our full sample of loans despite some observations being censored. We write the probability of default at t given that the loan is still active as

$$h(t|L_i, x_i) = \exp\left(L_i\delta_L + x_i'\delta_x\right) h_0(t). \tag{6}$$

In the usual formulation, t is time, but here we specify t as the fraction of the loan payments made, extending from $t = 0$ when the loan initiates to $t = 1$ when the loan is fully repaid. This transformation provides a simple way to account for the fact that the term of the loan varies somewhat across borrowers in our sample. The remainder of the model is straightforward: L_i is the size of the loan, x_i is a vector of individual and car characteristics, as well as financing terms such as interest rate and loan term, and the baseline hazard $h_0(t)$ is an arbitrary function. The model's main assumption is that changes in covariates shift the hazard rate proportionally without otherwise affecting the pattern of default. This is a strong assumption, but in our case appears to be fairly innocuous; we have experimented extensively with alternative specifications, always with similar conclusions.[19]

[19] This might be expected given the striking similarity in the timing of default patterns across different risk groups, as shown in Figure 1(a). For instance, we get similar results using probit or logit models of default, estimated either

We use the default model to address the central empirical implication of the theory, the relationship between default and loan size. Both moral hazard and adverse selection imply a positive cross-sectional correlation between these variables conditional on priced characteristics. Moral hazard yields a correlation because an individual buyer's probability of default increases with loan size. Adverse selection does so because buyers who are more likely to default take larger loans. We saw earlier that it is precisely the positive correlation between loan size and default that leads to credit limits. For this reason, identifying a correlation is itself a useful goal. Ideally, however, one wants to disentangle the effects of moral hazard and adverse selection, in part because the institutions for overcoming them are so different. Credit scoring and risk-based pricing can mitigate adverse selection, while tools such as improved collection or repayment incentives are needed to address moral hazard.

How does one separately identify moral hazard and adverse selection in our setting? The most obvious specification of the default model is to view the probability of default as a function of loan size and other observable default drivers, such as the interest rate, the loan term and individual and car characteristics. The estimated coefficient on loan size will then pool the causal effect of having a larger loan on the probability of default (i.e. moral hazard) with the correlation induced by observably equivalent borrowers, who nevertheless face different risks, taking different loans (i.e. adverse selection). To isolate moral hazard, we need to fully control for factors that affect both loan size and default. In particular we want to recognize that despite our rich individual controls buyers may have private information at the time of purchase that affects both their down payment and their later default behavior.[20]

Our solution is to model loan size jointly with default. The down payment model above can be re-written as a model of loan size, where:

$$L_i = \begin{cases} L_i^* = p_i - x_i'\gamma - \varepsilon_i & \text{if } L_i^* \leq p_i - d_i \\ p_i - d_i & \text{if } L_i^* > p_i - d_i \end{cases} . \tag{7}$$

Identification of the two equations (6) and (7) comes from an exclusion restriction. The key as-

using the uncensored loans in our data or using all loans with the dependent variable being a dummy equal to one if the loan defaults in the first quarter of the loan term. We have also experimented with default models where the cumulative distribution of defaults, rather than the hazard, is separable in duration and covariates, and with using calendar time rather than the fraction of payments as the dependent variable.

[20] As an indication of how rich our set of observable characteristics is, an ordinary least squares regression of loan size on our standard covariates returns an R^2 of 0.92.

sumption is that conditional on all relevant observables, repayment behavior depends on the size of a buyer's loan, but not directly on the pricing mark-up or the minimum down payment. These matter only insofar as they influence the size of the buyer's loan.

To implement estimation, we use a two-stage "control variable" approach in which we include the estimated residual from the down payment (equivalently, loan size) model as a control variable in estimating the default model. We define the down payment residual for individual i as:

$$\xi_i = \begin{cases} D_i - x_i'\hat{\gamma} & \text{if } D_i > d_i \\ \mathbb{E}\left[\varepsilon_i \mid \varepsilon_i \leq d_i - x_i'\hat{\gamma}\right] & \text{if } D_i = d_i \end{cases} . \tag{8}$$

For buyers who pay more than the required minimum, we observe the residual $D_i - x_i'\hat{\gamma}$ exactly. For buyers who pay the minimum, we have an upper bound and take the conditional expectation. The constructed residual ξ_i contains the buyer's private information as pertains to her choice of down payment and hence loan size. When we include it in estimating the default equation, as well as the other observed covariates, the remaining variation in loan size is due entirely to variation in the margin on the car and the minimum down payment.[21]

Note that the logic of the approach is simply to isolate variation in loan size that is independent of default risk conditional on our controls. While the idea is straightforward, we can think of several potential concerns. One is that the variation in car price may not translate into sufficient variation in loan size. This concern is mitigated by our earlier estimate that changes in price translate almost entirely into changes in loan size. Moreover, unlike in Section 3, the analysis here focuses only on purchasers so there is no issue of missing data. This provides much better variation in car price, allowing us to exploit more efficiently the observed changes in the pricing schedule and the non-linearities in the margin formula. We return to this point below in discussing the robustness of the estimates.

A second concern is that the negotiated price may incorporate information about default risk that we do not observe directly and that is not reflected in the choice of down payment. Presumably, this would be information available to the salesperson but not to headquarters or the borrower herself. Below we discuss versions of the estimation where we estimate the negotiated price jointly

[21] Note that the rationale for using a two-stage estimation approach, rather than estimating equations (6) and (7) jointly (or equations (6), (7) and (2) jointly) is mainly convenience. Using the control variable approach permits a standard partial likelihood method to estimate the Cox model, rather than having to develop an estimation strategy for a semi-parametric simultaneous equations model.

with the down payment equation and include the estimated price residual as an additional control in estimating default. In that alternative, the remaining variation in loan payment is solely due to variation in the company's centrally set margin schedule and the minimum down payment, the exogeneity of which was discussed in Section 3. The qualitative results are similar, suggesting that this type of loan size endogeneity is not empirically important.

A third concern about our empirical approach is that binding minimum down payments will prevent us from accurately recovering borrower's private information. We report later results from several specifications that indicate that this is not a problem. We also defer to the end of this section a discussion of car selection, which might confound our analysis if substitution across cars was important.

Our empirical strategy also provides a simple test for whether there is adverse selection on characteristics about which the parties are asymmetrically informed at the time of purchase. The argument is the following. Conditional on loan size, a borrower's down payment is sunk; it should not directly affect default. But it should also reflect all the buyer's relevant information about default at the time of purchase. Therefore, a negative correlation between the down payment residual and the probability of default, conditional on loan size and observed characteristics, indicates that buyers who made higher down payments for unobservable reasons are also those who are more likely to default for unobservable reasons. This is precisely the notion of adverse selection arising from asymmetric information about default risk.

Moreover, a straightforward extension of our approach allows us not just to test for adverse selection but to quantify the degree of self-selection on different dimensions. To do this, we use the proportional hazard model to estimate correlations between loan size and default rate conditioning on increasingly sparse subsets of individual characteristics. When we include the full set of controls, the coefficient on loan size gives the pure moral hazard effect of loan size on default. Omitting the down payment residual ξ from the set of controls, the estimated coefficient on loan size pools the moral hazard effect and the correlation between loan size and default that is driven by observably identical buyers self-selecting into different loan sizes. By subtracting off the estimated moral hazard effect, we obtain an estimate of the latter.

This idea extends further. If we omit observed buyer characteristics that are not directly priced, the estimated loan coefficient pools in the correlation that is driven by buyers who are observably different but face the same prices selecting different loan sizes. This allows us to assess the amount of adverse selection that is present under the existing pricing scheme. Finally, by dropping even controls for credit category, we can assess a key prediction of the theoretical model, that the use of risk-based pricing lowers the cross-sectional correlation between loan size and default.

Some readers may find it useful to relate our approach to the empirical literature on insurance markets, which recently has focused on similar issues. This literature uses the observed correlation between insurance coverage and insurance claims (the insurance analogues of loan size and default) to provide evidence that a market is or isn't characterized by some combination of moral hazard and adverse selection (e.g., Chiappori and Salanie, 2000). Often, however, the argument is made that the two forces cannot be separately identified. The reason we can separate moral hazard from adverse selection in our setting is that we have two sources of variation in the loan size. The first, which accounts for adverse selection, is the buyer's endogenous choice of down payment. The second, which accounts for moral hazard, is exogenous pricing variation. In the insurance setting, the corresponding variation would come from observing both endogenous choices of coverage and exogenous changes in the menu of coverage options.

5.3 Estimates of Moral Hazard and Adverse Selection

We report our estimates of the proportional hazard default model in Table 4. The first column reports the richest specification, including the full set of observable characteristics and the estimated down payment residual as controls. This specification isolates the effect of moral hazard in the coefficient on loan size and also contains our test for adverse selection arising from asymmetric information. The other columns report simpler specifications as discussed above. All display a large and significant positive relationship between loan size and default.

Our estimate of the causal effect of loan size on default indicates that a 1,000 dollar larger loan leads to a 17 percent higher default rate. The estimated model implies that the expected revenue from loan payments does not increase monotonically in loan size. To capture this, we fix all variables other than loan size and credit category at their sample means and use the default model to calculate expected loan revenue as a function of loan size for each credit category. We plot the relationship in Figure 6(a). Depending on the risk group, expected loan payments peak at loan sizes of between ten and twelve thousand dollars. Though it may not be immediately apparent, this figure (rotated 90 degrees) is essentially the empirical analogue to the iso-profit lines for lending to different risk groups that are depicted in Figure 4(d).[22] The main point in both figures is that marginal dollars loaned eventually become unprofitable and this occurs sooner for high risk borrowers.

While we focus primarily on loan size, we also find interesting effects arising from variation in the other financing terms. Both loans with higher interest rates and those with longer terms are more likely to default. The former is easy to understand. A higher interest rate implies that for a given loan size, monthly payments are higher. Consistent with moral hazard, we estimate that a one point increase in the annual interest rate increases default by 2.2 percent, a substantial effect. A longer loan term need not have an obvious effect on default. On the one hand, a longer loan term lowers the size of each monthly payment. On the other hand, it stretches the repayment period, potentially allowing more default-generating events to happen within the duration of the loan. Our estimates suggest that the latter is the more relevant. In particular, a one month increase in loan term increases the default rate by 1.5 percent.

Observable buyer characteristics also significantly affect default rates. Credit categories in particular have remarkable power in predicting default. Buyers classified as high risks are more than twice more likely to default than buyers classified as low risks, with medium risk buyers in between. Within credit category, buyers who have higher incomes, have bank accounts, do not live with their parents and have higher raw credit scores are all less likely to default. As we discuss below, however, the fact that these characteristics predict default and are not directly priced does not necessarily imply a serious adverse selection problem in financing choices. Indeed, self-selection on some observed characteristics is advantageous. For example, buyers who live with their parents tend to make larger down payments but have a greater likelihood of default later on.

The last variable of interest is the down payment residual. As discussed above, our test for adverse selection due to asymmetric information is based on the conditional correlation between this constructed variable and the default rate. The estimated correlation is negative and highly significant, consistent with the presence of adverse selection.

The second column of Table 4 drops the down payment residual, our control for privately known borrower characteristics. Without this control, the coefficient on loan size combines both the moral hazard effect of loan size on default and the cross-sectional correlation due to borrowers who are

[22] Note that Figure 6(a) does not factor in recoveries following default. Doing so raises the curve on the y-axis and shifts its peak slightly to the right, but does not change the basic shape or our qualitative point.

at higher risk of default for privately known reasons taking larger loans. Here we find that a 1,000 dollar larger loan is associated with a 27 percent higher rate of default. We estimated that 17 percentage points were due to moral hazard, leaving adverse selection on unobservables to explain the other ten. Roughly speaking, this implies that moral hazard is nearly twice as important from the lender's perspective than ex ante asymmetric information about default risk.

In the third column of Table 4, we omit the individual characteristics that are not directly priced, so that the coefficient on loan size pools the moral hazard effect with self-selection on both unobserved and observed but unpriced individual characteristics. Our estimate of the relationship between loan size and default differs minimally from the prior column. This indicates that to the extent that financing choices within credit category are characterized by adverse selection, the effect is almost totally due to selection on unobservables. It also suggests that credit categories, despite being coarse indicator of individual risk, nevertheless capture much of the predictable variation in default risk, or at least much of the predictable correlation between loan demand and default risk.

The final column of Table 4 does not control even for credit category so that the coefficient on loan size includes the correlation across credit categories as well as within credit categories. This change leads to a significant decline in the coefficient on loan size: the unconditional effect of loan size on default is 21 percent, compared to 25 percent when credit categories are included. This is evidence of point D from the beginning of this section: risk-based pricing forces riskier individuals, who given the same options would make smaller down payments, to pay more down, mitigating the potential for adverse selection. We explore this point in detail next.

5.4 Minimum Down Payments and Risk-Based Pricing

Our theoretical model emphasized several equilibrium responses to moral hazard and adverse selection. One was minimum down payment requirements, meaning that sufficiently risky buyers would not be allowed to finance their entire car purchase. This pattern is clear in our data. Even buyers with the highest possible credit category face a positive down payment requirement. And risky buyers can face minimum down payments on the order of 1,500 to 2,000 dollars — i.e., 25-30 percent of the cost of the car.

The remaining equilibrium predictions concern risk-based pricing. Risk-based pricing can take two forms: better financing terms for observably lower risks and better financing terms for buyers who effectively signal their lower risk. We observe both in our data. To display them graphically,

Figure 6(b) plots empirical "offer curves" at a particular set of dealerships during a particular period in our data (the choice is somewhat arbitrary). The horizontal axis represents the down payment and the vertical axis the loan liability, that is loan principal plus future interest payments. There are four curves, corresponding to different buyer categories and different car prices. As the picture shows, interest rates decrease with the down payment – each offer curve is convex rather than linear. The decrease, however, is fairly small and made smaller in practice because few buyers make down payments far above the minimum. The more substantive form of risk-based pricing is therefore differences in minimum payments across categories. In Figure 6(b) for instance, buyers classified as high risk cannot put down less than 1,400 dollars, while buyers classified as low risk can put down as little as 400.

We have already suggested that risk-based minimum payments play a substantial role in mitigating adverse selection in financing choices, exactly as predicted by the theory. To crystallize this point, we re-estimated both default rates and desired loan sizes as a function of credit category, controlling for car and contract characteristics but omitting the remaining buyer characteristics such as income and age. These estimates imply that for an average car priced at an average price, buyers classified as the best risk category desire to make a 854 dollar down payment while buyers classified as the worst credit category desire to make a 685 dollar down payment, a difference of 179 dollars (25 percent). Moreover, conditional on both groups making an *average* down payment, 24 percent of the buyers in the lowest risk category will default compared to 64 percent of the buyers in the highest credit category.

Figure 7(a) plots the results of these calculations, with the size of each dot representing the frequency of each credit category in the borrower population. What the figure indicates is a strong propensity for risky borrowers to self-select into smaller down payments. In the model, one effect of risk-based minimum payments is to prevent this, which is precisely what we observe in the data. In particular, Figure 7(b) plots the same default rates against the minimum of the desired loan size and the average loan limit for each category. As the picture makes clear, high risk buyers are heavily constrained by loan caps relative to low risk buyers. Indeed, once risk-based loan caps are factored in, the overall correlation between loan size and default is negative (the picture makes it clear that the relationship is in fact non-monotone). Consistent with the model, and our results above, the effective use of credit scoring forces riskier buyers to make larger up-front payments, mitigating adverse selection.

We have focused on the role that risk-based pricing plays in financing decisions, but it plays another role as well. Because purchase decisions are sensitive to required down payments, raising the minimum down payment screens out some applicants. To the extent that these applicants are ex post more likely to default, the effect of risk-based pricing on the purchasing dimension may be at least as large as on the financing dimension. This effect, however, is more difficult to quantify because it requires one to estimate a joint distribution of purchase and default probabilities, and to overcome the fact that the financing and repayment are not observed for non-purchasers. We tackle these problems in ongoing work.

5.5 Robustness of the Estimates

In this section we investigate the sensitivity of the results to some of the specific modeling assumptions. Detailed estimates are provided in the Appendix. As in Section 3, the main point to convey is that our basic findings are robust to a wide range of alternative specifications.

Sources of identification/Endogeneity of car price

Just as in Section 3, our baseline specification combines several sources of exogenous variation in loan size. To isolate these separate sources, we re-estimated the model using samples limited to narrow windows around changes in the margin schedule and around changes in the minimum down payment schedule, and on samples limited to applicants matched with cars in narrow windows around jumps in the margin schedule. The estimated effect of loan size on default is similar across these alternative specifications, ranging from 9 to 24 percent (compared to a 17 percent effect implied by Table 4), each of which focuses on a particular source of exogenous variation.

As noted above, one specific concern is that the negotiated price might incorporate information about default risk that we do not observe directly, and that is not incorporated into the down payment decision. Although the direction of bias is not a priori obvious, this could confound our estimate of the moral hazard effect. The stability of the results when we focus on limited samples around company-wide policy changes largely mitigates this concern, but to address it explicitly we incorporated the pricing equation, equation (2), and considered a model where the negotiated price, loan size and default are determined jointly. We again divide estimation into two steps, starting with estimation of the pricing and down payment equations, followed by estimation of the default model where we include both the pricing and down payment residuals as controls. With these controls, the effect of loan size on default is identified solely by variation in the company-set margin schedule and minimum down payment. This approach also gives qualitatively similar results; indeed it yields an estimated moral hazard effect that is somewhat larger than our baseline specification (27 percent).

Functional form and truncation

A more technical concern about our estimation strategy is that slightly less than half of the buyers decide to pay exactly the required minimum. This limits our ability to fully span the range of types at the lower end of the control variable we use. We explored this in two ways. First, we used the fact that a fraction of buyers who make the minimum down payment also choose to defer a fraction of it. We estimated a probit model for this deferral decision and constructed a second residual using the same method as above. Our idea here was that if the truncation prevented us from recovering all of the relevant private information, the deferral decision would contain additional information. It turns out, however, that including this additional residual has essentially no effect on our estimate of moral hazard.

A second approach is to fully parametrize the individual heterogeneity, and estimate loan size and default jointly in a maximum likelihood framework. The Cox default model makes it a bit awkward to model correlation with other latent variables, so instead we use a tobit model for the repayment process, where the fraction of repayments t_i is given by:

$$t_i = \min \left\{ t_i^* = \exp \left(L_i \delta_L + x_i' \delta_x + \eta_i \right), 1 \right\}. \tag{9}$$

where η_i is a normally distributed error term. We use maximum likelihood to estimate the tobit default equation, the tobit down payment equation and the pricing equation jointly under varying assumptions about correlation of the errors (including specifications that allow for rich forms of price endogeneity as discussed above).

The detailed results are reported in the Appendix. The estimated effects of loan size on default are close to what we obtain in the Cox model. The model also generates similar implications for adverse selection. We take this as evidence that our baseline findings are not driven by specific functional form assumptions, nor affected much by truncation of the control variable.

Car choice and non-purchase

By focusing only on purchasers, and abstracting from car choice as in Section 3, our empirical strategy arguably leaves the door open for selection effects: changes in price may affect the sample of purchasers or their chosen cars and not just loan size. The fact that the purchase decision appears quite insensitive to price provides some assurance on the former. Indeed, we considered a version of the tobit default model discussed above, where we estimate pricing, purchase, loan size *and* default jointly (Einav, Jenkins and Levin, 2007). Our estimates of the default process are not very different when we account for the purchase decision.

In terms of car choice, it is perhaps useful to envision the set of conditions under which substitution would confound our interpretation of the results. Recall that we control for individual and car characteristics. So the exercise is to compare the default behavior of two individuals who are identical on observable and unobservable characteristics (as captured by their down payment decision, and in some specifications their negotiated price), and also matched with an identical car. For substitution to be an issue, changes in car price would have to lead individuals who differ in some additional unobserved predictor of default — one not reflected in their choice of loan size — to be matched with different cars. Simple arithmetic also suggests that the amount of substitution would have to be large to have a material effect on our results. As we discussed in detail in the end of Section 3, however, both discussions with the company and examination of the data around the largest pricing change (Figure 3) strongly suggest that substitution across cars in response to price variation is not a major issue. For these reasons we doubt that explicitly modeling and accounting for car choice would have much effect on our results.

6 Conclusion

The notion that consumers may be liquidity constrained is an important theme in recent research on consumption, taxation, and social insurance. Our results provide fresh evidence on the role of liquidity in driving purchasing behavior, at least for people at the lower end of the income distribution. We view our primary contribution, however, as providing a snapshot of low-income credit markets, and especially of the informational problems that might characterize these markets and give rise to liquidity constraints. In particular, we have highlighted the substantial moral hazard and adverse selection problems faced by lenders serving the subprime population. Interestingly, it appears that modern credit scoring can go a significant distance toward mitigating adverse selection problems in the credit market, suggesting that innovations in this area may be an important cause of the rise of subprime lending that has occurred over the last decade. Such credit scoring is less likely to mitigate moral hazard problems, still restricting credit to subprime borrowers.

The paper has focused on market equilibrium and problems facing lenders, for which a statistical model of consumer behavior is sufficient. The focus on positive analysis of the subprime market allows us to abstract from a specific model of individual consumer behavior, but such a model would be necessary to any normative statements about the market, and understanding which behavioral model accurately characterizes the behavior of borrowers in our data would have broad implications for policy and welfare. We view this as a promising avenue for future work. We note that in the context of data like ours, plausible identification among various consumer models will benefit from information regarding borrowers' activities that are external to the company providing the data.

References

Attanasio, Orazio, Pinelopi Goldberg and Ekaterini Kyriazidou, "Credit Constraints in the Market for Consumer Durables: Evidence from Micro Data of Car Loans," Yale University Working Paper, 2006.

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.

Card, David, Raj Chetty and Andrea Weber, "Cash-on-Hand and Competing Models of Intertemporal Behavior: New Evidence from the Labor Market," *Quarterly Journal of Economics,* 122(4), November 2007, forthcoming.

Carroll, Christopher, "A Theory of the Consumption Function, with and without Liquidity Constraints," *Journal of Economic Perspectives*, 15(3), Summer 2001, 23-45.

Chetty, Raj, "Moral Hazard vs. Liquidity in Unemployment Insurance," U.C. Berkeley Working Paper, 2007.

Chiappori, Pierre-Andre and Bernard Salanie, "Testing for Asymmetric Information in Automobile Insurance," *Journal of Political Economy*, 108(1), February 2000, 56-78.

Deaton, Angus, "Savings and Liquidity Constraints," *Econometrica*, 59(5), September 1991, 1221-1248.

Edelberg, Wendy, "Risk-Based Pricing of Interest Rates in Household Loan Markets," Federal Reserve Working Paper, 2003.

Edelberg, Wendy, "Testing for Adverse Selection and Moral Hazard in Consumer Loan Markets," Federal Reserve Working Paper, 2004.

Einav, Liran, Mark Jenkins and Jonathan D. Levin, "Contract Pricing in Consumer Credit Markets," mimeo, Stanford University, 2007.

Gross, David B. and Nicholas S. Souleles, "Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data?" *Quarterly Journal of Economics*, 117(1), February 2002, 149-185.

Jaffee, Dwight and Thomas Russell, "Imperfect Information, Uncertainty, and Credit Rationing," *Quarterly Journal of Economics,* 90, November 1976, 651-666.

Jaffee, Dwight and Joseph E. Stiglitz, "Credit Rationing," *Handbook of Monetary Economics*, Vol. 2, ed. Benjamin Friedman and Frank Hahn, Elsevier, 1990, 837-888.

Jappelli, Tullio, "Who is Credit Constrained in the U.S. Economy?" *Quarterly Journal of Economics*, 105, 1990, 219-234.

Johnson, David, Jonathan Parker and Nicholas Souleles, "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, 96(5), December 2006, 1589-1610.

Juster, F. Thomas and R. P. Shay, "Consumer Sensitivity to Finance Rates: An Empirical and Analytical Investigation," NBER Occasional Paper no. 88, 1964.

Karlan, Dean and Jonathan Zinman, "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment," Yale University Working Paper, 2007.

Karlan, Dean and Jonathan Zinman, "Credit Elasticities in Less Developed Countries: Implications for Microfinance," *American Economic Review*, forthcoming.

Laibson, David, Andrea Repetto and Jeremy Tobacman, "A Debt Puzzle," in *Knowledge, Information and Expectations in Modern Economics: In Honor of Edmund S. Phelps*, eds. Phillipe Aghion, Roman Frydman, Joseph Stiglitz, Michael Woodford, Princeton: Princeton University Press, 2003, 228-266.

Parker, Jonathan, "The Reaction of Household Consumption to Predictable Changes in Social Security Taxes," *American Economic Review*, 89, 1999, 959-973.

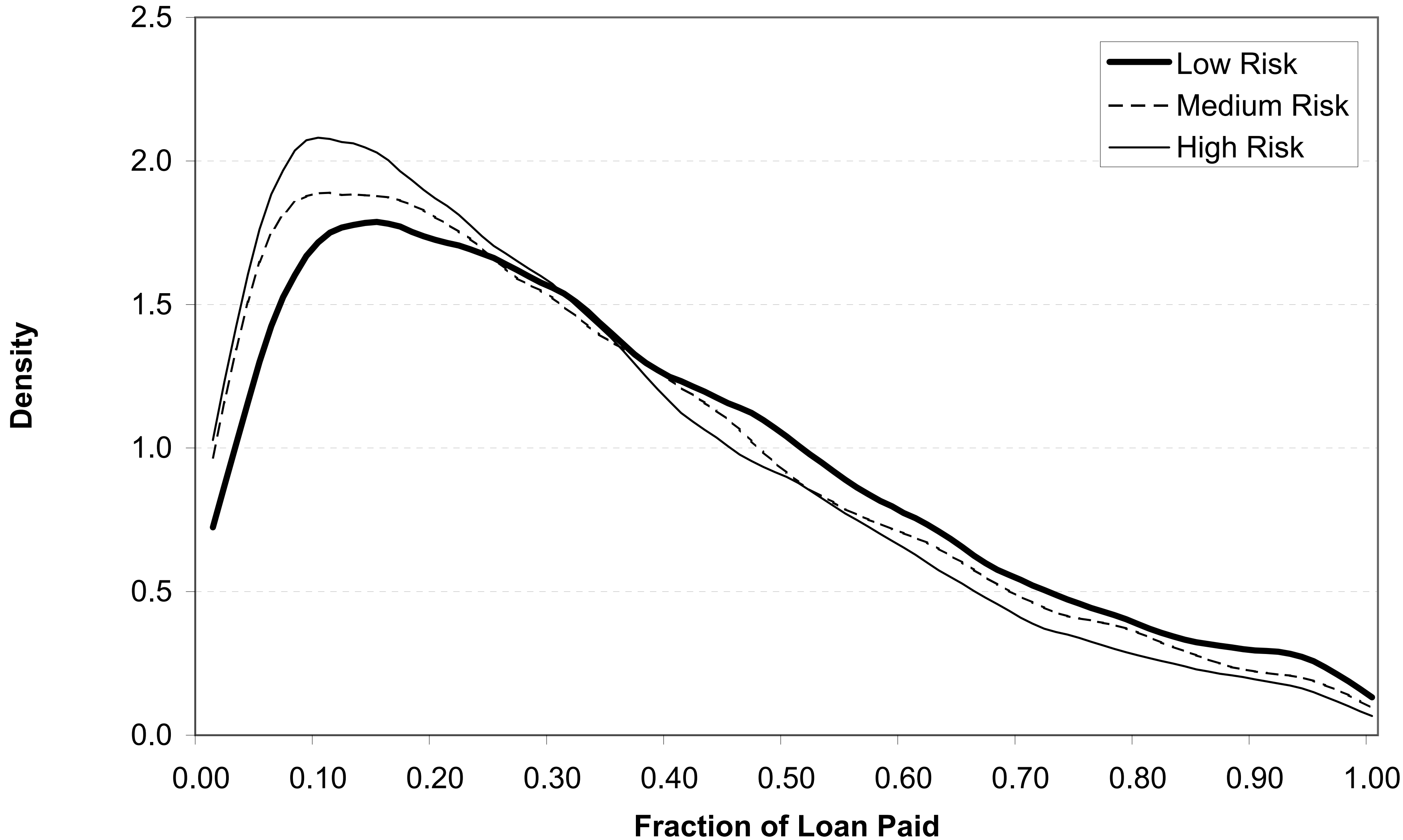
Rothschild, Michael, and Joseph E. Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, 90, 1976, 630-649.

Souleles, Nicholas, "The Response of Household Consumption to Income Tax Refunds," *American Economic Review*, 90, 1999, 947-958.

Stiglitz, Joseph and Andrew Weiss, "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 71(3), June 1981, 393-410.

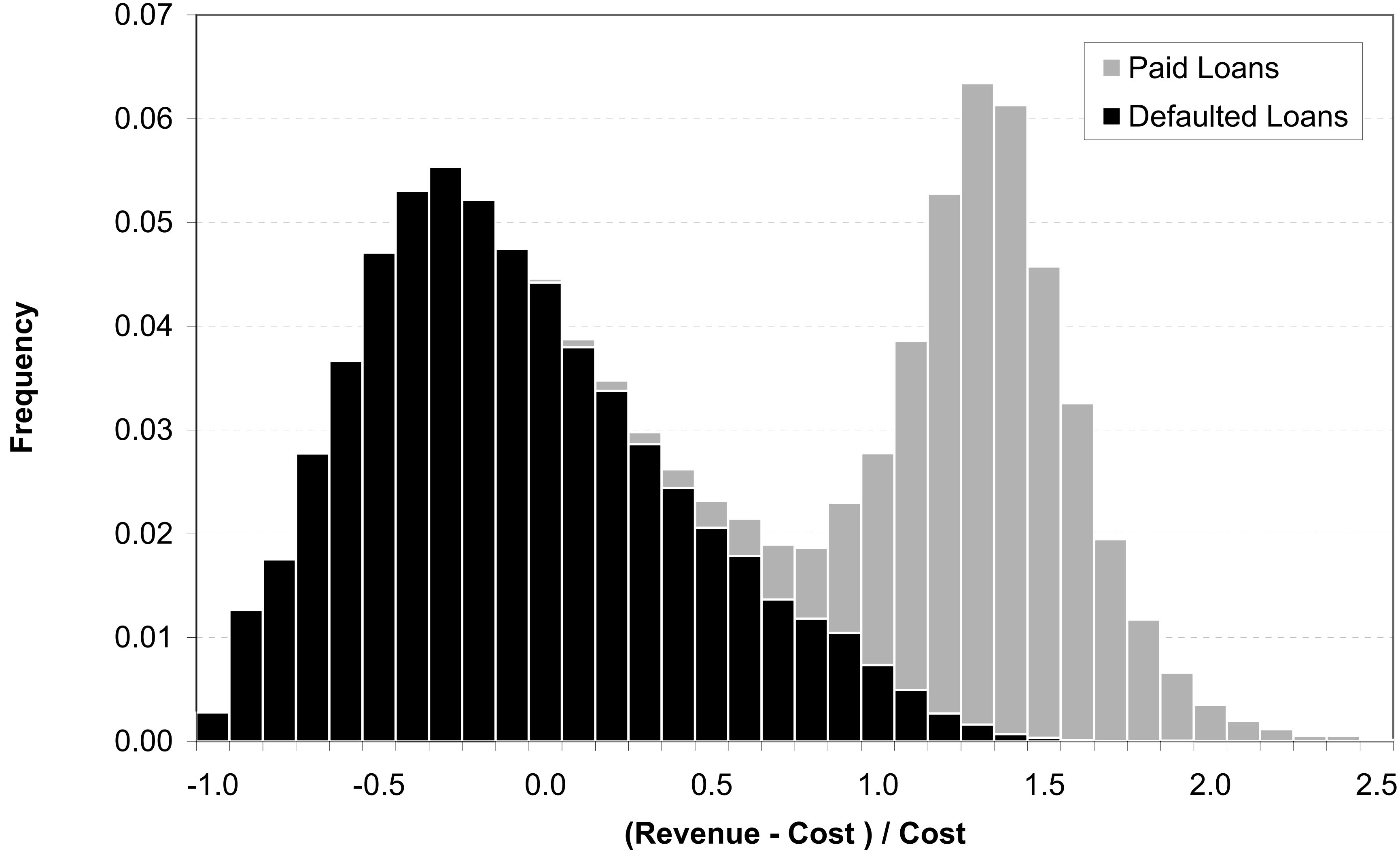
Zeldes, Stephen, "Consumption and Liquidity Constraints: An Empirical Investigation," *Journal of Political Economy*, 97(2), April 1989, 305-346.

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



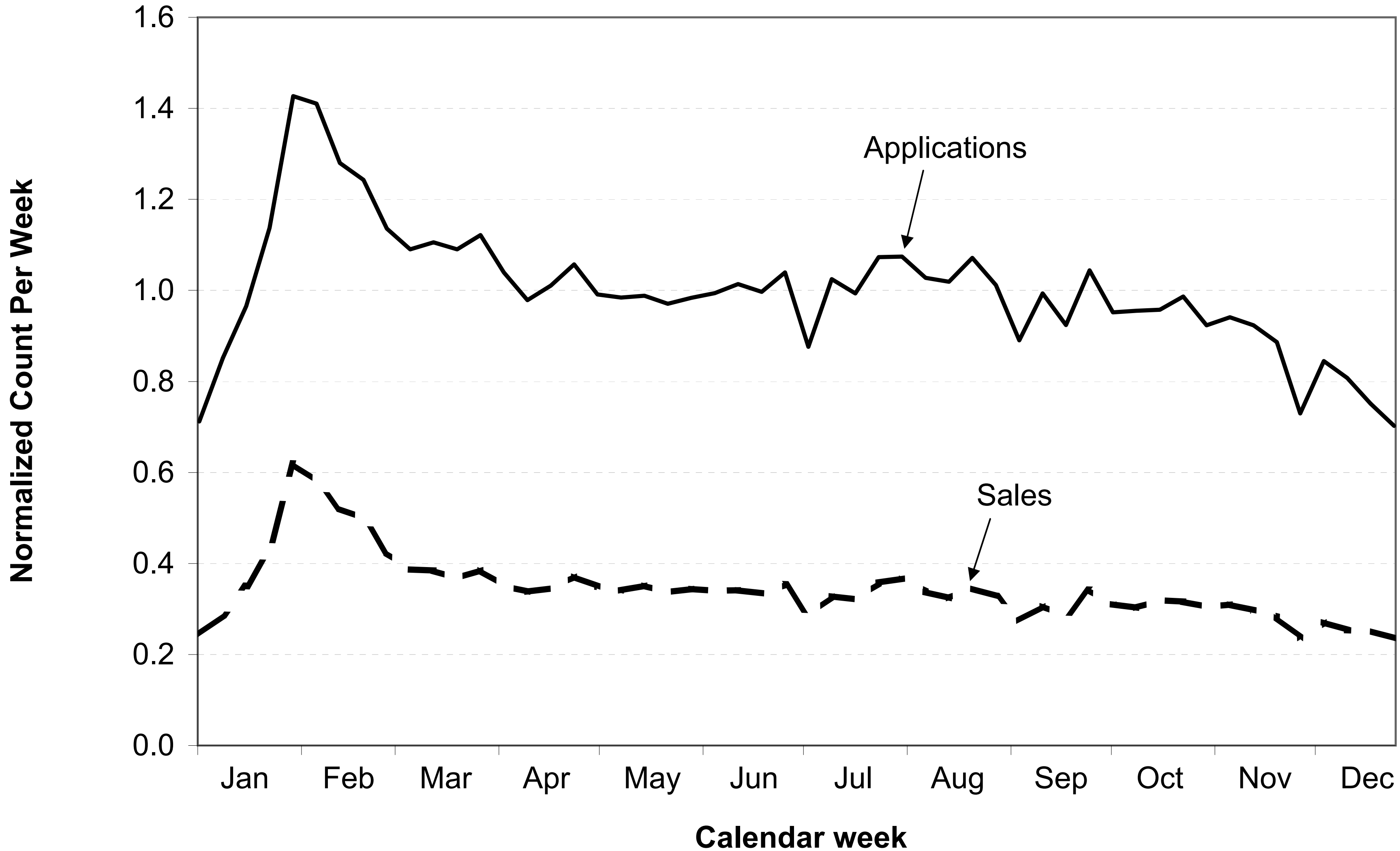
Notes: Based on data from uncensored loans that ended in default.

Figure 1(b): Rate of Return Histogram



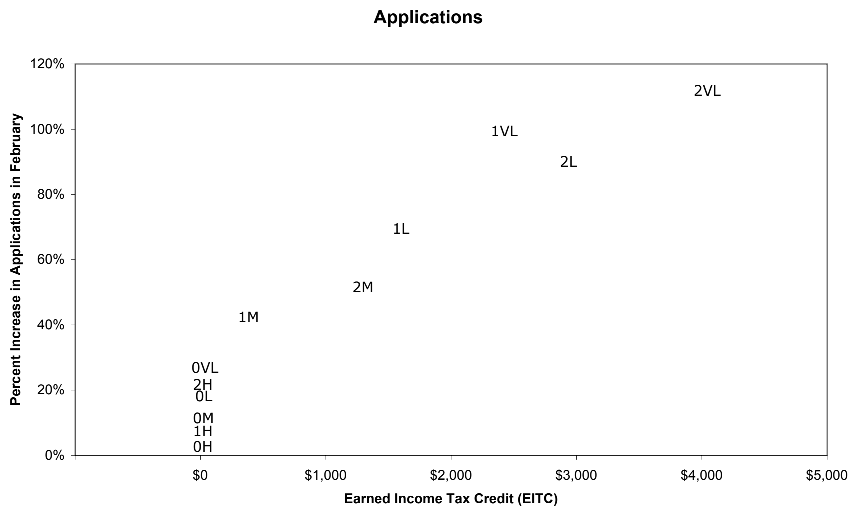
Notes: Based on data from uncensored loans. Revenue is calculated as down payment + PV of loan payments + PV of recovery, assuming an internal firm discount rate of 10 percent.

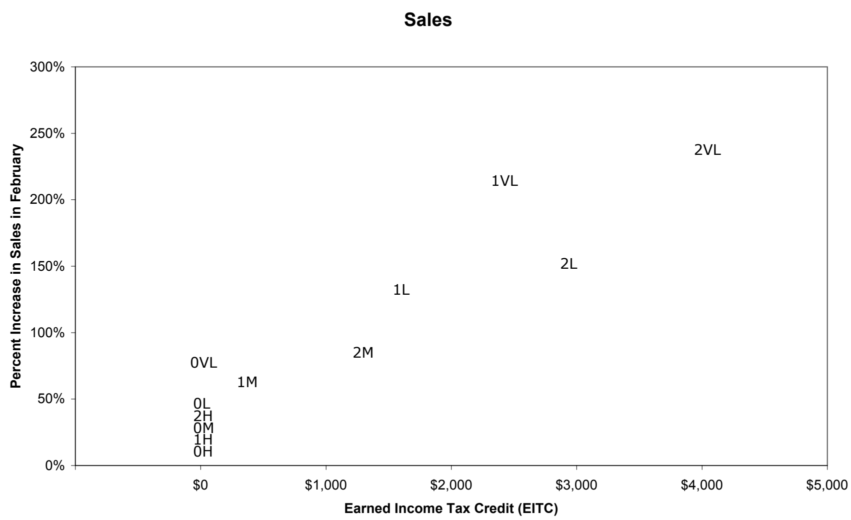
Figure 2(a): Seasonality in Applications and Sales



Notes : Based on data from all applications. Number of applications and sales are normalized by the average number of applications per week.

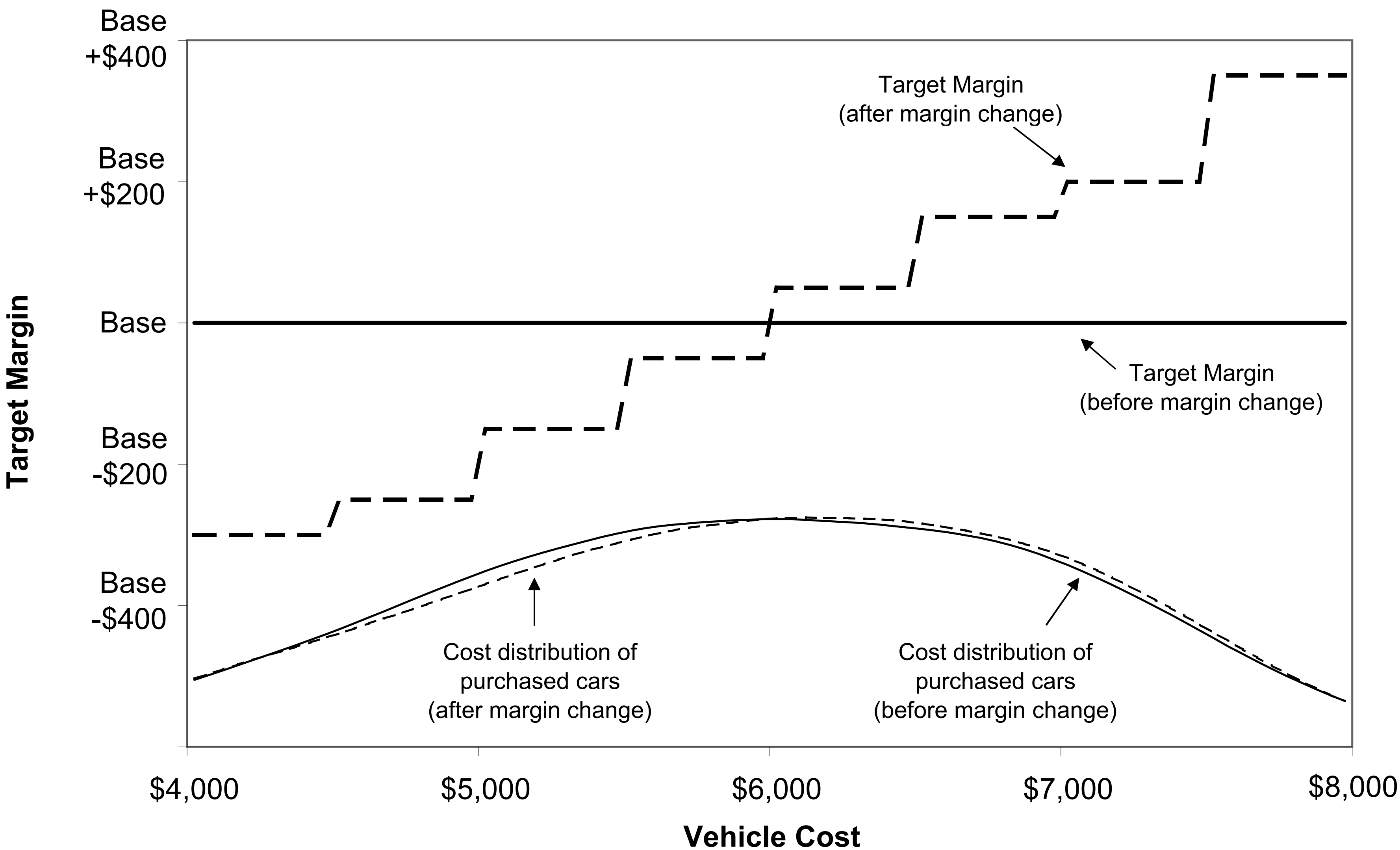
Figure 2(b): Tax Credit Effects on Applications and Sales





Notes: Based on data from all applications. Each point represents a group of applicants with a given income level and number of dependents. Labels for number of dependents are: 0 = No Dependents, 1 = 1 dependent, and 2 = 2 or more dependents. Lables for income level are: VL = Less than $1,500 per month, L = Between $1,500 and $2,000 per month, M = Between $2,000 and $3,000 per month, and H = More than $3,000 per month.

Figure 3: The Effect (or Lack Thereof) of Prices on Car Choices



Notes: Margin schedules based on firm's pricing policy before and after a policy change from a fixed markup to a cost-based markup. Target margin represents the markup over vehicle cost used to determine the vehicle's list price. To preserve confidentiality, we do not report the level of the target margin, but do show changes in margin resulting from the policy change. The cost distribution curves show the cost distribution of purchased cars in one month periods before and after the margin change. The curves are created by estimating a kernel density of vehicle costs with a $400 bandwidth in each period. A Komolgorov-Smirnov test fails to reject that the distributions are equal at the 90 percent confidence level.

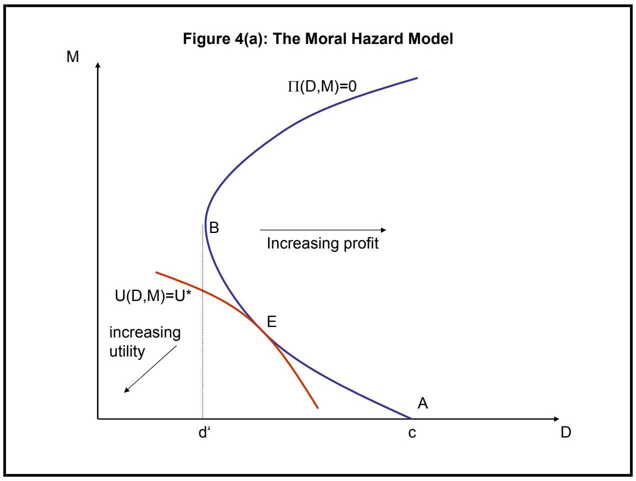


Figure 4(a): The Moral Hazard Model

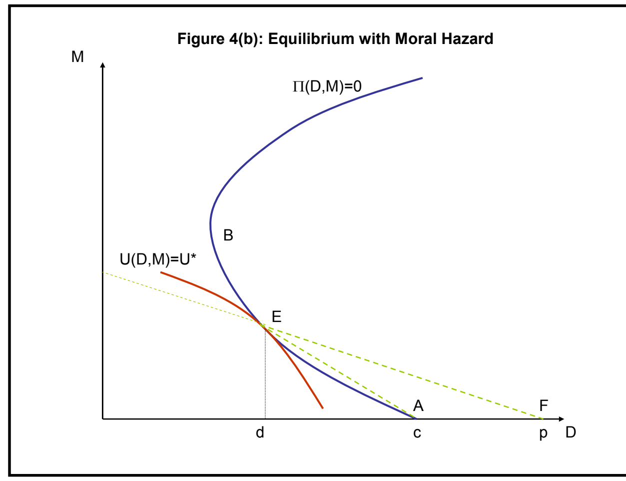


Figure 4(b): Equilibrium with Moral Hazard

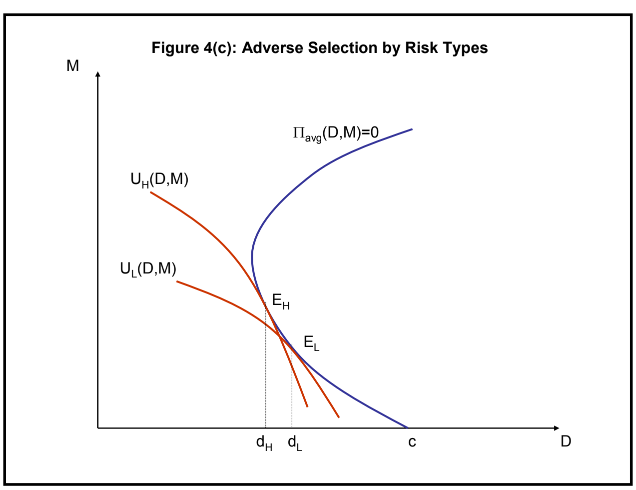


Figure 4(c): Adverse Selection by Risk Types

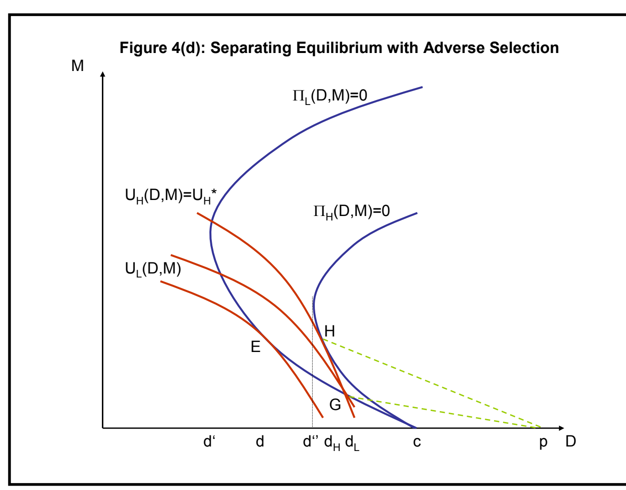
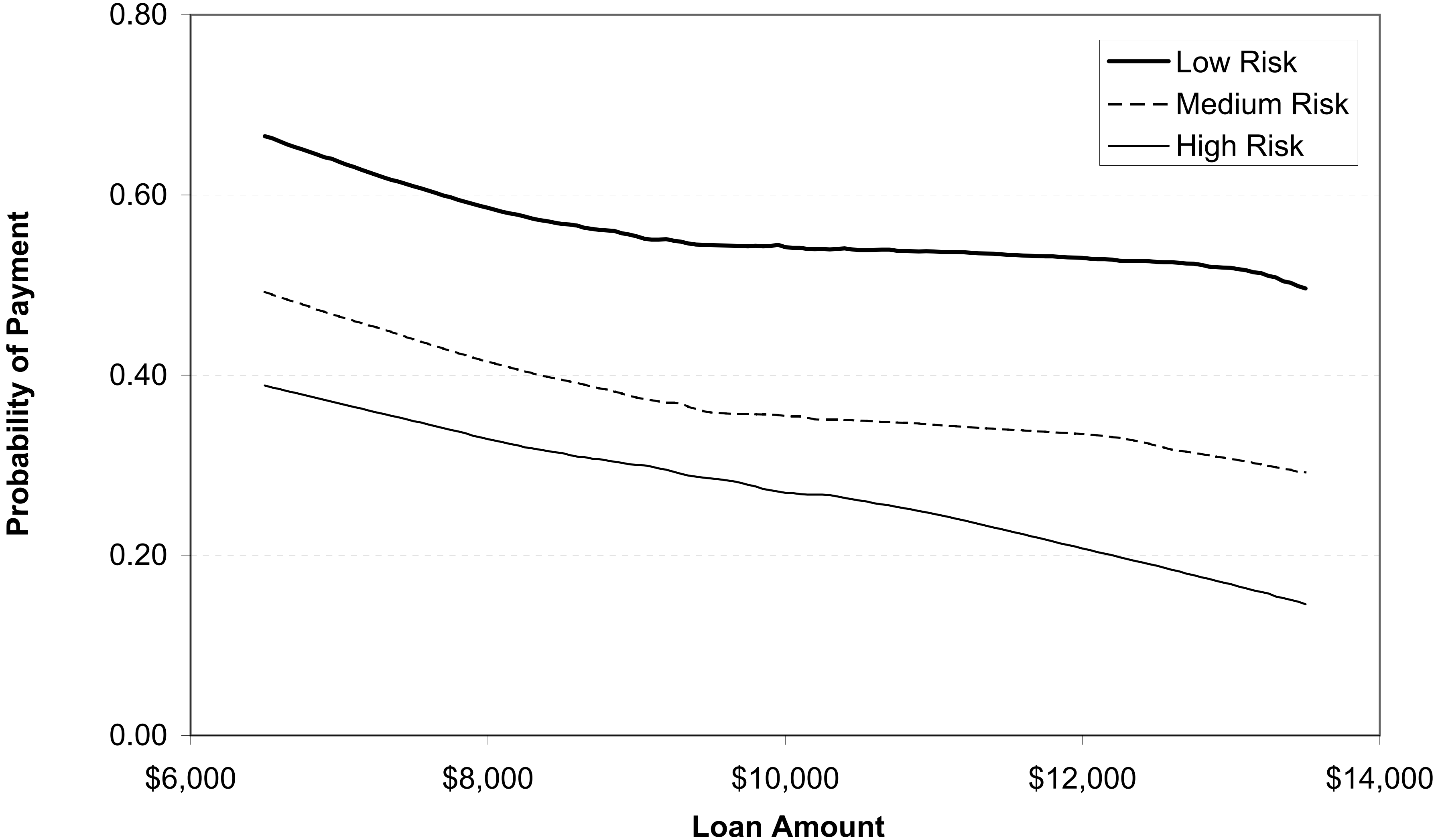


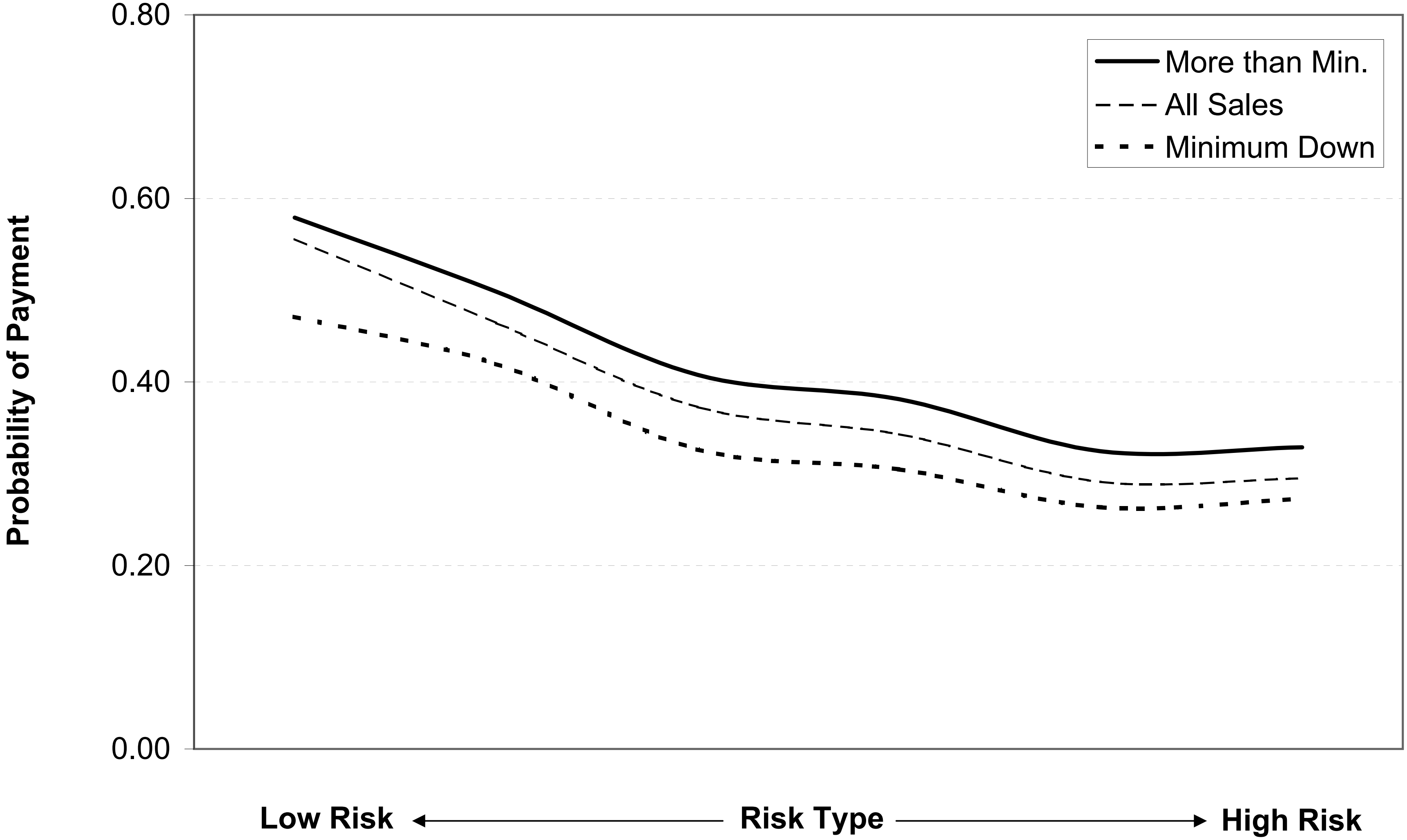
Figure 4(d): Separating Equilibrium with Adverse Selection

Figure 5(a): Probability of Payment vs. Loan Amount



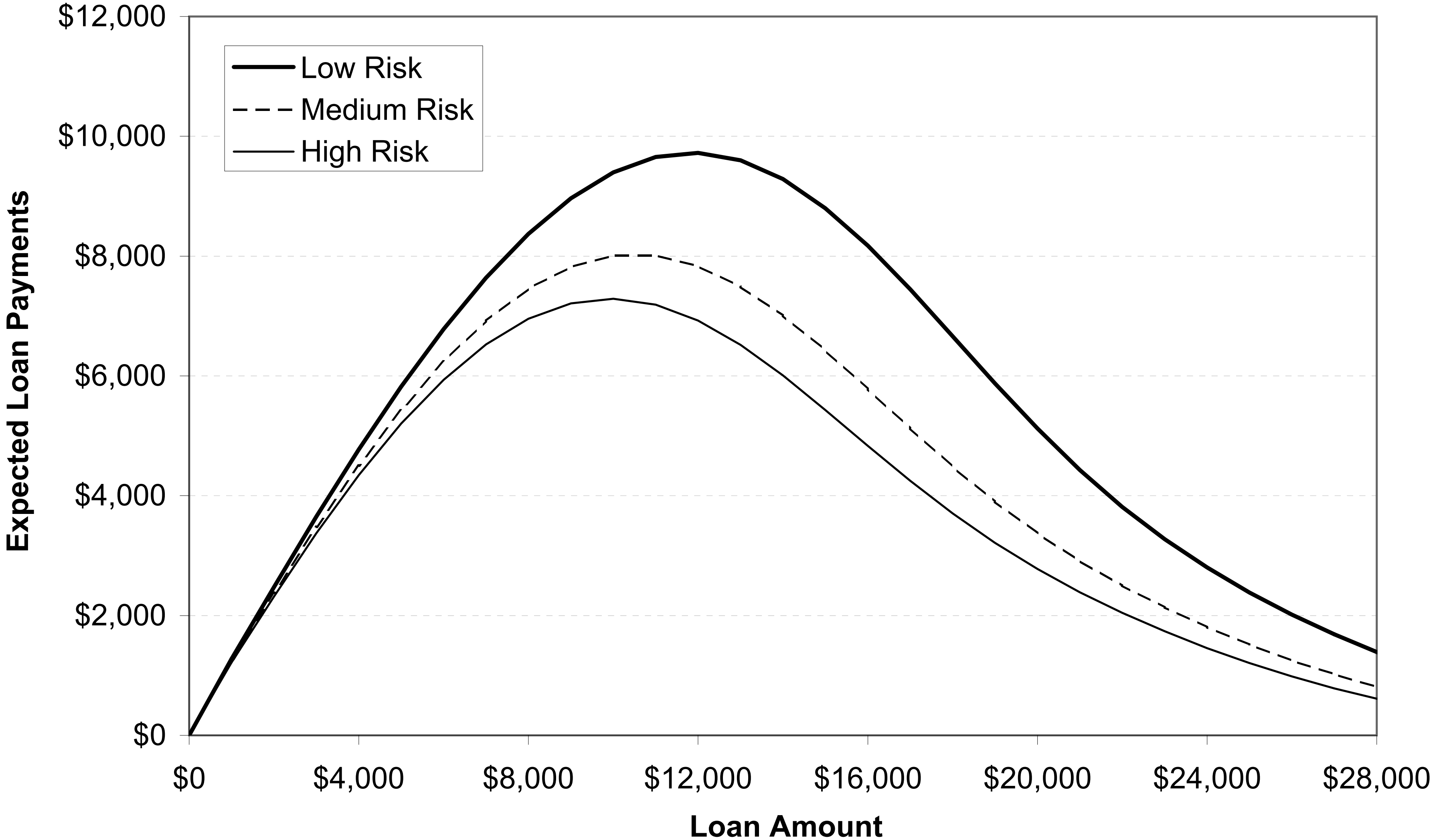
Notes: Based on data from uncensored loans. The x-axis represents the size of the loan at the time of origination, not including finance charges. The y-axis represents the probability that the loan is repaid in full. The "Low Risk" line (solid, dark), "Medium Risk" line (dashed), and "High Risk" line (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 5(b): Probability of Payment by Risk Type and Down Payment



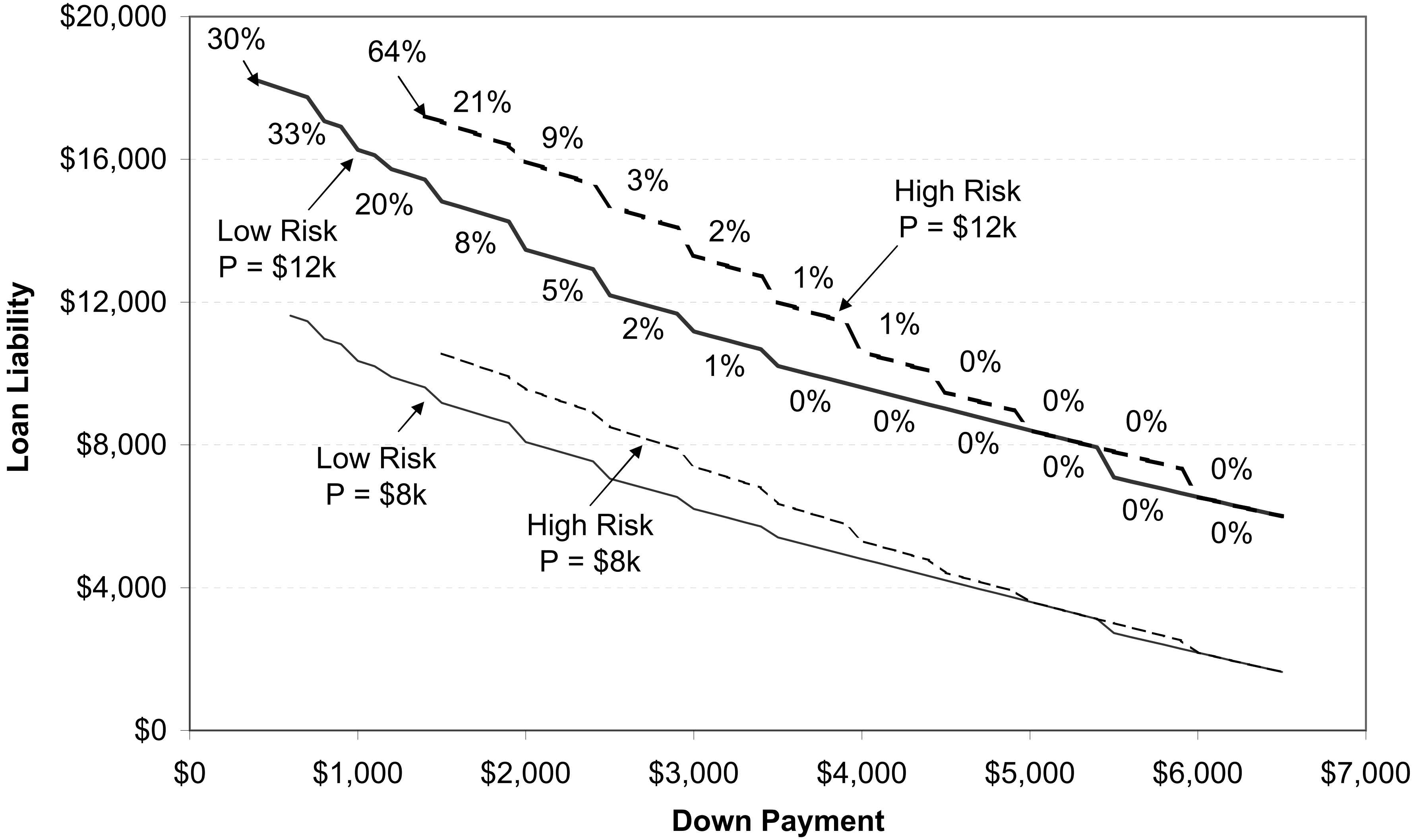
Notes: Based on data from uncensored loans. The x-axis represents a measure of buyer riskiness based on internal company credit scoring. The y-axis represents the probability that buyers repay the loan in full. The "More than Min." line (solid) shows the relationship between credit score and the probability 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 payment probabilities for each group, and smoothing.

Figure 6(a): Expected Loan Payments vs. Loan Amount



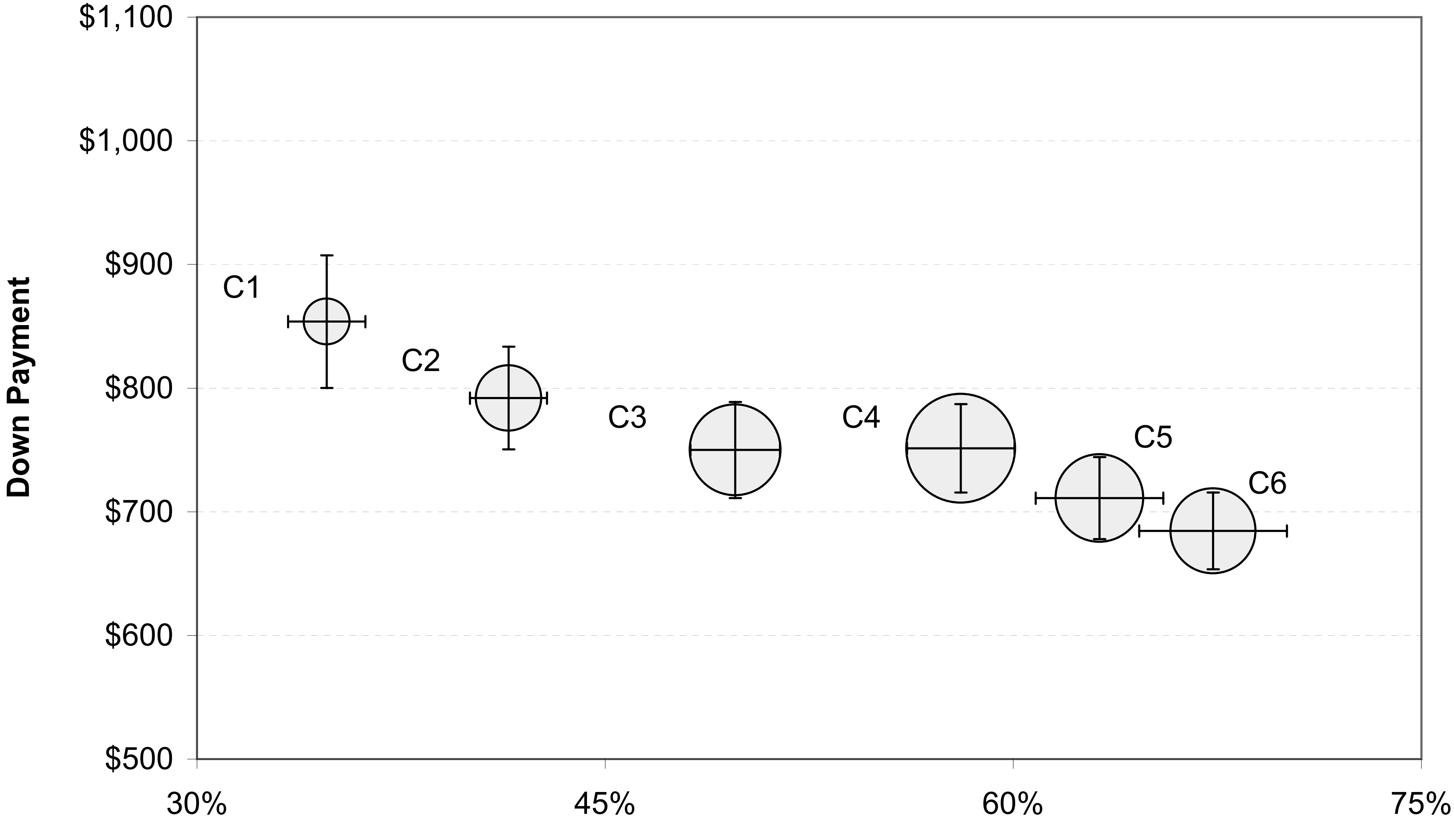
*Notes: B*ased on data from all loans. The x-axis represents the size of the loan at the time of origination, not including finance charges. The y-axis represents the expected NPV of payments on the loan, assuming payments are made monthly and discounted at an annual rate of 10 percent. Expected payment curves are constructed by estimating a Cox prop. hazard model of the fraction of loan payments made on financing terms (including loan amount), individual and car characteristics, and a down payment residual, and using the estimated model to calculate the probability of making each monthly payment for an average buyer purchasing an average car at varying loan amounts, assuming a 29.9% APR and 42 month loan term. The Low Risk, Medium Risk, and High Risk lines are created through proportional shifts in the payment hazard function determined by estimated risk category fixed effects.

Figure 6(b): Offer Curves



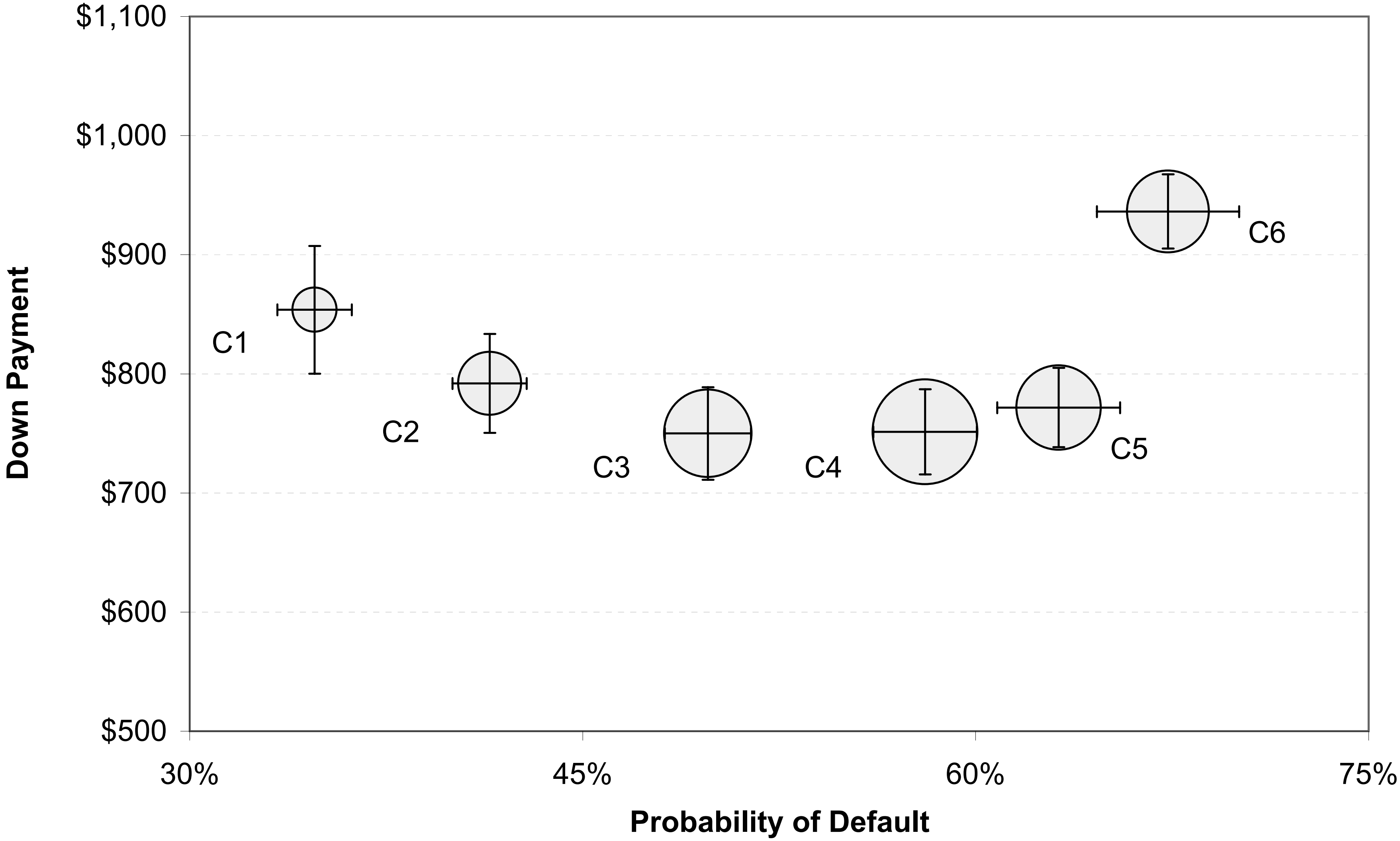
Notes: Based on firm's pricing policy in effect from 4/15/03 through 8/24/04. Each curve represents the trade-off between down payment and future loan payments offered to buyers in a given risk group for a car with a given price. Loan liability, on the y-axis, is defined as the initial loan amount (price minus down payment) plus finance charges. The offer curves are convex, rather than linear, because the firm offers lower APRs to buyers who make higher down payments. The percentages at each point along the curve show the portion of buyers in each risk group that select that point on the offer curve. The large point masses at the left end of each curve represent buyers making the minimum down payment.

Figure 7(a): Scatter Plot of Desired Down Payment vs. Default Probability by Credit Category



Notes : Based on data from all loans. Desired down payments calculated from a Tobit regression of down payment on transaction and car characteristics and time and city fixed effects, but no buyer characteristics, with left-censoring at the minimum down payment. Default probabilities calculated from a Cox proportional hazard model of fraction of payments made on the same set of explanatory variables. All estimates based on averge car price of $10,777 and average car characteristics given in Table 1. Size of bubbles represents number of sales to each credit category. Bars represent 95 percent confidence intervals.

Figure 7(b): Scatter Plot of Actual Down Payment vs. Default Probability by Credit Category



Notes: Based on data from all loans. Down payments are equal to the greater of desired down payment and the minimum down payment required by the firm, where desired down payments are calculated as in Figure 7(a). Default probabilities are also calculated as in Figure 7(a). All estimates based on averge car price of $10,777 and average car characteristics given in Table 1. Size of bubbles represents number of sales to each credit category. Bars represent 95 percent confidence intervals.

Table 1: Summary Statistics

	Obs*	Mean	Std. Dev.	5%	95%
Applicant Characteristics					
Age	N	32.8	10.7	19	53
Monthly Income	N	2,414	1,074	1,299	4,500
Home Owner	N	0.15	-	-	-
Live With Parents	N	0.18	-	-	-
Bank Account	N	0.72	-	-	-
Risk Category					
Low	N	0.27	-	-	-
Medium	N	0.45	-	-	-
High	N	0.29	-	-	-
Car Purchased	N	0.34	-	-	-
Buyer Characteristics					
Age	0.34N	34.7	10.8	20	55
Monthly Income	0.34N	2,557	1,089	1,385	4,677
Home Owner	0.34N	0.17	-	-	-
Live With Parents	0.34N	0.16	-	-	-
Bank Account	0.34N	0.76	-	-	-
Risk Category					
Low	0.34N	0.35	-	-	-
Medium	0.34N	0.47	-	-	-
High	0.34N	0.17	-	-	-
Car Characteristics					
Acquisition Cost	0.34N	5,213	1,358	3,205	7,240
Total Cost	0.34N	6,096	1,372	4,096	8,213
Car Age (years)	0.34N	4.3	1.9	2	8
Odometer	0.34N	68,776	22,090	31,184	102,300
Lot Age (days)	0.34N	33	44	1	122
Car Price	0.34N	10,777	1,797	8,095	13,595
Transaction Characteristics					
Min. Down Payment (applicants)	N	750	335	400	1,400
Min. Down Payment (buyers)	0.34N	648	276	400	1,200
Interest Rate (APR)	0.34N	26.2	4.4	17.7	29.9
Loan Term (months)	0.34N	40.5	3.7	35	45
Down Payment	0.34N	962	602	400	2,000
Loan Amount	0.34N	10,740	1,802	7,982	13,560
Monthly Payment	0.34N	395	49	314	471
Default (uncensored obs. only)	0.13N	0.61	-	-	-
Recovery Amt. (uncen. defaults)	0.08N	1,382	1,386	0	3,784

* To preserve the confidentiality of the company that provided the data, we do not report the exact number of observations, N >> 50,000.

Table 2: Purchasing Estimates

	Probit Estimates of Individual-Level Purchasing (dep. var. = sale indicator)								*Cell-Level Estimates (dep. var. = log(sales))*			
	(1)		(2)[a]		(3)[b]		(4)[c]		(5)[d]		(6)[d]	
	dF/dx	**Std. Err.**	**dF/dx**	**Std. Err.**	**dF/dx**	**Std. Err.**	**dF/dx**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**
Offer Variables												
Negotiated Price ($100s)	-0.0002	(0.0002)	-0.0010	(0.0011)	-0.0022	(0.0007)	-0.0032	(0.0006)	-0.0061	(0.0016)	-0.0102	(0.0063)
Minimum Down ($100s)	-0.0301	(0.0006)	-0.0299	(0.0006)	-0.0298	(0.0006)	-0.0303	(0.0006)	-0.0895	(0.0039)	-0.0889	(0.0039)
Maximum Interest Rate (APR)	-0.0010	(0.0004)	-0.0013	(0.0006)	-0.0016	(0.0005)	0.0003	(0.0006)	-0.0033	(0.0033)	-0.0049	(0.0041)
Term (months)	-0.0008	(0.0004)	-0.0002	(0.0008)	0.0006	(0.0006)	0.0025	(0.0006)	-0.0001	(0.0025)	0.0021	(0.0044)
Car Characteristics												
Car Cost ($100s)	0.0005	(0.0002)	0.0014	(0.0012)	0.0025	(0.0007)	0.0034	(0.0006)	0.0071	(0.0018)	0.0113	(0.0067)
Premium (Cost > $7,500)	0.0040	(0.0032)	0.0038	(0.0035)	0.0036	(0.0034)	0.0006	(0.0035)	0.0990	(0.0376)	0.0966	(0.0379)
Car Age (years)	0.0008	(0.0007)	0.0008	(0.0006)	0.0008	(0.0006)	-0.0001	(0.0006)	0.0086	(0.0067)	0.0088	(0.0066)
Odometer (10,000s)	-0.0008	(0.0004)	-0.0008	(0.0004)	-0.0008	(0.0004)	-0.0008	(0.0004)	-0.0139	(0.0042)	-0.0144	(0.0041)
Lot Age (months)	-0.0019	(0.0007)	-0.0034	(0.0022)	-0.0055	(0.0015)	-0.0071	(0.0010)	-0.0232	(0.0068)	-0.0308	(0.0133)
Individual Characteristics												
Income ($1,000s/month)	0.0245	(0.0008)	0.0250	(0.0010)	0.0258	(0.0008)	0.0284	(0.0009)	0.0983	(0.0060)	0.1014	(0.0075)
Age	0.0084	(0.0003)	0.0084	(0.0003)	0.0085	(0.0003)	0.0082	(0.0003)	0.0732	(0.0053)	0.0747	(0.0056)
Age squared	-0.0001	(3.7E-06)	-0.0001	(3.7E-06)	-0.0001	(3.8E-06)	-0.0001	(3.8E-06)	-0.0008	(6.7E-05)	-0.0008	(7.2E-05)
Bank Account	0.0271	(0.0014)	0.0270	(0.0014)	0.0269	(0.0014)	0.0281	(0.0014)	-0.0010	(0.0296)	-0.0019	(0.0296)
House Owner	-0.0320	(0.0018)	-0.0321	(0.0018)	-0.0321	(0.0018)	-0.0408	(0.0016)	-0.0171	(0.0367)	-0.0192	(0.0367)
Lives with Parents	0.0091	(0.0021)	0.0090	(0.0021)	0.0089	(0.0021)	0.0097	(0.0021)	0.0391	(0.0387)	0.0339	(0.0379)
Credit Category Fixed Effects												
Representative Low Risk	0.0269	(0.0070)	0.0264	(0.0070)	0.0256	(0.0070)	0.0239	(0.0070)	0.1333	(0.0517)	0.1304	(0.0519)
Representative Medium Risk	0.0394	(0.0063)	0.0397	(0.0063)	0.0402	(0.0063)	0.0381	(0.0064)	0.2974	(0.0434)	0.2975	(0.0435)
Representative High Risk	0.0043	(0.0050)	0.0045	(0.0050)	0.0048	(0.0050)	0.0038	(0.0051)	0.0489	(0.0494)	0.0524	(0.0493)
Month Fixed Effects												
February (tax season)	0.1603	(0.0044)	0.1594	(0.0047)	0.1581	(0.0045)	0.1592	(0.0044)	0.5900	(0.0190)	0.5862	(0.0199)
Other Fixed Effects	Year, Month, City, Credit Category		Year, Month, City, Credit Category		Year, Month, City, Credit Category		Year, Month, Credit Category		Year, Month, City, Credit Category		Year, Month, City, Credit Category	
Instrument for Price	-		List Price		Cost Bucket Dummies		State Dummies		-		List Price	

<u>Notes</u>

Sample for individual-level purchasing estimates is all applications; sample size is N >> 50,0000 (see Table 1). Sample size for cell-level estimates is ~0.03N.

a Instruments are list price (equal to zero if not available) and indicator equal to one if list price is not available. List prices are available for approximately 80 percent of the observations.

b Instruments are dummy variables corresponding to each of 11 cost buckets (see Figure 3 for illustration).

c Instruments are dummy variables corresponding to two states with APR caps below 29.9 percent.

d Cell-level regressions are weighted by number of apps and include log(apps) as an explanatory variable. The coefficient on log(apps) is 0.931 (0.009) in the OLS specification and 0.930 (0.009 with IV.

Cell-level "fixed-effects" represent the average of a given dummy variable within a cell and may take on values between 0 and 1.

Individual-level and cell-level estimates can be roughly compared by dividing the latter by 3, since the probability of sale is 0.34.

Omitted fixed effects are Very High Risk for credit categories and December for months.

All standard errors are bootstrap estimates based on 50 random samples.

Table 3: Tobit Estimates of Down Payment

Dependent Variable: Down Payment ($100s) Conditional on Purchase

	(1)		(2)	
	Coeff.	**Std. Err.**	**Coeff.**	**Std. Err.**
Offer Variables				
Negotiated Price ($100s)	0.049	(0.004)	0.177	(0.002)
Maximum Interest Rate (APR)	0.203	(0.011)	0.242	(0.013)
Term (months)	-0.413	(0.009)	-0.503	(0.018)
Car Characteristics				
Car Cost ($100s)	0.228	(0.004)	0.100	(0.002)
Premium (Cost > $7,500)	3.769	(0.078)	3.759	(0.079)
Car Age (years)	0.061	(0.016)	0.065	(0.016)
Odometer (10,000s)	-0.026	(0.012)	-0.022	(0.012)
Lot Age (months)	-0.559	(0.016)	-0.331	(0.044)
Individual Characteristics				
Income ($1,000s/month)	-0.164	(0.019)	-0.260	(0.026)
Age	-0.169	(0.010)	-0.186	(0.010)
Age squared	0.002	(1E-04)	0.002	(1E-04)
Bank Account	0.202	(0.046)	0.235	(0.047)
House Owner	0.226	(0.055)	0.241	(0.055)
Lives with Parents	0.264	(0.055)	0.266	(0.056)
Credit Category Fixed Effects				
Representative Low Risk	4.734	(0.119)	4.961	(0.126)
Representative Medium Risk	3.215	(0.107)	3.270	(0.108)
Representative High Risk	0.718	(0.119)	0.733	(0.120)
Month Fixed Effects				
February (tax season)	3.171	(0.098)	3.259	(0.162)
Other Fixed Effects	Year, Month, City, Credit Category		Year, Month, City, Credit Category	
Instrument for Price*	-		List Price	

<u>Notes</u>

Sample is all sales; sample size is ~0.34N, where N >> 50,000 (see Table 1). All results are based on Tobit regressions with actual down payment minus minimum down payment as the dependent variable and left-censoring at zero. Omitted fixed effects are Very High Risk for credit categories and December for months.

* The results in Column (2) are based on the joint maximum likelihood estimation of a negotiated price equation with list price as an explanatory variable and a down payment equation, assuming joint normality of the errors.

Table 4: Proportional Hazard Model Estimates of Default

Dependent Variable: Fraction of loan payments made

	(1)		(2)		(3)		(4)	
	Haz. Rat.	**Std. Err.**	**Haz. Rat.**	**Std. Err.**	**Haz. Rat.**	**Std. Err.**	**Haz. Rat.**	**Std. Err.**
Transaction characteristics								
Amount Financed ($100s)	1.016	(0.001)	1.024	(0.001)	1.023	(0.001)	1.019	(0.000)
Maximum Interest Rate (APR)	1.022	(0.002)	1.026	(0.002)	1.025	(0.002)	1.022	(0.002)
Term (months)	1.015	(0.002)	1.006	(0.002)	1.008	(0.002)	1.008	(0.002)
Down Payment Residual ($100s)	0.982	(0.001)	-	-	-	-	-	-
Car Characteristics								
Car Cost ($100s)	0.981	(0.001)	0.975	(0.001)	0.974	(0.001)	0.976	(0.001)
Premium (Cost > $7,500)	0.867	(0.015)	0.888	(0.015)	0.887	(0.015)	0.819	(0.014)
Car Age (years)	1.028	(0.003)	1.028	(0.003)	1.027	(0.003)	1.021	(0.003)
Odometer (10,000s)	1.012	(0.002)	1.012	(0.002)	1.012	(0.002)	1.015	(0.002)
Lot Age (months)	1.055	(0.003)	1.065	(0.003)	1.067	(0.003)	1.062	(0.003)
Individual Characteristics								
Income ($1,000s/month)	0.955	(0.004)	0.948	(0.004)	-	-	-	-
Age	0.996	(0.002)	0.993	(0.002)	-	-	-	-
Age squared	1.000	(2E-05)	1.000	(7E-03)	-		-	-
Bank Account	0.818	(0.007)	0.823	(0.007)	-	-	-	-
House Owner	0.998	(0.011)	1.004	(0.011)	-	-	-	-
Lives with Parents	1.059	(0.011)	1.060	(0.011)	-	-	-	-
Credit Category Fixed Effects								
Representative Low Risk	0.518	(0.011)	0.509	(0.011)	0.461	(0.009)	-	-
Representative Medium Risk	0.801	(0.013)	0.789	(0.013)	0.748	(0.012)	-	-
Representative High Risk	0.994	(0.018)	0.990	(0.018)	0.963	(0.017)	-	-
Month Fixed Effects								
February (tax season)	1.071	(0.022)	1.090	(0.022)	0.973	(0.023)	-	-
Other Fixed Effects	Year, Month, City, Credit Category		Year, Month, City, Credit Category		Year, Month, City, Credit Category		Year, Month, City, Credit Category	

Notes

Sample is all sales; sample size is ~0.34N, where N >> 50,000 (see Table 1). The down payment residual in column (1) is computed from the Tobit regression presented in column (2) of Table 3. Omitted fixed effects are Very High Risk for credit categories and December for months.