The Impact of Information Technology on Consumer Lending*

Liran Einav, Mark Jenkins, and Jonathan Levin†

July 2009

Abstract

We study the adoption of automated credit scoring at a large auto finance company and the changes it enabled in lending practices. Credit scoring appears to have increased profits by roughly a thousand dollars per loan. We identify two distinct benefits of risk classification: the ability to screen high-risk borrowers and the ability to target more generous loans to lower-risk borrowers. We show that these had effects of similar magnitude. We also explore whether increased reliance on hard information led to convergence in profitability across dealerships by substituting for varying qualities of soft local information. We find that all dealerships appear to have become more profitable, but little evidence that profits converged.

KEYWORDS: Information technology, Credit scoring, Consumer lending.

JEL classification: D82, G21, L86.

*We thank Luke Stein for excellent research assistance, Will Adams for his contributions to this project, and Chris Knittel, Ulrike Malmendier, and participants in IO Fest, the AEA meetings, and the NBER IO Summer meeting for helpful comments. We acknowledge support from the Stanford Institute for Economic Policy Research, the National Science Foundation (Einav and Levin), and the Center for Advanced Study in the Behavioral Sciences (Levin).

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

1 Introduction

Over the last thirty years information technology has revolutionized consumer lending. Automated credit scoring and underwriting have replaced traditional interview procedures to screen borrowers, and loan pricing has become increasingly sophisticated. This transformation has impacted virtually every consumer loan market, from mortgages to auto financing to unsecured lending such as credit cards. While the near universal adoption of these techniques indicates their value to lenders, there is relatively little specific evidence on exactly how benefits are realized, the size of the effects, and their organizational impacts.

We describe in this paper a natural case study of the changes in consumer lending. We analyze the implementation of automated credit scoring at an auto finance company. The company specializes in the low-income, high-risk consumer market — a market that is particularly well-suited for studying informational problems facing lenders. Default risk is high and recovery values are low, so profitability hinges on identifying better risks in the applicant pool (Adams, Einav and Levin, 2009; Einav, Jenkins and Levin, 2008). Loan applicants also vary substantially in their prospective risk of default, and their characteristics and credit histories provide prospective information about this risk. The potential to stratify borrowers can be seen in the fact that the top third of borrowers in terms of predicted risk is more than forty percent as likely to default as the bottom third.

We find that the adoption of credit scoring, and the changes it enabled in lending, increased profits by roughly a thousand dollars per loan. The effect is substantial: at the time, the average loan principal was around nine thousand dollars. We also identify two distinct channels through which better information improved loan profitability. First, credit scoring allowed the lender to set different down payment requirements for different applicants, creating a higher financing hurdle for severe default risks. Second, credit scoring allowed the lender to target more generous loans to low risks, increasing the quality of cars these consumers could purchase. We trace out these two channels — the screening out of marginal borrowers and the improved targeting of credit to infra-marginal borrowers — and show that together they explain the overall increase in profitability quite well.

The availability of detailed transaction-level data from before and after the adoption of scoring allows for a straightforward empirical approach. We first classify potential borrowers by assigning each loan applicant to a credit category using a rule that mirrors the lender's assignment following adoption. We then construct measures of profitability and related performance metrics — "close rates" on auto purchases, car choices, financing decisions, repayment behavior and recoveries — and compare how these metrics changed, both on aggregate and for the stratified groups, with the advent of credit scoring. Finally, we translate these changes into dollar terms by decomposing profits into separate components: the probability the applicant becomes a borrower, the "size" of investment in each borrower, and the return in terms of loan payments actually made.

We observe very different changes in lending patterns to high and low risk applicants. Following the adoption of credit scoring, high risk applicants saw their required down payment increase by more than 25 percent, and the close rates for this group fell notably. Default rates for high risk borrowers also fell, consistent with the idea that "marginal" borrowers were screened out by the higher down payment. When we translate these effects into dollar terms, we find that the improved loan repayment was largely responsible for what we measure to be about 1,200 dollar increase in profit per high-risk loan.

The increase in profitability on loans to low risk applicants is of similar size, but the mechanism is different. We find that required down payments and close rates changed little for low risk applicants. Instead, we observe a substantial increase in the average loan sizes and car quality. Even though default rates did not change much, we show that the larger loans had a substantial profit impact due to the high interest rates charged in the market. The increased "size" of each investment is largely responsible for the dollar increase in profit per low-risk loan. Both the changes we document in lending behavior are consistent with how firms theoretically should respond to the availability of better information about borrower risk. We elaborate on this point below by laying out a simple profit-maximizing model of lender behavior that captures the key features of the market we consider.

A useful feature of the episode we study is that most salient features of the lending environment, such as advertising, car pricing, salesperson incentives, and the composition of the applicant pool, remained stable during the periods before and after credit scoring was adopted. This makes for a relatively clean observational setting. While we cannot rule out every possible confounding change in the environment, particularly idiosyncratic shocks at specific dealerships, a variety of robustness checks support the general story we outlined. We show that the inclusion of controls for applicant quality and local economic conditions has little effect on any qualitative conclusions one might draw, and if anything slightly increases the estimated quantitative effects. We also show the effects we identify are observed *at every* dealership, consistent with the centralized change in lending practices we describe. Our analysis of the effects of down payment requirements and loan sizes is also consistent with results reported in Adams, Einav and Levin (2009) and Einav, Jenkins and Levin (2008). Those papers use more recent data from the same lender and exploit sharp changes in the pricing schedule to estimate the effects of alternative pricing on loan originations and subsequent loan performance.

The last part of the paper looks at the differential impact of credit scoring across dealerships in order to gauge its organizational implications. Research by Stein (2002) and others suggests that automated loan underwriting might involve a trade-off, with the increased use of "hard" information crowding out the production and use of "soft" information (see also Berger et al., 2004). This general line of thinking indicates that credit scoring might be a substitute for local managerial inputs, particularly if in the absence of scoring dealers differ in their ability to appropriately tailor loan terms to buyers. On the other hand, if dealers differ primarily in their ability to close sales and encourage repayments, one might expect credit scoring to be complementary to managerial ability. The difference has direct organizational implications. In both cases, increased use of credit scoring could increase the amount of centralized decision making, but under the latter hypothesis, local management's role could be even more (rather than less) important.[1]

We show that prior to credit scoring, there was in fact dramatic variation across dealerships in profitability, related primarily to differences in default rates. The advent of credit scoring did not collapse this variation, as one might expect if the initial variation was due to differences across dealerships in their ability to target loan offers. At the same time, we describe how more intensive use of data increased the fine-tuning done by headquarters. Therefore our findings provide an example in which the adoption of information technology

[1]Bloom et al. (2008) provide an interesting analysis of the multiple possible effects on information technology adoption on organizations.

increased the scope of headquarters decision-making without lessening performance variation at the local level.

This paper relates to a significant practitioner literature on credit scoring models. Much of this research focuses on statistical methods for predicting default (e.g. Hand and Henley, 1997; Straka, 2000). We focus on the complementary question of how credit scoring ultimately gets used. In this sense, the paper is closer to a smaller academic literature on the effect of information technology on lending. Much of this work has focused on bank practices in lending to small businesses (see, e.g. Petersen and Rajan, 2002; Frame, Srinivasan and Woosley, 2001; and Akhavein, Frame and White, 2005). Several of our own papers (Adams, Einav and Levin, 2009; Einav, Jenkins and Levin, 2008; and Jenkins, 2008) analyze the subprime auto market in more detail and use similar data, although they focus on different aspects of the market.

2 Data and Environment

2.1 The Lending Environment

The company we study specializes in making auto loans to consumers with low incomes or poor credit records. During the period we study, the company's average loan applicant had an annual household income of around 28,000 dollars. Almost a third of the applicants had no bank account, and only 14 percent owned their own home. A large majority of loan applicants had a FICO score below 600. Low FICO scores frequently reflect a history of loan delinquencies or defaults, which is consistent with the credit histories of the loan applicants in our data. Indeed, over the six months prior to their loan application, more than half of the company's applicants were delinquent on at least 25 percent of their debt. Such credit histories make it highly unlikely that the applicants could obtain a standard, "prime" auto loan, leading such applicants to seek other sources of credit. Indeed, the credit report for the average applicant recorded about nine credit inquiries.[2]

The lending process in the market operates as follows. Consumers fill out an application

[2] A credit inquiry is recorded on one's credit report when he or she attempts – successfully or unsuccessfully – to obtain credit. Most inquiries stay on a credit report for up to two years.

when they arrive at a dealership. They work with a sales representative and the dealership manager to select a vehicle and discuss financing terms. About forty percent of the loan applicants we observe purchase a car. The purchased cars typically are five to seven years old, with odometer readings in the 65,000 to 100,000 range. The average sale price is eight or nine thousand dollars, which represents a notable markup over the dealer cost (see Table I). Buyers are required to make a down payment, but usually finance about ninety percent of the purchase price. The financing terms are relatively standard across our sample. Buyers are expected to make regular payments at the dealership for a fixed term, typically around three years, and interest rates on the loans are high reflecting the risk of the borrower pool. Annual interest rates average close to thirty percent in our sample.

A central feature of the market is that consumers tend to be tightly cash-constrained. In earlier work, we used abrupt changes in the pricing schedule to estimate demand elasticities (Adams, Einav and Levin, 2009). A striking finding was that a loan applicant's probability of purchasing falls sharply when faced with a higher required down payment. We estimated that every hundred dollar increase in the minimum down payment reduces the purchase probability by two to three percentage points. Moreover, more than forty percent of buyers pay exactly the minimum amount down, and these "marginal" purchasers represent substantially worse default risks than buyers who pay more than the minimum down (Einav, Jenkins and Levin, 2008).

The role of the down payment in screening out marginal buyers is important for understanding the way in which risk-based pricing affects loan originations. In the period prior to the adoption of credit scoring, all buyers were required to make a down payment of at least 600 dollars. After credit scoring was put in place, minimum down payments were held constant or even modestly decreased for lower risk borrowers, but increased to as much as 1,500 dollars for high risks. As we will see, this increase helps explain why the fraction of applicants purchasing a car, and the subsequent default rate, fell in the period after credit scoring was adopted.

As can be seen in Table I, defaults during the repayment period are common and tend to occur relatively early in the repayment period. About 35 percent of loans default during the first year of repayment. Less than forty percent are repaid in full.[3] Following a default, the lender attempts to recover the car, and generally succeeds, but frictions in the recovery process result in a relatively low dollar value of recoveries after expenses are netted out (Jenkins, 2008). The average recovery in our sample was around 1,200 dollars, or around 25 percent of the original dealer cost of the car prior to the transaction.[4]

The combination of early defaults and low recoveries means that transaction outcomes have a bimodal pattern. Early defaults tend to result in losses, whereas fully paid loans can be quite profitable. Figure 1 documents this pattern by showing the distribution of transaction-level returns. For each sale, we computed the present value of borrower payments — the down payment, loan payments and recovery in the event of default — discounted back to the date of sale. We use a ten percent discount rate, which seems to be in line with industry standards. Neither the calculation here nor similar calculations later in the paper are very sensitive to using a somewhat higher or lower number.[5] We then divided the present value of borrower payments by the dealer cost of the car, providing a highly accurate overall rate of return on each transaction. The striking bimodal distribution of returns presented in Figure 1 illustrates the benefits of being able to identify the more credit-worthy applicants from those who are relatively more likely to default.

2.2 Implementation of Credit Scoring

The lender we study adopted credit scoring in the end of June 2001.[6] Prior to this time, the company did not use the credit bureau histories of prospective borrowers. Employees at the

[3]These are significantly higher default rates than those reported by Heitfield and Sabarwal (2004) in their study of securitized subprime auto loans, reflecting the relatively poor credit quality of the borrowers in our sample even compared to other subprime populations.

[4]This is for several reasons. In more than a quarter of defaults, for instance, it is hard to find the borrower, leading to a lengthy and costly recovery process. About a third of defaults are directly associated with a decrease in car value, such as mechanical breakdowns, car theft, and accidents (without maintaining appropriate insurance). See Jenkins (2008) for more details.

[5]Spcifically, we ran all the analyses in the paper using discount rates of 5 and 15 percent, and the results hardly change.

[6]To the best of our knowledge (which relies on conversations with the company's executives), there was nothing particularly special about the timing of implementation. In fact, many executives associate the company's idea to adopt automated credit scoring with the hiring of a senior executive who had quantitative background (and affection) in the late 1990s. Devloping, testing, and implementing the idea has taken several years.

dealership were responsible for eliciting information from applicants during the sales process and much of this information was not formally recorded. Prospective buyers were asked for basic information about their income, family and work status, scheduled debt payments and so forth, and as noted above all buyers were required to make at least a 600 dollar down payment. This traditional approach to lending was typical of the high-risk auto loan market at that time.

With the adoption of credit scoring, the company began to pull information from the major credit bureaus and use a proprietary algorithm to assess each applicant's risk profile. The scoring algorithm achieves impressive risk stratification. If we look at loans made in the first year after credit scoring began, borrowers in the top third of the applicant pool in terms of expected risk were 1.65 times as likely to repay a loan in full as borrowers in the bottom third (50.3 compared to 30.5 percent, respectively). A natural question is why the company uses its own scoring algorithm rather than a potentially cheaper metric available from the credit bureaus. Our understanding from discussions with experts is that a specialized scoring model may have particular value for niche markets such as this one. Standard credit models are designed to broadly assess the entire range of consumers, while those in our data are clustered at the low end of the credit spectrum. Lending to this part of the distribution requires separating consumers with transitory bad records from persistently bad risks, as opposed to simply identifying red flags in a consumer's history.[7]

The company uses the assigned credit score in several ways. As described above, a primary use of scoring is to establish a schedule for minimum down payments. Each applicant is required to pay at least some fixed dollar amount down; the amount depends on the applicant's credit score but not on the car being purchased. The credit scores are also used to match customers with appropriate cars. An applicant deemed a better risk is eligible to obtain financing for a larger range of vehicles, in particular newer, lower mileage cars that are more expensive. Applicants with better credit scores, however, do not qualify for any kind of automatic price discount. Finally, borrowers at a given dealership pay similar interest rates

[7] Indeed, beyond the standard and generally used FICO score, the credit bureaus also sell lenders more specialized scores, associated with default risks in specific markets, such as mortgages or auto lones. Presumably the benefit from a proprietary and customized algorithm is higher when the product and/or the customer base is more unique.

regardless of their credit score, as the rates are constrained by usury laws, and are clustered at, or close to, the relevant state interest rate cap.

2.3 Data

We focus our analysis on the pre-credit scoring period from January 2000 through December 2000, and the post-scoring period from July 2001 to June 2002. We drop the first half of 2001, when the company adopted a simple income cut-off to set minimum down payments in anticipation of credit scoring.[8] Finally, we include applications and sales data only from dealerships for which we have complete data for both the pre- and post-scoring periods.[9]

We compare full year periods rather than shorter pre-and-post windows for two reasons. First, the market has strong seasonality patterns: business peaks from February to April when many prospective buyers receive income tax rebates that facilitate down payments (Adams, Einav and Levin, 2009), and there is a slowdown around the December holidays. Second, although we can point to a specific date in late June 2001 on which dealers were required to use applicant credit scores in lending decisions, the practical day-to-day adjustments required for a successful implementation started earlier and continued later, which makes it more interesting to analyze changes over a moderate time period than a very narrow window.[10]

On the other hand, one reason to focus on a single year rather than longer run effects of credit scoring is that we are able to consider a period where other features of the lending environment remained constant. During the period we study, the sales and financing process and the incentive structure for salespeople and dealership managers were stable.[11] We also have little reason to believe that the inflow of prospective buyers into dealerships was affected

[8] We have looked at this period in some detail though we do not report the analysis. Perhaps not surprisingly, this intermediate approach led to intermediate outcomes.

[9] In Adams, Einav and Levin (2009) and Einav, Jenkins and Levin (2008), we used data from the post-scoring period, allowing us to expand the number of dealerships, applicants and borrowers in the post-period by roughly 50 percent relative to the (already large amount of) data we use here.

[10] We looked at time-series pictures around the implementation date, but between the seasonality and month-to-month variability it is hard to draw very sharp conclusions about the exact pace and timing of outcome changes.

[11] In fact, in late June 2002 the company significantly altered the incentive structure that governs loan origination. Thus, using data on loans originated after June 2002 would potentially confound the effects of credit scoring and incentives.

by the implementation of credit scoring. The company did not change its marketing and customers have little way of knowing the specific financing terms for which they qualify without visiting the dealership and filling out the loan application. This stability can be seen in Table I. Applicant characteristics are similar before and after credit scoring went into effect.

One qualification to this is that the number (but not the composition) of loan applicants was somewhat lower in the year after credit scoring, only 88% of the number in the year before scoring.[12] We are not aware of notable changes in the competitive environment, but one possible explanation is broader macroeconomic changes. Economic growth was fairly strong through the first half of 2000, but then slowed until the fourth quarter of 2001. To account for this in our analysis, we use data on local unemployment rates and local housing prices as controls in our empirical specifications. We also focus our analysis on the screening of applicants, the characteristics of loans made to borrowers, and their subsequent performance rather than try to explain the flow of customers into dealerships.

Table I shows significant changes in these basic operating metrics between the pre-scoring and post-scoring periods. The fraction of applicants who became buyers (the "close rate") dropped by about 15 percent, the average quality of cars sold increased (for example, the average odometer read was 7,000 lower after credit scoring), transaction prices and down payments were significantly higher, defaults were lower and loan revenues substantially increased. Overall, the firm's profitability increased markedly over the period, both on a per-transaction and a per-applicant basis.

3 A Model of Loan Origination

A simple model of loan origination helps to frame our empirical analysis. Consider a dealer-lender that faces customers of varying creditworthiness. Let θ denote the underlying credit quality of a given applicant. The company can make each applicant an offer consisting of a car at some quoted price, and financing terms (minimum down payment, interest rate and

[12] Note that to preserve the company's confidentiality, we do not report the exact number of loan applicants in Table I. Instead we report numbers of applicants and buyers as fractions of the number of loan applicants in 2000. For statistical inference purposes, these numbers are all quite large.

length of loan).[13] To keep things simple, we focus on the choice of minimum down payment and car quality as these seem to be the primary levers adjusted in our data.[14]

Let d denote the minimum down payment and q the quality of the car. Suppose that a car of quality q costs the firm $c(q)$ and is priced at $p(q)$. Assuming buyers pay d down and finance the remainder of the purchase by borrowing $L = p(q) - d$, the expected profit from a given applicant with creditworthiness θ can be written as:

$$\Pi(d, q, \theta) = G(d, q, \theta) \cdot [d + M(L, q, \theta) - c(q)]. \tag{1}$$

Here $G(d, q, \theta)$ is the probability of purchase and $M(L, q, \theta)$ is the present value of loan payments, which might depend on what car the borrower is driving as well as the size of her loan L and her creditworthiness θ.

We expect both the required down payment and car quality to affect the quantity of loans originated and their characteristics. Holding the car quality fixed, a reduction in the down payment is likely to increase the probability of sale and automatically will increase the amount borrowed. Similarly, holding the down payment fixed, offering a more valuable car is likely to increase sales, and the amount borrowed.[15] Intuitively, both effects are more desirable if the applicant is a better credit risk, suggesting the firm should offer better cars and lower down payments to more credit-worthy applicants.[16]

The preceding logic has immediate implications for the use of credit scores. Prior to scoring, the lender set a single down payment for all applicants, and dealers had to match

[13] The firm could also choose to offer, and in practice does offer each applicant a *menu of contracts*, e.g. a choice of cars, a choice of loan sizes, etc. Our basic conclusions about the value of information could be extended to this case (basically by interpreting $a(s)$ as a menu of offer vectors rather than a single vector), but it seems simplest to convey the ideas without introducing this complication.

[14] In the periods both before and after credit scoring, the majority of loans were offered at state interest rate ceilings, loan lengths were relatively standard, and car prices were mainly driven by a simple markup rule.

[15] A more valuable car may also increase the borrower's incentives to make loan payment. It is not entirely clear how this might interact with a borrower's risk type, so we do not focus on it in thinking about differential offers across risk groups.

[16] The easiest way to formalize this intuition is to make assumptions about the profit function $\Pi(d, q, \theta)$. Suppose that Π is strictly concave in (d, q) so it has a unique maximum for each θ. The intuitive idea that if the firm could observe θ it would set lower down payments and sell better cars to higher θ customers — i.e. that $d(\theta)$ will be decreasing in θ, and $q(\theta)$ increasing — is implied by an assumption that either Π or $\log \Pi$ are supermodular in $(-d, q, \theta)$, e.g. that $\Pi_{q\theta} \geq 0 \geq \Pi_{d\theta}, \Pi_{dq}$.

applicants with cars based only on an informal assessment of creditworthiness. To model this, we can imagine that dealers have access to some noisy signal t about creditworthiness θ and pre-scoring the firm was limited to a policy $d, q(t)$. With scoring, the firm also has the credit score s, and so it can use a more finely targeted policy $d(s), q(s,t)$.

We then can write the firm's problem in the absence of credit scoring as:

$$\max_{q(t),d} \mathbb{E}_t \left[\int \Pi(d, q(t), \theta) f(\theta, t) dF(\theta|t) \right], \tag{2}$$

and its problem with credit scoring as

$$\max_{q(s,t),d(s)} \mathbb{E}_{s,t} \left[\int \Pi(d(s), q(s,t), \theta) dF(\theta|s,t) \right]. \tag{3}$$

The discussion above suggests that the firm will want to use the available information to match better risks with higher quality cars and lower down payment requirements. Assuming that higher values of s and t are indicative of higher credit quality, we should expect the optimal policy with credit scoring to have $q(s,t)$ increasing in both s, t, and $d(s)$ decreasing in s. Absent credit scoring, the firm has to use a coarser policy that strikes a balance between the optimal offers for different risk groups. Nevertheless, we would still expect $q(t)$ to be increasing in t to match better risk with better cars.[17]

We can summarize the empirical content of the preceding discussion as follows. First, holding fixed the applicant pool, credit scoring should increase profits due to the firm being able to target offers more accurately. Second, applicants with better credit characteristics should be offered better cars both before and after credit scoring. Third, the differential in car quality between high and low-risk borrowers should increase with the advent of credit scoring. Fourth, credit scoring should lead to minimum down payments that are lower for low risk applicants and higher for high risk applicants. Fifth, after scoring is adopted close rates should increase for low risk applicants and decrease for high risk applicants. These are the main empirical patterns we will document and quantify in the next sections.

[17] Again, it is straightforward to formalize this discussion. Suppose that in addition to the assumption about Π in the prior footnote, (s, t, θ) are affiliated random variables with joint distribution F. Then the optimal policies $q(s,t), d(s)$, and in the no scoring case $q(t)$, will have the stated properties.

4 Empirical Strategy

4.1 Constructing Matched Applicant Pools

The adoption of credit scoring allowed the company to make systematically different offers to loan applicants with different risk profiles. Our basic analysis therefore relies on comparing the experiences of different types of loan applicants in the periods before and after scoring was adopted. For the period subsequent to adoption, we observe the credit score assigned by the company and the relevant information on which it was based, although not the exact algorithm. For the period prior to adoption, the lender collected less detailed data; we observe basic financial and demographic information for each applicant rather than a complete credit history.

To construct comparable risk groups in the two periods, therefore, we construct a risk measure that classifies applicants into low, medium and high risk using variables that are in the data for both periods, and then use this risk classification for *both* periods. To do this, we model each applicant's risk as a function of his or her household income and debt-to-income ratio. We assign each applicant to a cell based on the decile of his or her household income and debt-to-income ratio. We then assign each cell a risk category in a way that minimizes the distance in the post-scoring period between our assignment and the company's, subject to the constraint that our classification be monotone in both household credit variables. The Appendix provides details on the procedure. We note that our findings are not particularly sensitive to the exact classification scheme; we experimented with several and obtained nearly identical results for the analysis that follows.

Table II provides summary statistics for each risk category in the periods before and after the credit scoring . Low and medium risk applicants were much more likely to become buyers than high risk applicants, and this difference increased in the post-scoring period. Low risk buyers also tended to purchase more expensive cars in both periods. This difference also increased in the later period. Finally, despite taking larger loans, the lower risk applicants have lower default rates.

One point to emphasize is that our risk classification is imperfect. Ideally, we would have access to full credit histories for all applicants and construct risk groups by applying the company's algorithm retrospectively to the pre-scoring applicants. Relative to this approach, our construction may classify as low risk some applicants that the company treated as high risk, and vice versa. As a result when we look at the *differential* effect of credit scoring on low and high risk applicants, our estimates may underestimate the impact of credit scoring. As we will see, however, the differential effects we observe are quite large even with our current classification scheme.

4.2 Measuring the Effect

We measure the effect of credit scoring by estimating the change in different outcome variables between the pre-period (January-December 2000) and the post-period (July 2001 - June 2002).

The results we report rely on regressions of the following form:

$$y_i = \alpha_{R(i)} + \beta_{R(i)} D_i + X_i \gamma + \varepsilon_i, \tag{4}$$

where i is an individual, y_i is an outcome variable of interest, $R(i)$ is the individual's risk category (low, medium, or high), D_i is a dummy variable equal to one if the individual appeared at the dealership following the advent of credit scoring (that is, in the post-period), and X_i is a set of controls.

From this model, we can define:

$$y_{pre,r} = \mathbb{E}\left[y_i | D_i = 0, R(i) = r\right] = \alpha_r + \mathbb{E}\left[X_i | D_i = 0, R(i) = r\right] \gamma, \tag{5}$$

$$y_{post,r} = \mathbb{E}\left[y_i | D_i = 1, R(i) = r\right] = \alpha_r + \beta_r + \mathbb{E}\left[X_i | D_i = 1, R(i) = r\right] \gamma, \tag{6}$$

so that $y_{pre,r}$ is the expected outcome for an applicant of risk type r with average characteristics in the pre-period, and $y_{post,r}$ is the equivalent quantity for the post-period.

Their difference, $\Delta y_r = y_{post,r} - y_{pre,r}$, is:

$$\Delta y_r = \beta_r + \left(\mathbb{E}\left[X_i | D_i = 1, R(i) = r\right] - \mathbb{E}\left[X_i | D_i = 0, R(i) = r\right]\right) \gamma. \tag{7}$$

That is, the change in outcomes for risk group r can be decomposed into the estimated coefficient β_r, which we interpret as the effect of credit scoring, and the effect of changes in observable covariates within the risk group.

If both the pool of applicants and broader economic conditions were identical before and after the policy change, the second component of Δy_r will be zero, and β_r will reflect the same differences between the average outcomes for group r across the time periods observed in our earlier summary statistics. To the extent that the applicant pool and economic conditions changed, Δy_r will differ from β_r. Below we report estimates of β_r for regressions that gradually add more controls, allowing us to see the contribution of observable shifts in applicant characteristics and economic conditions. We discussed above that changes in the applicant pool were limited; this is reflected below in the fact that controlling for the composition of the applicant pool has little effect on our estimates of β_r. The estimates we report do suggest that economic conditions may have led to modest decreases in loan performance during the post-period, so that the effects of credit scoring may be a bit understated in the raw numbers.

One limitation to our observational data approach is that we cannot rule out some unobserved change in the lending environment that might have contributed to, or even independently generated, the effects we document below. We believe the latter is highly unlikely. The inclusion of observed controls does not attenuate the estimated effects, and the set of confounding events required to generate all the predicted effects we observe would need to be quite special. It is possible that there was some broad ongoing trend in the attitude of borrowers that we do not account for. If so, one might expect it to have had a fairly uniform effect on the risk groups we construct. In this case, the differences (across risk categories) between the β_r's that we emphasize below will still be informative about the impact of credit scoring. Many of the other unaccounted for changes that naturally come to mind (a large layoff, or the opening of a local competitor) would likely to have had a targeted effect at certain dealerships. The inclusion of dealership dummies accounts for these possibilities to some extent, and we also will see in Section 6 that essentially *all* dealerships experienced similar qualitative changes between the two period, something we might not expect if there were important local, risk-group specific, unobserved trends.

4.3 Profitability and Other Outcomes of Interest

To assess the effect of credit scoring, it is useful to identify several measures of profitability. In the short run, it seems natural to take the flow of applicants as given, and to view the firm's objective as maximizing per-applicant profits.

We can write the operating profits from applicant i as

$$\Pi_i = Sale_i \cdot [DP_i + LP_i + REC_i - C_i]\,. \tag{8}$$

Here $Sale_i$ is an indicator variable equal to 1 if i buys a car, DP_i is the down payment, C_i is the cost of the car offered to i, LP_i is the present value of loan payments and REC_i is the present values of recoveries in the event of default (or zero if the loan is fully repaid).[18] In our data, LP_i depends primarily on the transaction price (which after subtracting the down payment, determines the loan principal), and whether and when default occurs. More generally, it depends on the loan length and the interest rate, but as these did not change much with credit scoring, we do not discuss them separately.

In the longer run, and particularly in obtaining external financing, one may be more interested in the rate of return on capital. Restricting attention to buyers rather than applicants, we can define the return on sale i as:

$$\Pi_i/C_i = DP_i/C_i + LP_i/C_i + REC_i/C_i - 1. \tag{9}$$

Below, we report regressions where the outcomes of interest are per-applicant profit and its components, regressions where the sample is restricted to buyers and the relevant outcomes are per-sale profit and its components, and regressions where the sample is buyers but the dependent variables are rate of return and its components. As we will see, the approaches yield similar insights, but a comparison is useful to facilitate interpretation.

[18] As mentioned earlier, we use an annual interest rate of 10 percent to value the stream of payments and recoveries, and also experimented (in unreported regressions) with rates of 5 and 15 percent and verified that this assumed rate doesn't drive any of the results.

5 Empirical Results

We report our regression results in Table III. In the first panel (Table III(a)), the unit of observation is an applicant, and we measure profit and its components in dollar terms. In the second panel, the unit of observation is a buyer, and we look at changes in profit per buyer measured in dollar terms. The third panel is also at the buyer level, with the dependent variables being rates of return.

Each panel has a similar structure. For each outcome of interest, we report in the left-most column its grade-specific average before credit scoring, while the remaining columns report estimates of the effect of credit scoring, β_r. Column (1) presents these estimates with no additional controls (essentially replicating the summary statistics of Table II). In column (2), we add dealership and calendar month fixed effects, and the household total (monthly) income, residual income, and debt-to-income ratio of each applicant or buyer. In column (3), we also include measures of local economic conditions (at the MSA in which the dealership is located) at the time of sale and over the initial twelve months of the loan.[19] The first set of covariates is intended to control for compositional changes in the applicant or buyer pool within a given credit category. The economic indicators are intended to account for macroeconomic changes that might impact close rates or borrower repayment.

5.1 The Effect of Credit-Scoring on Profitability

In all of our specifications, we find a very strong effect of credit scoring on profitability. We estimate that profits per transaction increased by over 1,000 dollars for each risk category, with the rate of return on capital increasing by an order of 15-20 percent depending on the exact specification (Tables III(b) and III(c)). At a per-applicant level, we find that profits increased by almost 600 dollars for lower risk applicants and by 546 dollars for medium risk applicants. We find a slight decrease in profitability per applicant for high risks, reflecting the fact that the close rate declined substantially for this group and we calculate transactions

[19] Specifically, we construct eight variables to capture local economic conditions. Six are related to local unemployment rates: the average level, the average change, and the standard deviation of (monthly) local unemployment rates in the previous six months and in the subsequent 12 months. The last two are the annual change in the (quarterly) local housing price index for the previous 6 months and subsequent 12 months.

in this category to have been profitable prior to the advent of credit scoring.

This last conclusion depends somewhat on how we account for the fixed costs associated with selling, handling, and collection activities associated with each loan. The company estimates these costs as being in the ballpark of 1,000 dollars. If we were to include this as a cost for every transaction, high risk sales would have been only marginally profitable prior to credit scoring, and we would conclude that profits per applicant increased by 105 dollars per applicant for the highest risk category.[20] This adjustment also makes the rate of return effects even more dramatic, implying more than doubling of the company's overall rate of return.

5.2 How did Profits Increase?

To understand the source of the profitability gains, it is useful to look at the separate components of profit. Here we focus discussion mainly on the estimates in the first column of Tables III(a) and III(b). What we want to emphasize is the very different channels through which profits increased for the better and worse risk groups.

The conventional story of credit scoring is apparent for high risks. Before credit scoring almost one in four applicants in our high-risk category became a buyer; with credit scoring this was cut by more than half (Table III(a)). A likely cause of this change was the required down payment, which increased from 600 dollars prior to scoring to more than 1,000 dollars for the highest risk applicants. As noted above, increases in the down payment requirement have a remarkably large impact on purchasing decisions, and also lead to a better selection — that is, buyers who are just able to come up with the minimum down payment turn out to be substantially worse risks than buyers for whom this constraint is not binding (Einav, Jenkins and Levin, 2008). The results in Table III(b) are consistent with this selection effect. Default rates for buyers in the highest risk category fell from 70 to 62 percent, leading to about a 1,000 dollar increase in repayments.

Credit scoring had a very different effect on the lower risk applicants. For applicants

[20] This adjustment has little impact on the change in profitability from low and medium risk applicants because close rates for these groups hardly changed. Specifically, with the adjustment we estimate the effect on profits for low and medium risks to be 598 and 566 dollars per applicant, respectively (compared to 595 and 546 reported earlier).

with better risk scores, the company did not raise the minimum down payment requirement, and indeed close rates remained virtually the same (Table III(a)). Nevertheless profitability increased dramatically. Here the biggest factor appears to have been that lower risk applicants were allowed to take larger loans, leading them to purchase better cars, and leading the company to raise its markups on these cars. The incentive for the company to do this can be seen clearly in Table III(c). Prior to credit scoring, the transaction rate of return was significantly higher for lower risk buyers than for higher risk buyers (38 percent vs. 26-29 percent). With the ability to identify these buyers, it was possible to extend them more credit. Table III(b) shows the significant increase in car cost for the lowest risk buyers (431 dollars), an even greater increase – due to increased markups – in the price of these cars (1,125 dollars), and also the increase relative to buyers in higher risk categories.

To see how these different effects aggregate into an overall change in profit per-buyer, consider the high risk buyers first. Their down payments increased by 309 dollars, and loan payments by 1,021, from which we need to subtract a modest 184 dollar increase in car costs. Incorporating a small increase in recoveries leads to the 1,205 increase in profit per buyer reported in Table III(b). For the low risk buyers, car costs and car prices increased much more, by 431 dollars and 1,125 dollars respectively, and also loan sizes because the increase in down payments (of 234 dollars) did not increase enough to offset it. The increase in profitability of 1,059 can be therefore attributed almost entirely to the larger stream of loan payments received on the larger loans, almost 1,000 per buyer, plus a 306 increase in recoveries reflecting the initially higher quality of the cars.

5.3 Potential Confounding Factors

The preceding discussion focused on the first column of Tables III(a)-(c), in which we make no attempt to control for compositional or macroeconomic changes that might impact our results. Column (2) adds dealer and calendar month fixed effects, as well as individual characteristics. As we describe below, dealership performance varies substantially, and we have already mentioned the seasonality effects in the data. Nevertheless, the inclusion of these variables has virtually no effect on our estimates. This basically reflects the fact that within each of our credit categories, the composition of applicants and buyers did not change very much during the evaluation period, neither across dealers, nor across months, nor in terms of individual characteristics.[21]

Column (3) of Tables III(a)-(c) reports on specifications where we control for local (MSA-level) economic indicators related to unemployment and housing prices (see footnote 19). Repayment in the subprime market can be quite sensitive to employment shocks (Jenkins, 2008), and unemployment rates were somewhat higher during the post-scoring period. This suggests that our estimates of how much credit scoring increased profitability might be attenuated by adverse macroeconomic changes. Indeed when we control for these macroeconomic factors — interacting these controls with the risk category to allow for different risk groups to be affected differentially — we find somewhat larger effects at the per-buyer level (Table III(b)) and on a rate of return basis (Table III(c)). We estimate an increase in profit per buyer of 1,430 dollars for low risks and 1,277 for high risks when we include the full set of controls, compared to 1,059 and 1,205 in the baseline specification. There are corresponding changes in the estimates of the profit components, but nothing in the results leads us to revisit the qualitative interpretations above.

6 Differential Effects across Dealerships

We now investigate the differential impact of credit scoring across dealerships. We document that the basic changes in lending behavior and the increase in loan profitability described above were not limited to a specific set of dealerships, but occurred at all dealerships. We also explore the extent to which the adoption of credit scoring narrowed or widened the variation in performance across dealerships. A starting point for the latter analysis is the observation, described below in more detail, that prior to the advent of credit scoring there was a high degree of variation across dealerships in profitability and the differences are not easily explained by differences in the composition of applicants or observable economic conditions.

An interesting hypothesis in the organizational economics literature is that the adoption

[21] The results do not change noticeably if we leave out the individual characteristics (household income and debt-to-income ratio), or if we add additional characteristics (that we have only for buyers) such as the number of dependents or the time that the buyer has been living in his current address.

of "hard information" technologies such as quantitative risk assessment may crowd out the use of "soft information" obtained at the dealership level (Stein, 2002). For instance, think of the credit score s in our simple model (Section 3) as being strictly more informative than the local information t, rendering it of little value. If dealers differ in their ability to assess customers and act on their assessments, we would expect credit scoring to have the biggest effect at the poorest-performing dealerships and compress any variation in performance.

Of course there is also an alternative hypothesis that the most capable managers are the best positioned to benefit from a new technology. In our setting, the adoption of credit scoring was accompanied by changes in lending practices, such as guidelines for matching cars to buyers, and steeper down payments for high risk applicants, that effectively shifted control away from the local level. But to the extent that high-performing dealers were best able to implement these changes, we might still expect to see a widening in performance gaps between dealers. Finally, it is possible that the benefits enabled by credit scoring were largely orthogonal to the features of dealerships that make them differentially profitable.

6.1 Heterogeneity across Dealerships

We start by describing the heterogeneity across dealerships. Table IV, which follows a similar structure to that of Tables I and II, presents summary statistics for "high" and "low" performing dealerships. To construct the table we rank dealerships by their profit per sale in the pre-credit scoring period. We then report summary statistics separately for the top third of dealerships (Table IV(a)) and for the bottom third (Table IV(b)).

The table highlights two important facts. First, dealerships in the top third were dramatically more profitable than dealerships in the bottom third, earning about 1,000 dollars more per sale. The difference is not driven by differences in the composition of the applicant pools, which are similar in terms of applicant credit categories. We make this point more rigorously below in the context of a regression model for profitability that includes dealership fixed effects along with controls for applicant quality. Table IV also shows that there are few systematic differences between the top-performing and bottom-performing dealerships in the characteristics of their transactions: car costs, loan sizes, interest rates or loan terms.

The primary difference across dealerships is in their borrower repayment rates. Top performing dealerships have significantly lower default rates and a higher fraction of payments made, particularly for medium and high risk borrowers. This in turn generates a substantial difference in loan payments received and profitability. Again the difference in repayment rates does not appear to be generated by differences in observable features across dealerships such as the quality of the applicant pool. It seems likely, therefore, that higher performing dealerships are either more effective in their collection efforts or that the borrowers at these dealerships are more inclined to repay their loans for reasons that we cannot account for even with the rich individual-level borrower characteristics in our data.

6.2 Measurement

To measure how the adoption of credit scoring affected individual dealerships, we use a similar strategy to the one described in Section 4. We rely on a regression of the following form

$$y_i = \alpha_{d(i)} + \beta_{d(i)} D_t + \delta_{R(i)} + X_i \eta + \upsilon_i, \tag{10}$$

where we now allow the effect of credit scoring to vary across dealerships. As before, i is an individual, y_i an outcome variable of interest, $d(i)$ is the dealership involved in the transaction, $R(i)$ is the individual's risk category (low, medium, or high), D_i is a dummy variable which takes the value of 1 in the post-scoring period, and X_i is a set of other controls. We separate the credit category dummies from the rest of the controls because we vary the set of Xs but always control for credit category.

In this specification, the coefficient α_d represents the dealership effect prior to credit scoring, while the coefficient β_d represents the dealership-specific effect of credit scoring. The sum $\alpha_d + \beta_d$ is the dealership effect after credit scoring. A narrowing of the performance gap across dealerships would imply that $\alpha_d + \beta_d$ is less dispersed than α_d. Another approach is to consider the correlation between α_d and β_d. To the extent that our estimates have some sampling error, classical regression to the mean is likely to imply some degree of negative correlation, but a finding of positive correlation would indicate that the impact of credit scoring was largest at the top performing dealerships.

6.3 Profitability Differences across Dealerships

We focus our results on two specifications, where in both cases the outcome of interest is the dollar profit per sale. (An analysis of rate of return regressions, as in Table III(c), reveals an almost identical pattern.) Figure 2 shows kernel density plots of the cross-sectional distribution in the estimated dealership effects before (solid) and after (dashed) credit scoring, i.e. it depicts the distribution of the α_d's and the $\alpha_d + \beta_d$'s. Figure 3 presents scatter plots of the dealership-specific credit scoring effect β_d against the pre-credit scoring dealership effect α_d. In both figures, panel (a) comes from a regression specification without controls, and panel (b) from a specification that uses a full set of controls (as in column (3) of Table III).

The scatterplots indicate that *every dealership* appears to have experienced an increase in profitability from the pre-scoring to post-scoring period. The kernel density plots also show that the cross-sectional variation in profitability remained fairly constant, or if anything, increased slightly. The addition of controls does little to reduce the heterogeneity across dealerships, indicating that the differences in profitability do not seem to be driven by (measurable) differences in buyer characteristics or in local economic conditions.

The covariation in the scatterplots provides more detail into how the effect of credit scoring varied across dealerships. The estimates obtained without controls for economic conditions or buyer characteristics show a slight negative correlation between dealership performance and the effect of credit scoring. This is what one would expect if there were some underlying regression to the mean in dealership performance, or if lower performing dealerships systematically benefited the most from the advent of credit scoring. The fact that the overall dispersion in profitability remained constant, however, suggests the former interpretation. Moreover, when we add controls the negative correlation disappears. Specifically, Figure 3 shows essentially no relationship between a dealership's performance in the pre-scoring period and the effect of credit scoring on profit per sale. Because the dealerships did not experience uniform changes, this lack of correlation explains the increased cross-sectional dispersion in profitability shown in Figure 2.

6.4 Lending Changes at the Dealership Level

We emphasized above that credit scoring affected profits through two distinct channels: better selection and screening of high-risk applicants, and more generous financing for low risk applicants. The dealership-effect model described above allows us to confirm that these effects were not limited to a subset of dealerships, but were realized broadly. Rather than report an overwhelming number of regression coefficients, we focus on two key ones — dealership-specific changes in close rates for high-risk applicants, and changes in the quality of cars sold to low-risk buyers.

Figure 4(a) plots for each dealership the change in close rate for high risk applicants against the base close rate for these applicants. The plot shows that close rates for high risks declined at every dealership, to roughly 50 percent of the pre-scoring rate. Figure 4(b) plots for each dealership the change in car costs for low risk applicants against the average cost in the base period. Again the plot shows that virtually all dealerships increased the quality of cars sold to low risk customers. The fact that the firm-wide effects we described in Section 5 occurred at each dealership provide further confidence that they can be attributed to the lending changes enabled by credit scoring rather than being the product of omitted changes in the environment, many of which are likely to be highly local in nature.

How can we reconcile the uniform direction of lending changes across dealerships with the persistent variation in dealership performance? The likely explanation, as suggested above, is that dealers did not differ all that much in their allocation of credit across risk groups, in a way that might have been affected by credit scoring. Rather, dealerships seem to have different success in collecting on loans, and these differences appear to have survived the changes in loan origination. In this sense, we do not find much evidence of subtle within-organization changes brought on by the adoption of information technology.

7 Conclusions

In this paper, we reported detailed results on the adoption of automated credit scoring and the changes it enabled in lending at a large auto finance company. A primary conclusion is that the adoption of new credit scoring technology led to a large increase in profitability. Lending to the highest risk applicants contracted due to more stringent down payment requirements, and lending to lower risk borrowers expanded driven by more generous financing for higher quality, and more expensive, cars. We find that these effects were remarkably consistent across dealerships, and perhaps surprisingly that the impact of credit scoring did little to reduce the large performance differences across dealerships.

Several aspects of our analysis may be interesting to follow up in other contexts. Much of the academic and practitioner literature emphasizes how better information about customers enables more efficient screening of marginal borrowers; our work highlights how improved information technology also allows better customization of contract terms to infra-marginal borrowers. A related point is that in our setting the relevant margins of adjustment following the advent of credit scoring was not the interest rate, but rather the down payment and maximum loan sizes — i.e. the amount of leverage borrowers were allowed to take on. It has become increasingly clear that this leverage aspect of consumer borrowing, particularly in regards to the subprime market, deserves much more attention than it has generally received.

References

Adams, William, Liran Einav and Jonathan Levin, "Liquidity Constraints and Imperfect Information in Subprime Lending," *American Economic Review* 99(1), 49-84, March 2009.

Akhavein, Jalal, W. Scott Frame and Lawrence J. White, "The Diffusion of Financial Innovations: An Examination of the Adoption of Small Business Credit Scoring by Large Banking Organizations," *Journal of Business* 78(2), 577-596, March 2005.

Berger, Allen N., Nathan H. Miller, Mitchell A. Petersen, Raghuram G. Rajan, and Jeremy C. Stein, "Does Function Follow Organizational Form? Evidence From the Lending Practices of Large and Small Banks," *Journal of Financial Economics* 76(2), 237-269, May 2002.

Bloom, Nicholas, Luis Garicano, Raffaella Sadun and John van Reenen, "The Distinctive Effects of Information Technologies and Communication Technologies on Firm Organization," manuscript, Stanford University, 2008.

Einav, Liran, Mark Jenkins, and Jonathan Levin, "Contract Pricing in Consumer Credit Markets," manuscript, Stanford University, 2008.

Frame, W.S., A. Srinivasan, and L. Woosley, "The Effect of Credit Scoring on Small-Business Lending," *Journal of Money, Credit and Banking* 33(3), 813-825, August 2001.

Hand, D.J., and Henley, W.E., "Statistical Classification Methods in Consumer Credit Scoring: a Review," *Journal of the Royal Statistical Society: Series A* 160(3), 523-541, 1997.

Heitfield, Eric and Tarun Sabarwal, "What Drives Default and Prepayment on Subprime Auto Loans?," *Journal of Real Estate Finance and Economics* 29(4), 457-477, December 2004.

Jenkins, Mark, "Bad Luck or Bad Behavior: What Drives Default in Subprime Lending?," manuscript, Stanford University, 2008.

Petersen, Mitchell A. and Raghuram G. Rajan, "Does Distance Still Matter? The Information Revolution in Small Business Lending," *Journal of Finance* 57(6), 2533-2670, December 2002.

Stein, Jeremy, "Information Production and Capital Allocation: Decentralized versus Hierarchical Firms," *Journal of Finance* 57(5), 1891-1922, October 2002.

Straka, John, "A Shift in the Mortgage Landscape: the 1990s Move to Automated Credit Evaluations," *Journal of Housing Research* 11(2), 207-232, 2000.

Appendix

We describe the process we followed to construct the matched applicant pool. Recall that the main challenge arises because the company did not credit score applicants in the pre-period, and, moreover, did not collect all the individual characteristics which are used as inputs for the (proprietary) credit scoring algorithm. Therefore, to construct our matched applicant pools, we need to construct our own credit scoring algorithm, which relies on the individual characteristics that are observed both before and after credit scoring, income and debt-to-income ratio. To do so, we assume that applicants can be of one of three risk categories – high, medium, or low – and use the actual risk classification from the post-period as a guide.

Formally, the problem we try to solve is to find a function $f : \mathbb{R}^2_+ \rightarrow \{high, medium, low\}$, which maps applicants' income and debt-to-income ratio into one of the three risk categories. A naive approach (which turns out to do reasonably well) is to use the post credit scoring period, and in particular the high/medium/low risk category each applicant in the post period is classified to (by the company), and run an ordered probit regression of this classification on income and debt-to-income. Since the goal is to predict well, we allow for flexible functional form by generating ten decile dummies for income and debt-to-income, and fully interacting them. Given the estimation results, we then compute the predicted values for the predicted latent variable, order them over the 100 cells, and assign a risk category to each cell accordingly, in order to match the overall distribution of high, medium, and low risk categories in the post-period data (which are 29, 46, and 25 percent, respectively). We then assign each applicant in the pre-period data a credit category based on his or her income and debt-to-income cell. The top panel of Table A.I presents the results. It shows that the risk category is close to monotone in both income and debt-to-income ratio. That higher-income applicants are generally lower risk is intuitive. It turns out that, in our data, higher debt is also associated with lower risk. Presumably, for this population higher debt is associated with the extension of credit by other lenders, which serves to indicate creditworthy behavior, and this underlying correlation dominates any likely effect of debt burden on default risk.

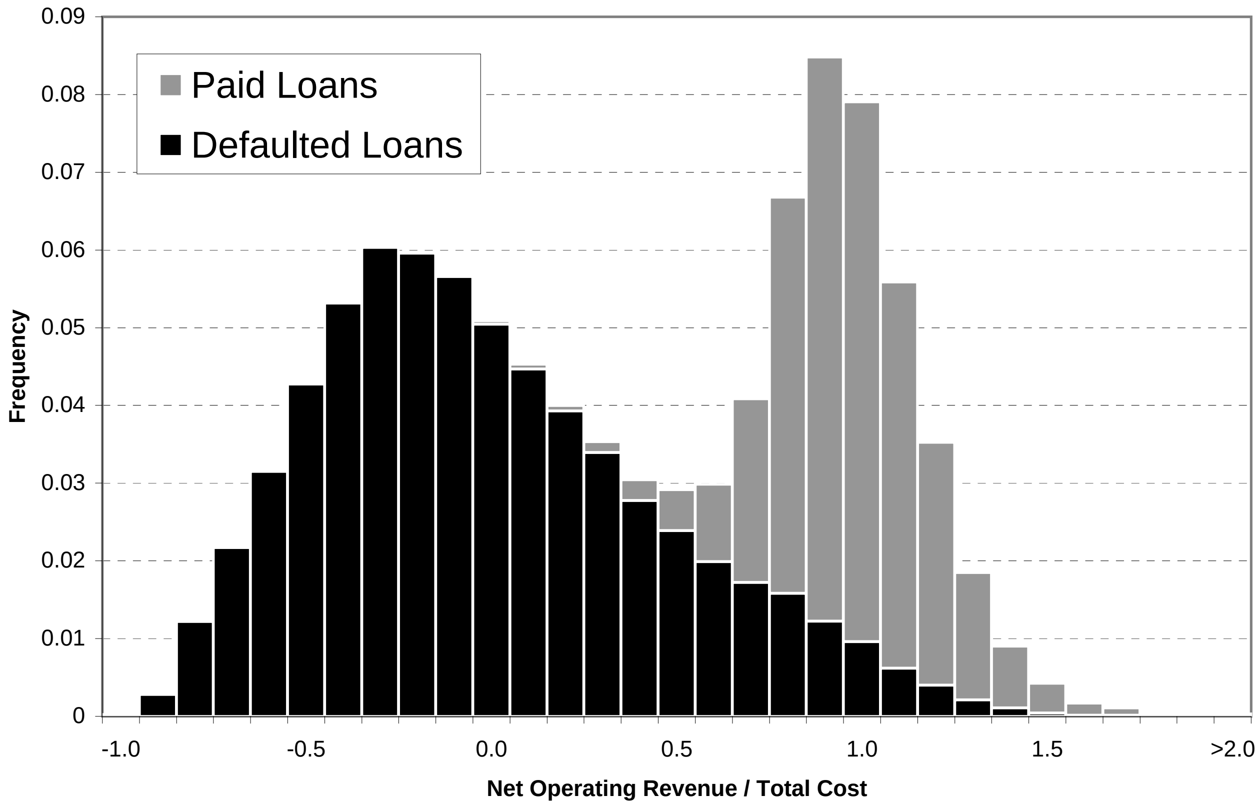
Our actual risk categorization is a small modification of the above described procedure. Motivated by the few cases of non-monotonicities in the top panel of Table A.I – which are likely driven by sampling errors – we reran this prediction model, under the restriction that $f(\cdot)$ is (weakly) monotone in both income and debt-to-income ratio, again characterizing each individual by the interaction of his income and debt-to-income decile dummy variables. Among the set of monotone mappings, we seek a mapping that meets two objectives: it matches the individuals' actual credit score, and it accurately predicts the fraction in the population of each risk category (as classified by the company in the post-period). Let $s_i \in \{H, M, L\}$ be applicant i's actual credit category and $f(x_i) \in \{H, M, L\}$ be individual i's predicted credit category. We then parametrize a loss function over prediction models, so that the optimal prediction model $f(\cdot)$ (within the set of monotone models) minimizes

$$\begin{aligned} &\sigma_1 \sum_i I\left(s_i \neq f(x_i)\right) + \sigma_2 \sum_i \left(I\left(s_i = L, f(x_i) = H\right) + I\left(s_i = H, f(x_i) = L\right)\right) + \\ &+\omega \sum_{j \in \{H,M,L\}} \left| \sum_i I\left(f(x_i) = j\right) - \sum_i I\left(s_i = j\right) \right|, \end{aligned} \tag{11}$$

where ω, σ_1, and σ_2 are non-negative parameters. That is, the first component in the loss function penalizes for wrong predictions, the second component increases the penalty for "really bad" predictions (predicting high risk although actual score is low risk, and vice versa), and the third component penalizes against deviation from the overall mix of high, medium, and low risks in the population.

We solve this constrained optimization problem numerically, by searching over the entire set of monotone functions. Based on many different trials, it seems that the prediction model is largely insensitive to the exact values of the weights σ_1, σ_2, and ω. The results presented in the paper use weights of $\sigma_1 = 1$, $\sigma_2 = 3$, and $\omega = 8$. The bottom panel of Table A.I reports its predictions. As one can see, it is similar to the results obtained from the ordered probit model (top panel), but it imposes monotonicity, and is slightly different for some marginal cells.

Figure 1: The Distribution of per-loan Rate of Return



Net Opertaing Profits = Down payment + PV of loan payments + PV of recoveries - Total cost
The histogram uses all observations used in the subsequent analysis, pooling th epre-period and post-periodn (see Table I).

Figure 2(a): Distribution of dealer average profit per sale (with no controls)

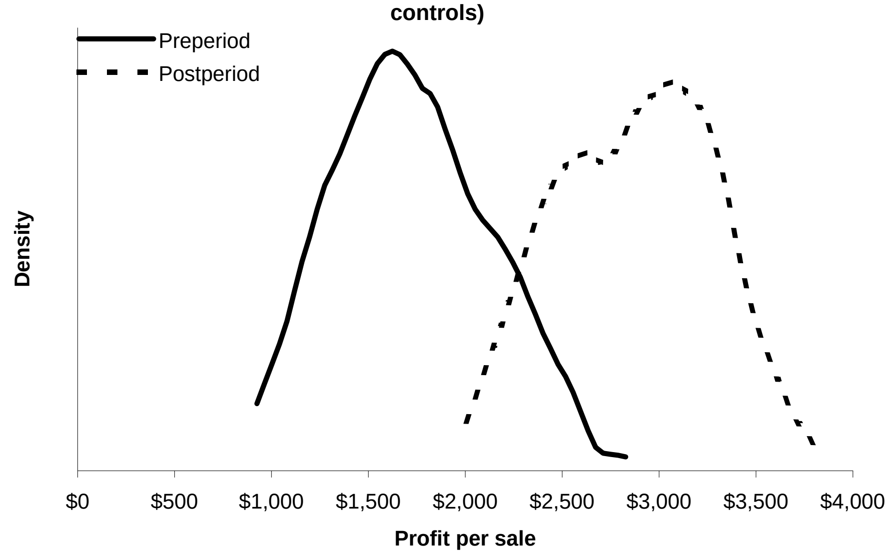
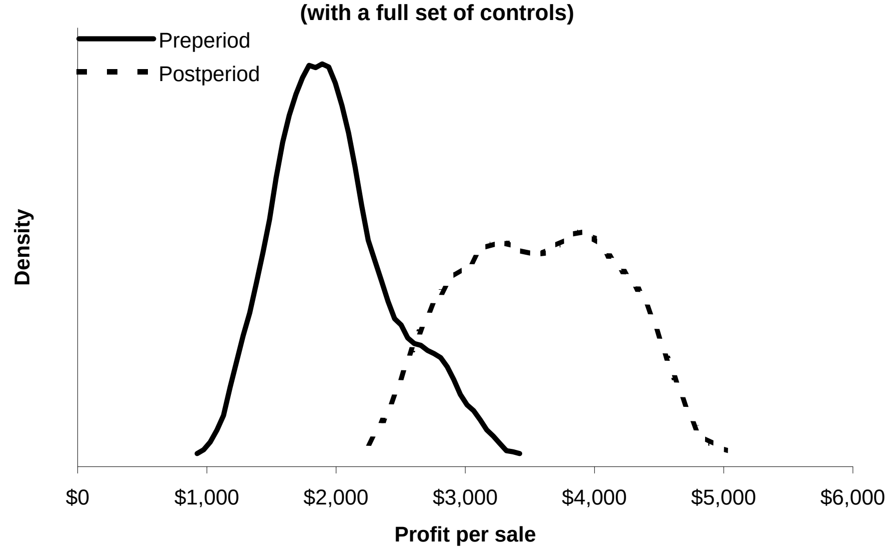
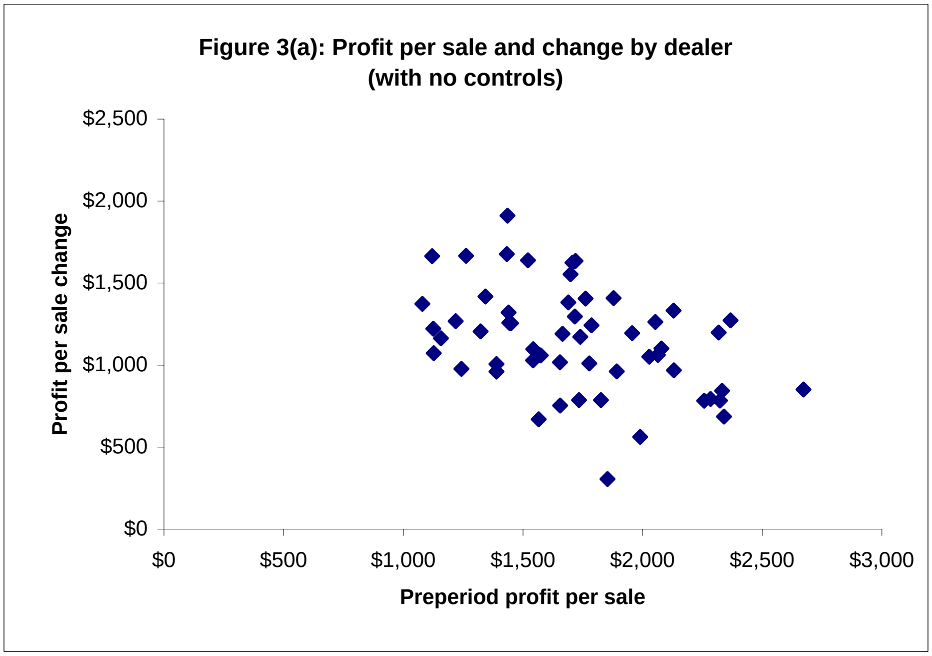
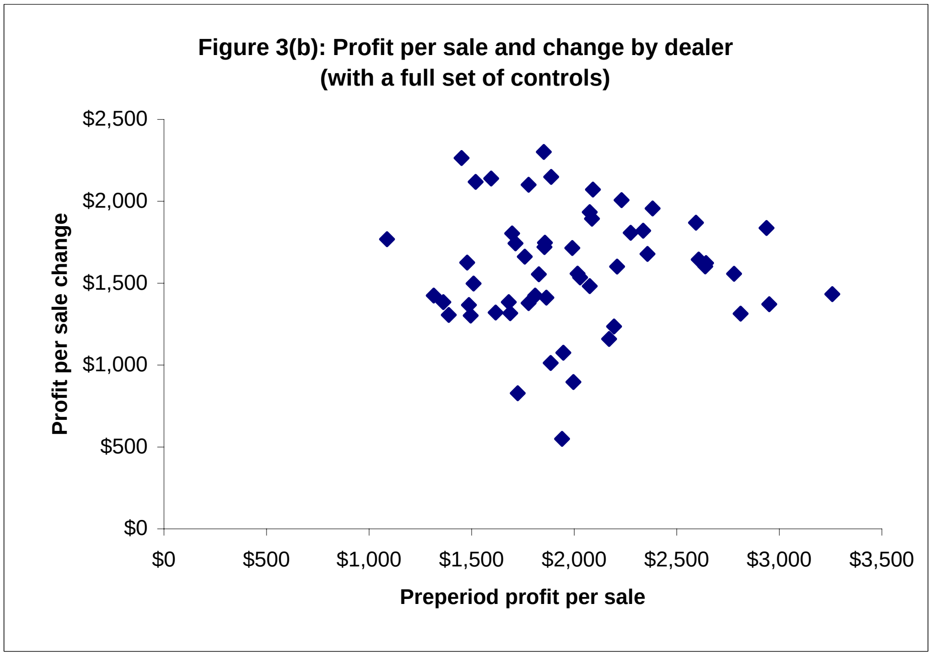


Figure 2(b): Distribution of dealer average profit per sale (with a full set of controls)



Each of the graph presents estimates from a regression of the form of equation (10) in the paper, with profit per sale as the dependent variable. The pre-period graph plots a kernel density of the estimated alpha's, and the post-period graph plots a kernel density of the estimated alpha+beta's. The top panel uses no other controls (except credit grade fixed effects), while the bottom panel uses a full set of controls (as in column (3) of Table III),





Each of the scatterplot presents estimates from a regression of the form of equation (10) in the paper, with profit per sale as the dependent variable. Each point reflects the delaership-specific alpha (pre-period profitability, on the horizontal axis) and beta (the change in profitability due to credit scoring, on the vertical axis). The top panel uses no other controls (except credit grade fixed effects), while the bottom panel uses a full set of controls (as in column (3) of Table III),

Figure 4(a): High-risk close rate and change by dealer

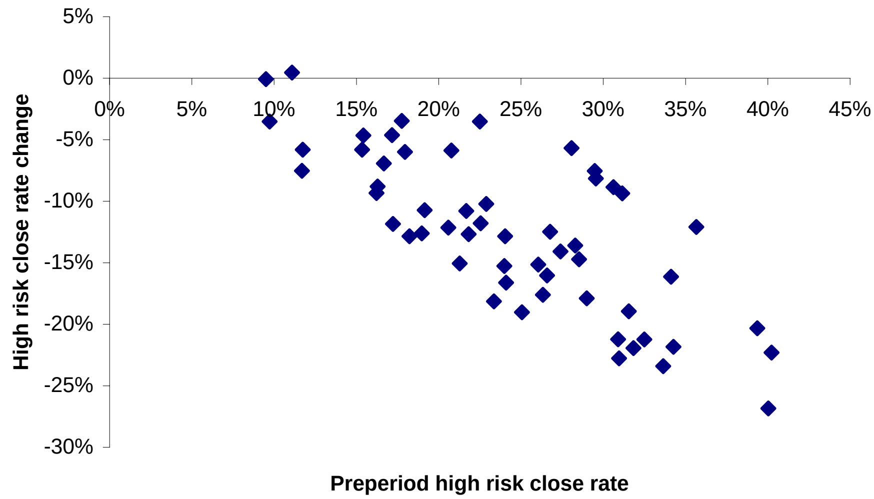
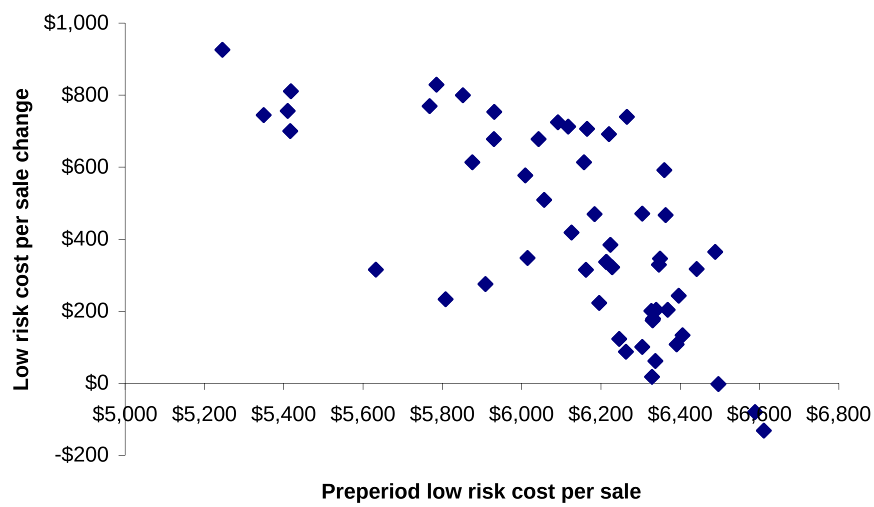


Figure 4(b): Low-risk cost per sale and change by dealer



Each of the scatterplot presents estimates from a regression of the form of equation (10) in the paper, with close rate of high-risk applicants (top panel) and car cost of low-risk buyers (bottom panel) as the dependent variables. Each point reflects the delaership-specific alpha (pre-period close rate / car cost, on the horizontal axis) and beta (the change in due to credit scoring, on the vertical axis). The regressions use no other controls (except credit grade fixed effects), but results with full sets of controls (as in column (3) of Table III) are very similar,

Table I: Summary Statistics

	January – December 2000				July 2001 – June 2002			
	Mean	Std. Dev.	5%	95%	Mean	Std. Dev.	5%	95%
Applicant characteristics	*N = 1.00*				*N=0.88*			
Applicant demographics								
Monthy income	2,214	973	1,204	4,000	2,256	975	1,238	4,000
Residual monthly income after debt payments	1,715	985	748	3,525	1,843	1,024	824	3,750
Debt-to-income ratio	0.26	0.16	0.03	0.48	0.25	0.12	0.10	0.45
Car purchased	0.43				0.37			
Local economic indicators								
Local unemployment rate: prior 6 months	0.036	0.009	0.021	0.051	0.050	0.009	0.034	0.063
Local unemployment rate: following 12 months	0.037	0.008	0.022	0.049	0.056	0.007	0.041	0.066
Local housing price change: prior 2 quarters vs. year earlier	0.063	0.016	0.035	0.085	0.078	0.023	0.038	0.114
Local housing price change: following 4 quarters vs. year earlier	0.072	0.017	0.045	0.098	0.077	0.033	0.037	0.140
Transaction characteristics	*N = 0.43*				*N=0.32*			
Buyer characteristics								
Monthy income	2,319	973	1,300	4,088	2,410	984	1,360	4,286
Residual monthly income after debt payments	1,723	1,079	753	3,800	1,859	1,122	790	4,018
Debt-to-income ratio	0.32	0.13	0.15	0.49	0.32	0.10	0.16	0.47
Car characteristics								
Car cost	4,954	863	3,571	6,346	5,273	1,015	3,717	6,944
Car age (years)	6.4	1.8	4	9	5.5	1.7	3	9
Odometer	88,668	17,822	57,746	113,856	81,810	18,048	50,242	108,381
Inventory age (days)	68	62	13	178	72	63	13	184
Lot age (days)	40	57	1	145	43	58	1	152
Purchase Characteristics								
Sale price	8,370	930	6,907	9,795	9,368	1,297	7,307	11,495
Down payment	740	451	200	1,500	1,003	502	600	1,900
Loan term (months)	34.1	3.0	30.0	37.0	36.6	3.9	32.0	42.0
APR	0.288	0.019	0.259	0.299	0.284	0.026	0.219	0.299
Monthly equivalent payment	362	65	298	421	374	42	306	442
Loan performance								
Outcomes								
Default	0.67				0.62			
Fraction of payments made	0.57	0.37	0.05	1.00	0.59	0.37	0.06	1.00
Loan payments excluding down payment	6,113	3,916	653	11,837	7,146	4,441	766	13,636
Recovery amount (all sales)	691	951	0	2,530	923	1,216	0	3,224
Recovery amount (all defaults)	1,032	999	1	2,848	1,483	1,243	73	3,665
Components of Profits								
Gross operating revenue	7,557	3,530	2,284	12,706	9,084	3,901	3,013	14,744
Total cost	5,810	965	4,301	7,378	6,193	1,099	4,518	8,012
Net operating revenue	1,746	3,401	-3,434	6,144	2,891	3,727	-3,005	7,620

Note: To preserve confidentiality of the company that provided the data, the number of observations is normalized by the number of applicants in 2000, *N* (*N* >> 10,000).

1. Loan payments, Recovery amount, Gross operating revenue, and Net operating revenue are PV.
2. Gross operating revenue = Down payment + PV of loan payments + PV of recovery.
3. Total cost includes car cost, taxes and fees, and shortfalls when value of trade-in does not cover down payment.
4. Net operating revenue = Gross operating revenue - Total cost

Table II: Summary Statistics by Applicants' Predicted Credit Grade

	January – December 2000			July 2001 – June 2002		
	Low Risk	Medium Risk	High Risk	Low Risk	Medium Risk	High Risk
Applicant characteristics						
Number of applicants*	*N=0.22*	*N=0.40*	*N=0.38*	*N=0.18*	*N=0.34*	*N = 0.35*
Applicant demographics						
Monthy income	3,528	2,130	1,557	3,620	2,152	1,646
Residual monthly income after debt payments	2,776	1,569	1,270	2,915	1,639	1,483
Debt-to-income ratio	0.26	0.30	0.22	0.24	0.29	0.20
Car purchased	0.57	0.55	0.23	0.57	0.53	0.12
Local economic indicators						
Local unemployment rate: prior 6 months	0.0370	0.0358	0.0357	0.0506	0.0496	0.0494
Local unemployment rate: following 12 months	0.0382	0.0372	0.0372	0.0567	0.0561	0.0560
Local housing price change: prior 2 quarters vs. year earlier	0.0646	0.0630	0.0608	0.0797	0.0779	0.0763
Local housing price change: following 4 quarters vs. year earlier	0.0744	0.0726	0.0703	0.0798	0.0765	0.0749
Transaction characteristics						
Number of buyers*	*N=0.12*	*N=0.22*	*N=0.09*	*N=0.10*	*N=0.18*	*N=0.04*
Buyer characteristics						
Monthy income	3,424	2,042	1,453	3,459	2,032	1,387
Residual monthly income after debt payments	2,670	1,461	1,042	2,718	1,479	1,318
Debt-to-income ratio	0.28	0.33	0.34	0.27	0.34	0.37
Car characteristics						
Car cost	5,235	4,949	4,569	5,602	5,212	4,707
Car age (years)	6.3	6.4	6.7	5.4	5.6	5.8
Odometer	89,593	88,735	87,198	81,924	81,823	81,471
Inventory age (days)	63	67	75	64	74	84
Lot age (days)	35	40	47	36	45	55
Purchase Characteristics						
Sale price	8,703	8,391	7,851	9,828	9,302	8,504
Down payment	762	725	746	996	995	1,055
Loan term (months)	34.2	34.1	34.1	37.1	36.5	36.0
APR	0.288	0.287	0.288	0.283	0.284	0.285
Monthly equivalent payment	380	363	334	391	372	339
Loan performance						
Outcomes						
Default	0.62	0.68	0.70	0.59	0.64	0.62
Fraction of payments made	0.63	0.56	0.54	0.62	0.58	0.59
Loan payments excluding down payment	6,912	5,979	5,319	7,864	6,914	6,340
Recovery amount (all sales)	710	709	620	1,016	926	679
Recovery amount (all defaults)	1,146	1,036	881	1,710	1,449	1,088
Components of Profits						
Gross operating revenue	8,400	7,424	6,695	9,890	8,845	8,085
Total cost	6,134	5,807	5,364	6,565	6,126	5,548
Net operating revenue	2,267	1,617	1,331	3,325	2,719	2,536

Note: To preserve the confidentiality of the company that provided the data, the number of observations is normalized by the number of applicants in 2000, *N* (*N* >> 10,000).

1. Loan payments, Recovery amount, Gross operating revenue, and Net operating revenue are PV.
2. Gross operating revenue = Down payment + PV of loan payments + PV of recovery.
3. Total cost includes car cost, taxes and fees, and shortfalls when value of trade-in does not cover down payment.
4. Net operating revenue = Gross operating revenue - Total cost

Table III(a): The Effect of Credit Scoring (Applicant level analysis)

			(1)		(2)		(3)	
		Pre-period Average	Est. Change	Std. Err.	Est. Change	Std. Err.	Est. Change	Std. Err.
Sample: All applicants								
Close rate (pct.)	Low risk	57.3	-0.4	(1.2)	0.7	(1.2)	-5.8	(4.2)
	Med. risk	54.5	-2.0	(1.3)	-2.0	(1.1)	-7.1	(3.6)
	High risk	23.5	-11.6	(1.0)	-10.8	(0.9)	-23.5	(3.4)
Price ($US)	Low risk	4,990	608	(122)	703	(119)	197	(383)
	Med. risk	4,577	309	(119)	317	(106)	-156	(315)
	High risk	1,844	-832	(75)	-764	(75)	-1,726	(294)
Default (pct.)	Low risk	35.5	-1.7	(1.0)	-0.9	(1.0)	-8.7	(3.8)
	Med. risk	37.3	-3.8	(1.0)	-3.7	(0.9)	-10.1	(3.0)
	High risk	16.5	-9.1	(0.7)	-8.4	(0.7)	-17.8	(2.9)
Down payment ($US)	Low risk	437	130	(11)	139	(10)	162	(37)
	Med. risk	396	127	(11)	125	(9)	144	(33)
	High risk	175	-50	(7)	-44	(7)	-76	(30)
Loan payments ($US)	Low risk	3,963	517	(83)	584	(76)	487	(249)
	Med. risk	3,261	370	(84)	365	(74)	267	(197)
	High risk	1,249	-495	(55)	-467	(49)	-1,177	(187)
Recovery ($US)	Low risk	407	172	(19)	181	(19)	173	(58)
	Med. risk	387	100	(17)	103	(16)	64	(49)
	High risk	146	-65	(7)	-53	(7)	-104	(49)
Gross ($US)	Low risk	4,817	817	(102)	902	(95)	803	(313)
	Med. risk	4,050	596	(103)	593	(91)	459	(258)
	High risk	1,572	-610	(66)	-565	(60)	-1,370	(238)
Cost ($US)	Low risk	3,517	223	(80)	283	(77)	-47	(249)
	Med. risk	3,168	50	(80)	52	(70)	-279	(212)
	High risk	1,260	-600	(51)	-558	(49)	-1,220	(202)
Profit ($US)	Low risk	1,300	595	(36)	618	(33)	851	(121)
	Med. risk	882	546	(36)	541	(34)	738	(90)
	High risk	313	-11	(22)	-7	(18)	-150	(83)
Controls								
Dealer fixed effects					yes		yes	
Calendar month dummies					yes		yes	
Applicant characteristics					yes		yes	
Local indicators * risk category							yes	

All regressions are of the form of $y_i = \alpha_{P(i)} + D_i{*}\beta_{R(i)} + X_i\gamma + \varepsilon_i$, where D is a post-period dummy and y is on the left column. Only the estimated beta coefficients are reported. Individual characteristics include monthly income, debt-to-income ratio, and residual monthly income. Standard errors (clustered by dealer) in parentheses.

Table III(b): The Effect of Credit Scoring (Buyer level analysis)

			(1)		(2)		(3)	
		Pre-period Average	Est. Change	Std. Err.	Est. Change	Std. Err.	Est. Change	Std. Err.
Sample: All buyers								
Price ($US)	Low risk	8,703	1,125	(56)	1,107	(52)	1,068	(108)
	Med. risk	8,391	911	(52)	900	(48)	697	(104)
	High risk	7,851	653	(61)	621	(54)	175	(148)
Default (pct.)	Low risk	61.9	-2.5	(0.9)	-2.8	(0.9)	-7.3	(3.4)
	Med. risk	68.4	-4.5	(0.7)	-4.4	(0.6)	-8.8	(2.6)
	High risk	70.4	-8.0	(0.9)	-7.2	(1.1)	-10.9	(3.4)
Down payment ($US)	Low risk	762	234	(16)	229	(15)	351	(48)
	Med. risk	725	269	(13)	261	(12)	362	(48)
	High risk	746	309	(20)	307	(18)	394	(53)
Loan payments ($US)	Low risk	6,912	952	(70)	969	(71)	1,277	(288)
	Med. risk	5,979	934	(47)	909	(43)	1,128	(218)
	High risk	5,319	1,021	(101)	890	(108)	826	(295)
Recovery ($US)	Low risk	710	306	(23)	297	(22)	348	(80)
	Med. risk	709	217	(23)	217	(21)	205	(69)
	High risk	620	59	(25)	76	(22)	1	(86)
Gross ($US)	Low risk	8,400	1,490	(67)	1,493	(68)	1,939	(268)
	Med. risk	7,424	1,421	(43)	1,388	(40)	1,659	(209)
	High risk	6,695	1,389	(92)	1,272	(101)	1,192	(286)
Cost ($US)	Low risk	6,134	431	(37)	416	(36)	509	(84)
	Med. risk	5,807	319	(39)	301	(34)	290	(82)
	High risk	5,364	184	(49)	150	(41)	-84	(114)
Profit ($US)	Low risk	2,267	1,059	(60)	1,077	(59)	1,430	(248)
	Med. risk	1,617	1,102	(48)	1,087	(43)	1,369	(200)
	High risk	1,331	1,205	(87)	1,122	(89)	1,277	(258)
Sample: Defaulters only								
Recovery (per default)	Low risk	1,146	564	(26)	557	(26)	787	(102)
	Med. risk	1,036	413	(26)	409	(24)	514	(81)
	High risk	881	207	(31)	214	(26)	213	(105)
Controls								
Dealer fixed effects					yes		yes	
Calendar month dummies					yes		yes	
Applicant characteristics					yes		yes	
Local indicators * risk category							yes	

All regressions are of the form of y= $\alpha_{P(i)} + D_i{*}\beta_{R(i)} + X_i\gamma + \varepsilon_i$, where D is a post-period dummy and y is on the left column. Only the estimated beta coefficients are reported. Individual characteristics include monthly income, debt-to-income ratio, and residual monthly income. Standard errors (clustered by dealer) in parentheses.

Table III(c): The Effect of Credit Scoring (Buyer level analysis; Rate of return)

			(1)		(2)		(3)	
		Pre-period Average	Est. Change	Std. Err.	Est. Change	Std. Err.	Est. Change	Std. Err.
Sample: All buyers								
Down payment/cost (pct.)	Low risk	12.5	2.9	(0.3)	2.8	(0.2)	4.1	(0.8)
	Med. risk	12.6	3.9	(0.2)	3.9	(0.2)	5.2	(0.8)
	High risk	14.0	5.6	(0.4)	5.6	(0.4)	7.6	(1.0)
Loan payments/cost (pct.)	Low risk	113.8	7.4	(1.1)	8.0	(1.0)	10.8	(4.5)
	Med. risk	104.1	10.2	(1.1)	10.1	(0.9)	13.2	(3.7)
	High risk	100.3	15.8	(1.9)	14.4	(1.8)	16.5	(5.0)
Recovery/cost (pct.)	Low risk	11.5	3.8	(0.3)	3.7	(0.3)	4.0	(1.2)
	Med. risk	12.1	2.8	(0.3)	2.8	(0.3)	2.2	(1.0)
	High risk	11.4	0.6	(0.4)	0.9	(0.4)	-0.9	(1.3)
Gross/cost (pct.)	Low risk	138.1	14.0	(1.0)	14.4	(0.9)	18.4	(4.1)
	Med. risk	129.0	16.9	(0.9)	16.8	(0.8)	20.0	(3.5)
	High risk	125.9	21.9	(1.7)	20.9	(1.6)	22.9	(4.6)
Profit/cost (pct.)	Low risk	38.1	14.0	(1.0)	14.4	(0.9)	18.4	(4.1)
	Med. risk	29.0	16.9	(0.9)	16.8	(0.8)	20.0	(3.5)
	High risk	25.9	21.9	(1.7)	20.9	(1.6)	22.9	(4.6)
Sample: Defaulters only								
Recovery/cost (pct.)	Low risk	18.6	7.1	(0.4)	7.1	(0.4)	9.7	(1.4)
	Med. risk	17.7	5.6	(0.3)	5.6	(0.3)	6.4	(1.1)
	High risk	16.3	3.1	(0.5)	3.2	(0.4)	2.2	(1.5)
Controls								
Dealer fixed effects					yes		yes	
Calendar month dummies					yes		yes	
Applicant characteristics					yes		yes	
Local indicators * risk category							yes	

All regressions are of the form of $y_i = \alpha_{P(t)} + D_i * \beta_{R(i)} + X_i\gamma + \varepsilon_i$, where D is a post-period dummy and y is on the left column. Only the estimated beta coefficients are reported. Individual characteristics include monthly income, debt-to-income ratio, and residual monthly income. Standard errors (clustered by dealer) in parentheses.

Table IV(a): Summary Statistics for High Pre-period Profit Dealers

	Predicted Grade: Low Risk			Predicted Grade: Medium Risk			Predicted Grade: High Risk		
	Pre	Post	Change	Pre	Post	Change	Pre	Post	Change
Applicant characteristics									
Number of applicants*	*N=0.075*	*N=0.061*		*N=0.124*	*N=0.108*		*N=0.127*	*N=0.115*	
Applicant demographics									
Monthy income	3,578	3,701	123	2,148	2,154	6	1,538	1,642	105
Residual monthly income after debt payments	2,806	3,015	209	1,595	1,660	65	1,255	1,488	233
Debt-to-income ratio	0.26	0.24	-0.02	0.30	0.29	-0.01	0.23	0.21	-0.02
Car purchased	0.60	0.59	-0.01	0.59	0.57	-0.02	0.29	0.14	-0.14
Local economic indicators									
Local unemployment rate: prior 6 months	0.041	0.052	0.011	0.039	0.050	0.011	0.038	0.049	0.011
Local unemployment rate: following 12 months	0.041	0.057	0.016	0.039	0.055	0.016	0.038	0.055	0.017
Local housing price change: prior 2 quarters vs. year earlier	0.064	0.089	0.025	0.061	0.085	0.024	0.059	0.083	0.024
Local housing price change: following 4 quarters vs. year earlier	0.074	0.101	0.026	0.071	0.095	0.024	0.068	0.092	0.024
Transaction characteristics									
Number of buyers*	*N=0.045*	*N=0.036*		*N=0.074*	*N=0.062*		*N=0.036*	*N=0.017*	
Buyer characteristics									
Monthy income	3,481	3,567	86	2,081	2,047	-34	1,451	1,395	-56
Residual monthly income after debt payments	2,721	2,862	141	1,524	1,530	6	1,051	1,330	280
Debt-to-income ratio	0.28	0.26	-0.02	0.32	0.33	0.01	0.33	0.36	0.03
Car characteristics									
Car cost	5,202	5,584	382	4,842	5,187	345	4,434	4,568	134
Car age (years)	6.6	5.6	-1.0	6.8	5.7	-1.1	7.1	5.9	-1.2
Odometer	89,676	83,016	-6,660	88,689	82,703	-5,986	87,501	81,452	-6,049
Inventory age (days)	62	66	3	66	74	8	71	78	7
Lot age (days)	35	38	4	38	47	8	43	51	8
Purchase Characteristics									
Sale price	8,625	9,679	1,054	8,274	9,186	912	7,731	8,364	632
Down payment	751	984	233	695	984	288	703	1,024	320
Loan term (months)	34.0	37.7	3.6	33.6	37.0	3.5	33.4	36.2	2.8
APR	0.296	0.295	-0.002	0.296	0.294	-0.002	0.295	0.289	-0.005
Monthly equivalent payment	377	385	9	362	367	5	333	332	-1
Loan performance									
Outcomes									
Default	0.58	0.57	-0.01	0.63	0.60	-0.02	0.66	0.59	-0.07
Fraction of payments made	0.66	0.65	-0.02	0.61	0.61	0.00	0.59	0.62	0.03
Loan payments excluding down payment	7,285	8,166	880	6,451	7,286	835	5,708	6,502	795
Recovery amount (all sales)	605	927	322	578	826	248	499	611	112
Recovery amount (all defaults)	1,050	1,637	587	920	1,366	446	758	1,032	275
Components of Profits									
Gross operating revenue	8,658	10,093	1,434	7,738	9,105	1,367	6,927	8,148	1,221
Total cost	6,060	6,494	434	5,627	6,037	410	5,157	5,356	199
Net operating revenue	2,599	3,599	1,000	2,111	3,068	957	1,769	2,792	1,022

1. Includes dealers in top third by pre-period net operating revenue per sale.
2. Loan payments, Recovery amount, Gross operating revenue, and Net operating revenue are PV.
3. Gross operating revenue = Down payment + PV of loan payments + PV of recovery.
4. Total cost includes car cost, taxes and fees, and shortfalls when value of trade-in does not cover down payment.
5. Net operating revenue = Gross operating revenue - Total cost

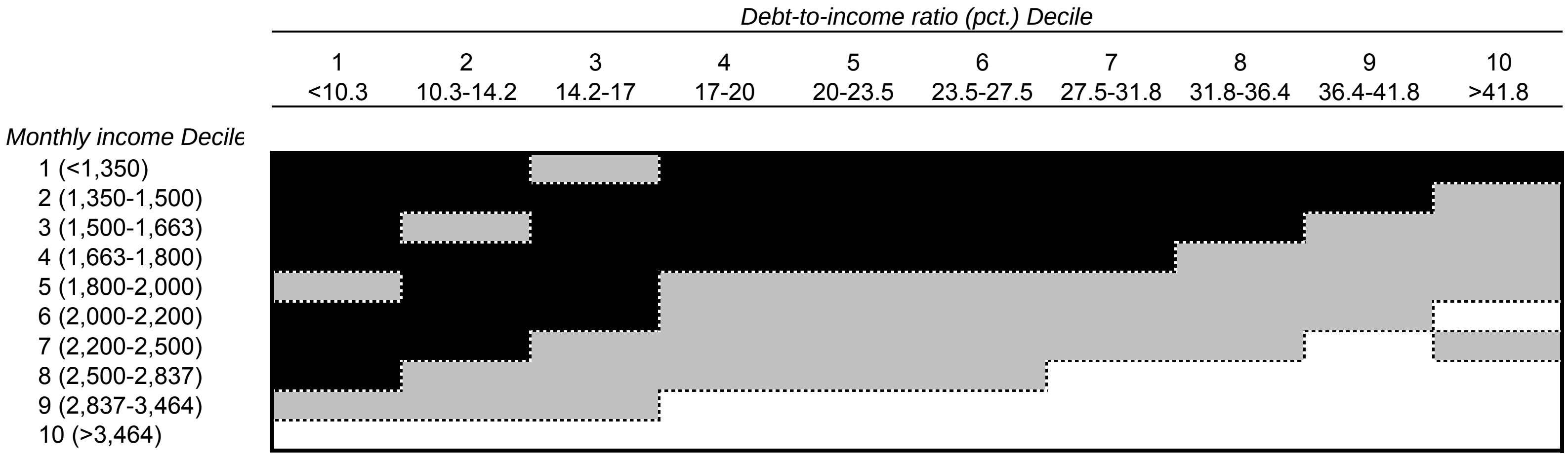
Table IV(b): Summary Statistics for Low Pre-period Profit Dealers

	Predicted Grade: Low Risk			Predicted Grade: Medium Risk			Predicted Grade: High Risk		
	Pre	Post	Change	Pre	Post	Change	Pre	Post	Change
Applicant characteristics									
Number of applicants*	*N=0.062*	*N=0.057*		*N=0.126*	*N=0.112*		*N=0.116*	*N=0.120*	
Applicant demographics									
Monthy income	3,516	3,568	52	2,112	2,148	36	1,567	1,639	72
Residual monthly income after debt payments	2,787	2,867	80	1,550	1,630	80	1,284	1,497	214
Debt-to-income ratio	0.25	0.25	-0.01	0.30	0.29	-0.01	0.22	0.20	-0.02
Car purchased	0.58	0.55	-0.03	0.56	0.51	-0.05	0.23	0.11	-0.12
Local economic indicators									
Local unemployment rate: prior 6 months	0.033	0.049	0.016	0.034	0.049	0.015	0.034	0.049	0.015
Local unemployment rate: following 12 months	0.036	0.056	0.020	0.036	0.056	0.020	0.036	0.056	0.020
Local housing price change: prior 2 quarters vs. year earlier	0.064	0.072	0.008	0.063	0.071	0.008	0.060	0.067	0.007
Local housing price change: following 4 quarters vs. year earlier	0.073	0.063	-0.011	0.072	0.061	-0.010	0.069	0.058	-0.011
Transaction characteristics									
Number of buyers*	*N=0.036*	*N=0.032*		*N=0.070*	*N=0.062*		*N=0.027*	*N=0.013*	
Buyer characteristics									
Monthy income	3,404	3,399	-6	2,018	2,016	-2	1,467	1,381	-86
Residual monthly income after debt payments	2,666	2,650	-15	1,428	1,452	24	1,062	1,386	324
Debt-to-income ratio	0.28	0.28	0.00	0.33	0.34	0.01	0.33	0.37	0.04
Car characteristics									
Car cost	5,216	5,625	408	5,014	5,241	226	4,673	4,840	167
Car age (years)	6.2	5.2	-0.9	6.2	5.4	-0.7	6.4	5.6	-0.8
Odometer	90,219	80,612	-9,607	89,072	81,053	-8,019	87,724	81,657	-6,067
Inventory age (days)	65	63	-2	70	73	3	82	92	9
Lot age (days)	37	35	-2	42	43	1	54	61	7
Purchase Characteristics									
Sale price	8,731	9,975	1,244	8,458	9,438	980	7,907	8,593	686
Down payment	739	999	260	727	985	257	761	1,072	311
Loan term (months)	34.8	37.1	2.3	34.8	36.7	1.9	34.9	36.1	1.1
APR	0.280	0.273	-0.006	0.281	0.276	-0.004	0.283	0.283	0.001
Monthly equivalent payment	369	393	24	356	374	18	328	341	12
Loan performance									
Outcomes									
Default	0.67	0.62	-0.05	0.73	0.68	-0.05	0.76	0.66	-0.10
Fraction of payments made	0.58	0.60	0.02	0.51	0.55	0.04	0.47	0.57	0.09
Loan payments excluding down payment	6,468	7,634	1,166	5,547	6,627	1,080	4,788	6,131	1,343
Recovery amount (all sales)	800	1,091	290	801	997	196	719	722	2
Recovery amount (all defaults)	1,202	1,763	561	1,098	1,473	375	952	1,096	144
Components of Profits									
Gross operating revenue	8,028	9,738	1,710	7,086	8,622	1,536	6,274	7,929	1,655
Total cost	6,119	6,612	492	5,891	6,179	288	5,488	5,670	182
Net operating revenue	1,909	3,126	1,217	1,195	2,443	1,248	786	2,259	1,473

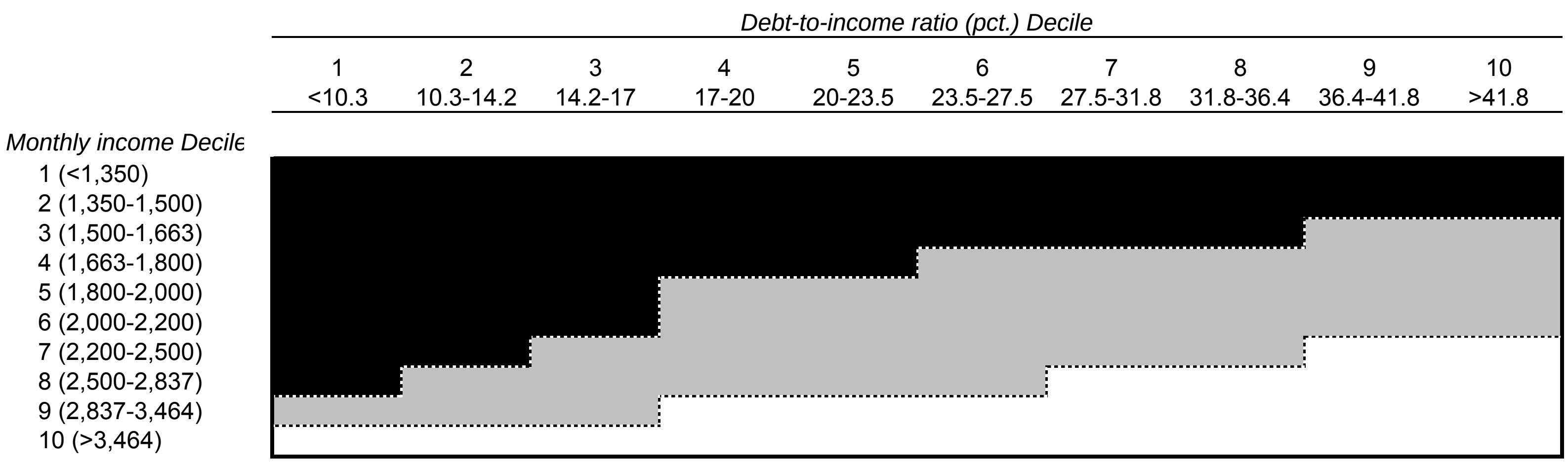
1. Includes dealers in bottom third by pre-period net operating revenue per sale.
2. Loan payments, Recovery amount, Gross operating revenue, and Net operating revenue are PV.
3. Gross operating revenue = Down payment + PV of loan payments + PV of recovery.
4. Total cost includes car cost, taxes and fees, and shortfalls when value of trade-in does not cover down payment.
5. Net operating revenue = Gross operating revenue - Total cost

Table A.I: Results from risk prediction model

A. Results based on an orderd probit model



B. Results based on the full model



Legend:

Predicted Low Risk

Predicted Medium Risk

Predicted High Risk