
Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

# The Effectiveness of Field Price Discretion: Empirical Evidence from Auto Lending

Robert Phillips, A. Serdar Şimşek, Garrett van Ryzin

MANAGEMENT SCIENCE
Vol. 61, No. 8, August 2015, pp. 1741–1759
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2014.2084


# The Effectiveness of Field Price Discretion: Empirical Evidence from Auto Lending

Robert Phillips
Columbia Business School, New York, New York 10027; and Nomis Solutions, San Bruno, California 94066, rp2051@columbia.edu

A. Serdar Şimşek
School of Operations Research and Information Engineering, Cornell University, Ithaca, New York 14853, as2899@cornell.edu

Garrett van Ryzin
Columbia Business School, New York, New York 10027, gjv1@columbia.edu

In many markets, it is common for headquarters to create a price list but grant local salespeople discretion to negotiate prices for individual transactions. How much (if any) pricing discretion headquarters should grant is a topic of debate within many firms. We investigate this issue using a unique data set from an indirect lender with local pricing discretion. We estimate that the local sales force adjusted prices in a way that improved profits by approximately 11% on average. A counterfactual analysis shows that using a centralized, data-driven pricing optimization system could improve profits even further, up to 20% over those actually realized. This suggests that centralized pricing—if appropriately optimized—can be more effective than field price discretion. We discuss the implications of these findings for auto lending and other industries with similar pricing processes.

*Keywords*: customized pricing; sales force price discretion; price sensitivity estimation; endogeneity; consumer lending
*History*: Received May 28, 2013; accepted September 17, 2014, by Serguei Netessine, operations management. Published online in *Articles in Advance* March 20, 2015.



## 1. Introduction

The extent to which sales staff should be able to set or negotiate prices for individual deals is a source of considerable tension within many firms. The primary source of tension is between the capability of headquarters to leverage enterprise-wide information and advanced analytics ("pricing optimization") to control prices versus the ability of field sales staff to utilize deal-specific information to negotiate a better price. Most research on the appropriate allocation of pricing authority between headquarters and the field has focused on the two ends of the continuum: fully centralized fixed pricing, in which all prices are determined ex ante by headquarters, versus discretionary pricing with no limits, in which local sales staff set (or negotiate) prices based on the characteristics of individual deals.

In reality, there is little consistency in how discretionary pricing is applied among industries and even within the same industry. A 1977 survey of wholesale medical supply and equipment companies found that 29% of the responding firms had fixed pricing, 48% had discretionary pricing with discretion limits, and 23% had discretionary pricing with no limits (Stephenson et al. 1979). A 2011 survey found that between 30% and 40% of unsecured loans and lines of credit originated or renewed at major Canadian banks involved some level of local price discretion (Phillips 2012b). A 2012 Oliver Wyman survey noted that local discretion was used for more than 50% of the unsecured loans and more than 70% of the secured loans offered by major European banks (Oliver Wyman and Efma 2012). This wide variation of policies both between markets and within the same market reflects the fact that there is no universally accepted approach to pricing decentralization and discretion.

In this paper, we distinguish between *customized pricing* and *discretionary pricing*.[1] There are many selling situations in which pricing is customized for individual customers based on differences in cost to serve. For example, industrial contracts for parts are based on the complexity of the part, the order quantity and quality, and the delivery requirements of the customer. In our auto loan context, the price (interest

[1] To be clear in our categorization of pricing types, note that we consider two dimensions: customized versus uniform pricing and discretionary versus fixed pricing.

rate) for a loan is adjusted based on the creditworthiness of the buyer, the amount and duration of the loan, the loan-to-value ratio, and the type of vehicle financed. In situations where such idiosyncratic factors affect the cost to serve, pricing is naturally customized to reflect inherent cost differences. We distinguish this sort of cost-based customized pricing from *discretionary pricing*, in which pricing is varied based on the salesperson's (or system's) assessment of the customer's willingness to pay (WTP). Here, the basis for differential pricing is not cost-based, but rather based on an assessment of differences in WTP. Although the distinction between these two rationales for price differences is not always sharp (e.g., volume price discounts can be both cost related and WTP related), the distinction is nevertheless helpful in understanding the fundamental drivers of pricing differences.

Our focus here is on *customized pricing with discretion*—that is, a setting where prices are inherently customized because of cost differences but where salespeople are allowed further discretion to negotiate prices with customers. Under customized pricing with discretion, pricing takes place in three steps: (1) Headquarters establishes a price list based on objective and observable factors related to each deal. (2) Headquarters then establishes limits on the adjustments that field sales staff can apply to the list prices for individual deals. (3) Field sales staff can then negotiate a price within the discretion limits applied to the list price for that deal. This method is the most common approach to pricing in many business-to-business (B2B) markets. It is also common in some business-to-consumer markets such as insurance, automobile sales, and consumer lending.

Although the determinants of the trade-offs between fixed and discretionary pricing in the uniform pricing case have been studied, it is not immediately clear whether these same conditions apply to the customized pricing setting. In particular, many of the papers addressing the trade-off between uniform fixed pricing and uniform discretionary pricing state that, although negotiation is costly, it provides better price discrimination when there is a wide variation of willingness to pay among customers and customers have little bargaining power (Roth et al. 2006, Kuo et al. 2013). In a customized pricing setting, on the other hand, firms are already utilizing a higher level of price discrimination by taking each customer's individual characteristics into account. In this case, if the centralized prices are set optimally using all available information, this can significantly reduce the potential gain from local negotiations. Therefore, it is not obvious whether the conditions that favor discretion over uniform fixed pricing also apply to customized pricing.

To study customized pricing with discretion, we use a proprietary data set from an automotive lender who offered loans exclusively through dealers. In the auto lending industry, price negotiation is the norm. As noted, the cost of loans varies for each individual transaction, so prices are inherently customized. In addition, auto lending has characteristics that favor negotiation: high variation in customer willingness to pay, low cost of bargaining (compared with the value of the product sold), and lack of comparability resulting in low customer bargaining power (Caufield 2012). Yet many lenders now have the capability to use centralized analytic customized pricing systems to price discriminate based on factors such as size of loan, credit score, term, etc. (Phillips 2005, 2012a; Agrawal and Ferguson 2007).

The analyzed auto lender generated a price list that specified rates for various types of loans; however, the dealers had the authority to change the rate within limits for individual deals. We address two main questions in this setting: (1) Did the local sales staff increase or reduce profits through the adjustments they made to the list prices? And (2) could additional profit have been achieved from (fully centralized) customized fixed pricing if the list prices had been optimized? Our answer is "yes" in both cases. This provides evidence for the benefits of pricing discretion as well as an explanation for the shift away from discretionary pricing to fixed pricing that has occurred in many industries.

The steps of our analysis are as follows:

*Step* 1. We estimated the effect of price response on the take-up rate for customers in a single-stage probit model using rate as one of the explanatory variables. If the local sales force were not using local information *systematically* to influence rates, this approach would provide an unbiased estimate of price response.

*Step* 2. We then used a two-stage *control function approach*. In stage 1, we used a regression to estimate the rates that would have been offered for each deal in the absence of local negotiation. In stage 2, we estimated a take-up model using the difference between the actual offered rates and the predicted rates from stage 1 as one of the explanatory variables. We found that the rate residuals were highly significant, and the changes in estimated coefficients for the rate variables from stage 1 were consistent with the hypothesis that local sales staff were using local information to negotiate rates that were closer to the customers' willingness to pay.

*Step* 3. We then used the coefficients for rate estimated in stage 2 to model how customers would respond to different offered prices. To the extent that our stage 2 model accurately captured price response, we found that by using their price discretion to negotiate better deals, the local sales force increased profitability by an average of 11%. However, by setting optimal prices produced by a centralized, data-driven profit maximization procedure, which utilizes the data of all nationwide sales, prices could have been set that would have increased profitability by an additional 20%.

The magnitude of our conclusions is dependent on the assumption that the difference between the price-coefficient estimates in Steps 1 and 2 are due to the use of pricing discretion on the part of field sales. Although it is impossible to prove that this is the only source of the differences, we performed a similar analysis on data from an online lender that set all prices centrally (i.e., did not use price discretion). In this case, there was no significant difference in the coefficients estimated in Steps 1 and 2, lending support to the hypothesis that the difference we found in estimated coefficients from the indirect lender were due to local sales discretion.

### 1.1. Fixed vs. Discretionary Pricing

Simonetto et al. (2012, p. 846) note, "Generally speaking, decentralized (pricing) organizations offer greater speed, flexibility, and responsiveness, while centralized organizations offer greater consistency, standardization, and control." To understand the tension between fixed and discretionary pricing in more detail, it is useful to list the main arguments on each side.

**1.1.1. Fixed Pricing.** Potential benefits for fixed pricing include the following:

- *Data pooling*—Headquarters has access to data from all corporate sales. To the extent that customer characteristics and responses are consistent across regions, corporate-wide sales data can provide more accurate estimation of customer preferences, price sensitivity, and better market segmentation than any single salesperson or sales region.
- *Control*—Fixed pricing enables headquarters to change prices quickly in response to changing economic environments, cost changes, or corporate goals. It also enables the company to pursue different goals in different markets or for different product lines. If the sales force is setting deal-specific prices, it can be difficult to pursue such mixed goals (Baker et al. 2010).
- *Consistency*—There are a number of reasons why a company might wish to achieve price consistency. For one, the company may wish to advertise prices nationally. For another, the company may wish to avoid being perceived as "unfair" for selling the same product to different customers at different prices. Finally, management may be concerned that price discretion could lead to discrimination—real or perceived—against protected groups such as women or minorities. This concern is very real in the auto lending industry where there have been a number of class-action lawsuits against lenders claiming such discrimination (Cohen 2007).
- *Variation in sales force skill*—In a large sales force, there is usually a wide variation of skills. Skilled salespeople may excel at selling a product on its value and thereby capturing higher prices. On the other hand, less skilled or less experienced sales staff may "sell on price" and drop quickly to the lowest possible price point in order to make the sale. Limited discretion is often justified internally as enabling the firm to gain some benefit from the skills of its more skilled sales staff while limiting the underpricing damage from less experienced sales staff.

**1.1.2. Discretionary Pricing.** There are three common arguments for granting price discretion to local sales staff:

- *Selling tradition*—In many industries, price negotiation is customary. In this case, customers may abandon a seller who tries to unilaterally impose "take-it-or-leave-it" pricing. Gelber (2008, pp. 66–67) discusses the role of custom in the persistence of negotiation over new car prices and the failure of companies such as CarMax to eliminate negotiations from the buying process.
- *Perception of a bargain*—It has been observed across many industries as well as in experimental settings that presenting a price as a discount from a higher list price will often result in higher sales than presenting the same price as the undiscounted list price (see Özer and Zheng 2012 for a survey). Therefore, allowing local sales staff to present a price as a discount from a list price may result in higher sales than posting the lower price and treating it as fixed.
- *Local information*—Local sales staff have information on individual deals that is not available to headquarters and is not incorporated in the list price. This information may include customer-specific information such as the customer's expressed eagerness to purchase or insistence on a low price as well as local competitive information. Most existing research has focused on the advantage of local information on the part of field sales.

The increasing use of pricing optimization systems in many companies is leading to more centralized pricing. The idea is that, by using richer sources of data and structural analysis, such systems can outperform local sales staff in estimating the price sensitivity of individual customers (Simonetto et al. 2012). There is an interesting parallel with credit scoring. Prior to the 1960s, evaluation of the creditworthiness of prospective borrowers was performed entirely at the local level. But scoring methodologies that analyzed data about each customer's credit history and current financial situation turned out to be more efficient and effective predictors of default than individual judgment. As a result, evaluating the credit risk of prospective borrowers is now almost entirely centralized at major financial institutions (Poon 2007). The same phenomenon may occur in customized pricing settings as well; hence careful analysis of such settings is warranted.

## 2. Literature Review

Starting with Weinberg (1975), a number of researchers have modeled pricing discretion as a principal–agent problem. The problem for the principal (headquarters) is to design a contract that specifies how much commission the agent (salesperson) will be paid as a function of sales and price. The problem for the salesperson is how much effort to expend given the contract and the sales prospects she faces given that effort is expensive (in terms of utility). Weinberg (1975) shows that if sales staff are paid a commission based on gross margin, discretionary pricing will maximize both their own and their company's profits. Lal (1986) shows that in this model, optimal profits are the same under centralization and delegation if information is symmetric. However, if the agent has private information about deals, pricing delegation can generate higher profits for the firm. Mishra and Prasad (2004) show that, even when the agent possesses private information, if the salesperson first observes the private information and then signs the contract, a company can set prices centrally and capture the full benefits available from discretionary pricing by using an appropriate contract structure—though compensation may be a complicated function of quantity and price. They then analyze competitive markets and show that, under asymmetric information, an equilibrium always exists where all firms use centralized pricing, regardless of the intensity of competition (Mishra and Prasad 2005).

Joseph (2001) develops a model that explicitly incorporates the salesperson's superior information about customers' willingness to pay and the possibility of suboptimal trade-offs between price and effort (using discounting rather than expending effort on selling). In his model, centralization of pricing authority is sometimes preferable to full decentralization, depending on the effort cost of following a high-quality prospecting strategy and the structure of customer segments. Bhardwaj (2001) reaches a similarly mixed conclusion in the case of competing firms when principals and agents have the same information. Roth et al. (2006) consider the case of bargaining versus fixed (or posted) price for customized services and show that, if bargaining costs are low, bargaining can be a preferred approach for the seller. Kuo et al. (2011) present a model that incorporates the interactions among dynamic pricing, negotiation, and inventory and show that negotiation is an effective tool to achieve price discrimination, particularly when the inventory level is high and/or the remaining selling season is short.

In the operations management (OM) literature, Taylor (2002) shows that a properly designed target rebate and returns contract achieves supply chain coordination and win-win outcomes in cases where demand is influenced by retailer sales effort. Other types of sales force incentives have been studied as well (see Coughlan 1993 for a comprehensive review). Kalkanci et al. (2011) study a two-tier supply chain with a single supplier and a single buyer to characterize the impact of contract complexity and local demand information on performance using human subjects, and they show that simpler contracts are sufficient for a supplier in this setting. Kuo et al. (2013) examine the choice of pricing policy (posted pricing versus negotiation) toward end customers in a supply chain and show that the retailer prefers negotiation at lower wholesale prices and posted pricing at higher wholesale prices. They also show that, in cases where product availability or the cost of negotiation is moderate, the manufacturer may offer a substantial discount to persuade the retailer to negotiate.

These papers are primarily theoretical (for an exception, see Kalkanci et al. 2011). There are fewer empirical studies of the effectiveness of pricing discretion. Stephenson et al. (1979) compare the performance of 108 medical equipment suppliers. They found that, for this sample, firms granting the highest level of pricing discretion generated the lowest sales and profits. Frenzen et al. (2010) compare a sample of 181 companies from the industrial machinery and electrical engineering industry in Germany and report a positive effect of price delegation on firm performance, which is amplified under high market uncertainty and information asymmetry. Homburg et al. (2012) use a survey of 124 companies from various B2B industries in Germany and find a nonlinear, inverted-U-shaped relationship between the vertical delegation of pricing authority and profitability. Therefore, both theoretical and empirical studies suggest that there are factors at play other than simply superior information on the part of the salesperson. Our paper contributes to this literature by developing a method for quantifying the value of price discretion in a specific organization.

There are several recent empirical OM papers that estimate demand parameters from sales data and connect the resulting estimated customer valuations to pricing strategies. Vulcano et al. (2010) offer an expectation–maximization (EM) algorithm to estimate a choice model using sales data from an airline market and compute counterfactual revenue performance using the calibrated demand model. Veeraraghavan and Vaidyanathan (2012) develop an estimation method for measuring the seat value in stadiums and theaters and use the estimates to provide recommendations for pricing strategies. Li et al. (2014) provide empirical evidence for the existence of strategic consumers and also perform counterfactual analyses of different pricing strategies.

The auto industry has been the focus of many empirical papers in OM. For example, Olivares and Cachon (2009), Cachon and Olivares (2010), and Cachon et al. (2013) empirically investigate the impacts of competition on inventories, drivers of finished goods inventories, and the effects of various types of inventory on sales, respectively, in the U.S. auto industry. Moreno and Terwiesch (2013) and Moreno and Terwiesch (2014) analyze the benefits and costs of maintaining a broader product line and empirically analyze the relationship between production flexibility and responsive pricing, respectively.

## 3. Auto Lending Sales and Pricing Processes

Our results are based on a data set of approved loan applications from an auto lender in the U.S. market. "Customers" in the auto lending market wish to borrow money to purchase a car, which serves as collateral for the loan. Loans differ in terms of the amount borrowed (the size of the loan) and the loan duration (the *term*). The "price" of an automobile loan is determined by its *annual percentage rate* (APR). All of the loans in the data set are simple loans with equal monthly payments. The monthly payment $p$ for each loan can be calculated as a function of the initial principal $P$, monthly APR $r$, and term in months $n$ according to the standard calculation

$$p = Pr(1+r)^n/[(1+r)^n - 1]. \quad (1)$$

We call the auto lender the *indirect* lender, because it offers loans exclusively through automobile dealerships. In particular, loans are sold during a face-to-face interaction at a dealership. When a customer indicates that she is interested in purchasing a vehicle, she is typically sent to the finance and insurance (F&I) department of the dealership to discuss financing. A prospective borrower fills out a loan application specifying her identity, information on the car she wishes to purchase, and the amount and term of the loan she wants. The F&I manager enters each application into an automated system, such as RouteOne, which transmits the application to prospective lenders. Each lender that receives the application first determines whether it wishes to extend a loan based on an estimate of the default risk. To estimate this risk, the lender will obtain information about the applicant's credit history from a credit agency such as TransUnion or Experian. A lender that accepts an application will communicate the APR that it will charge—the so-called *lender rate*. Assuming that the applicant has been approved for the loan by at least one lender, the F&I representative will decide on an APR to quote to the borrower. The rate paid by the customer—the so-called *customer rate*—is determined through negotiation between the F&I representative and the customer.[2] Usually (but not always), the customer rate is higher than the lender rate, in which case the dealer profits from the difference between the two rates. In some cases, to make a sale, the dealership may agree to a customer rate that is lower than the lender rate, in which case the dealership will lose money on the financing, which it will presumably make up through the margin on the sale of the car. Financing auto loans is an important source of profit for a dealership, so F&I representatives are usually trained salespeople who are incentivized to increase dealer profits from lending.

## 4. Price-Response Model Estimation

Our analysis is based on a data set of approved auto loans from an indirect lender that has requested anonymity. The data set includes information on all approved applications during a multiyear period starting January 2009.[3] There are 2,138,691 approvals in this data set of which 1,473,786 (69%) were taken up.

We are interested in predicting how the probability of take-up changes as a function of the APR. We use a control function approach (Petrin and Train 2010) to estimate the consumers' price-response function. In the first step, we use probit regression on the data with take-up as the target variable. We then use a control function to estimate what rates would have been offered in the absence of deal-specific information on the part of the sales force. We find that the residuals between these estimated rates and actual rates are highly correlated with take-up, which confirms our hypothesis that deal-specific knowledge does influence the final rates. We then run a regression that includes the rate differentials as explanatory variables to estimate the underlying price sensitivity.

[2] The term "negotiation" is potentially misleading because there may not be any of the "back and forth" between the F&I representative and the customer usually implied by the term. Some customers are not aware that the rate is potentially negotiable and may accept it as a "take it or leave it" rate.

[3] The indirect lender requested that the exact period under consideration be withheld to preserve confidentiality.

**Table 1 Descriptions and Summary Statistics for Continuous Variables**

| Variable | Comment | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|---|
| *FICOScore* | Credit score | 721.84 | 87.06 | 330 | 900 |
| $\widehat{FICO}$ | Normalized FICO score | 0 | 54.3 | −385 | 330 |
| *Term* (months) | Term of the loan | 61.37 | 11.61 | 6 | 96 |
| *Amount* ($1,000) | Size of the loan | 24.50 | 11.07 | 1 | 120 |
| *CustomerRate* (%) | APR of the deal | 5.99 | 4.90 | 0 | 38.79 |
| *PrimeRate* (%) | One-month LIBOR at time of approval | 0.27 | 0.07 | 0.18 | 0.56 |
| *CustomerCash* ($) | Cash incentive | 629 | 1,064 | 0 | 9,000 |

*Note.* LIBOR, London Interbank Offered Rate.

**Table 2 Descriptions and Summary Statistics for Categorical Variables**

| Variable | Comment | Category | Frequency | % |
|---|---|---|---|---|
| *Tier* | Risk-based classification of borrowers | 1 | 1,053,411 | 49.25 |
| | | 2 | 430,434 | 20.13 |
| | | 3 | 382,349 | 17.88 |
| | | 4 | 111,530 | 5.22 |
| | | 5 | 160,967 | 7.52 |
| *VehicleType* | New or used | New | 1,683,542 | 78.72 |
| | | Used | 455,149 | 21.28 |
| *SubventionOffer* | Significantly reduced promotional APR | Not offered | 378,976 | 17.72 |
| | | Offered | 1,759,715 | 82.28 |
| *Outcome* | Customer's decision | Not take | 664,905 | 31.09 |
| | | Take | 1,473,786 | 68.91 |
| *VehicleModel* | Civic, Corolla, Focus, etc. | Too many | Too many | |
| *DealerID* | Dealer-specific source of the transaction | Too many | Too many | |

### 4.1. Base Price-Response Model

Tables 1 and 2 list the continuous and categorical variables, respectively, that were used in the final base model along with some summary statistics. The lender categorized applications into five risk tiers based on estimated default risk. The lender used a proprietary methodology to estimate risk and assign applications to tiers. The estimate of risk increased with each tier—i.e., tier 1 applications were considered the least risky and tier 5 the most risky.

We included both *PrimeRate*, the prime rate, and $\Delta Rate$, defined as the difference between the APR of a deal (*CustomerRate*) and the current prime rate, as explanatory variables. The prime rate is a measure of the current cost of funds common to all lenders, whereas the APR is set by the lender and is potentially different for every deal. Using *PrimeRate* and $\Delta Rate$ enables us to separate the effect of industry-wide cost changes from deal-specific pricing decisions and, since $\Delta Rate = CustomerRate - PrimeRate$, it is equivalent to including both *PrimeRate* and *CustomerRate*.

The data set included the FICO score associated with each approved application. The FICO score is an industry standard individual risk score, which ranges from 300 to 850, with higher values representing lower risk. A FICO score is highly correlated with risk tier; however, the lender used additional information such as the size of the loan or number of credit cards that the borrower owns (which can be determined through the credit agencies) to classify potential borrowers into risk tiers. Therefore, we first calculated the average FICO score for each risk tier. Then, for each application, we calculated a normalized FICO score, $\widehat{FICO}$ as the difference between the FICO score on that application and the average FICO score of the application's risk tier. This gives a measure of individual application risk that is not correlated with risk tier.

There are two promotion-based variables that are relevant to the two types of promotions that were offered at various times.[4] Under a "customer cash" deal, the vehicle manufacturer offered a cash rebate to the customer. The variable *CustomerCash* measures the size of this rebate (if it was offered). Under a "subvention" deal, the vehicle manufacturer subsidized the lender to offer a reduced APR. For example, a manufacturer might offer 1.0% financing for all sales of a particular model in the month of May. In this case, the manufacturer pays the lender for the difference

[4] Both types of promotions are controlled by the manufacturer, and neither the lender nor the local sales staff has any discretion on them.

in expected profit between pricing the loan at the list APR and at the reduced APR. This payment is called a *subvention*. The availability of a subvened rate is recorded for borrowers who did not take up a loan as well as for those who took up the loan.

To group loans for the same vehicle model in the same month in the base model, we used the categorical variable *VehicleModelByMonth*, which is defined by crossing *VehicleModel* with *Month*. This variable is included to account for differences in *CustomerRate* resulting from common demand shocks as a result of vehicle-model-based pricing and/or promotions. There were initially 2,055 categories in the *VehicleModelByMonth* variable; however, we excluded categories with fewer than 10 observations (0.06% of the data), which resulted in 1,628 categories.

*DealerID* indicates which dealer originated the loan. *DealerID* is included in the base price-response model as a categorical variable to incorporate dealer-specific effects, such as differences in competitive intensity. The data set included loans originated from 7,730 distinct dealers. We excluded dealers with fewer than 10 transactions, corresponding to 1.6% of the transactions, resulting in 6,103 distinct dealers in our final data set. Region indicators are not included in the models since the dealer-specific dummies capture regional effects as well.

The majority of loans fell into a relatively small number of terms (specified in months). For example, there are many 60-month loans but no 59-month or 61-month loans. We created a categorical variable *TermClass* with possible values (0–36], (36–48], (48–60], or $>60$.

We used maximum-likelihood estimation to estimate coefficients for probit price-response function. This is a standard approach for price-response estimation in similar situations (Phillips 2005, 2012a; Agrawal and Ferguson 2007). Using the probit model allows us to apply well-developed approaches for detecting and adjusting for endogeneity (Wooldridge 2002). Using a logit model for the base model did not make a significant difference in either the model fit statistics or the signs of the coefficients. We used standard variable selection techniques and tried different variable transformations as well as crossing the rate variable with other variables. We performed holdout sample validation by estimating the model on a randomly chosen training set and testing the take-up on the remaining validation (or test) set. We performed this validation test while adding and removing each explanatory variable one at a time (we used a combination of forward inclusion and backward elimination techniques by testing at each step for variables to be added or eliminated). The models shown have the highest predictive performance on the validation set

**Table 3 Base and Endogeneity Corrected Choice Model (Probit) Estimates**

| Explanatory variable | Base | | Corrected | |
|---|---|---|---|---|
| | Estimate | Std. error | Estimate | Std. error |
| *Intercept* | 0.91*** | 0.1001 | 3.62*** | 0.2033 |
| *Tier* (Base level = 1) | | | | |
| 2 | −0.11*** | 0.0048 | 0.03*** | 0.0104 |
| 3 | −0.01*** | 0.0050 | 0.31*** | 0.0127 |
| 4 | 0.17*** | 0.0087 | 0.81*** | 0.0246 |
| 5 | 0.01 | 0.0082 | 1.02*** | 0.0296 |
| *TermClass* (Base level = 36) | | | | |
| 48 | −0.43*** | 0.0065 | −0.25*** | 0.0141 |
| 60 | −0.62*** | 0.0051 | −0.41*** | 0.0114 |
| >60 | −0.75*** | 0.0056 | −0.38*** | 0.0132 |
| *VehicleType* (Base level = New) | | | | |
| Used | 1.35*** | 0.0058 | 1.67*** | 0.0146 |
| *SubventionOffer* (Base level = Offered) | | | | |
| Not offered | −1.53*** | 0.0055 | −0.78*** | 0.0206 |
| $\widehat{FICO}$(/100) | −0.18*** | 0.0021 | −0.26*** | 0.0050 |
| log(*Amount*) | 0.11*** | 0.0038 | −0.14*** | 0.0099 |
| *PrimeRate* | 9.67 | 9.4355 | −9.94 | 19.895 |
| *CustomerCash* ($1,000) | 1.12*** | 0.0027 | 1.25*** | 0.0082 |
| $\Delta$*RateforTier1* | −7.54*** | 0.0714 | −22.84*** | 0.3978 |
| $\Delta$*RateforTier2* | −5.07*** | 0.0819 | −20.93*** | 0.1715 |
| $\Delta$*RateforTier3* | −4.53*** | 0.0778 | −20.80*** | 0.1701 |
| $\Delta$*RateforTier4* | −7.32*** | 0.1092 | −23.73*** | 0.2259 |
| $\Delta$*RateforTier5* | −4.33*** | 0.0938 | −21.10*** | 0.2073 |
| *ResidualforRate* | NA | | 19.53*** | 0.4111 |
| Concordance | 87% | | 87.1% | |
| Log-likelihood | −893,781 | | −888,320 | |

*Note.* Both models include *DealerID* and *VehicleModelByMonth* fixed effects.
$^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

(in terms of concordance[5] and log-likelihood values) of all the models we tested based on this procedure. Alternative models that use different specifications and definitions of some of the variables are discussed in §6.

The second column of Table 3 shows the coefficients for the explanatory variables in the final base model. For brevity, we have not shown the coefficients for *DealerID* and *VehicleModelByMonth*. Note that every variable was significant at $p<0.01$ with the lone exception of *PrimeRate*, which is highly correlated with *VehicleModelByMonth*. A total of 1,216 of the 1,628 *VehicleModelByMonth* categories were significant at the 0.05 level. The base model showed a high level of concordance: 87%. The base model is similar in structure and in coefficient values to models that have been used by the auto lender to predict take-up. For the most part, the coefficient signs are intuitive. All of the coefficients for $\Delta$*Rate* by tier are negative,

[5] The *concordance* of a model is defined as the percentage of distinct pairs $i, j$: $i \neq j$ with $\hat{y}_i \neq \hat{y}_j$ such that $\text{sgn}[\hat{y}_j - \hat{y}_i] = \text{sgn}[y_j - y_i]$, where $\hat{y}_j$ and $y_j$ are the model estimate and the actual value of the response variable for customer $j$, respectively. It is also equal to the area under the receiver operating characteristic curve.

indicating that higher APRs lead to lower take-up, as expected. However, lenders generally observe that, all else being equal, riskier customers tend to be less price sensitive and have higher overall take-up rates. This is related to the phenomenon of *price-dependent risk* (Phillips and Raffard 2011). This does not appear in the coefficients for the base model, suggesting the possibility of bias in the base model estimation.

### 4.2. Endogeneity Corrected Price-Response Model

A key question is the extent to which the base price-response model coefficients suffer from endogeneity. If the rates offered for loans are influenced by factors correlated with customer price sensitivity that are not captured in the data, then the base model coefficients may be severely biased. To correct for endogeneity, we use a control function approach in which the endogenous variable (in this case, *CustomerRate*) is regressed against exogenous instruments, and the residual from this regression is entered as an additional explanatory variable in the price-response models. Details of the implementation of this method and the justification of its selection among other endogeneity correction methods can be found in Appendix A.

As noted, we used the control function approach (Petrin and Train 2010) to test and correct for possible endogeneity. As an instrumental variable, we calculated, for each loan application, the average of the interest rates that were offered for similar applications in other regions during the same month. We denote this variable by $\widehat{Rate}$. To calculate $\widehat{Rate}$, we clustered loans based on *Amount*. For each observation, $\widehat{Rate}$ is the average rate offered for loans in the same *Amount* cluster with the same *TermClass*, *VehicleType*, and subvention eligibility condition in other pricing regions[6] during the same month.

This *Hausman-type* variable (Hausman 1997) is an appropriate instrument since it shares the same marginal cost characteristics of the endogenous *CustomerRate*. Also, it averages out the unrecorded customer characteristics, so the instrument is uncorrelated with the error term. Additionally, it does not reflect common demand shocks such as unobserved local advertising or regional promotions. Finally, even though the aggregated auto sales might be correlated across regions, the effectiveness/skill of the sales force in using pricing discretion and the promotions/advertising strategy of the financing deals for the cars are geographically variable enough so that the *CustomerRate* exhibits sufficient regional variation.[7] To illustrate this, we analyzed the variation of the offered rates for "similar" loans in the same month among 22 pricing regions. There were 2,724 different groups of loans in total, but we analyzed the ones that have more than 100 observations to provide the necessary denseness, which corresponds to 1,221 loan categories. For each category, we calculated the coefficient of variation (CV = standard deviation/mean) and the range-to-mean (RM) ratios of the offered rates in each region. The mean of the CVs and RMs among 1,221 loan categories turned out to be 0.2 and 0.7 (with standard deviations of 0.13 and 0.49), respectively, which indicates significant regional variation.

**Table 4 First-Stage Linear Regression Coefficients**

| Explanatory variable | Estimate | Std. error |
|---|---|---|
| *Intercept* | 0.161*** | 1.68E–03 |
| *Tier* (Base level = 5) | | |
| 1 | −0.055*** | 1.94E–04 |
| 2 | −0.049*** | 1.93E–04 |
| 3 | −0.039*** | 1.93E–04 |
| 4 | −0.021*** | 1.99E–04 |
| *TermClass* (Base level = > 60) | | |
| 36 | −0.009*** | 7.71E–05 |
| 48 | −0.004*** | 7.47E–05 |
| 60 | −0.002*** | 4.42E–05 |
| *VehicleType* (Base level = Used) | | |
| New | −0.016*** | 7.08E–05 |
| *SubventionOffer* (Base level = Offered) | | |
| Not offered | 0.011*** | 9.95E–05 |
| $\widehat{FICO}$(/100) | −0.006*** | 2.92E–05 |
| log(*Amount*) | −0.007*** | 5.16E–05 |
| *PrimeRate* | −0.448*** | 1.33E–01 |
| *CustomerCash* ($1,000) | 0.014*** | 1.92E–05 |
| $\widehat{Rate}$ | 0.513*** | 1.36E–03 |
| $R^2$ | 79% | |
| $F$-value | 1,035.96 | |

*Note*. The model includes *DealerID* and *VehicleModelByMonth* fixed effects.
$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Using this instrument, we applied the procedure described in Appendix A. For the first stage, we ran an ordinary least squares (OLS) regression of *CustomerRate* on the exogenous variables and the instrument ($\widehat{Rate}$). Table 4 shows the coefficient estimates for this linear regression except for the *DealerID* and *VehicleModelByMonth* variables, which are omitted for brevity. The first-stage $R^2$-value is 0.79, so the *CustomerRate* is not fully predicted by the exogenous variables and the instruments.[8] We argue that the additional variation in the indirect lender prices is mostly due to the discretion applied by local sales force.

For the second stage, we took the residuals of this first-stage regression, denoted *ResidualforRate*, and ran a probit regression of the *Outcome* on all the exogenous variables, the endogenous variable ($\Delta Rate$), and

[6] The indirect lender had 22 pricing regions across the United States.

[7] We used various other weak instrument tests (e.g., $F$-test) as well to validate our choice of instrument.

[8] The $R^2$-value of the first-stage linear regression model without the instrument was 0.77, so including our instrument, $\widehat{Rate}$, explained a significant amount of the variation in *CustomerRate*.

**Table 5** **Estimated Price Elasticities**

| | Uncorrected | 95% CI | With control function | 95% CI |
|---|---|---|---|---|
| Tier 1 | −0.38 | (−0.39, −0.37) | −1.10 | (−1.15, −1.04) |
| Tier 2 | −0.26 | (−0.26, −0.25) | −0.96 | (−0.98, −0.94) |
| Tier 3 | −0.22 | (−0.23, −0.21) | −0.85 | (−0.87, −0.83) |
| Tier 4 | −0.35 | (−0.37, −0.34) | −0.80 | (−0.83, −0.78) |
| Tier 5 | −0.21 | (−0.22, −0.20) | −0.54 | (−0.56, −0.52) |

*ResidualforRate*. Table 3 shows the results of this regression (fourth and fifth columns) along with the coefficients from the base model. These estimates are normalized as described in Appendix A. The endogeneity corrected model's level of concordance was 87.1% for the entire data set. We also randomly split the data set into training and test samples of equal size. The concordance was 86.7% when we estimated the coefficients using the training set and predicted the customer's decision on the test set. The total log-likelihood was −453,641 in the test set, and the average predicted acceptance probability was 69.14% compared with the true acceptance rate of 68.96%.

As shown in Table 3, all of the variables that entered the base model with high significance ($p < 0.01$) also entered with high significance in the corrected model. Of the 1,628 *VehicleModelByMonth* variable categories, 1,104 were significant at the 0.05 level for the corrected model. The *ResidualforRate* variable was highly significant ($p < 0.0001$). This is consistent with the hypothesis that the local sales force is systematically adjusting rates based on customer willingness to pay.

Most of the coefficients not associated with rate did not change significantly when the model is corrected for endogeneity. The sign of the coefficient of log(*Amount*) changed from positive to negative, consistent with the expectation that, all else being equal, take-up rates are smaller for larger loans. On the other hand, many of the coefficients related to APR changed significantly. In particular, the $\Delta Rate$ by tier variables increased in magnitude by factors ranging from 203% to 387%. This implies that the base model significantly underestimated price sensitivity.

Table 5 shows the corrected and uncorrected price elasticities with their 95% confidence intervals (CIs).[9] We used the mean levels of each continuous variable and the most common levels of each categorical variable[10] to calculate these estimates. We also averaged out the intercept adjustments of the dealers to report the average elasticities over all customers instead of the average elasticity of a single dealer's customers. Correcting for endogeneity increased the estimates of elasticity by factors ranging from 126% to 285%.[11] Figure 1 shows the endogeneity-corrected price-response models together with the uncorrected (base) ones for each risk tier.

## 5. Quantifying the Value of Field Price Discretion

We used the corrected model to estimate the benefits of sales force discretion. We considered only the nonsubvened loans in this analysis (43.02% of the data), because dealers have very limited authority to adjust rates for subvened loans. On the other hand, dealers exhibited significant price discretion for the nonsubvened transactions. Indeed, as shown in Table 6, the customer rates differ from the lender rates for 60.34% of the nonsubvened transactions (41.5% of customer rates are more than the lender rates and 18.84% are less than the lender rates). Table 6 also shows the extent of price discretion along with the average price change (in absolute and relative terms) for each risk tier. Dealers used relatively more price discretion for less risky customers; while they increased the rates for tier 1 customers by 14.27%, they decreased rates for the other tiers, on average. The amount of this decrease is more significant for the riskier customers (tiers 4 and 5). This difference can also be explained by the hypothesis that although headquarters increased the rates for risky customers, the dealers did not internalize this risk directly and so reduced them back to their unadjusted levels.

### 5.1. Fixed Pricing Benchmark Methods

Estimating the implications of field price discretion requires answering two counterfactual questions: (1) What rate other than the actual rate would have been quoted in the absence of field price discretion? And (2) how would customers have responded to these rates? To answer the latter question, we used the endogeneity corrected price-response model to estimate customer responses to the nominal centralized rates[12] (that is, the rate that would be used absent

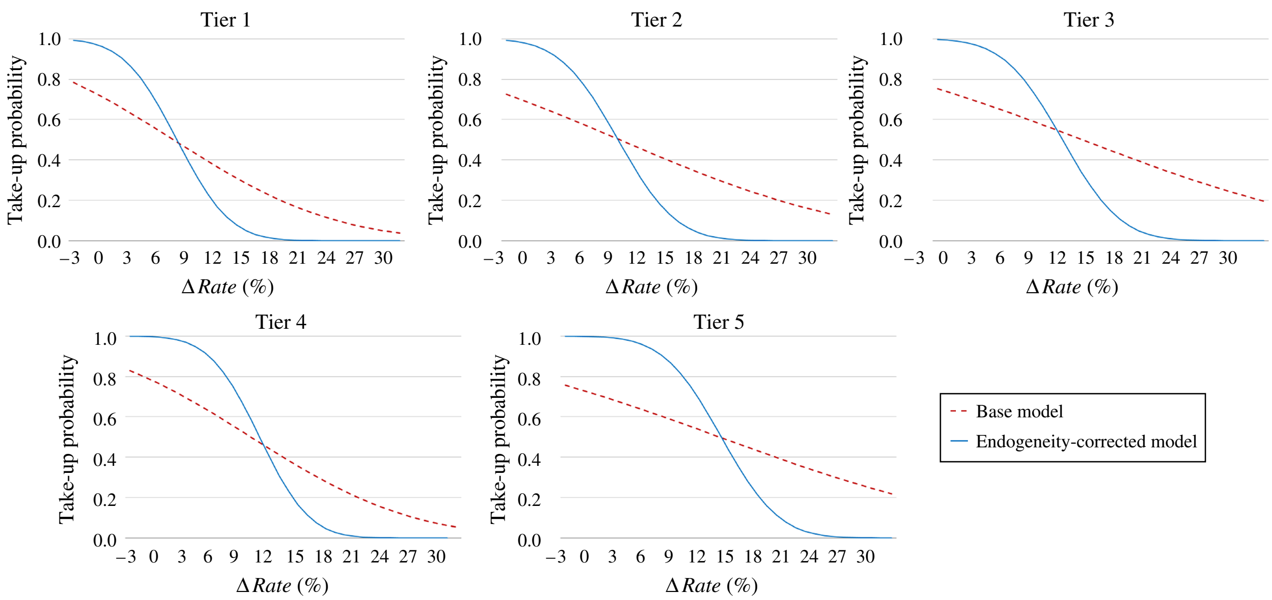
[9] Confidence intervals are computed using the 95% confidence intervals of the $\Delta Rate$ coefficients.

[10] One exception is that we used the nonsubvened level of the *SubventionOffer* variable in the indirect model, since the mean rate that we used was more consistent with a nonsubvened offer.

[11] The magnitude of this endogeneity correction in the indirect lending data is greater than the average changes as a result of endogeneity correction reported for various uniform pricing settings in Bijmolt et al. (2005). They found that incorporating price endogeneity leads to an average increase in estimated price elasticity of 50%. We believe that endogeneity is stronger in this case because of the use of deal-specific information on the part of the sales force.

[12] Endogeneity correction is further crucial for such a counterfactual analysis. Ebbes et al. (2011) state that it is necessary to correct for endogeneity when the goal is a better explanation of the customer behavior in a different environment instead of better predictions in the same environment.

**Figure 1** (Color online) Estimated Price-Response Functions for Each Risk Tier




any dealer discretion). To address the former question, we used the following four approaches to construct benchmark fixed rates.

**5.1.1. Status Quo Fixed Pricing–Projected Rate Method.** In the status quo fixed pricing–projected rate method (SQFP-PR), we used the projected rates from the first-stage (linear) regression of the control function approach as the nominal fixed rates. The justification for using these projected rates comes from the assumptions of the control function approach. As explained in Equation (A1) in Appendix A, the endogenous rate variable can be written as the sum of two parts: a function of the available exogenous variables and instruments and an unobserved component, which is an indicator of the unrecorded transaction characteristics affecting the pricing decision. We estimate these two parts in the first-stage (linear) regression and take the projected rates as the nominal fixed rates and the residuals as the adjustments due to field price discretion. We note that, by using this method, we implicitly assume that the pricing function ($W$ in Equation (A1)) does not change between the two scenarios.

**Table 6** Amount of Price Discretion Use for Each Risk Tier

| | Data frequency (%) | Discretion use ratio (%) | Mean relative change[a] (%) | Mean absolute change[a] (%) |
|---|---|---|---|---|
| Tier 1 | 58.94 | 66.90 | 14.27 | 0.69 |
| Tier 2 | 19.20 | 56.79 | −1.05 | −0.03 |
| Tier 3 | 14.63 | 46.89 | −0.16 | 0.003 |
| Tier 4 | 3.29 | 42.93 | −4.11 | −0.57 |
| Tier 5 | 3.93 | 43.82 | −6.02 | −0.91 |
| Total | | 60.34 | 7.82 | 0.34 |

[a]Over the entire population of the tier.

**5.1.2. Status Quo Fixed Pricing–Lender Rate Method.** Using the lender rates as the nominal fixed rates provides another fixed pricing benchmark, which we call the status quo fixed pricing–lender rate method (SQFP-LR). However, we note that the indirect lender calculates the lender rates knowing that field price discretion will be applied by the dealers. Hence, the lender rates may not be the fixed rates that the lender would charge in the absence of field price discretion. This benchmark still provides valuable insights.

**5.1.3. Optimized Fixed Pricing.** Our third approach to estimating fixed rates is optimized fixed pricing (OPT-FP), where we use the coefficients estimated for the corrected take-up model to determine the price that would maximize expected profit for each deal. We divided the data into two parts based on the transaction date: we used the first part, which covers the period up to the last 12 months (65.3% of the data), as a "training set" and estimated the endogeneity corrected probit model using this data. Then, for the remaining "test set," which covers the last 12 months (34.7% of the data), we calculated the rates that maximize expected profit based on the estimated model and used these as the estimates of the fixed rates.[13] Specifically, for each customer $j$, we found the

[13] The concordance of the endogeneity corrected model in the test set was 83.4%, and the total log-likelihood was −343,651. The average predicted acceptance probability was 74.57% compared with the true acceptance rate of 70.47%. We also tried alternative time periods for the test set as a robustness check. This did not change our main insights.

rates $r_j$ that solve

$$\max_{r_j} \; \{TakeUp\_Probability_j(r_j) \cdot [PoP_j \cdot (TotalPayment_j(r_j) - CapitalCost_j) - (1 - PoP_j) \cdot LGD \cdot CapitalCost_j]\}$$
$$\text{s.t.} \quad r_j \le l_j + 2\% \tag{2}$$

and substituted these rates in Equation (4) to calculate the counterfactual profits. Here, $l_j$ denotes the lender rate for this transaction; *PoP* denotes the "probability of payment," which is calculated by the indirect lender based on their estimates of the default risk of each loan; and *LGD* denotes the "loss-given-default" ratio. We set $LGD = 0.5$, which is the value used by the indirect lender. The take-up probability in problem (2) is calculated using the choice model estimated for the training set. We used pricing regions instead of *DealerID* variables, since the lender sets list rates by region but not by dealer. Additionally, since the lender cannot utilize the *VehicleModelByMonth* dummies in its optimization, we averaged out the fixed effects originating from these variables. The $TotalPayment_j$ is calculated by multiplying the monthly payments (calculated using Equation (1)) by the term of the loan, and $CapitalCost_j$ is obtained by calculating the total payment at the prime rate.[14] We also imposed the constraint that the customer rate for any transaction cannot exceed the lender rate by more than 2.00% (200 basis points). This is consistent with the policy of the lender. The objective function of this problem calculates the expected profit at the rate $r_j$ and is unimodal. For every transaction, we solved this problem using line search in MATLAB.

**5.1.4. Optimized Discretionary Pricing.** In the optimized discretionary pricing (OPT-DP) method, we added the residuals from the first-stage (linear) regression of the control function method ($\mu_n$ in Equation (A1)) to the profit-maximizing rates calculated for the previous benchmark (OPT-FP) and imposed the constraint that the customer rate cannot differ by more than 2% from the OPT-FP rate. We then used these rates as our estimates of the offered rates. Because these residuals are estimates of the adjustments due to field price discretion, using this method enables us to approximate the effects of field price discretion when centralized, data-driven profit-maximizing rates are given to the dealers.

## 5.2. Results

For each benchmark method, we compared the profits of the method with the actual profits. For each transaction $j$, we estimated the profitability as

$$ActualProfit_j = [PoP_j \cdot (TotalPayment_j - CapitalCost_j) - (1 - PoP_j) \cdot LGD \cdot CapitalCost_j] \cdot Booked\_Indicator_j, \tag{3}$$

$$BenchmarkProfit_j = [PoP_j \cdot (TotalPayment_j - CapitalCost_j) - (1 - PoP_j) \cdot LGD \cdot CapitalCost_j] \cdot Conditional\_TakeUp\_Probability_j, \tag{4}$$

where *ActualProfit* is the profit that was achieved from the acceptance and *BenchmarkProfit* is our counterfactual estimate of the expected profit that would have been achieved if the dealer had used rates from the benchmark method. The conditional take-up probability is defined as

$$Conditional\_TakeUp\_Probability_j(\hat{r}) = P(\text{taking up the loan at rate } \hat{r} \mid \text{customer's decision at rate} = CustomerRate). \tag{5}$$

If the lender offers a nonnegotiable rate that is lower than an accepted rate (respectively, higher than a rejected rate), we assume that the same customer would accept (respectively, reject) the loan offer. For the other cases, we calculate the conditional take-up probability using the endogeneity corrected price-response model. Profit changes are calculated as $(BenchmarkProfit_j - ActualProfit_j)/ActualProfit_j$.

The second column of Table 7 shows the means of the benchmark rates, and the sixth column shows the mean of the offered rates in the current system, which we call *status quo discretionary pricing* (SQDP). Mean SQDP rates for all nonsubvened customers and for the nonsubvened customers of the test set are 8.25% and 6.75%, respectively. Hence the optimal rates in OPT-FP are higher than the actual rates in SQDP, on average.[15] This can potentially be explained by the phenomenon derived by Kuo et al. (2013): retailers prefer negotiation at lower wholesale prices, and in some cases (when the cost of negotiation is moderate), the manufacturer may offer a substantial discount from the optimal price to the retailer to give the retailer more room for price discrimination via negotiation.

[14] The cost of capital experienced by the lender over this period was actually somewhat higher than the prime rate. As discussed in §5.2, using a higher capital cost within realistic limits does not change our results.

[15] If we had used the actual capital cost of the indirect lender (which is higher than the prime rate), we would have obtained higher optimal rates. As a sensitivity test, we estimated the outcomes using capital costs 0.25% and 0.50% (25 and 50 basis points) higher than the prime rates. Under these assumptions, the optimal rates turned out to be 7.81% and 7.84%, the means of the *ActualProfit* were \$2,500 and \$2,374, and the means of the *BenchmarkProfit* were \$3,065 and \$2,949, respectively. In short, the results were not significantly changed by the assumption of higher capital costs.

**Table 7 Summary Statistics for the Benchmark Methods**

| Benchmark method | Mean rate (%) | Mean conditional take-up prob. (%) | Mean profit ($) | Mean profit's 95% CI ($) | SQDP Mean rate (%) | SQDP Take-up ratio (%) | SQDP Mean profit ($) | Benchmark profit diff. (%) |
|---|---|---|---|---|---|---|---|---|
| SQFP-PR | 7.80 | 77.61 | 2,562 | (2,513, 2,603) | 8.25 | 77.94 | 2,888 | −11.31 |
| SQFP-LR | 7.91 | 75.63 | 2,632 | (2,603, 2,654) | 8.25 | 77.94 | 2,888 | −8.87 |
| OPT-FP | 7.72[a] | 78.10[a] | 3,150[a] | (3,061, 3,224)[a] | 6.75[a] | 82.09[a] | 2,626[a] | 19.95 |
| OPT-DP | 7.87[a] | 76.23[a] | 3,167[a] | (3,043, 3,269)[a] | 6.75[a] | 82.09[a] | 2,626[a] | 20.60 |

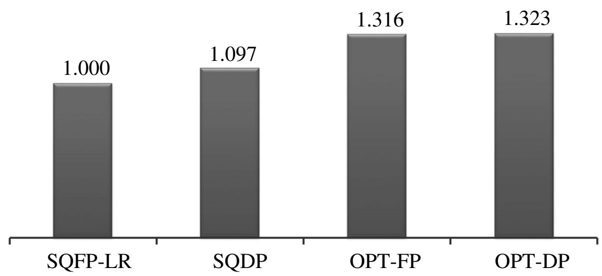
[a]Test set customers only.

Table 7 also shows the average profit of the status quo, each benchmark method, and the difference between them. For each transaction, the 95% confidence interval around the benchmark profit is computed using the 95% confidence interval of the conditional take-up probability derived from the endogeneity corrected probit model. Their means are also reported in Table 7. The estimated profits of the first two benchmarks, SQFP-PR and SQFP-LR, are significantly lower than the actual profit (−11.31% and −8.87%, respectively). This suggests that dealers were, in fact, adjusting rates in a way that improved profitability.

When we use OPT-FP, the counterfactual profit is higher than the status quo profit by 19.95%. This result suggests that even though the dealers were improving profits relative to the nominal rates set by the lender, there was an additional opportunity to improve profitability further through optimized prices. Finally, the highest levels of profit are achieved when a combination of analytical, data-driven profit maximization procedure and field price discretion is used. In particular, OPT-DP increased the profit over OPT-FP another 0.65% relative to the status quo profits. But this is substantially less than the 8%–11% improvement relative to the status quo centralized prices. Hence, it suggests that even though dealer price discretion improved profits generated using lender rates or projected rates, the profit improvement as a result of field price discretion is much less significant—almost negligible—once lender rates have been optimized. Note that this slight profit increase may not be sufficient to justify the cost of highly trained sales staff. Figure 2 summarizes the normalized profit performance of the pricing benchmarks, taking SQFP-LR as the base case.

We also analyzed the distribution of these profits derived from different benchmarks across pricing regions and customer groups.[16] Table 8 shows the price discretion use ratios and its average size

**Figure 2 Normalized Profit Performances of Three Pricing Benchmarks**




(in absolute and relative terms) for each pricing region, and Figure 3 shows the percentage change from the actual profits when the SQFP-LR and OPT-FP methods are used. We categorized the pricing regions into five groups based on their geography. There were no significant patterns in the frequency of price discretion use across pricing regions. The average amount of price discretion and profit changes relative to different benchmarks are close to each other within each group; however, they exhibit significant variation across groups, which suggests that there are systematic regional variations in the aggregate performance of the local sales staff. For example, sales staff in the northeast area changed the lender rates by 10.6%, on average, which led to a 13.3% profit increase relative to the lender rates. Sales staff in the southeast area, on the other hand, changed the rates by 5.8%, on average, which led to a 6.2% profit increase relative to the lender rates.

Furthermore, Table 8 and Figure 3 together suggest that the additional profits generated by the dealers' use of price discretion increased with the size of that discretion, which supports the claim that dealers were offering rates that were closer to the customers' willingness to pay than the lender rates. Figure 3 also suggests that when the increase in profit stemming from the use of price discretion is higher, then the additional profitability that can be obtained using profit-maximizing rates decreases, which is quite intuitive. OPT-FP still generated higher profits in every region.

[16] We disguised the name of the pricing regions for anonymity considerations.

**Figure 3** (Color online) Estimated OPT-FP and SQFP-LR Profits by Pricing Region

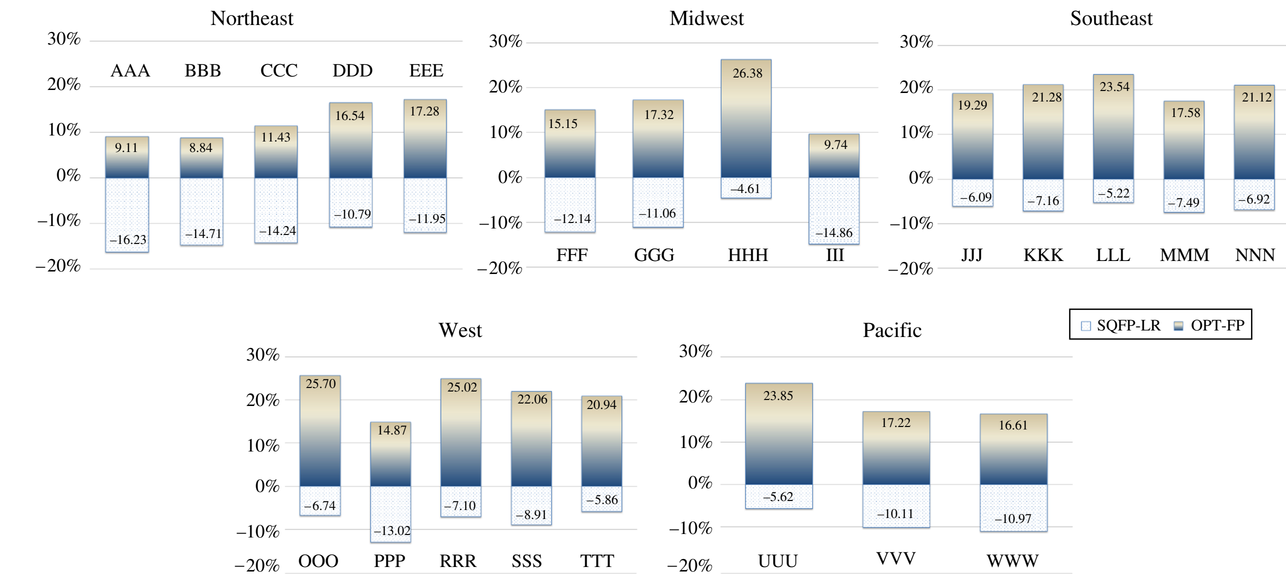



Table 9 shows the percentage change from the actual profits when the SQFP-LR and OPT-FP methods are used for each risk tier. The status quo average profit is significantly higher than the lender rate profits only for tier 1, which is consistent with the amount of the price discretion used for each tier presented in Table 6. Therefore, again consistent with the hypothesis that dealers were doing better than the lender rates but were undercharging customers with respect to the optimal rates, the potential for profit increase by using the profit-maximizing rates increased for riskier customers, since the dealers did not increase the rates sufficiently for riskier customers and decreased them even more for tier 5 customers.

We make two final observations from these experiments regarding the endogeneity correction of the price-response models. First, the profit increase under field price discretion is significantly underestimated when the uncorrected (base) model is used to calculate profits, since customers are estimated to be less price sensitive in that model. Second, when we use the uncorrected model in the profit maximization procedure, the resulting profit-maximizing rates turned out to be much higher than the ones reported in Table 7 (the mean was approximately 30%). This is again intuitive because of the estimation of less price-sensitive customers. Hence, we conclude that

**Table 8** Amount of Price Discretion Use for Each Pricing Region

| | Data frequency (%) | Discretion use ratio (%) | Mean relative change[a] (%) | Mean absolute change[a] (%) |
|---|---|---|---|---|
| Northeast | | | | |
| AAA | 2.89 | 65.60 | 14.70 | 0.77 |
| BBB | 3.09 | 49.97 | 8.40 | 0.58 |
| CCC | 2.53 | 53.73 | 11.01 | 0.62 |
| DDD | 4.18 | 58.59 | 9.37 | 0.43 |
| EEE | 3.41 | 64.08 | 10.48 | 0.54 |
| Midwest | | | | |
| FFF | 5.58 | 57.35 | 9.50 | 0.45 |
| GGG | 4.38 | 61.38 | 11.11 | 0.51 |
| HHH | 6.21 | 41.92 | 2.28 | 0.11 |
| III | 4.53 | 63.98 | 14.02 | 0.68 |
| Southeast | | | | |
| JJJ | 3.97 | 60.45 | 5.36 | 0.18 |
| KKK | 3.82 | 59.98 | 6.46 | 0.23 |
| LLL | 8.07 | 60.19 | 5.12 | 0.11 |
| MMM | 1.83 | 62.92 | 5.77 | 0.30 |
| NNN | 5.32 | 64.71 | 6.67 | 0.29 |
| West | | | | |
| OOO | 9.38 | 62.58 | 6.87 | 0.23 |
| PPP | 2.33 | 76.48 | 12.63 | 0.67 |
| RRR | 7.64 | 61.41 | 7.08 | 0.27 |
| SSS | 7.11 | 58.06 | 7.74 | 0.31 |
| TTT | 2.42 | 71.06 | 6.95 | 0.25 |
| Pacific | | | | |
| UUU | 3.91 | 60.29 | 3.87 | 0.12 |
| VVV | 2.08 | 63.44 | 8.25 | 0.41 |
| WWW | 2.44 | 70.39 | 9.30 | 0.52 |

[a]Over the entire population of the region.

**Table 9** Mean Profit Levels for the SQDP, SQFP-LR, and OPT-FP Methods for Each Risk Tier

| | SQDP[a] ($) | OPT-FP[a] ($) | Difference (%) | SQDP ($) | SQFP-LR ($) | Difference (%) |
|---|---|---|---|---|---|---|
| Tier 1 | 2,589 | 3,027 | 16.92 | 2,889 | 2,486 | −13.95 |
| Tier 2 | 3,120 | 3,939 | 26.24 | 3,340 | 3,325 | −0.45 |
| Tier 3 | 2,596 | 3,026 | 16.57 | 2,747 | 2,646 | −3.67 |
| Tier 4 | 2,206 | 2,827 | 28.14 | 2,216 | 2,139 | −3.49 |
| Tier 5 | 1,704 | 2,462 | 44.51 | 1,756 | 1,797 | 2.34 |
| Total | 2,626 | 3,150 | 19.95 | 2,888 | 2,632 | −8.87 |

[a]Test set customers only.

not accounting for endogeneity would lead to substantial mispricing for this lender.

## 6. Robustness Analysis

We tested the robustness of the base model and corrected model by using different variable transformations and model specifications, as well as by including alternative variables.

We tested different specifications of the *CustomerRate*, *PrimeRate*, and $\widehat{FICO}$ variables in the models. The $\widehat{FICO}$ variable is the absolute value of the difference between a customer's FICO score and the average FICO score of that customer's risk tier. Replacing the absolute value with the percentage difference did not significantly change any of the coefficient estimates, nor did it improve the model fit. We tested the percentage difference of the *CustomerRate* compared with *PrimeRate* in our choice models as well. When we include both types of rate variables ($\Delta Rate$ and the percentage rate difference) in the model, we obtained similar results. Using the percentage rate difference instead of $\Delta Rate$ introduced a slight bias into this variable's coefficient estimation (in the positive direction), but this small bias was not significant enough to change our interpretations of the results.

We included the *VehicleModelByMonth* dummies to correct for the endogeneity in *CustomerRate* resulting from common demand shocks. We also tried using different dummy variables (day-of-week, week, quarter, etc.). However, we found no evidence of consistent seasonal effects on APR.

Finally, we tested different model specifications such as logistic regression and observed very similar model estimates. We also used two-stage least squares (2SLS) regression. In particular, we used linear regression both for the base model and for the second stage of the endogeneity correction procedure. In this case, using the residuals from the first-stage (linear) regression as an extra control variable in the second stage and using the predicted rates instead of the actual rates in the second stage are equivalent. Table 10 shows the resulting parameter estimates for both the base and the corrected (by 2SLS) models. We also included pricing region-month interactions in these models to clear possible endogeneity arising from region-specific temporal pricing fluctuations. Adding these interactions did not have a significant effect on the model coefficient estimations, which suggests that such temporal fluctuations were not present at a significant level. Figure 4 illustrates the price-response change for tier 1 customers (other tiers exhibit similar behavior). Clearly, customers' price sensitivities are significantly underestimated in the base model in this case as well.

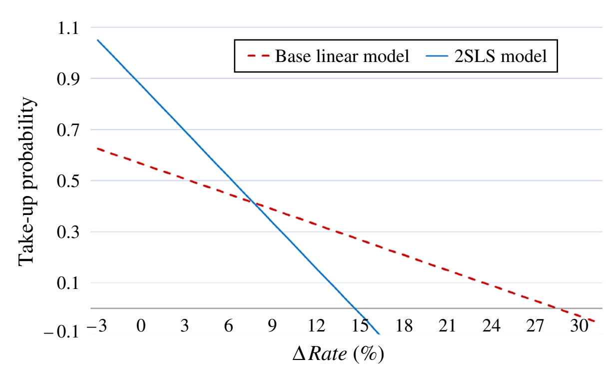
We note that although Hausman-type instruments are commonly used in the literature, they are not without their critics (Breshanan 1997). As stated in Petrin and Train (2010), aggregate demographics such as average income, household size, etc., do not enter the disaggregate models, but they affect market price. Therefore, they can serve as extra instruments for demand estimation. However, it is difficult to find additional instruments in our case, since the only available demographic is the FICO score of each customer, but the average FICO scores for each pricing region do not exhibit much variation. Therefore, we rely on $\widehat{Rate}$, which is a robust and efficient instrument and can possibly be effective in other customized pricing environments. In fact, verifying efficient use of

**Table 10 Base Linear Regression and 2SLS Method Estimates**

| Explanatory variable | Base | | 2SLS | |
|---|---|---|---|---|
| | Estimate | Std. error | Estimate | Std. error |
| *Intercept* | 0.77*** | 0.030 | 2.22*** | 0.033 |
| *Tier* (Base level = 5) | | | | |
| 1 | 0.02*** | 0.011 | −0.54*** | 0.013 |
| 2 | −0.02 | 0.011 | −0.55*** | 0.012 |
| 3 | 0.01 | 0.011 | −0.47*** | 0.012 |
| 4 | 0.06*** | 0.011 | −0.34*** | 0.012 |
| *TermClass* (Base level = > 60) | | | | |
| 36 | 0.16*** | 0.001 | 0.08*** | 0.001 |
| 48 | 0.09*** | 0.001 | 0.04*** | 0.001 |
| 60 | 0.05*** | 0.001 | 0.01*** | 0.001 |
| *VehicleType* (Base level = Used) | | | | |
| New | −0.28*** | 0.001 | −0.37*** | 0.002 |
| *SubventionOffer* (Base level = Offered) | | | | |
| Not offered | −0.41*** | 0.001 | −0.24*** | 0.002 |
| $\widehat{FICO}$ (/100) | −0.05*** | 0.001 | −0.07*** | 0.001 |
| log(*Amount*) | 0.03*** | 0.001 | −0.03*** | 0.001 |
| *PrimeRate* | 0.77 | 2.239 | −3.88* | 2.239 |
| *CustomerCash* ($1,000) | 0.18*** | 0.000 | 0.24*** | 0.001 |
| Δ*RateforTier1* | −1.99*** | 0.045 | −5.97*** | 0.045 |
| Δ*RateforTier2* | −1.25*** | 0.046 | −5.32*** | 0.046 |
| Δ*RateforTier3* | −1.14*** | 0.045 | −5.29*** | 0.045 |
| Δ*RateforTier4* | −1.89** | 0.048 | −6.13*** | 0.048 |
| Δ*RateforTier5* | −1.04*** | 0.045 | −5.30*** | 0.045 |
| *ResidualforRate* | NA | | 4.46*** | 0.044 |

*Note.* Both models also include *DealerID*, *VehicleModelByMonth*, and region × month fixed effects.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

**Figure 4 (Color online) Estimated Linear Price-Response Functions for Tier 1**



this instrument in other settings is a promising future research direction.

We also tested the robustness of our counterfactual profit analysis to some of the cost assumptions. One possible explanation for the difference between the profit-maximizing rate profits and the status quo profits could be that incentives of the dealers and lending company are not aligned. In fact, dealers profit from the difference between the lender rates and customer rates but not from the difference between the prime rates and customer rates; i.e., a dealer's cost for the loan is not the *CapitalCost* but the total payment calculated at the lender rate.

To test this argument, we calculated the profit-maximizing rates using the lender rates as the cost. The gap between the estimated benchmark profit and its status quo level was smaller in the unconstrained optimization, i.e., when we do not restrain the final rate to be at most 2% more than the lender rate (as in problem (2)). However, it did not change our results significantly for the constrained optimization problem. The reason for this is the optimal levels of the rates for both cases are usually more than their upper bounds, so—for both cases—the constrained optimum is achieved in the upper bound. Therefore, different incentives alone cannot explain the profit gap between the profit-maximizing rate benchmark (OPT-FP) and the status quo (SQDP).

### 6.1. Online Lender Data Set

As another robustness check, we performed a similar analysis using a different data set from an online lender who sets all prices centrally.[17] The loans in this data set were from a different period, so the results are not directly comparable. However, we use this analysis primarily to support our hypothesis that the endogeneity observed in the indirect lender data was due to deal-specific information possessed by the field sales staff, and not to any intrinsic feature of the auto-lending market in general.

This second auto lender, which we call the *online lender*, offered loans exclusively through the Internet with no local discretion. Details of the online channel, variable descriptions, and analysis can be found in Appendix B.

Table 11 presents the coefficients for the base and corrected models.[18] The *ResidualforRate* variable was not statistically significant for the online lender[19]

**Table 11 Base and Endogeneity Corrected Choice Model (Probit) Estimates for the Online Lender Data**

| Explanatory variable | Base | | Corrected | |
|---|---|---|---|---|
| | Estimate | Std. error | Estimate | Std. error |
| *Intercept* | 14.76*** | 0.1457 | 14.78*** | 0.1528 |
| *Tier* (Base level = 1) | | | | |
| 2 | −0.62*** | 0.0690 | −0.60*** | 0.0717 |
| 3 | −0.94*** | 0.0710 | −0.92*** | 0.0814 |
| 4 | −0.80*** | 0.1021 | −0.75*** | 0.1378 |
| *Partner* (Base level = Direct) | | | | |
| Partner A | −0.40*** | 0.0176 | −0.39*** | 0.0182 |
| Other | −0.24*** | 0.0098 | −0.24*** | 0.0098 |
| *TermClass* (Base level = 36) | | | | |
| 48 | 0.27*** | 0.0201 | 0.29*** | 0.0209 |
| 60 | 0.71*** | 0.0170 | 0.71*** | 0.0179 |
| > 60 | 1.51*** | 0.0219 | 1.52*** | 0.0294 |
| *VehicleType* (Base level = New) | | | | |
| Used | 1.42*** | 0.0118 | 1.43*** | 0.0189 |
| $\widehat{FICO}$(/100) | −0.30*** | 0.0221 | −0.30*** | 0.0230 |
| log(*Amount*) | −1.39*** | 0.0131 | −1.39*** | 0.0133 |
| *PrimeRate* | 5.15** | 2.1489 | 5.08** | 2.1519 |
| Δ*RateforTier1* | −80.35*** | 1.7096 | −81.62*** | 2.8615 |
| Δ*RateforTier2* | −55.25*** | 1.9073 | −56.54*** | 1.9078 |
| Δ*RateforTier3* | −42.34*** | 1.8394 | −43.63*** | 1.8396 |
| Δ*RateforTier4* | −38.00*** | 2.0047 | −39.29*** | 2.0048 |
| *ResidualforRate* | NA | | 1.14 | 2.5438 |
| Concordance | 87.3% | | 87.3% | |
| Log-likelihood | −45,297 | | −45,297 | |

*Note.* Both models also include month fixed effects.
$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

($p = 0.58$), which is consistent with the hypothesis that endogeneity was not present in the online lender data. This supports the hypothesis that the endogeneity observed in the indirect lender data set was due to field price discretion and is not an industry-wide effect.

We also performed a similar counterfactual analysis to the online data set. In particular, we compared the actual mean profit with the benchmark mean profit, where we used two benchmarks: the *projected rate* (SQFP-PR) and *profit-maximizing rate* (OPT-FP) methods. The current profits were only 0.76% more than the SQFP-PR benchmark. This difference is small as expected because the online lender was already using a centralized pricing system.[20]

For the latter benchmark, we calculated the profit-maximizing rates using a similar estimation and optimization procedure as in the indirect lending data.[21] It turned out that the online lender was significantly

[17] This data set is available from the Columbia Center for Pricing and Revenue Management website at http://www8.gsb.columbia.edu/cprm/research/datasets#Online_Auto-Lending (last accessed February 27, 2015).

[18] The coefficients for the month variable are not shown for brevity.

[19] We ruled out the possibility that this was the result of multicollinearity using the condition index test (Lesaffre and Marx 1993).

[20] We note that, in these calculations, we did not take the default risk into account because of lack of *PoP* estimates in the data set; however, calculations with estimated average *PoP*'s for each tier did not change the results significantly.

[21] We did not incorporate the *PoP* and *LGD* values and the constraint of the problem (2).

overpricing (as opposed to the indirect lender who was underpricing), since the mean of the unconstrained optimal rates for the test set (last 12 months) was 4.9% compared with the mean of the actual rates of 6.2%.[22] This benchmark's mean profit was 38% higher than the mean actual profit ($674 versus $488). These observations are not surprising since the actual take-up rate was only 20.1% for the test set.[23] Using the optimal rates increased the mean conditional take-up probability to 33%. This analysis also supports the argument that there is a potential for increasing the profits by implementing a more structured and data-driven profit maximization procedure.

## 7. Discussion and Conclusion

We empirically studied the trade-off between fixed and discretionary pricing policies for a customized pricing environment. There are many theoretical papers that investigate this trade-off in uniform pricing environments. These studies show that discretionary pricing policies, in which the price is determined via a negotiation between the salesperson and the customer, dominates fixed pricing policies in cases where the wholesale prices are low, bargaining costs are low (for both retailers and consumers), there are substantial variation in customers' willingness to pay, and customers have low bargaining power. In these cases, the seller has the opportunity to price discriminate through the use of negotiation. On the other hand, it is not clear from the literature whether these conditions lead to similar results for customized pricing settings. With the emergence of advanced statistical pricing tools that can utilize large amounts of data, a centralized fixed pricing system may better segment potential customers based on willingness to pay and risk. These data-driven capabilities have the potential to provide better price discrimination opportunities compared with the field sales force's negotiation skills.

To empirically estimate fixed and discretionary pricing policies, we utilized a unique data set from an indirect auto lender that used field negotiation to determine prices. First, using a control function approach, we estimated the price response of the customers by correcting for the price endogeneity potentially caused by the negotiations with customers. We then developed a method for quantifying the value added through field price discretion using a counterfactual analysis, in which we compared status quo profits to different benchmark profits.

Our main empirical findings can be stated as follows:

1. Applying their discretion to negotiate, local sales staff charged "better" prices than the list rates. Using their pricing discretion, the sales force increased profits by 8.9% compared to the SQFP-LR benchmark and 11.3% compared to the SQFP-PR benchmark.[24] This supports the hypothesis that the field sales staff can indeed use local information to negotiate more profitable prices.
2. However, the list rates used by the indirect lender are suboptimal (this may partially be intentional, since the indirect lender generates its prices knowing that the dealers will use their discretion to adjust them). And using rates generated by a data-driven profit maximization procedure (OPT-FP), which utilizes data of all nationwide sales, can significantly increase profitability, about 20% (compared with the rates determined after the use of field price discretion—SQDP).
3. The combination of optimized list rates and field price discretion (OPT-DP) generated the highest level of profits. However, the incremental value of price discretion is not significant in this case (0.64%), since most of the value is already captured by the data-driven optimization process; this left very little room for improvement to subsequent negotiations.

Therefore, fixed pricing (OPT-FP) provides better price discrimination compared with the field sales force's negotiation skills under the conditions of the analyzed indirect lender; hence it should be favorable to discretionary pricing policy (SQDP) in this particular setting (assuming the marginal profit increase in OPT-DP does not justify the cost of field price discretion). These empirical results suggest that conclusions about the benefits of fixed versus discretionary pricing policies in uniform pricing environments may not extend to customized pricing settings. More theoretical research in this area is needed.

Finally, as is the case in most empirical studies, our work has limitations that stem from data availability problems. For example, we do not have reliable data to represent the competition in the auto lending industry during the analyzed period of time, and we do not have access to the operating costs of the lender (hiring and training sales force, for instance) or the customer demographics that could be useful in our estimation. It is also possible that the endogeneity observed does not stem entirely from field price

[22] We note that this difference may also be a result of incomplete cost information. As is the case of indirect lender data, we do not have reliable capital cost information for the online lender data—we used the prime rate for this optimization as well. So using a higher cost level would increase the mean of the optimal rates.

[23] During the period of the online lender data, many dealerships were offering heavily discounted (subvened) deals for auto loans (e.g., 0% financing, 1% financing). These deals were not available through the online lender, which meant that, for many new cars at least, the rates offered by the online lender were uncompetitive.

[24] We note that we did not incorporate the bargaining costs associated with the decentralized pricing policy. So this increase should be interpreted as an upper bound.

discretion but from field price discretion plus other unidentified sources. We showed that similar residual correlation did not occur in an online auto lender with no field price discretion, which shows that at least the residual correlation is not intrinsic to the auto lending industry. However, the data were from a different time period, and the possibility of change in the industry cannot be fully ruled out. Some of these limitations suggest future research directions. In particular, the relationship between the automobile sales and auto lending processes would be worth investigating both empirically and analytically. The effect of automobile inventory levels on the financing process also deserves investigation.

**Acknowledgments**
The authors acknowledge the very helpful suggestions and comments of Marcelo Olivares and Oded Netzer from the Columbia Business School, and Robin Raffard and his colleagues from Nomis Solutions. The department and associate editors and three anonymous referees also made a number of suggestions that resulted in a much improved paper. This research was supported by a grant from the W. Edwards Deming Center at Columbia Business School.

## Appendix A. Endogeneity Correction Method

To correct for endogeneity, we use a control function approach. Although there are other methods available to control for endogeneity, the control function approach is best suited to our situation. Note that 2SLS regression is commonly used in linear models, but it cannot be easily extended to nonlinear models such as logit or probit (Foster 1997).[25] Furthermore, the control function approach is better suited to nonlinear models with continuous endogenous variables than the generalized method of moments approach. In particular, when some endogenous variables are continuous, as is the case in our situation, the control function approach makes it much easier to test the endogeneity of a variable as well as compute the average partial effects of each variable (Wooldridge 2002).

Following Petrin and Train (2010), we assume that the utility that consumer $n$ would achieve from accepting a loan offer can be written as $U_n = V(y_n, x_n, \beta) + \epsilon_n$, where $y_n$ is the observed endogenous variable, $x_n$ is a vector of observed exogenous variables, $\beta$ is a vector of parameters, and $\epsilon_n$ denotes the unobserved component of utility. Endogeneity occurs when $\epsilon_n$ is correlated with $y_n$. The binary response model assumes that the customer accepts the offer if her utility from doing so is positive. This approach assumes that $y_n$ can be written as

$$y_n = W(x_n, z_n; \gamma) + \mu_n, \tag{A1}$$

where $z_n$ denotes the instrumental variables that do not enter utility directly but affect $y_n$, $\gamma$ denotes parameters of $W$, and $\mu_n$ denotes the unobserved component. Note that $\mu_n$ and $\epsilon_n$ are independent of $x_n$ and $z_n$ but are not independent of each other. The source of the dependence between $y_n$ and $\epsilon_n$ is the fact that $\mu_n$ affects $y_n$ and is correlated with $\epsilon_n$. Under these assumptions, conditional on $\mu_n$, $\epsilon_n$ is independent of $y_n$ (Petrin and Train 2010).

Decomposing $\epsilon_n$ into the part that can be explained by a general function of $\mu_n$ and the residual yields $\epsilon_n = CF(\mu_n; \lambda) + \tilde{\epsilon}_n$, where $CF(\mu_n; \lambda)$ denotes the control function with parameters $\lambda$. Several parametric forms might be assumed for $CF(\mu_n; \lambda)$. We assume that the errors in both equations, $\mu_n$ and $\epsilon_n$, are jointly normal, which is equivalent to assuming that the control function is linear in $\mu_n$ (the analysis goes through if we use a nonlinear approximation provided we have sufficient identifying assumptions), since $CF(\mu_n; \lambda) = \mathbf{E}(\epsilon_n \mid \mu_n) = \lambda\mu_n$ and the deviations $\tilde{\epsilon}_n = \epsilon_n - CF(\mu_n; \lambda)$ are independent of $\mu_n$ and all other regressors. Then the utility function becomes $U_n = V(y_n, x_n, \beta) + \lambda\mu_n + \tilde{\epsilon}_n$, where $\tilde{\epsilon}_n$ are independent and identically distributed (i.i.d.) normal with zero mean. Since we assume fixed $\beta$ for every consumer, the model becomes an independent probit with the residual entering as an extra variable. Hence, this model specification coincides with a price-response model commonly used in customized pricing environments, which indeed motivated our selection.[26] Inclusion of the residuals $\mu_n$ "controls" for the endogenous $y_n$ in the original equation (with sampling error, since $\mu_n \neq \hat{\mu}_n$).

Following the assumptions stated in Wooldridge (2002), we use the two-step approach developed by Rivers and Vuong (1988): (1) Perform an OLS regression of the endogenous variable $y$ on the exogenous variables $x$ and instruments $z$ to obtain the residuals $\hat{\mu}$. (2) Perform a probit regression of the take-up variable on the exogenous variables $x$, endogenous variable $y$, and $\hat{\mu}$ to estimate the (unnormalized) coefficients.

Testing the null hypothesis that $y$ is exogenous is straightforward. Asymptotically, a simple $t$-test on $\hat{\mu}$ is valid to test $H_0$: $\lambda = 0$. If this $t$-statistic is sufficiently large, meaning $\hat{\mu}$ is significant in the probit regression, we can reject the null hypothesis and conclude that $y$ is endogenous.

We made two adjustments to the estimation results: First, we normalized the step 2 coefficient estimates by the factor $(\hat{\lambda}^2\hat{\tau}^2 + 1)^{1/2}$, where $\hat{\lambda}$ is the coefficient of $\hat{\mu}_n$ and $\hat{\tau}^2$ is the error variance estimator from the first-stage regression of $y$ on $(x, z)$[27] (Wooldridge 2002). Second, we used the bootstrapping procedure for estimating the true standard errors.[28] In particular, we repeatedly estimated the first-stage regression with bootstrapped samples, obtained the residuals, and estimated the second-step model with the new residuals. The variance in the estimates over the bootstrapped first-stage samples was added to the sampling variance to estimate the true standard error (Petrin and Train 2002).

[25] We still applied this method in §6 and observed similar behavior of the estimated models.

[26] Another commonly used assumption is that $\epsilon_n = \epsilon_n^1 + \epsilon_n^2$, where $\epsilon_n^1$ and $\mu_n$ are jointly normal and $\epsilon_n^2$ is i.i.d. extreme value. This specification leads to a mixed logit model as specified in Villas-Boas and Winer (1999).

[27] We need to do this adjustment since $\text{Var}(\tilde{\epsilon}_n) < \text{Var}(\epsilon_n)$. As noted in Petrin and Train (2010), normalizing by setting $\text{Var}(\tilde{\epsilon}_n) = 1$ increases the magnitude of coefficients with respect to the normalization $\text{Var}(\epsilon_n) = 1$.

[28] The standard errors and test statistics calculated in the second-step regression are not accurate because the second step uses an estimate of $\mu_n$ from the first step, as opposed to the true $\mu_n$, which induces additional error.

**Table B.1 Summary Statistics for Online Lender Data Variables**

| Variable | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|
| *FICOScore* | 727.34 | 45.43 | 594 | 854 |
| *Term* (months) | 58.03 | 10.53 | 36 | 72 |
| *Amount* ($) | 28,551 | 10,791 | 4,770 | 100,000 |
| *CustomerRate* (%) | 5.47 | 1.52 | 2.45 | 13.9 |
| *PrimeRate* (%) | 1.32 | 0.24 | 1.02 | 1.84 |
| *Competitor'sRate* (%) | 4.71 | 0.58 | 2.99 | 6.45 |

| Variable | Category | Frequency | % |
|---|---|---|---|
| *Tier* | 1 | 74,890 | 48.96 |
| | 2 | 32,047 | 20.95 |
| | 3 | 28,465 | 18.61 |
| | 4 | 17,561 | 11.48 |
| *VehicleType* | New | 115,531 | 75.53 |
| | Used | 37,432 | 24.47 |
| *Partner* | Direct | 62,982 | 41.17 |
| | Partner A | 20,905 | 13.67 |
| | Other partners | 69,076 | 54.84 |
| *Outcome* | Not take | 126,641 | 82.79 |
| | Take | 26,322 | 17.21 |

## Appendix B. Details of the Online Lender Analysis

The online channel and the indirect channel differ both in the selling and pricing process and in the information available to the lender at the time of pricing. In the online channel, a prospective borrower fills out an online application, which includes information similar to that required for the indirect lender application. Based on this information, the Internet lender first determines whether it wishes to extend a loan, based on the default risk, which is again estimated similar to the indirect lenders. If an applicant is approved, then the lender determines which price (APR) to offer. After being approved and learning the price, the applicant has 45 days to accept the loan. Applicants who do not accept the loan by the end of this period are considered lost. The data set for the online lender includes all approved applications from July 2002 to November 2004. This consists of 152,963 approved applications, of which 26,322 (17%) were taken up and resulted in loans.

The variables used in the online base model are shown in Table B.1, along with their summary statistics. The online lender categorized borrowers into four risk tiers.[29] Loan applicants for the online lender either arrived directly to the lender's website or were directed to the website from another site. The online lender had an exclusive commission agreement with one site ("partner A") for referrals. There were no subvened deals offered by the online lender during the data period, and customer cash information was not available for online lender applications. Additionally, since there were no model-based promotions offered by the online lender, we only included the month of each observation in the online model.

We applied the same procedures for the base model estimation and endogeneity correction. Table B.2 presents the coefficients for the first-stage linear regression. The first-stage $R^2$-value is 0.82.[30]

**Table B.2 First-Stage Linear Regression Coefficients for the Online Lender Data**

| Explanatory variable | Estimate | Std. err. |
|---|---|---|
| *Intercept* | $-1.168^{***}$ | 0.0941 |
| *Tier* (Base level = 4) | | |
| 1 | 0.012 | 0.0354 |
| 2 | 0.025 | 0.0286 |
| 3 | 0.013 | 0.0212 |
| *TermClass* (Base level = > 60) | | |
| 36 | −0.001 | 0.0102 |
| 48 | 0.011 | 0.0083 |
| 60 | −0.006 | 0.0071 |
| *VehicleType* (Base level = Used) | | |
| New | $-0.014^{*}$ | 0.0076 |
| *Partner* (Base level = Other) | | |
| Direct | $0.022^{***}$ | 0.0036 |
| Partner A | $0.109^{***}$ | 0.0055 |
| $\widehat{FICO}$(/100) | $-0.271^{***}$ | 0.0082 |
| log(*Amount*) | −0.005 | 0.0047 |
| *PrimeRate* | $0.950^{***}$ | 0.0076 |
| $\widehat{Rate}$ | $1.001^{***}$ | 0.0092 |
| $R^2$ | 82% | |
| *F*-value | 28,056.6 | |

*Note.* Month fixed effects are omitted for brevity.
Significance levels: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

## References

Agrawal V, Ferguson M (2007) Bid-response models for customized pricing. *J. Revenue Pricing Management* 6(3):212–228.

Bhardwaj P (2001) Delegating pricing decisions. *Marketing Sci.* 20(2): 143–169.

Baker WL, Marn MV, Zawada CC (2010) *The Price Advantage* (John Wiley & Sons, Hoboken, NJ).

Bijmolt THA, van Heerde HJ, Pieters RGM (2005) New empirical generalizations on the determinants of price elasticity. *J. Marketing Res.* 42(2):141–156.

Breshanan T (1997) The Apple-Cinnamon Cheerios war: Valuing new goods, identifying market power, and economic measurement. Working paper, Department of Economics, Stanford University, Stanford, CA.

Cachon GP, Olivares M (2010) Drivers of finished-goods inventory in the U.S. automobile industry. *Management Sci.* 56(1):202–216.

Cachon GP, Galliano S, Olivares M (2013) Does adding inventory increase sales? Evidence of a scarcity effect in U.S. automobile dealerships. Columbia Business School Research Paper 13-60, Columbia University, New York. http://papers.ssrn.com/abstract=2286800.

Caufield S (2012) Consumer credit pricing. Özer Ö, Phillips R, eds. *The Oxford Handbook of Pricing Management* (Oxford University Press, Oxford, UK), 138–153.

Cohen MA (2007) Imperfect competition in auto lending: Subjective markup, racial disparity, and class action litigation. Vanderbilt Law and Economics Research Paper 07-01, Vanderbilt University, Nashville, TN.

[29] Each lender used its own methodology to estimate risk, so the tiers are not directly comparable.

[30] The online lender performed a number of pricing tests and experiments during the period covered by our data.

Coughlan AT (1993) Salesforce compensation: A review of MS/OR advances. Eliashberg J, Lilien GL, eds. *Marketing*, Handbook in Operations Research and Management Science, Vol. 5 (Elsevier, Amsterdam), 611–651.

Ebbes P, Papies D, van Heerde HJ (2011) The sense and non-sense of holdout sample validation in the presence of endogeneity. *Marketing Sci.* 30(6):1115–1122.

Foster EM (1997) Instrumental variables for logistic regression: An illustration. *Soc. Sci. Res.* 26(4):487–504.

Frenzen H, Hansen A, Krafft M, Mantrala MK, Schmidt S (2010) Delegation of pricing authority to the sales force: An agency-theoretic perspective of its determinants and impact on performance. *Internat. J. Res. Marketing* 27(1):58–68.

Gelber SM (2008) *Horse Trading in the Age of Cars: Men in the Marketplace* (Johns Hopkins University Press, Baltimore).

Hausman JA (1997) Valuation of new goods under perfect and imperfect competition. Bresnahan TF, Gordon R, eds. *The Economics of New Goods* (University of Chicago Press, Chicago), 209–248.

Homburg C, Jensen O, Hahn A (2012) How to organize pricing? Vertical delegation and horizontal dispersion of pricing authority. *J. Marketing* 76(5):49–69.

Joseph K (2001) On the optimality of delegating pricing authority to the sales force. *J. Marketing* 65(1):62–70.

Kalkanci B, Chen K-Y, Erhun F (2011) Contract complexity and performance under asymmetric demand information: An experimental evaluation. *Management Sci.* 57(4):689–704.

Kuo C-W, Ahn H-S, Aydın G (2011) Dynamic pricing of limited inventories when customers negotiate. *Oper. Res.* 59(4):882–897.

Kuo C-W, Ahn H-D, Aydın G (2013) Pricing policy in a supply chain: Negotiation or posted pricing. *Production Oper. Management* 22(3):626–641.

Lal R (1986) Delegating pricing responsibility to the salesforce. *Marketing Sci.* 5(2):159–168.

Lesaffre E, Marx BD (1993) Collinearity in generalized linear regression. *Comm. Statist. Theory Methods* 22(7):1933–1952.

Li J, Granados N, Netessine S (2014) Are consumers strategic? Structural estimation from the air-travel industry. *Management Sci.* 60(9):2114–2137.

Mishra BK, Prasad A (2004) Centralized pricing versus delegating pricing to the salesforce under information asymmetry. *Marketing Sci.* 23(1):21–27.

Mishra BK, Prasad A (2005) Delegating pricing decisions in competitive markets with symmetric and asymmetric information. *Marketing Sci.* 24(3):490–497.

Moreno A, Terwiesch C (2013) The effects of product line breadth: Evidence from the automotive industry. Working paper, Northwestern University, Evanston, IL. http://ssrn.com/abstract=2241384.

Moreno A, Terwiesch C (2014) Pricing and production flexibility: An empirical analysis of the U.S. automotive industry. Working paper, Northwestern University, Evanston, IL. http://ssrn.com/abstract=2188246.

Olivares M, Cachon GP (2009) Competing retailers and inventory: An empirical investigation of General Motors' dealerships in isolated U.S. markets. *Management Sci.* 55(9): 1586–1604.

Oliver Wyman, Efma (2012) Pricing at European retail banks. Report, Oliver Wyman, New York, and Efma, Paris.

Özer Ö, Zheng K (2012) Behavioral issues in pricing management. Özer Ö, Phillips R, eds. *The Oxford Handbook of Pricing Management* (Oxford University Press, Oxford, UK), 415–464.

Petrin A, Train K (2002) Omitted product attributes in discrete choice models. Working paper, Department of Economics, University of California, Berkeley, Berkeley.

Petrin A, Train K (2010) A control function approach to endogeneity in consumer choice models. *J. Marketing Res.* 47(1):370–379.

Phillips RL (2005) *Pricing and Revenue Optimization* (Stanford University Press, Palo Alto, CA).

Phillips RL (2012a) Customized pricing. Özer Ö, Phillips R, eds. *The Oxford Handbook of Pricing Management* (Oxford University Press, Oxford, UK), 465–491.

Phillips RL (2012b) Why are prices set the way they are? Özer Ö, Phillips R, eds. *The Oxford Handbook of Pricing Management* (Oxford University Press, Oxford, UK), 13–44.

Phillips RL, Raffard R (2011) Price-driven adverse selection in consumer lending. Working paper, Center for Pricing and Revenue Management, Columbia University, New York. http://www8.gsb.columbia.edu/cprm/sites/cprm/files/files/Working%20Papers/2011-3-Price-Driven-Adverse-Selection.pdf.

Poon M (2007) Scorecards as devices for consumer credit: The case of Fair, Isaac & Company Incorporated. Callon M, Millo Y, Muniesa F, eds. *Market Devices* (Blackwell, Malden, MA), 284–306.

Rivers D, Vuong QH (1988) Limited information estimators and exogeneity tests for simultaneous probit models. *J. Econometrics* 39(3):347–366.

Roth S, Woratschek H, Pastowski S (2006) Negotiating prices for customized services. *J. Service Res.* 8(4):316–329.

Simonetto M, Montan L, Meehan J, Kaji J (2012) Structuring and managing an effective pricing organization. Özer Ö, Phillips R, eds. *The Oxford Handbook of Pricing Management* (Oxford University Press, Oxford, UK), 825–853.

Stephenson RP, Cron WL, Frazier GL (1979) Delegating pricing authority to the salesforce: The effects on sales and profit performance. *J. Marketing* 43(Spring):21–28.

Taylor TA (2002) Supply chain coordination under channel rebates with sales effort effects. *Management Sci.* 48(8):992–1007.

Veeraraghavan S, Vaidyanathan R (2012) Measuring seat value in stadiums and theaters. *Production Oper. Management* 21(1): 49–68.

Villas-Boas JM, Winer RS (1999) Endogeneity in brand choice models. *Management Sci.* 45(10):1324–1328.

Vulcano G, van Ryzin G, Chaar W (2010) Choice-based revenue management: An empirical study of estimation and optimization. *Manufacturing Service Oper. Management* 12(3):371–392.

Weinberg CB (1975) An optimal commission plan for salesmen's control over price. *Management Sci.* 21(8):937–943.

Wooldridge JM (2002) *Econometric Analysis of Cross Section and Panel Data* (MIT Press, Cambridge, MA).