

# Retail Pricing under Contract Self-Selection: An Empirical Exploration

Yuanfang Lin, Lianhua Li

School of Business, University of Alberta, Edmonton, Canada  
Email: Yuanfang.lin@ualberta.ca, lianhua@ualberta.ca

Received 2012

## ABSTRACT

Using cross-sectional survey data on prices, station and market characteristics for 730 gasoline stations in the Greater Saint Louis area, we estimate a switching regression model of station decisions. We employ a binary probit choice model to study a station's decision to enter a contract relationship with greater control from the upstream refinery, or a contract relationship with greater degree of independence, as a function of market and station characteristics. We then estimate stations' pricing decisions with self-selectivity corrections for the station's contract decision. We show that incorrect inferences about retail gasoline station's pricing behavior would result if the endogeneity in the choice of contract type were treated as exogenous condition in the estimation.

**Keywords:** Retail Gasoline Markets; Marketing Channel; Pricing Strategy; Selection Bias; Switching Regression

## 1. Introduction

As a channel intermediary, a retailer obtains product from manufacturers and resell to the end consumers in the market. When selecting a certain contract relationship with the upstream manufacturer, a retailer considers and compares the benefits versus the levels of control from the manufacturer on its marketing activities, associated with different types of contract relationship. If a retailer agrees to become a lower level division or a franchise of the manufacturer, the retailer's marketing activities in terms of daily prices, promotional events, etc will be largely controlled by the manufacturer. On the other hand, if the retailer maintains its independence by only paying wholesale prices to get product from manufacturer, the above mentioned marketing activities can then be largely determined by the retailer itself or least with greater degree of flexibility.

This paper examines a retailer gasoline station's pricing decision after accounting for its endogenous, comparatively longer-term decision on the contract relationship with upstream refinery. The model and analysis presented here are inspired by work done by Iyer and Seetharaman (2003) which examines a firm's incentive to price discriminate after self selecting the product offering. We model a gasoline station's decision to be in a more "controlled" or "independent" relationship with upstream refinery as a function of the market and station characteristics. We then model the retail price set by the gasoline station conditional on its prior endogenous contract deci-

sion. The study allows us to explicitly and comprehensively investigate the interaction between two important Marketing "4Ps", namely, price and place (channel) in the retail gasoline industry.

## 2. Literature Background

When it comes to the marketing "4Ps" about a retail gasoline station, the "Price" and "Place" (including both geographic location and contract choice) have been of strong interests in existing economic studies.<sup>1</sup> The following we briefly review a few key references before stating the contribution of this paper.

Shepard (1993) argues that gasoline refiners will choose contractual forms with strong performance incentives, i.e., lessee-dealerships or open dealerships, at gasoline stations where unobservable effort is more important (e.g. auto repair, full service), and contractual forms, i.e. company-owned, that allow more direct control but offer weaker performance incentives, at stations where observable effort is more important (e.g. convenient store). Pricing regression using cross-sectional data on contractual forms and other characteristics from 1,527 gasoline stations in Eastern Massachusetts showed that company-owned stations indeed have lower prices than other stations, all else being equal.

Slade (1996) argues that when there is a high degree of

---

<sup>1</sup> Interested readers are referred to Lin and Seetharaman (2012) for details.

complementarity between gasoline and other activities in the station (e.g. convenient store), contract with high salary and low commission must be offered, while when there is a high degree of substitutability between gasoline and other station activities (e.g. auto repair), contract with low salary and high commission must be offered. Slade (1998) tests whether strategic reasons could explain why manufacturers choose to remain separate from their retailers in some markets. Using a binary probit model on cross-sectional data on contractual forms and other characteristics collected from 96 branded gasoline stations in Vancouver during Fall 1991, the author demonstrates that a station's likelihood of being a lessee-dealership increases as the predicted difference in the price-cost margins between vertical separation and vertical integration increases.

Pinkse and Slade (1998) assess whether gasoline stations of a given contract form (e.g. vertically integrated) cluster together in geographic space. Using spatial statistics, and six different measures of geographic closeness, the authors recovered positive spatial correlations, i.e., firms with similar contract forms are found to cluster. Regarding stations' contract choice, the authors found pattern consistent with Shepard (1993) and Slade (1996).

Our paper contributes to the literature by further distinguishing comparatively longer versus shorter term marketing factors. A gasoline station's retail prices could be adjusted due to demand and competitive situations rather frequently within a short time frame. On the contrary, the investment of land acquisition, asset purchase and contract selection tends to be decided earlier and sustained throughout a longer period. Following this rationale, we propose that when studying retail stations' pricing behavior in competition, their prior choice of entering certain contract with upstream refinery needs to be taken into consideration. Further the contract type of a station cannot be merely treated as an exogenous variable in the pricing equations, rather it has to be treated as an endogenous decision of the gasoline station separately and then incorporate into the later analysis of pricing strategies. A close reference of our paper is Iyer and Seetharaman (2003) where the authors investigate a gasoline station's incentive to price discriminate by self selecting to sell full-service as well as self-service gasoline.

### 3. Econometric Model

We employ the "switching regression with endogenous switching" (Trost 1977) to estimate a retail gasoline station's pricing decision conditional upon its endogenous contract choice. Other econometric applications of this model have been in the context of explaining discrete/continuous choice decisions of households (e.g.

Hanemann 1984; Dubin and McFadden 1984; Chintagunta 1993; etc). Details of the model specification are described in the following two-step procedure:

*Step1:* we estimate a binary probit model of the gasoline station's decision of contract relationship with upstream refinery, which is represented by the following choice probabilities:

$$\Pr_{ctl} = 1 - \Phi[-(\alpha_0 + z_1\alpha_1 + z_2\alpha_2)] \quad (1)$$

$$\Pr_{idp} = \Phi[-(\alpha_0 + z_1\alpha_1 + z_2\alpha_2)] \quad (2)$$

where  $\Pr_{ctl}$ ,  $\Pr_{idp}$  stand for the probability of a gasoline station choosing a contract relationship which receives stronger control from the refinery, or maintains more self independence, respectively.  $\Phi$  is the cdf of a standard normal distribution,  $z_1$  is a vector of variables representing market condition with  $\alpha_1$  being the corresponding coefficients,  $z_2$  is a vector of variables representing station characteristics with  $\alpha_2$  being the corresponding coefficients, and  $\alpha_0$  is the intercept term.

*Step2:* We estimate a linear regression for prices that explicitly accounts for the effects of contract self-selection as shown below.

$$P = \beta_0 + \beta_1 X + \beta_2 I_{ctl} + \beta_3 SS_1 I_{ctl} + \beta_4 SS_0 I_{idp} + \varepsilon \quad (3)$$

where  $X$  is the vector of exogenous variables representing the relevant market and station characteristics with  $\beta_1$  being the corresponding coefficients.  $I_{ctl}$  is an indicator variable which equals 1 for stations with contract of strong "control" from the upstream refinery and 0 otherwise.  $I_{idp}$  is an indicator variable which equals 1 for stations with contract of strong "independence" from the upstream refinery and 0 otherwise. The variable  $SS_1$  is a self-selectivity correction for the "control" contract regime, while the variable  $SS_0$  is a self-selectivity correction for the "independent" contract regime. Incorporating these variables in the price regression corrects for the self-selectivity bias that would arise in the parameters of a pricing model that ignores the station's endogenous contract choice. The self-selectivity correction terms are computed as follows based on Maddala (1983).

$$SS_1 = \phi(-Y)/(1 - \Phi(-Y)) \quad (4)$$

$$SS_0 = \phi(-Y)/\Phi(-Y) \quad (5)$$

where  $\phi$  is the pdf of the standard normal distribution and  $Y$  represents the estimates from the binary probit model in step 1.

$$Y = \alpha_0 + z_1\alpha_1 + z_2\alpha_2 \quad (6)$$

For comparison purpose, a pricing regression without correcting for the contract self-selection can be specified as

$$P = \gamma_0 + \gamma_1 X + \gamma_2 I_{ctl} + \varepsilon \quad (7)$$

where the station's contract choice is merely treated as an exogenous variable, same as those in the vector  $X$ .

## 4. Data and Estimation

We employ a survey data<sup>2</sup>, collected during 1999, which covers 730 retail gasoline stations in the Greater Saint Louis area. The survey data contain information on retail prices and various service and local market characteristics pertaining to the 730 gasoline stations. We also employ the 2000 U.S. census records and information from the Missouri Census Data Center for demographic characteristics of the local markets where these gasoline stations operate.

### 4.1. Empirical Measures

We use the Type-of-Operation recorded in the station survey to construct the dependent variable for the binary probit model at step 1. Stations in the survey are classified into four different types of operation: 1) owned by refinery; 2) franchise of refinery; 3) independent retailer; and 4) local jobbers. We group the first two types of operation (company-owned, franchise) into the “control”-type contract relationship (*Ctl*), while the remaining two are grouped into the “independent”-type contract relationship (*Idp*). This practice is consistent with the literature studies on marketing channel coordination (e.g. McGuire and Staelin 1983). Consistent with the literature, we include the following variables in the estimation of contract choice model and pricing regression, which contains station and market characteristics, as well as demographics of the local area where the station is operated.

1) *Wash*, i.e., a dummy variable that takes the value 1 if the station offers car wash and 0 otherwise.

2) *Full*, i.e., a dummy variable that takes the value 1 if the station sells full-service gasoline and 0 otherwise.

3) *Conv*, i.e., a dummy variable that takes the value 1 if the station has convenience store and 0 otherwise.

4) *Day*, i.e., a dummy variable that takes the value 1 if the station opens 24 hour a day and 0 otherwise.

5) *Serv*, i.e., a dummy variable that takes the value 1 if the station has service station and 0 otherwise.

6) *Brd*, i.e., a dummy variable that takes the value 1 if the station is a brand-name station (Amoco, Shell, or Exxon-Mobil) and 0 otherwise.

7) *Noz*, a discrete variable, whose values range from 2 to 8, that captures the number of pumping nozzles at the gasoline station.

8) *Cmp*, a non-negative discrete variable that captures the number of gasoline stations (other than the focal one) that operate in the same local market (defined by census track).

9) *Mhi*, a non-negative continuous variable that cap-

tures the median household income of the census track where the station operates.

10) *Pop*, a non-negative discrete variable that captures the size of population for the census track where the station operates.

The station survey includes, for each gasoline station, the prices of three grades – 87, 89 and 93 octane levels. Since 87-octane level is the most commonly sold grade of gasoline in retail gasoline markets, and is available at all the stations in our dataset, we operationalize the dependent variable (*Rup*) for the pricing regression at step 2 using the observed (posted) price of self-service gasoline in cents per gallon. And the following variables are included in the pricing regression to assess the potential impact from local demand and competitive condition.

11) *Cvis*, a dummy variable that takes the value 1 if there is another station visible from the location of the focal station and 0 otherwise.

12) *Neaf*, a dummy variable that takes the value 1 if the station is within 1 mile distance from a highway entrance and 0 otherwise.

13) *Neac*, a dummy variable that takes the value 1 if the station is near a business entity (e.g. grocery store, shopping plaza, etc).

Table 1 reports the descriptive statistics for all the variables.

**Table 1. Descriptive Statistics**

Variable	Mean	Std Dev	Min	Max
<i>Ctl</i>	0.34	0.48	0	1
<i>Wash</i>	0.22	0.42	0	1
<i>Full</i>	0.12	0.33	0	1
<i>Conv</i>	0.91	0.29	0	1
<i>Day</i>	0.69	0.46	0	1
<i>Serv</i>	0.20	0.40	0	1
<i>Brd</i>	0.42	0.49	0	1
<i>Noz</i>	16.56	9.46	1	66
<i>Cmp</i>	3.22	3.11	0	17
<i>Mhi</i>	45255.00	21444.00	0	200001
<i>Pop</i>	1635.00	1058.00	0	7667
<i>RUP</i>	960.32	54.04	19	1159
<i>Cvis</i>	0.36	0.48	0	1
<i>Neaf</i>	0.21	0.41	0	1
<i>Neac</i>	0.20	0.40	0	1

### 4.2. Contract Choice Results

Table 2 reports the probit model estimation results from step 1. Stations are more likely to choose a “control”-type contract with a major brand refinery. Stations are also likely to be in a “control”-type contract if offering car wash or full-service gasoline product. Also longer opening hours and greater number of pumping nozzles are both more likely associated with a “control”-type contract. Consistent with the literature (e.g. Shepard

<sup>2</sup> The survey data was collected by New Image Marketing Ltd.

1993), stations are more likely to be in a “independent”-type contract if offering auto repair service where unobservable effort is more important. Different from the literature, we did not find significant impact on the “control”-type contract from the presence of convenience store, which could be explained by the increasing co-existence of convenience stores inside gasoline station in recent years. Stations are less likely to choose a “control”-type contract when operating in a local area with higher medium household income. Finally the presence of competing stations in the same local area marginally reduces a station’s likelihood of choosing a “control”-type contract as remaining independence would help the station react more promptly in competitive situation such as promotional decisions.

**Table 2. Probit Model Results**

Variable	Coeff	Std. Error	T-Stat	P-Value
<i>Const</i>	-1.098	0.245	-4.489	0.000
<i>Wash</i>	0.301	0.144	2.093	0.036
<i>Full</i>	0.827	0.191	4.321	0.000
<i>Conv</i>	-0.190	0.194	-0.978	0.328
<i>Day</i>	0.415	0.147	2.827	0.005
<i>Serv</i>	-0.017	0.182	-0.093	0.926
<i>Brd</i>	0.949	0.123	7.689	0.000
<i>Noz</i>	2.501	0.383	6.534	0.000
<i>Cmp</i>	-0.028	0.018	-1.588	0.112
<i>Mhi</i>	-0.881	0.513	-1.718	0.086
<i>Pop</i>	-0.482	0.471	-1.023	0.306

### 4.3. Pricing Regression Results

Table 3 reports the estimation results of the pricing regression from Step 2.

**Table 3. Pricing Model Results**

Variable	Coeff	Std. Error	T-Stat	P-Value
<i>Const</i>	0.000	0.007	-0.018	0.985
<i>Day</i>	-0.020	0.007	-3.085	0.002
<i>Brd</i>	0.007	0.014	0.535	0.592
<i>Mhi</i>	0.033	0.021	1.569	0.107
<i>Pop</i>	-0.023	0.021	-1.125	0.261
<i>Cvis</i>	0.003	0.007	0.402	0.687
<i>Neac</i>	-0.014	0.006	-2.513	0.012
<i>Neaf</i>	-0.006	0.006	-0.969	0.333
<i>Ctl</i>	0.045	0.037	1.221	0.222
<i>SS1</i>	-0.080	0.046	-1.750	0.080
<i>SS0</i>	0.023	0.050	0.468	0.640

The visibility of another station does not seem to significantly impact the focal station’s pricing decision, a finding prompts us to further consider a better measure

of the competitive intensity in the local market. Among all the variables indicating the potential demand from local market, closeness to another business entity (e.g. grocery store, shopping plaza) has the most significant impact. In particular, a retail station tends to lower its gasoline price in order to attract traffics likely going to the nearby business locations. The covariance between the “control”-type contract choice and the pricing ( $\beta_3$ ) is significantly negative, i.e., if a station is more likely to be in a “control”-type contract relationship with upstream refinery, it is also more likely to charge a lower price.

### 4.4. Effect of Endogeneity Correction

The main objective of this paper is to demonstrate the importance of incorporating a gasoline station’s endogenous contract choice when studying its pricing strategy. For this purpose, we also separate estimate the pricing equation specified in Equation 7 and the results are reported in the following Table 4.

**Table 4. Pricing without Contract Self-Selection**

Variable	Coeff	Std. Error	T-Stat	P-Value
<i>Const</i>	0.008	0.005	1.508	0.132
<i>Day</i>	-0.018	0.006	-3.186	0.001
<i>Brd</i>	0.021	0.007	3.157	0.002
<i>Mhi</i>	0.024	0.019	1.267	0.205
<i>Pop</i>	-0.025	0.019	-1.294	0.196
<i>Cvis</i>	0.002	0.007	0.295	0.768
<i>Neac</i>	-0.014	0.006	-2.545	0.011
<i>Neaf</i>	-0.006	0.006	-0.974	0.330
<i>Ctl</i>	-0.001	0.006	-0.076	0.939

When a station’s contract type (*Ctl*) is merely treated as an exogenous variable in the pricing regression, the estimate coefficient is statistically insignificant, which implies that there is no covariance between a station’s contract relationship with upstream refinery and its pricing decision. This is different from the result of step 2 pricing regression as reported in Table 3. A likelihood-ratio test is further conducted with the null hypothesis being the pricing model with contract self-selection correction (Table 3) and the alternative hypothesis being the pricing model without contract self-selection correction (Table 4). Test statistics favors the self-selection correction model with high significance ( $P < .001$ ).

## 5. Conclusion

This paper investigates a gasoline station’s endogenous decision to choose a specific contract relationship with upstream refinery and the corresponding pricing decision. We find that a pricing regression that does not endogenize the gasoline station’s contract decision to station and

local market characteristics leads to incorrect inferences in the pricing estimation. The most interesting direction to further expand the current study is to identify other station decisions that are also endogenously set. One such factor would likely be the geographic location of the station where a potential study will then involve a three-stage decision sequence where stations choose geographic location first, followed by contract decision with upstream refinery, and finally the competitive pricing behavior in the local market.

## 6. Acknowledgements

We thank the Social Sciences and Humanities Research Council of Canada for funding that supports this research project.

## REFERENCES

- [1] A. Shepard, "Contractual Form, Retail Price, and Asset Characteristics in Gasoline Retailing," *The RAND Journal of Economics*, Vol. 24, No. 1, 1993, pp58-77.
- [2] G. Iyer and P.B. Seetharaman, "To Price Discriminate or Not: Product Choice and the Selection Bias Problem," *Quantitative Marketing and Economics*, Vol. 1, 2003, pp155-178.
- [3] G.S. Maddala, "Limited-Dependent and Qualitative Variables in Econometrics," Cambridge University Press, 1983.
- [4] J. A. Dubin and D.L. McFadden, "An Econometric Analysis of Residential Appliance Holdings and Consumption," *Econometrica*, Vol. 52, No. 2, 1984, pp345-362.
- [5] J. Pinkse and M. E. Slade, "Contracting in Space: An Application of Spatial Statistics to Discrete-Choice Models," *Journal of Econometrics*, Vol. 85, No. 1, 1998, pp125-154.
- [6] M. E. Slade, "Multitask Agency and Contract Choice: An Empirical Exploration," *International Economic Review*, Vol. 37, No. 2, 1996, pp465-486.
- [7] M. E. Slade, "Strategic Motives for Vertical Separation: Evidence from Retail Gasoline Market," *Journal of Law, Economics and Organization*, Vol. 14, No. 1, 1998, pp84-113.
- [8] M.W. Hanemann, "Discrete/Continuous Models of Consumer Demand," *Econometrica*, Vol. 52, No. 3, 1984, pp541-562.
- [9] P. K. Chintagunta, "Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households," *Marketing Science*, Vol. 12, No. 2, 1993, pp184-208.
- [10] R.P. Trost, "Demand for Housing: A Model Based on Inter-Related Choices between Owning and Renting," Unpublished Ph.D dissertation, University of Florida, 1977.
- [11] T. McGuire and R. Staelin, "An Industry Equilibrium Analysis of Downstream Vertical Integration," *Marketing Science*, 2 (Spring), 161-192.
- [12] Y. Lin and P.B. Seetharaman, "The Gasoline Industry: A Review for Marketing Research," *The Marketing Review*, forthcoming, 2012.