THREE ESSAYS ON RETAIL PRICE COMPETITION

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

TAEHWAN KIM

Dr. Emek Basker, Dissertation Supervisor

JULY 2018

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled:

THREE ESSAYS ON RETAIL PRICE COMPETITION

presented by Taehwan Kim,

a candidate for the degree of Doctor of Philosophy and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Emek Basker

Professor Peter Mueser

Professor Shawn Ni

Professor Michael Sykuta

For my father

ACKNOWLEDGMENTS

I sincerely thank my advisor, Emek Basker, for her endless support, patience, and guidance. She was always willing to listen to me and discuss my academic and non-academic problems even though she has been one thousand miles away from me. While working with her for the past three years, I have learned a great deal of lessons useful to both my personal and professional careers. I thank her so much for giving me the option that I did not have to change my research fields during my time at Mizzou.

I am indebted to a number of the faculty members in the Mizzou Department of Economics. Peter Mueser always welcomed me anytime I asked for his advice. He provided many constructive comments that greatly improve my papers. Shawn Ni also gave me many insightful comments, which often enabled me to understand one problem with various angles. In addition, it was a great honor to learn Microeconomics theory from Daivd Mandy; he is the best lecturer ever. And I thank X.H. Wang, as well as the wonderful staff members (Lynne Owen, Sheila Akers, and Sue Leutschaft), for their generous support on me.

I would like to express my gratitude to my classmates for their conversation and encouragement. Seungyub Lee is a friend indeed that I made on campus. It was a great pleasure to live with him for the half of my time at Mizzou. I will miss our conversations on many different subjects ranging from class assignments to human futures. Ingul Baek was a person I turned to mostly while preparing my job market. He always inspired me with confidence, and I have learned a lot from his tolerance. Last but not least, it will be my precious memory forever to have time with my friends (too numerous to name) in Columbia.

Finally, I thank my family tremendously for trusting me. I know, to my mother, staying alone for six years was never easy. I am deeply grateful for her courage and patience. My sister did not(!) favor my challenge to earn a Ph.D. degree in the U.S., which motivated me much harder (she is very supportive more than anyone else now). My last gratitude goes to my father who sacrificed his everything for his family: I missed a number of things I ought to do while he was alive, for which I am so sorry. I dedicate this dissertation to my father, and I wish he would have been proud of my achievement.

TABLE OF CONTENTS

Three Essays on Retail Price Competition

by Taehwan Kim

Dr. Emek Basker, Dissertation Supervisor

Abstract

This dissertation consists of three chapters.

The first chapter examines price dispersion in retail gasoline and focuses on differentiation along the service dimension: full service versus self service. Consistent with more intensive search by self-service customers, I find that price dispersion always decreases with the number of nearby self-service stations, but does not decrease with the number of nearby full-service stations. When I segment the market by brand, I observe that the estimates are sensitive to how brands are separated into different types. These findings show that the market is more clearly segmented by service level than by brand type and also highlight the importance of product differentiation when modeling price dispersion.

In the second chapter, I examine product positioning and pricing strategies of sellers in a market undergoing a significant restructuring using data from the introduction of self-service technology in the Korean gasoline market in the 2000s. I show that the decision of full-service sellers to exit or switch to self service is positively correlated with the intensity of competition they face. The pricing strategies of sellers differ by product position: self-service sellers compete for price-sensitive consumers, whereas full-service sellers differentiate their product by offering a variety of bundled products and services, such as coffee, carwash or even a nail salon, to compete for less-price-sensitive consumers. Taken together, these patterns have led to an increase in the full-service premium during the market transition.

In the third chapter, I study the effect of a government contract on price. Since 2013, Korean government officials have been required to refuel at contracted gasoline stations, at about 5% discounts relative to the posted price. The initial contract terminated in November 2015 and a new group of sellers took over the contract. In this paper, I use this natural experiment to examine the impact of the government contract on gasoline prices, using a difference-in-difference analysis and price data on all gasoline stations in Seoul. I find that, all else equal, posted prices of contracted gasoline stations are about 2% higher than those of non-contracted stations. This finding is consistent with the prediction of models of price discrimination that prices decrease when the elasticity of demand falls. The effect on prices is not uniform across all stations, however. The contract leads to larger increases in full-service stations' posted prices than in self-service stations' prices, and larger increases at stations with fewer nearby competitors. The contract also decreases prices of non-contracted stations very close to contracted stations.

Chapter 1

Price Competition and Market Segmentation in Retail Gasoline: New Evidence from South Korea

1.1 Introduction

Price dispersion is universal even in markets for seemingly homogeneous goods. One possible explanation for price dispersion is that markets may be segmented in ways that are unobservable to the econometrician, so that seemingly homogeneous goods are in fact perceived by consumers as heterogeneous.

In this paper, I exploit the recent transition of gasoline stations in Seoul, South Korea, from full service to self service, so that I can jointly examine price dispersion and competition in the context of service differentiation. I find that price dispersion always decreases with the number of nearby self-service stations, whereas greater competition is not always related to low dispersion for full service. This finding suggests that the recognition of market segmentation is important when applying insights from models of price dispersion to real markets.

My analysis makes two contributions: First, I use service level to identify sellers of different types. Existing papers on gasoline pricing have often used brand type. For example, Hastings (2004) studies the price effects of brand-contract changes from an independent retailer to ARCO in Southern California, and Lewis (2008) shows that price dispersion varies depending on the brand composition of local competitors in San Diego. Both papers assume that the cross-price elasticity of demand across brands is low; but this may not be justified in retail gasoline. My analysis utilizes a difference between customers at full-service versus self-service stations to identify search intensity; full-service customers may well be less willing to engage in costly search for lower gasoline prices than are self-service customers because gasoline is just one component of the full package purchased by full-service customers.[1]

The paper's second contribution derives from the empirical finding: I find that the relationship between price dispersion and local competition varies across sellers of different service levels, but this evidence is a finding that does not have a clear corollary in search models. For example, the classic models of search, such as Varian (1980) and Stahl (1986), attempt to explain the presence of price dispersion and show that price dispersion is non-monotonic in consumers' search intensity; but these models are not designed for a market with differentiated sellers that attract buyers of different types. My setting provides a natural experiment for examining how price dispersion varies with number of sellers by service level and indirectly with consumers' search intensity.[2]

Specifically, I show that price dispersion varies systematically with the number of sellers, as well as with seller characteristics such as presence of a store and/or a carwash facility and service level. Considering market segmentation by service level,

[1]Korean full-service customers are reluctant to engage in costly behavior. Kim et al. (2010) conduct a survey of 1000 consumers who buy gasoline in three months, and document that 22.5% of the consumers report "I am not willing to use self-service stations." About 40% of those unwilling to buy self-serve gasoline report "inconvenience of getting in and out of the car" as the reason.

[2]Identifying search intensity empirically is not trivial. Marvel (1976) and Houde (2012) argue that commuters have lower search costs, and Lewis and Marvel (2011) use the number of visits to a website that provides gas price information as a proxy for consumer search.

I find that price dispersion among self-service stations is lowest when their density is high, but there is no evidence of such a relationship for full-service stations. Regarding competition across service levels, I find that the elasticity of self-service price dispersion with respect to the number of nearby full-service stations is approximately double the elasticity of full-service price dispersion with respect to the number of nearby self-service stations. These results indicate that self-service stations are responsible for lower dispersion, which is consistent with the claim that self-service buyers are more likely to engage in costly search for arbitrage opportunities.

This paper is in line with a small but growing literature that studies self-service gasoline stations (for a recent review, see Noel, 2016). Shepard (1991) estimates the self-service discount in four Massachusetts counties in 1987 and argues that service-level differentiation allows stations to price discriminate over and above cost-based price differences. Kim and Kim (2011) study the entry effects of self-service stations in the Korean market and find a four-cent-per-gallon reduction in price. Recently, using US data from the period 1977 to 1992, Basker et al. (2017) examine the effects of self-service adoption on station-level employment and find about a third fewer workers per pump at self-service stations.

The rest of this paper is organized as follows. I describe the data and provide an overview of the market in Section 1.2, and discuss my empirical methodology in Section 1.3. I present my main results, as well as several robustness tests, in Sections 1.4 and 1.5. Section 1.6 concludes.

1.2 Data and Market Overview

1.2.1 Data

My data cover the period from May 2010 to September 2014. Over this period, I have daily price information on all gasoline stations in Seoul, and each station's address, service level, brand, and shop name – also at a daily frequency. The data come from the Oil Price Information Network (OPINET) in Korea. OPINET is a website that is operated by the Korea National Oil Corporation to provide retail gasoline prices to the public for market transparency. Most stations upload price information to the website automatically based on transactions data.[3]

I use data from the Wednesday of each week because using the full dataset is computationally burdensome.[4] Information on stations' entry and exit is not explicitly available from the dataset; but the enforcement in Korea enables me to infer the information: Stations must report at least once a week even if there is no price change. To be conservative, I assume that a station that does not report prices for up to 30 days remains open, as long as its observed characteristics (such as name or service level) do not change.[5]

Using the address information, I geocode all stations in my dataset and count the number of full- and self-service stations within a mile, a mile and a half, and two miles.[6] A unique market feature in Korea is that stations are perfectly partitioned by service level at each point in time, with no station offering a combination of full-service

[3]If stations choose a manual update, they must report their price within 24 hours of a change in price.

[4]The loss of the data is minimal due to the fact that station-level prices change once every 12 days on average. I chose Wednesday because Tuesday is the modal day for price changes.

[5]There are 79 instances of stations that do not report prices for over a week without a change in station characteristics. In my dataset, when stations drop out of the sample for longer than seven days, observed station characteristics typically change, in which case I assume that the stations temporally closed to make that change.

[6]My estimation results are qualitatively similar using one, one-and-a-half, and two-mile radii. I report results using a 1.5-mile radius to be consistent with previous studies.

and self-service pumps.

Table 1.1: Summary Statistics, by Competition Variables and Station Characteristics [117,212 obs.]

Variable	Description	Mean	SD	Min	Max
Price	Price of gasoline (Korean won/liter)	1973	151	1587	2490
Competition Variables					
Num	Num. of stations within 1.5 miles	17.8	6.52	1	39
NumSS	Num. of self-service stations within 1.5 miles	2.78	1.95	1	11
NumFS	Num. of full-service stations within 1.5 miles	15	6.05	1	40
NumLO	Num. of low-brand stations within 1.5 miles	5.77	3.49	1	22
NumHI	Num. of high-brand stations within 1.5 miles	12	4.47	1	25
NumFB	Num. of follower-brand stations within 1.5 miles	10.4	4.48	1	30
NumLB	Num. of leader-brand stations within 1.5 miles	7.32	2.96	1	16
NumUB	Num. of unbranded stations within 1.5 miles	0.91	1.08	0	7
NumBB	Num. of branded stations within 1.5 miles	16.8	6.04	1	37
Brand Share	Share of same-brand stations within 1.5 miles	0.28	0.16	0	1
Station Characteristics (X)					
Pop[a]	Population density (unit: million)	12.4	0.11	11.4	12.5
Rent[a]	A housing rent index (=100 at Jan. 2006)	132	4.37	125	139
Full	Station serves as full-service	0.84	0.36	0	1
Intersection	Station is located at an intersection	0.14	0.35	0	1
Store	Station has a convenience store	0.1	0.3	0	1
Carwash	Station has an automatic carwash machine	0.69	0.46	0	1
Repair	Station has an auto-repair center	0.28	0.45	0	1
SK	Station brand: SK Energy	0.42	0.49	0	1
GS	Station brand: GS Caltex	0.28	0.45	0	1
Hyundai	Station brand: Hyundai Oilbank	0.14	0.34	0	1
S-Oil	Station brand: S-Oil	0.12	0.33	0	1
Alddle	Station brand: Alddle	0.02	0.13	0	1
Unbranded	Station brand: Unbranded	0.03	0.18	0	1
High-Brand	Station brand group: SK and GS	0.7	0.46	0	1
Follower-Brand	Station brand group: All brands expect for SK	0.58	0.49	0	1
Measure of Price Dispersion					
$\hat{u}$	Citywide residuals	0	0.02	-0.16	0.19
$\hat{v}$	Localized residuals	0	0.02	-0.15	0.18
Gas Stations	Number of stations in Seoul	588	18.5	558	628

a. Quarterly data and district level.

I restrict the sample by removing 43 stations that have no self-service stations within 1.5 miles over the sample period (because the log number of these stations is undefined in my specification in the next section). To ensure that my analysis is robust to the missing stations, I also define the number of full- and self-service stations within a five-mile radius.[7] Since my analysis is robust to this change, I use a 1.5-mile metric to be consistent with earlier literature in this field of studies.

To control for demand characteristics, I include population density and a housing rent index (as a proxy for household income), both of which are provided at the district level at a quarterly frequency and come from Seoul Statistics: a database that is operated by the Seoul government. There are 25 districts in Seoul, each with 30 stations on average. Population density is the actual level data.[8] The rent index is constructed from a repeat-sales model and is normalized to 100 for all districts in January 2006. Table 1.1 summarizes the data that are described in this section.

1.2.2 Market Overview

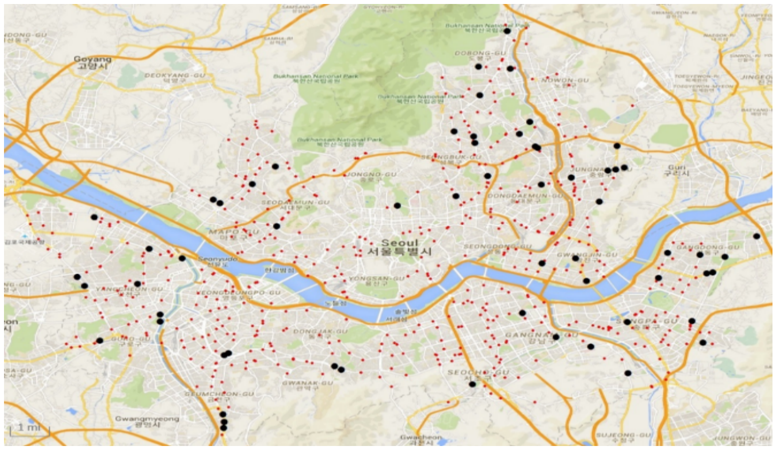
The first gas station with self-service pumps opened in Seoul in 1993, but at the end of 2007 more than 99% of stations in Seoul still offered only full-serve gasoline. Since 2008, the conversion of the market to self service has dramatically accelerated – possibly due to a large spike in crude oil prices. As of September 2014, self-service stations accounted for about 20% of Seoul's 583 stations.[9] One or two new self-service stations have opened every month on average since 2008. In many cases these are a result of switching rather than de novo entry. Self-service customers generally pay 2-4

There are seven brands that operate in the market. In descending order by market share, these are: SK Energy, GS Caltex, S-Oil, Hyundai Oilbank, Alddle, Nonghyup Oil, and Namhae Oil. In my analysis, I drop Nonghyup Oil and Namhae Oil. Each

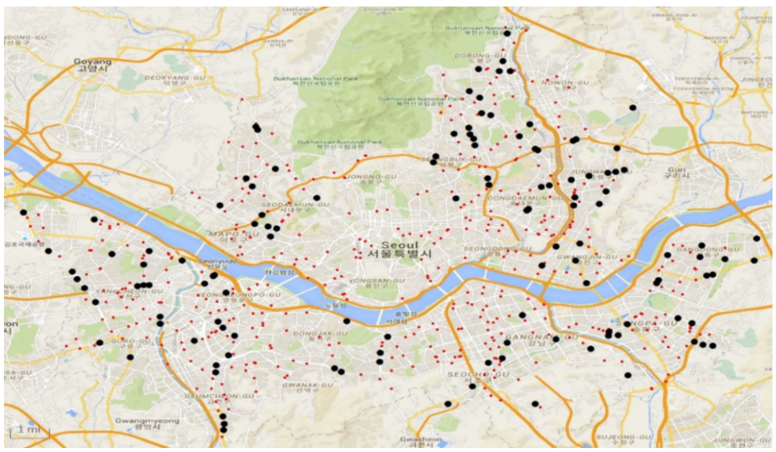
[7] When using a five-mile radius, no stations are dropped. I re-estimate the main regressions of this paper, and the results are shown in Table A6 in Kim (2017). In addition, using a logistic regression with odds ratios, I find that the dropped stations are about twice as likely to be full service and to have a convenience store, but are about a third less likely to have other station amenities such as a carwash or repair facility. They are also less likely to be located at an intersection. These results are shown in Table A7 in Kim (2017)

[8] Korean residents are required to report relocations as well as births and deaths to a local administrative office.

[9] The total number of gasoline stations decreases from 628 to 583 over the sample period. This phenomenon is called "station rationalization" and often observed in the retail gasoline industry. See Eckert and West (2005).



(a) May 05, 2010



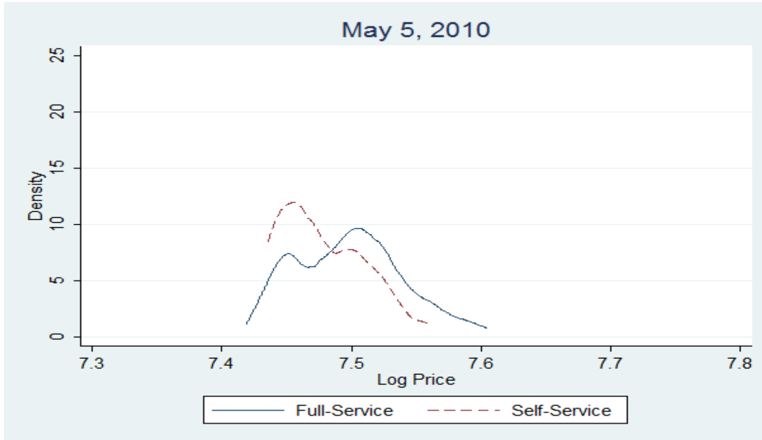
(b) September 24, 2014

Figure 1.1: Expansion of Self-Service Stations

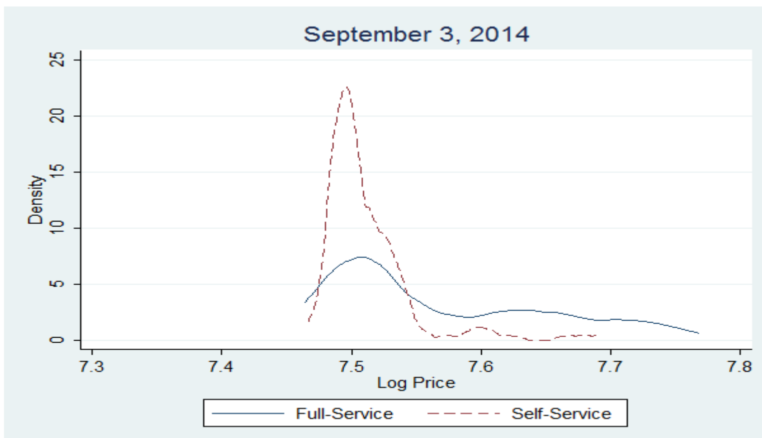
brand has only one station in the market; Nonghyup is a non-private company, and Namhae exited the market in June 2012.

Figure 1.1 shows the locations of gasoline stations on the first day of my sample (May 5, 2010) and the last day (September 24, 2015). Large dots indicate self-service stations, and small dots are full-service stations. Seoul is the largest city (234 square miles) in Korea and highly dense (43,600 residents per square mile), and is located in a small basin that minimizes concerns about stations near the city boundary.

Figure 1.2 shows price distributions by service level on the first Wednesday of May 2010 and September 2014, respectively. Gasoline prices are widely dispersed in this market. For example, the highest price was 20% above the lowest price in May 2010,



(a) May 05, 2010



(b) September 3, 2014

Figure 1.2: Price Distribution by Service Level in May 2010 and September 2014

and this gap was 37% in September 2014. More interestingly, the shape of price distributions changes in opposite directions for the two service levels: The distribution of self-service prices narrowed, while the distribution of full-service stations widened over time. These changes in distributions may suggest that full-service stations have been differentiated from self-service stations and possibly differentiated on other dimensions, such as extra services (e.g., carwash or coffee).

1.3 Empirical Methodology

1.3.1 Measure of Price Dispersion

I calculate price dispersion as the variance of residuals from the price equation, or "cleaned" prices, rather than actual prices. My analysis focuses on the price dispersion that cannot be explained by station characteristics, and a fixed-effects specification is a basic framework for this.[10] To decompose actual prices more precisely into explained and unexplained components, I also control for the number of nearby sellers by service level, which is very likely related to gasoline pricing but is time-variant in my setting, as well as a few more time-variant demand characteristics. The price specification is as follows:[11]

$$\begin{aligned}\ln Price_{it} = \beta_i^{FS} Full_{it} + \beta_i^{SS} Self_{it} + \theta \ln(NumFS_{it}) + \lambda \ln(NumSS_{it}) \\ + \delta \mathbf{X}_{it} + \phi_t + u_{it}. \end{aligned} \quad (1.1)$$

The station fixed effects β_i capture price differences due to observed and unobserved time-invariant station characteristics; Full_{it} and Self_{it} are indicators for full service and self service, respectively. I allow station fixed effects to change if the station switches from full service to self service; each station i can therefore have up to two fixed effects: β_i^{FS} if it offers full service at time t; and β_i^{SS} if it offers self service. NumFS_{it} and NumSS_{it} indicate the number of full-service and self-service stations within 1.5 miles, respectively. The vector $\boldsymbol{X}$ includes time-varying controls that are potentially correlated with gasoline pricing, such as the station's brand and district-level population density and rent; and the time fixed effects ϕ_t reflect changes in

[10] A fixed-effects approach has been used to measure price dispersion for this purpose in many other settings: e.g., Sorensen (2000) for prescription drugs, Lach (2002) for supermarkets, and Lewis (2008) for retail gasoline.

[11] Lewis (2008) measures localized dispersion by estimating a regression of citywide residuals on local average residuals. Both specifications produce qualitatively similar results in my setting.

average prices, mainly driven by wholesale gasoline prices. With the exception of service level, station characteristics are generally time invariant in my dataset.

From Equation (1.1), my primary focus is on the variance of the residuals. The residual $\hat{u}_{it}$ represents the price deviation of station i on day t from its average position relative to the city average, so that the variance of these residuals can be interpreted as a citywide measure of price dispersion. The validity of my approach to computing a "clean" measure of price dispersion depends on how well Equation (1.1) describes the expected price. The adjusted R^2 represents a relative measure of the variation in prices that is explained by the model, although it is generally not informative on an absolute magnitude of the variation that is left unexplained by the model: The adjusted R^2 of the regression is 0.90, and much of the variation is accounted for by fixed effects.[12]

I have an alternative approach to measuring dispersion, which captures the possibility that price competition is localized in the retail gasoline market. Lewis (2008) finds that local price dispersion is useful in studying competition in retail gasoline. I define local price variation for station i at time t as the simple difference between its citywide residual $\hat{u}_{it}$ and the average residual of stations within 1.5 miles of it:

$$\hat{v}_{it} = \hat{u}_{it} - \frac{\sum_{j \in J(it)} \hat{u}_{jt}}{N_{J(it)}}, \tag{1.2}$$

where J(it)={stations within 1.5 miles of station i at time t}. The adjusted residual $\hat{v}_{it}$ in Equation (1.2) represents the price deviation of station i on day t from its average position relative to the local average. The variance of the adjusted residuals can be interpreted as a localized measure of price dispersion.

Table 1.2 shows the results of a simple mean comparison of prices and dispersions between two service groups of stations. Specifically, the mean price of full-service

[12] I also examine how station fixed effects are correlated with station characteristics. Both results are shown in Table A1 and in Table A2 in Kim (2017), respectively.

Table 1.2: Mean Comparisons of Full-Service and Self-Service Prices and Dispersions

Variable	Full Service	Self Service	T-statics for the difference in means[a]
Price	1983.4 (155.69)	1910.8 (103.57)	64.9
$\hat{u}_{it}^2$ (Citywide Variation)	0.0006 (0.0012)	0.0004 (0.0013)	15.8
$\hat{v}_{it}^2$ (Local Variation)	0.0005 (0.0011)	0.0004 (0.0012)	12

Standard deviations in parentheses
a. All p values are < 0.01.

stations is higher than that of self-service stations; and the variance of full-service stations is larger than that of self-service stations. These differences are statistically significant.

1.3.2 Model of Price Dispersion

Following the earlier literature, I exploit a form of multiplicative heteroskedastic variance to define price dispersion as a function of the log number of nearby stations, station characteristics, and time fixed effects.[13] Equation (1.3) shows the explicit form of the model of price dispersion in this paper. These right-hand side variables are basically the same ones that are in the price equation: To specify the most complete model of dispersion, the variables that affect the mean of gasoline prices are very likely to affect the variance of gasoline prices as well.[14] This approach is traditional in the empirical literature when estimating a regression of variance (Genesove, 1995

[13]Harvey (1976) proposes a general form of a regression model with multiplicative heteroskedastic variance, and Genesove (1995) and Lewis (2008) estimate price dispersion using this functional form.

[14]Unlike a model of price level in Equation (1.1), I do not include station fixed effects when modeling price dispersion due to insufficient within-station variation in the number of nearby stations. Station fixed effects would absorb most of the variation in the log squared residuals across individual stations and result in lack of estimation power of the number of stations nearby on price dispersion in my data. Also, I separate the number of stations by service level in a modified specification in the result section.

and Barron et al., 2004).

$$VARIANCE(u_{it}) = e^{\gamma \ln(Num_{it}) + \eta \mathbf{X}_{it} + \xi_t + w_{it}}, \tag{1.3}$$

where E(u_{it})=E(w_{it})=0, ln(Num_{it}) is the log number of gasoline stations that are within 1.5 miles of station i at time t, and ξ_t are time fixed effects that capture differences in the overall level of observed price dispersion over time. The vector $\boldsymbol{X}$ includes Brand Share (the share of nearby stations within 1.5 miles of station i on day t that share the same brand as station i), station characteristics, district-level population density, and the housing-price index from Table 1.1.

The vector X is included primarily to help isolate the effect of ln(Num), although it is also interesting to see how the covariates are related to observed price dispersion. Brand Share is intended to control for the possibility that a local dealer leases or owns several stations under the same brand in the market; the presence of additional station amenities, location at an intersection, population density, and housing rent (a proxy for household income) can all shift gasoline demand.

Equation (1.4) is the empirical counterpart of Equation (1.3). I apply a two-step procedure for estimating the model of price dispersion with multiplicative heteroskedasticity. I first use the residuals that are estimated by OLS in Equation (1.1) for the error terms u_{it}, and then I examine price dispersion (citywide and local) with respect to my variables of interest. For example, the coefficient γ reveals how price dispersion, computed from the residuals that already account for time-invariant station characteristics and time-varying controls as well as time fixed effects, varies with the intensity of local competition:[15]

$$\ln(\hat{u}_{it}) = \gamma \ln(Num_{it}) + \eta \mathbf{X}_{it} + \xi_t + w_{it}, \tag{1.4}$$

where $\mathrm{E}(u_{it})$=$\mathrm{E}(\mathrm{w}_{it})$=0, w_{it}=$\rho_i \mathrm{w}_{it-1}$+u_{it}, $\mathrm{E}(\mu_t \mu_t^{'}) \equiv \mathbf{M}$, $E(\mu_{it})$=0, $\mathrm{E}(\mu_{it})$=0, and $\rho \in$ (-1,1). I allow for the error term w_{it} to be heteroskedastic, correlated across stations, and serially correlated with a station-specific AR(1) process.[16] Heteroskedasticity allows for the possibility that the variance of the log squared residuals may differ across stations within time t, and the AR(1) process permits price persistence of each station from time to time. In estimation, the assumption that the innovation part of w_{it} is stationary is important to obtain feasible residuals $\hat{w}_{it}$ and $\hat{w}_{it-1}$. All combined, I implement a Feasible Generalized Least Squares procedure.[17]

Stations that switched from full service to self service are special, so I briefly review their citywide and local price dispersion. There are 63 such stations in my dataset. I plot the residuals from each station's last full-service observation and first self-service observation in Figure 1.3. More than 65% of such stations close for longer than one Wednesday, so the "before" and "after" observations may not be in two consecutive weeks.[18]

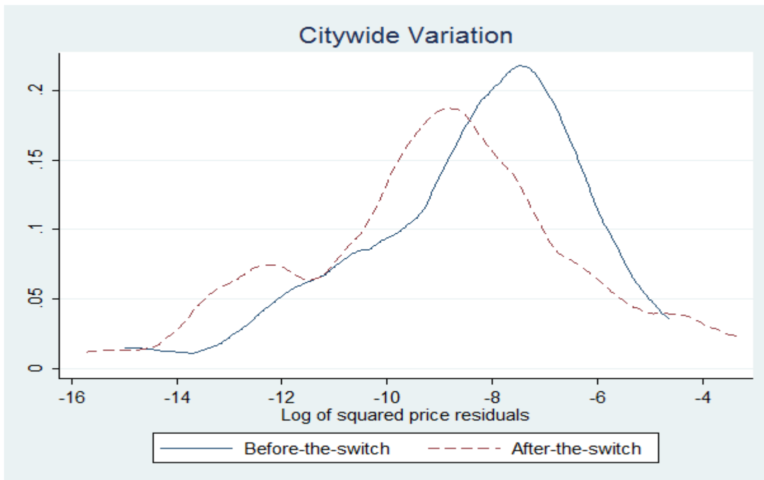
I take two lessons from this exercise: The distribution of the residuals shifts to the left; and the range of the distribution shrinks by about 20% (from -0.25 to -0.30,

[15]This specification explains the heteroskedastic variance of the squared residuals. Note that the residuals that are estimated by OLS are produced under the assumption that the variance of log price of stations is homogenous. However, this assumption does not change the residuals in the first-stage regression, although it changes the standard errors of the coefficient estimates. This paper focuses on the residuals rather than the coefficient estimates.

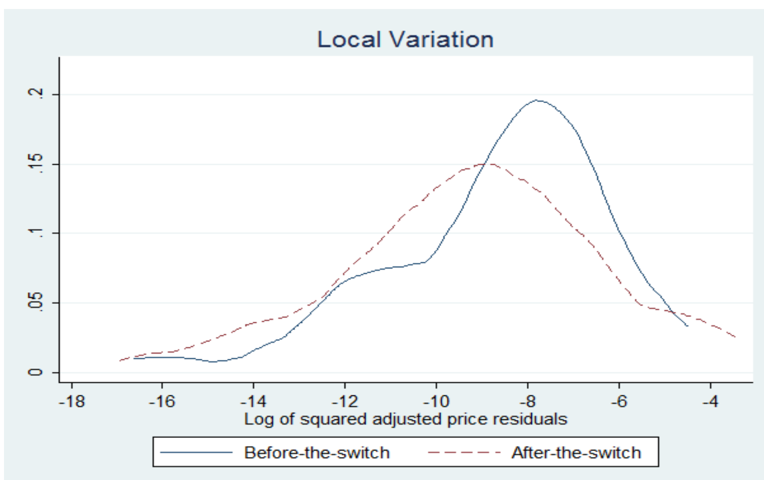
[16]These are conservative and data-dependent assumptions. The distribution of the station-specific coefficient estimates is shown in Figure 1.4 in Appendix 1.A, and the null hypothesis that all stations have common ρ is rejected at a significance level of 0.01.

[17]The coefficient estimates in this specification are recovered in three stages. First, I obtain the residuals $\hat{w}$ from the specification without an AR process. Second, I obtain an estimated vector of coefficients ρ in the AR process (on the assumption that μ is stationary). Finally, having and ρ and $\hat{w}$ enables me to construct the estimated variance-covariance matrix M.

[18]The average number of missing weeks is 17; the median is 12; the minimum is zero; the maximum is 125 weeks.



(a) Citywide Variation



(b) Local Variation

Figure 1.3: Price Dispersion of Stations that Switched from Full Service to Self Service during the Sample Period

Note: The before-the-switch distribution uses stations' last full-service observations and the after-the-switch distribution uses stations' first self-service observation.

approximately, for both the citywide and local measures).

In the next section, I analyze the correlates of price dispersion with local competition more formally using a regression framework.[19]

1.4 Results

1.4.1 Price Dispersion and Competition

Table 1.3 shows estimation results of Equation (1.4). In the baseline regression in Column (1), I find that a 10% increase in the number of existing stations within 1.5 miles is associated with a 1.64% reduction in price dispersion. This negative coefficient is consistent with previous studies of retail gasoline markets including Marvel (1976), Barron et al. (2004), and Lewis (2008), while counter to Chandra and Tappata (2011) and Lach and Moraga-González (2015).

Next, I test for market segmentation by service level and examine how price dispersion varies (indirectly) with search intensity. I assume that full-service customers have a lower search intensity than do self-service customers. For example, full-service buyers may be less willing to search just for lower gasoline prices because gasoline is one component of the full package that they buy – particularly because the second component, service, is perceived as heterogeneous. In contrast, self-service buyers are more likely to search for a low price because gasoline is perceived as homogeneous. As a result, in markets with many self-service stations, stations are exposed to more high-search intensity customers than are stations in markets with no self-service stations. This identification strategy derives from Salop and Stiglitz (1977)'s intuition, consumer heterogeneity in search costs, and Lewis (2008)'s inference although Lewis

[19]Similarly, I compare the price dispersion of stations that enter or exit to the market-level distribution of price dispersion. However, I do not find any interesting patterns. See Figure 1.5 in Appendix 1.A for details.

Table 1.3: Explaining Citywide Price Dispersion

	(1)	(2)	(3)	(4)	(5)
Competition Variables					
ln(Num)	-0.164***	-0.175***	-0.188***	-0.115***	-0.149***
	(0.032)	(0.068)	(0.039)	(0.038)	(0.038)
ln(Num)*Full		0.115			
		(0.108)			
ln(Num)*High-Brand			0.036		
			(0.032)		
ln(Num)*Leader-Brand				-0.142***	
				(0.047)	
ln(Num)*Branded					-0.019
					(0.023)
Brand Share	0.738***	0.758***	0.697***	0.799***	0.776***
	(0.129)	(0.129)	(0.134)	(0.131)	(0.136)
Station Characteristics					
Full	0.729***	0.197	0.731***	0.721***	0.727***
	(0.037)	(0.184)	(0.037)	(0.037)	(0.037)
Store	0.166***	0.167***	0.163***	0.172***	0.169***
	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)
Carwash	0.214***	0.214***	0.215***	0.214***	0.213***
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
Repair	-0.028	-0.027	-0.029	-0.032	-0.028
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
Intersection	-0.144***	-0.143**	-0.144**	-0.144**	-0.151**
	(0.038)	(0.039)	(0.038)	(0.038)	(0.038)
SK	0.148**	0.147**	0.072	0.553***	0.167**
	(0.065)	(0.065)	(0.093)	-0.145)	(0.074)
GS	-0.035	-0.034	-0.0114	-0.034	-0.013
	(0.060)	(0.060)	(0.091)	(0.060)	(0.067)
Hyundai	-0.142**	-0.139**	-0.122*	-0.147**	-0.116
	(0.063)	(0.063)	(0.066)	(0.064)	(0.072)
S-Oil	-0.036	-0.032	-0.024	-0.033	-0.009
	(0.065)	(0.065)	(0.067)	(0.066)	(0.074)
Alddle	-0.313***	-0.306***	-0.311***	-0.298***	-0.310***
	(0.094)	(0.094)	(0.094)	(0.094)	(0.095)
ln(Pop)	-0.471***	-0.469***	-0.474***	-0.464***	-0.469***
	(0.055)	(0.056)	(0.055)	(0.055)	(0.055)
ln(Rent)	-1.375***	-1.368***	-1.370***	-1.394***	-1.376***
	(-0.117)	(-0.118)	(-0.117)	(-0.117)	(-0.117)
Num. Observation	117212	117212	117212	117212	117212
Mean of Dep. Var	-9.274	-9.274	-9.274	-9.274	-9.274
SD of Dep. Var	2.441	2.441	2.441	2.441	2.441

All specifications also include time fixed effects.
Denote significance ***, **, * at 1%, 5%, 10% level, respectively.
Unbranded is omitted.

focuses on a brand dimension.

To estimate the relationship between price dispersion and local competition by service level, I add an interaction of ln(Num) and a full-service indicator to the baseline specification (Equation 1.4).[20] The results in Column (2) of Table 1.3 show how the relationship varies across sellers of different service levels. To be specific, ceteris paribus, 10% more stations within 1.5 miles reduce a self-service station's price dispersion by 1.75%. As for the interaction term, it picks up differential sensitivity to the degree of competition when a station is full service, and it is statistically insignificant in this specification.

I also test for segmentation by brand type. Based on market shares and average price by brand type, I first separate high-brand and low-brand similar to Lewis (2008). High-brand stations are SK Energy and GS Caltex, which, together, have about a 70% market share. Following Lewis (2008), I assume that low-brand customers are more likely to search and have better price information than do high-brand customers. In Korea, consumers rate SK Energy and GS Caltex as higher quality than the other brands (Kim et al., 2010); these stations charge about 3% more on average during the sample period. Second, I partition brands into a leader brand (SK Energy) and follower brands (all others), where SK Energy accounts for about half of the entire market. Lastly, I distinguish between branded and unbranded stations, following similar logic.

In Columns (3)-(5) of Table 1.3, I find that the coefficients on ln(Num) are negative across brand specifications and the interaction terms are negative or insignificant. The

[20] In addition to this pooled-sample regression, I also estimate separate-sample regressions for full-service and self-service subsamples, starting from a price-level regression in Equation (1.1). Chow-like F tests show no significant structural breaks between the subsamples, which may indicate that the market has been not clearly segmented by service level during my study period. The market conversion had just started at the beginning of my sample, and many low-price full-service stations charged prices similar to self-service stations during this period, so it seems that some full-service stations still competed with self-service stations on gasoline price in this market. Accordingly, I keep the pooled-sample regressions in this study. See Table A3 in Kim (2017) for detailed results of the Chow-like tests.

negative coefficient on the interactions does not support the claim that customers at high-type stations may be less informed than customers at low-type stations. I also find that the coefficient estimates on the interaction terms differ considerably across specifications. This finding suggests that either the Seoul market is not segmented by brand or the classification of brands in my study does not accurately reflect the segmentation. In either case, this finding shows that the estimated coefficients by brand type are sensitive to the criteria that are used to separate stations.

The coefficient on brand share is positive throughout, which indicates that the variability of prices is higher in markets with more stations of the same brand. One possible explanation for this finding is that stations may engage in "soft competition" within brands. This positive coefficient contrasts with Lewis (2008), who estimates a negative but insignificant coefficient on brand share.

In the lower part of Table 1.3, I present the coefficient estimates on station characteristics. Specifically, when ln(Num) is zero and holding all else equal, stations that are located at an intersection have 14% lower price dispersion than those that are not located at an intersection. Price dispersion for full-service stations is approximately double price dispersion for self-service gasoline, and stations with a store or a carwash facility have up to 18% and 23% higher price dispersion than do stations without those amenities. These findings suggest that consumers who buy gasoline at a full-service station or at a station with a store or a carwash facility may be less informed about prices or more interested in these extra amenities.

I estimate the same specifications with local price variation as the dependent variable (instead of citywide price variation). These results are shown in Table 1.4. I find similar patterns to those that were observed for the citywide price dispersion and an even clearer pattern in the service-level specification. For example, in Column (2) of Table 1.4, the coefficient on ln(Num) becomes larger in absolute value (more negative) and the interaction term becomes larger (more positive) and is significant

Table 1.4: Explaining Local Price Dispersion

	(1)	(2)	(3)	(4)	(5)
Competition Variables					
ln(Num)	-0.235***	-0.414***	-0.232***	-0.162***	-0.291***
	(0.033)	(0.074)	(0.04)	(0.04)	(0.041)
ln(Num)*Full		0.214***			
		(0.083)			
ln(Num)*High-Brand			-0.004		
			(0.033)		
ln(Num)*Leader-Brand				-0.156***	
				(0.046)	
ln(Num)*Branded					0.063**
					(0.026)
Full	0.681***	0.237***	0.681***	0.681***	0.685***
	(0.038)	(0.067)	(0.038)	(0.038)	(0.038)
Num. Observation	117212	117212	117212	117212	117212
Mean of Dep. Var	-9.352	-9.352	-9.352	-9.352	-9.352
SD of Dep. Var	2.466	2.466	2.466	2.466	2.466

All specifications also include brand share, station characteristics, and time fixed effects.
Coefficients on brands are not shown due to their insignificance.
Denote significance *** at 1% level.

at the 1% level. The positive coefficient on the interaction term implies that self-service price dispersion is much more elastic than is full-service price dispersion with respect to the number of nearby stations. As for market segmentation by brand type, the coefficients on the interaction terms remain unchanged except for the branded-unbranded setting, which changes sign and statistical significance relative to Table 1.3.

Overall my results are consistent with previous studies that show that price dispersion varies with the heterogeneity of sellers and buyers, although my setting is significantly different and my specifications extend the notion of differentiated sellers in retail gasoline (commonly proxied by brand) into service level.

1.4.2 Relative Effects by Seller Type

Recall that dispersion is negatively associated with the number of stations but the magnitude of the relationship differs by service level. To examine this pattern more closely, I partition the number of stations into the number of full-service stations and the number of self-service stations. By interacting these variables with the service-level indicator as in Equation (1.5), I identify four relative effects of local competition by service level: the elasticity of self-service price dispersion with respect to self-service stations (γ_1), the elasticity of self-service dispersion with respect to full-service stations (γ_3), the elasticity of full-service dispersion with respect to self-service stations ($\gamma_1+\gamma_2$), and the elasticity of full-service dispersion with respect to full-service stations ($\gamma_3+\gamma_4$). I estimate Equation (1.5) under the same error structure that I assume in the baseline specification:[21]

$$\begin{aligned}\ln Price_{it} = \gamma_1 \ln(NumSS_{it}) + \gamma_2 \ln(NumSS_{it}) * Full_{it} \\ + \gamma_3 \ln(NumFS_{it}) + \gamma_4 \ln(NumFS_{it}) * Full_{it} \\ + \eta \mathbf{X}_{it} + \xi_t + \varepsilon_{it}. \end{aligned} \tag{1.5}$$

Table 1.5 shows the estimates by service level. If I start with citywide variation, Column (1) shows that a 10% increase in the number of nearby self-service stations leads to a 3.6% reduction in self-service dispersion and reduces full-service dispersion by about 1.7%. These estimates support my claim that stations in markets with more self-service stations are likely to face high-search-intensity customers. One possible explanation for this may be that self-service customers purchase only gasoline, whereas gasoline is only one component of the full package that full-service customers purchase. In addition, a 10% increase in nearby full-service stations reduces price dispersion for both services: a 1% reduction in self-service dispersion and a 0.4% reduction

[21] I specify the empirical models of brand classifications (low-brand vs. high-brand, follower-brand vs. leader-brand, and unbranded vs. branded) similarly.

Table 1.5: Relative Effects of Different Sellers, by Service Level

	Citywide Variation (1)	Local Variation (2)
Self-Service vs. Full-Service		
ln(NumSS)	-0.363***	-0.481***
	(0.052)	(0.053)
ln(NumSS)*Full	0.194***	0.301***
	(0.056)	(0.057)
ln(NumFS)	-0.104**	-0.272***
	(0.051)	(0.057)
ln(NumFS)*Full	0.061	0.208***
	(0.060)	(0.065)
Quick Summary of Estimates[a]		
Elasticity of SS dispersion with (SS or FS) stations	(-0.363 or -0.104)	(-0.481 or -0.272)
Elasticity of FS dispersion with (FS or SS) stations	(-0.043 or -0.169)	(-0.064 or -0.180)
Num. Observation	117212	117212
Mean of Dep. Var	-9.274	-9.352
SD of Dep. Var	2.441	2.466

All specifications also include brand share, station characteristics, and time fixed effects. Denote significance *** at 1% level.

[a] The pair of numbers in each parentheses indicates how the elasticity of price dispersion with respect to number of stations varies depending on the service composition of nearby stations. For Elasticity of SS dispersion, the left number is the elasticity with respect to the number of self-service stations. number is the elasticity with respect to the number of full-service stations. For Elasticity of FS dispersion, the left number is the elasticity with respect to number of full-service stations. The right number is the elasticity with respect to the number of self-service stations.

in full-service dispersion, although the latter is not statistically different from the former.

The patterns I observe in local price variation mirror those for citywide price variation. In Column (2) of Table 1.5, self-service dispersion always decreases with local competition: It decreases by 4.8% with a 10% increase in self-service stations nearby and by 2.7% with a 10% increase in full-service stations nearby. Full-service dispersion decreases with the number of self-service stations nearby, but interestingly, the effect of full-service stations on full-service dispersion is statistically not different from zero. These results show that greater competition is not always related to less dispersion, as is often suggested in empirical studies in the gasoline literature.

I also find that the relationship between local price dispersion and number of sellers by service level is asymmetric across service levels. For example, focusing on local variation in Column (2) of Table 1.5, I find that the elasticity of self-service price dispersion with respect to the number of nearby full-service stations is one and a half times larger (in absolute value) than the elasticity of full-service price dispersion with respect to the number of nearby self-service stations. Overall, across specification, self-service price dispersion exhibits greater sensitivity to the number of sellers than does full-service dispersion.[22]

The primary difference between Lewis's (2008) results and mine is in the coefficient that captures the elasticity of low-type station dispersion with respect to low-type stations, where the low type refers to self-service stations in my paper and low-brand stations in Lewis's paper. I find a large and statistically significant elasticity, whereas Lewis found a small and insignificant coefficient. One possible explanation for this difference is the market transformation: In my setting, the conversion of many full-service stations to self-service stations may have increased price competition among self-service stations, so the large coefficient may be attributed to the changes in competitive conditions in this market.

Finally, I estimate a variant of Equation (1.5), replacing the full- and self-service interactions with interactions by brand types, and I show results for local variation in Table 1.6; results for citywide variation are shown in Table A4 in Kim (2017). In Columns (1)-(2) of Table 1.6, I generally find qualitatively similar results to those that I report in the service-level segmentation. One difference is that the relationship between price dispersion and local competition can even be positive depending on the composition of the existing stations. For example, in Column (1), price dispersion for low-brand stations decreases with the number of nearby low-brand stations but increases with the number of nearby high-brand stations.[23]

[22]Models of search generally predict a non-monotonic relationship between search intensity and dispersion, but my main specification estimates a monotonic and linear relationship between them. To test for non-monotonicity, I divide my sample into stations with above-average numbers of self-service stations and those with below-average numbers of self-service stations. Using Chow tests, I confirm that the elasticities are statistically indistinguishable across the two samples. This result is also qualitatively similar when I divide the sample based on the average number of full-service stations. I do not show the results in this paper.

[23]In the unbranded/branded setting in Column (3), the results are generally insignificant; the sample size is much smaller because many stations have no unbranded stations within 1.5 miles. I re-estimate the same specification using a five-mile radius in which most stations have both branded

Table 1.6: Relative Effects of Different Sellers, by Brand Type

	Local Variation		
	(1)	(2)	(3)
Low-Brand vs. High-Brand			
ln(NumLO)	-0.523***		
	(0.050)		
ln(NumLO)*High-Brand	0.036		
	(0.061)		
ln(NumHI)	0.405***		
	(0.059)		
ln(NumHI)*High-Brand	-0.049		
	(0.063)		
Follower-Brand vs. Leader-Brand			
ln(NumFB)		-0.406***	
		(0.046)	
ln(NumFB)*Leader-Brand		0.148**	
		(0.078)	
ln(NumLB)		0.318***	
		(0.048)	
ln(NumLB)*Leader-Brand		-0.337***	
		(0.083)	
Unbranded vs. Branded			
ln(NumUB)			-0.192
			(0.131)
ln(NumUB)*Branded			0.096
			(0.136)
ln(NumBB)			-0.170***
			(0.063)
ln(NumBB)*Branded			0.076*
			(0.039)
Quick Summary of Estimates[a]			
Elasticity of LT dispersion with (LT or HT) stations	(-0.523 or 0.405)	(-0.406 or 0.318)	(-0.192 or -0.170)
Elasticity of HT dispersion with (HT or LT) stations	(0.356 or -0.487)	(-0.019 or -0.258)	(-0.094 or -0.096)
Num. Observation	117212	117212	75819
Mean of Dep. Var	-9.352	-9.352	-9.3
SD of Dep. Var	2.466	2.466	2.451

All specifications also include brand share, station characteristics, and time fixed effects.
Denote significance *** at 1% level.
LT is low-type stations (low, follower, and unbranded) and HT is high-type stations (high, leader, and branded).
[a] The pair of numbers in each parentheses indicates how the elasticity of price dispersion with respect to number of stations varies depending on the brand composition of nearby stations. For Elasticity of LT dispersion, the left stations. For Elasticity of SS dispersion, the left number is the elasticity with respect to number is the elasticity with respect to the number of low-type stations. The right number is the elasticity with respect to the number of high-type stations. For Elasticity of HT dispersion, the left number is the elasticity with respect to the number of high-type stations. The right number is the elasticity with respect to the number of low-type stations.

To summarize the results from the brand-segmentation exercises, the coefficients on both local competition and the interaction of local competition and the brand-type indicator are sensitive to how brands are classified into different types, which suggests that service level is a cleaner dimension on which to segment the market.

1.4.3 Potential Bias with Coefficient Estimates

Neither the location of a gas-station nor the decision to convert a full-service station to a self-service station is exogenous. For example, full-service stations may prefer to locate in a district that is populated by customers with high search costs. In this case, the relationship between dispersion and station configuration is, at least in part, due to reverse causality; the coefficient on ln(NumFS) would therefore be biased upward in absolute value. By the same logic, the coefficient on ln(NumSS) could be biased downward in absolute value.

To mitigate potential endogeneity bias, I add district fixed effects to the main specification. I confirm that the earlier results in Table 1.5 are robust to controlling for time-invariant district effects.

1.4.4 Applying Search Model

Search-theoretic models have been developed to generate equilibrium price dispersion as a function of information costs generally with an assumption of one homogeneous product. For example, some homogenous-product models (such as Varian, 1980) predict, ceteris paribus, a non-monotonic relationship between price dispersion and search intensity; or a positive relationship between price dispersion and the number of sellers.[24]

and unbranded stations nearby, and I continue to find insignificant coefficients. The results are omitted in this paper.

[24]Search-based models differ with respect to assumptions such as the information acquisition channel and firm and consumer heterogeneity. As a result, model predictions about the sign of the

In this empirical study, I consider the co-movement of three variables: price dispersion; number of differentiated sellers; and search intensity. My findings are consistent with the notion that the self-service market has a higher fraction of searchers than does the full-service market, and that the higher fraction of searchers in the self-service market limits the extent of self-service price dispersion. Nevertheless, my setting is not well-suited to a direct test of search models because the co-movement of the three variables does not satisfy the *ceteris paribus* assumption. In this sense, the extant theories of dispersion may not be sufficient to explain my empirical results.[25]

1.5 Robustness Checks

I perform several robustness checks on all regressions above: using both citywide and local variation as LHS variables, and estimating the relative effects of both variation by service level and brand type. The results of the robustness checks on relative effects by service level are in Section 1.7.2 (Appendix 1.B).

1.5.1 Shorter Sample Period

An important assumption in the fixed-effects specification is that station fixed effects control for price differences that are driven by station characteristics across stations, so price dispersion in this paper is not a result of station characteristics. My sample is long (about 230 weeks), and it may be long enough to allow for significant changes in station characteristics. In Table 1.7, I use a subsample of 108 weeks (from September 5,

relationship between price dispersion and information costs vary – predicting, variously, a positive relationship (Reinganum, 1979), a negative relationship (MacMinn, 1980), or a non-monotonic relationship (Varian, 1980). Empirical counterparts of the relationship are also not uniform: Studies have found a positive relationship (Borenstein and Rose, 1994), a negative relationship (Gerardi and Shapiro, 2009), or a non-monotonic relationship (Chandra and Lederman, 2018). See Baye et al. (2006) for a review of a wide range of search models from both theoretic and empirical perspectives.

[25] Wildenbeest (2011) provides a model-based framework for studying price dispersion in markets with product differentiation and search frictions with simplification assumptions on consumer preference and firm's quality input factors.

2012, to September 24, 2014) to test robustness of the results. The observed patterns of estimates are generally robust to the subsample regression, even though I lose some observations on the stations that switch from full service to self service before September 2012, which weakens the statistical significance.

1.5.2 Alternative Measure of Price Dispersion

Measuring price dispersion as multiplicative heterogeneous variance imposes a symmetric relationship between price dispersion and local competition (i.e., residuals are squared). To investigate whether positive and negative residuals are differentially correlated with competitive conditions, I separate my sample into two subsamples: One contained only observations for which residuals are non-negative, and the other contained only the negative residuals. In a related exercise, I replace the variance with the absolute value of the residuals. Both sets of results are shown in Table 1.8 The main estimates are generally robust to these changes.[26]

1.6 Concluding Remarks

This paper argues on ex ante grounds that demand for gasoline is segmented by service level, and demonstrates that price dispersion patterns are consistent with this segmentation. Gasoline stations sell a nearly identical quality of gasoline. Because driving is costly to consumers, spatial differentiation is a first-order dimension along which the market is segmented. Lewis (2008) shows the importance of brand as another dimension on which the gasoline market is segmented. In addition to those dimensions of differentiation, I show that the service dimension – full service versus self service – is important, and I argue that it is more intuitive and robust than brand

[26] I present results only from local variation in Table 1.8 since results from citywide variation and those from local variation are qualitatively similar.

segmentation.

Using the transition of service level in the Korean gasoline market, I find that price dispersion always decreases with the number of nearby self-service stations, whereas the number of full-service stations does not predict the level of price dispersion for full service; a higher number of full-service stations can even exist with high dispersion. This result suggests that dimensions other than price are too important to ignore when studying price dispersion.

1.7 Appendices to Chapter 1

1.7.1 Appendix 1.A: Supplementary Results

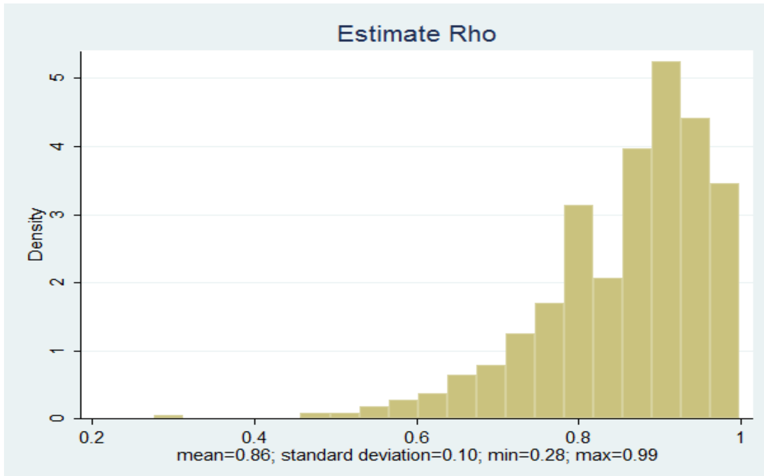
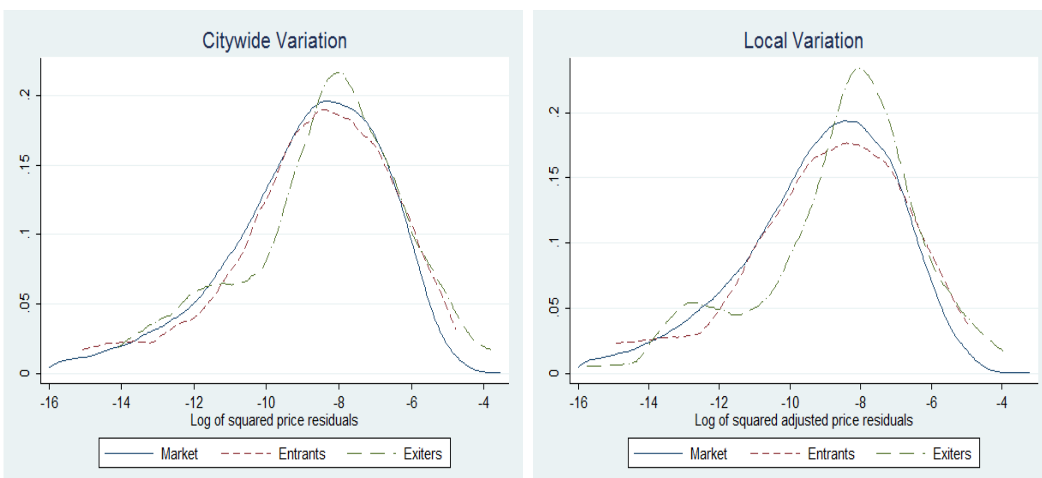


Figure 1.4: The Distribution of Station-Specific Estimated AR(1) Coefficients



(a) Citywide Variation

(b) Local Variation

Figure 1.5: Price Dispersion of Stations that Enter or Exit during the Sample Period.

1.7.2 Appendix 1.B: Robustness Checks

Table 1.7: Robustness Check: Subsample (September 5, 2012 – September 24, 2014). Relative Effects of Different Sellers, by Service Level

	Citywide Variation (1)	Local Variation (2)
Self-Service vs. Full-Service		
ln(NumSS)	-0.322*** (0.068)	-0.293*** (0.071)
ln(NumSS)*Full	-0.053 (0.076)	0.207*** (0.080)
ln(NumFS)	0.134** (0.064)	0.036 (0.078)
ln(NumFS)*Full	0.071 (0.060)	-0.072 (0.065)
Quick Summary of Estimates		
Elasticity of SS dispersion with (SS or FS) stations	(-0.322 or 0.134)	(-0.293 or 0.036)
Elasticity of FS dispersion with (FS or SS) stations	(0.205 or -0.375)	(-0.036 or -0.086)
Num. Observation	48121	48121
Mean of Dep. Var	-10.06	-10.09
SD of Dep. Var	2.518	2.506

All specifications also include brand share, station characteristics, and time fixed effects.
Denote significance *** at 1% level.

Table 1.8: Robustness Check: Alternative Measure of Price Dispersion. Relative Effects of Different Sellers, by Service Level

	Local Variation		
	(1) $\hat{v}$ >=0	(2) $\hat{v}$ <0	(3) $\lvert\hat{v}\rvert$
Self-Service vs. Full-Service			
ln(NumSS)	-0.720*** (0.075	-0.277*** -0.062	-0.240*** -0.026
ln(NumSS)*Full	0.549*** -0.08	0.196*** -0.068	0.150*** -0.028
ln(NumFS)	-0.199** -0.077	-0.218*** -0.063	-0.136*** -0.028
ln(NumFS)*Full	0.145 -0.089	0.140* -0.075	0.104*** -0.032
Quick Summary of Estimates			
Elasticity of SS dispersion with (SS or FS) stations	(-0.720 or -0.199)	(-0.277 or -0.218)	(-0.240 or -0.136)
Elasticity of FS dispersion with (FS or SS) stations	(-0.054 or -0.171)	(-0.078 or -0.081)	(-0.032 or -0.090)
Num. Observation	58738	58473	117211
Mean of Dep. Var	-9.341	-9.34	-4.67
SD of Dep. Var	2.431	2.486	1.229

All specifications also include brand share, station characteristics, and time fixed effects.
Denote significance *** at 1% level.

Chapter 2

Changing Market Structure and Evolving Ways to Compete: Evidence from Retail Gasoline

2.1 Introduction

Firms constantly innovate, and innovations force changes in the competitive landscape. The impact of technological change has been particularly remarkable in the retail sector.[1] As one example, automation has increased rapidly in various retail markets, altering the optimal allocation of inputs. There is a growing body of research studying employment effects of automation, but very little quantitative evidence on how automation impacts prices. This paper uses a case-study approach to study the competitive effect of an innovation from traditional to modern in a retail sector: the introduction of self-service technology in retail gasoline in Korea.

The retail gasoline industry is useful for examining the pricing behavior of sellers because the product is homogeneous and prices are clearly posted and observed

[1] See Bronnenberg and Ellickson (2015) for an overview of global retail markets and Hortaçsu and Syverson (2015) for an overview of the U.S. retail market.

by consumers. My setting is unique in that I have high-frequency, station-level data starting in 2010, when self-service sellers were very rare, and continuing through 2015, when they accounted for a quarter of the market. I analyze full-service sellers' adoption and exit decisions, pricing strategies, and product re-positioning, and find that gas stations that continued to offer full service differentiated themselves by offering bundled products and services and raised their prices during this period.

The price gap between full service and self service increased most dramatically in Seoul; the full-service premium was 2% in early 2010 and increased to 8% by 2015.[2] At the same time, the distribution of full-service prices became increasingly right-skewed and the variability of self-service prices was relatively stable. These stylized facts clearly suggest a different evolution of pricing strategies of stations by service level during the market conversion from full service to self service.

My findings are three-fold. First, I focus on full-service stations' decisions; they may close, switch to self service, or continue to sell full-serve gasoline after the market transformation accelerated after 2010. Using a multinomial logistic regression, I find that stations that have more self-service competitors nearby are more likely to close than to remain full service. Also, the probability of adopting self-service technologies is about 40% higher than that of remaining full service, at stations with more self-service competitors nearby, which implies the presence of strategic complementarities in service. This result is consistent with prediction of models of competition that intense price competition drives high-cost sellers out of a market.

My second finding is a confirmation that self-service stations not only offer competitive prices but also drive down competitors' prices. Difference-in-difference specifications show that self-service stations charge 5% less per gallon on average during my study period. In addition, *ceteris paribus*, having one more self-service station nearby decreases self-service prices by 1%, but does not statistically impact full-service prices.

[2] A weaker but similar pattern also appeared at the country level.

Overall, the price elasticity with respect to competition is much higher in absolute value when competitors are self service than when they are full service. This finding implies that self-service stations compete for price-sensitive customers who may want low-price gasoline rather than a bundle of gasoline and service, and that full-service stations are more differentiated than self-service stations.[3]

The third finding is descriptive evidence of subtle differentiation on one or more dimensions. For example, some gas stations combine with other types of business, including a dry cleaner, nail salon, or fast-food restaurant. Moreover, some stations offer extra products free of charge, such as coffee, carwash, or vacation packages, depending on the amount of gasoline purchased.[4] These unique features are generally provided by full-service stations whose gasoline prices are significantly higher than the market price for full service (e.g., 14% higher in December 2017). I interpret bundled products and services as new strategic choices by full-service stations in response to the new competitive landscape – i.e., the emergence of self-service competitors.

Using price data, I perform two analyses to supplement the descriptive evidence. First, given information on station's bundling in May 2017, I distinguish stations that later bundle products from those that do not, and then trace the price difference between stations that later bundle products and those that do not. I find that the price difference increases in the later period of my sample, suggesting that stations increasingly bundle their products to charge a premium during the market transition. Second, I examine the relative price stability of stations. After controlling for station characteristics, I show that high-priced stations have much more stable prices than low- or mid-priced stations. That high-end stations settle into fairly fixed prices implies that they are perceived by consumers as heterogeneous.[5]

---

[3]The change in the service composition of sellers alone is not enough to explain the increasing price difference between the two station types. I discuss this in Appendix 2.B.

[4]This strategy is a sort of product bundling. There has been a well-established literature on product bundling and its profitability both theoretically and empirically. See Chen and Riordan (2013) for an overview of the literature on this topic.

[5]A relative price of sellers for one homogeneous good should not change systematically in standard

Differentiated sellers enjoy less intense competition on price, which allows the sellers to charge high and stable prices.[6] Consistent with this implication, my descriptive evidence and supplementary analyses together support the view that the strategic choices of sellers evolve in different ways; higher-priced stations compete for less-price-sensitive consumers, while lower-priced stations that usually offer self-serve gasoline focus on price-sensitive consumers. Taken together, these features have led to an increase in the full-service premium during the introduction of self-service technologies in this market.

This paper contributes to a literature on product differentiation and its softening effect on competition. Mazzeo (2002) examines motel markets located along U.S. interstate highways, and finds that the effect of competition on price is insignificant when motels are differentiated. Basker and Noel (2009) show that the effect of Wal-Mart's entry on competitors' prices is greatest at low-end chains that compete for price-elastic consumers, and smaller at supermarkets that differentiate themselves from Wal-Mart. Ellison and Ellison (2009) provide evidence that online retailers offer a variety of add-on products to frustrate consumers' price search, and show that this practice is designed to mitigate intense competition. Ching (2010) studies evidence that entry of generic drugs causes branded drug price to increase, and applies insights from theories of consumer learning and heterogeneity in price sensitivity to interpret this effect.

My analysis also contributes to a vast literature on different aspects of retail gasoline pricing.[7] Shepard (1991) is the first study to document price discrimination by gasoline stations and estimate the premium charged for full-serve gasoline. Png and Reitman (1994) focus on service-time competition as one aspect of station quality and estimate the premium for a service-time reduction at gas stations. Houde (2012)

models of search – e.g., Varian (1980) and Stahl (1996). Consumers would otherwise know which seller charges the lowest price.

[6] This is an implication of a combination of price differentiation and search.

[7] see Noel (2016) for an overview of retail gasoline pricing.

models retail-gasoline demand that allows for spatial differentiation of stations. More recently, Remer (2015) contributes to a growing body of literature studying asymmetric pricing of retail gasoline along with consumer search.[8]

Within this literature, several other papers have studied the relationship between pricing and product differentiation. Borenstein (1991) examines why stations' markups differ for leaded vs. unleaded gasoline, and reviews some explanations for the difference, such as cost-based, purchase-size-based, and paying-method-based explanations. Lee et al. (2015) use Korean gasoline data and analyze a price gap between stations that sell only regular gasoline and those that sell both regular and premium gasoline. Similar to my setting, Soetevent and Bružikas (2017) exploit the transition of the Dutch retail gasoline market from self-service "staffed" stations to fully automated stations, and find no significant effects of automated stations on staffed stations' prices.

I describe the evolution of market segmentation in Section 2.2, and explain the data used in this paper in Section 2.3. I present my main results on product positioning and pricing strategies in Section 2.4. Section 2.5 concludes.

2.2 The Evolution of Market Segmentation

2.2.1 Full Service vs. Self Service

The first introduction of self-service technologies in the Korean retail-gasoline market was in 1993, but self-service stations failed to attract customers and disappeared before long. A gas station with self-service pumps opened again in 2003, but the self-service format was rare until the end of 2007, when it accounted for about 0.3% of the market. The market conversion from full service to self service accelerated in

[8]See, for example, Yang and Ye (2008), Tappata (2009), and Lewis (2011).

early 2008, possibly due to a sharp increase in global oil prices. One or two new self-service stations have opened every month on average since then. Self-service stations constituted 17% of all stations as of December 2015 at the country level. Most self-service stations are converted from full-service rather than being new entries.

The distinction between full service and self service is whether consumers must put their labor into the production function of stations.[9] Consumers who purchase self-serve gasoline fill their gas tank and pay at the pump themselves, whereas those buying full-serve gasoline wait for a service attendant in their car and request the amount of gasoline they want to buy. In the early 2000s, a few full-service stations in Korea used to provide “full services” such as cleaning windshield and checking under the hood of a car, similar to the U.S. market in the 1960s. However, such services disappeared.[10] Full-service consumers nowadays expect to receive only pumping and paying services, or at most, a windshield-cleaning service.

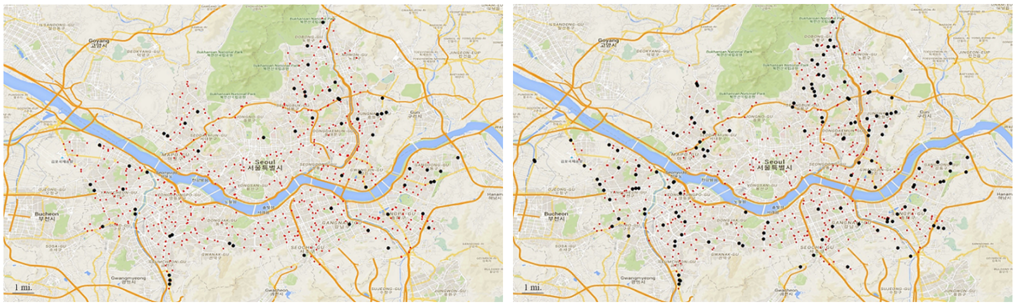
Gasoline stations in Korea are perfectly partitioned by service level at each point in time, with no mixed station offering both full- and self-service pumps, unlike in the U.S. market. I show the locations of gas stations in Seoul on May 2010 and December 2015 in Figure 3.2. Large dots are self service and small dots are full service. Self-service stations that entered later are often located near earlier self-service stations.

I plot the time series of price level and the number of stations by service level, also in Seoul.[11] Figure 2.2-(a) shows average log retail gasoline price by service level and wholesale price over time. The prices fell sharply in early 2015 due to an abrupt

[9]Basker et al. (2017) study the transformation of gasoline stations from full service to self service in the U.S. and examine the labor productivity of full- vs. self-service stations. Foster et al. (2006) more generally examine productivity changes from a massive restructuring in the U.S. retail sector in the 1990s.

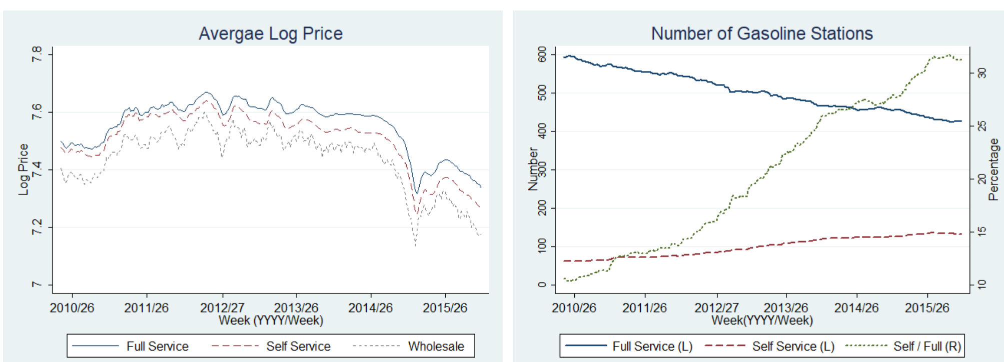
[10]There are two possible reasons for this. One reason is that modern cars requires less services so demand for full services falls. Another reason is due to intense price competition in the market. For example, the total number of gasoline stations more than doubled from 2002 to 2010. Moreover, the government implemented several polices to foster price competition in the retail gasoline market, such as publishing station-level prices online from 2008 and operating government-sponsored stations from 2012.

[11]I explain the dataset used in this study in the next section.



(a) May, 2010 (b) December, 2015

Figure 2.1: Locations of Gasoline Stations by Service Level
Note: Large dots are self service and small dots are full service



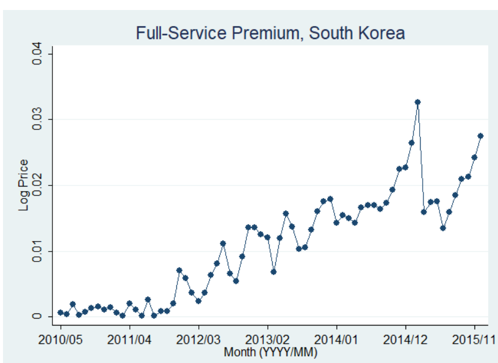
(a) Time Series of Market Prices (b) Time Series of the Number of Gasoline Stations

Figure 2.2: The Overview of Market Prices and Market Conversion, 2010-2015

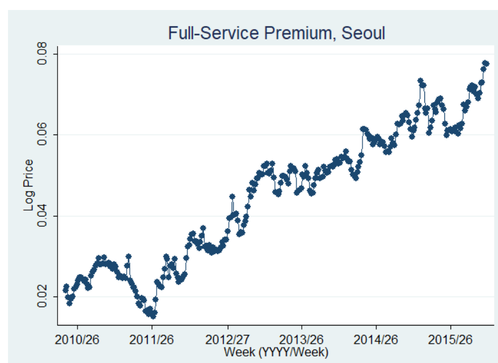
reduction in crude oil prices. In Figure 2.2-(b), the total number of stations falls by about 15% (from 658 to 559) over the period.[12]

2.2.2 The Evolving Full-Service Premium

My focus in this paper is on understanding the different pricing strategies of stations by service level during a period in which (formerly) full-service stations have adopted self-service technologies. Using price information on all gasoline stations in South Korea, I plot the time series of the log price difference between the two services from May 2010 to December 2015 in Figure 2.3-(a). The full-service premium (or self-service discount) started to increase from mid-2011. Narrowing down the market to the city of Seoul, there exists the similar and strong pattern in Figure 2.3-(b): the full-service premium was 2% on the first Wednesday of May 2010 and increased to 8% over the next five years. Herein, I focus on Seoul, for which I have daily price information on all stations, and where the adoption rate of self-service pumps is much higher than that in the remainder of the country.



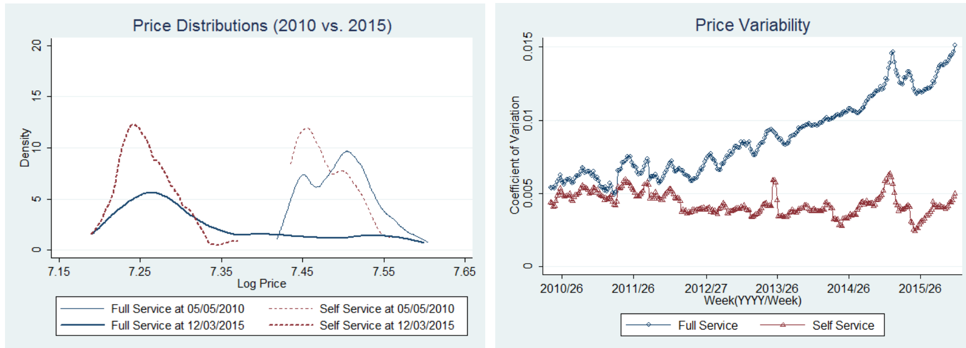
(a) South Korea, Monthly



(b) Seoul, Wedensday

Figure 2.3: The Evolving Full-Service Premium, 2010-2015

[12]The decline in the number of stations is a well-known phenomenon in retail gasoline industry, which is called "station rationalization" and observed while retail gasoline markets have matured. Eckert and West (2005) document a 63% decline in the number of stations in Canada over the period from 1972 to 2000 and a 47% fall in the U.S. over the same period.



(a) Price Distribution: 05/05/2010 vs. 12/03/2015

(b) The Time Series of Price Variability, by service level

Figure 2.4: The Different Patterns of Prices by Service Level, 2010-2015

Figure 2.4-(a) shows price distributions for full- and self-service stations on the first Wednesday of May 2010 and December 2015, and Figure 2.4-(b) presents the time series of coefficient of variation, also by service level. The shift of distributions shows a difference in price levels between two periods (recall Figure 2-(a)). The most important lesson from these figures is that the distribution of full-service prices has increasingly become *right-skewed* over time, while the variability of self-service prices has remained relatively stable. These features indicate the different evolution of pricing strategies of sellers by service level, and show that the increasing full-service premium is closely linked to stations that increasingly charge a premium for their *full-serve* gasoline. I discuss these stylized features more in the result section.

2.3 Data

My data are constructed from information available at the Oil Price Information Network (OPINET), a website operated by Korea National Oil Corporation. OPINET provides retail gasoline prices to the public at a daily frequency, and almost all prices are automatically collected based on transactions data.[13] OPINET also provides each station's non-financial information, such as address, brand, level of services, and name,

[13] If stations use manual updates, they must report their price within 24 hours of a change in price.

also at a daily frequency. I utilize a station-level panel dataset of all stations in Seoul that covers the period from May 2010 to December 2015. Figure 2.5 is one particular screenshot of OPINET to show how OPINET looks like.

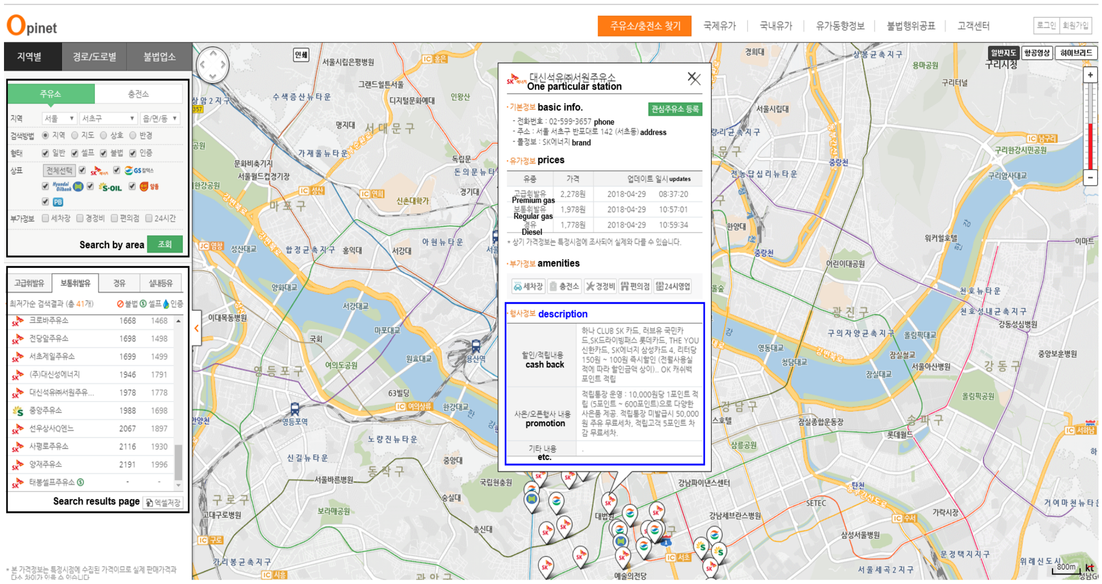


Figure 2.5: Web page of OPINET on April 29, 2018

I make two compromises in constructing the dataset. First, I use data from the Wednesday of each week because using all daily observations is burdensome to study the 5-year-period.[14] The loss of information is minimal: station prices change once every 11 days on average and Tuesday is the modal day for price changes in the full dataset.[15] Second, since entry and exit information are not explicitly available, I infer the information from the enforcement implemented on May 1, 2009 – all stations must report price at least once a week even in case of no price change. To be conservative, I drop the first year after the implementation of this rule and assume that a station remains open even if it does not report for up to four weeks, as long as its observed characteristics (e.g., name or service level) do not change.[16]

[14]Using one day of each week is also useful to reduce measurement errors on sellers' entry and exit. See my second compromise in this paragraph.

[15]See Figure 2.10 in Appendix 2.C for detailed information on the full dataset.

[16]When stations drop out of the sample for longer than one week, changes in station characteristics such as service level are often observed. In this case, I assume that these stations temporally closed

Using the address information, I geocode every station in my dataset and measure each station's competitive condition based on two conventional metrics. I first count the number of full- and self-service competitors within 0.5-mile and 1.0-mile radii, denoted $\text{Num}^{\text{FS0.5}}$, $\text{Num}^{\text{SS0.5}}$, Num^{FS1}, and Num^{SS1}, respectively. I also calculate the great-circle distance to the nearest stations of each service level, denoted Dist^{FS1} and Dist^{SS1}, respectively. In some regressions, I include the distance to the second-nearest stations, also by service level. Table 2.1 summarizes these variables. The average station has 2 full-service competitors and 0.4 self-service competitors within 0.5 miles; is located 0.3 miles from the nearest full-service competitors and 0.9 miles from the nearest self-service competitor.

Table 2.1: Summary Statistics: Numbers and Distances of Competitors [175,940 Obs.]

	Mean	SD	Min	Max	Mean	SD	Min	Max
Number of	**Full-Service Competitors**				**Self-Service Competitors**			
Within 0.5 mile (#)	1.98	1.54	0.00	10.0	0.41	0.67	0.00	4.0
Within 1.0 mile (#)	7.15	3.44	0.00	21.0	1.33	1.32	0.00	7.0
Distance to	**Full-Service Competitors**				**Self-Service Competitors**			
1st nearest (mi)	0.32	0.22	0.01	2.70	0.89	0.63	0.02	3.50
2nd nearest (mi)	0.50	0.24	0.03	2.71	1.31	0.70	0.12	4.12

Note: Average across all stations in all time periods.

I include three time-varying variables to help isolate competitive effects on price in my analyses. Table 2.2 summarizes the variables. First, I include indicators for station brands, which affects demand; brands are mostly time invariant but can change. There are five brands operating in this market. Second, I create the share of stations within 1.5 miles that have the same brand as each station in time, to control for the possibility that a dealer may operate multiple stations under the same brand in a local area.[17] Third, I also generate an indicator of whether a station sells regular gasoline only or premium gasoline together because the number of product lines affects optimal

to make that change. My compromise applies to 80 instances of a total of 175,940 observations. See Figure 2.11 in Appendix 2.C for details.

[17]Station brand is not the same as joint ownership, but brand share is often used as a proxy for an ownership in retail gasoline. As in Lewis (2008) and Chandra and Tappata (2011) that study retail gasoline pricing, I also control for it in my estimation.

pricing.[18]

I include some control variables only used for analysis of a discrete choice of full-service stations (e.g., decision to exit or convert to self service). Specifically, I have station amenities, such as presence of a convenience store and/or a carwash facility, available at OPINET. I control for monthly household income and the registration number of vehicles at the district level, both of which come from Seoul Statistics, a database operated by the Seoul government.[19] The monthly household income is based on survey data and the number of vehicles registered is from administrative data. These variables are observed only in May 2010 and May 2011, and the summary statistics on these variables are also shown in lower panel of Table 2.2.

Table 2.2: Summary Statistics: Control Variables [175,940 Obs.]

Variable	Description	Mean	SD	Min	Max
P	Price of gasoline (unit: KRW/liter)	1904.7	206.7	1317	2490
lnP	Log price of gasoline	7.54	0.11	7.18	7.82
Full	Station offering full-serve gasoline	0.83	0.37	0	1
SK	Station brand: SK Energy	0.36	0.48	0	1
GS	Station brand: GS Caltex	0.25	0.43	0	1
SO	Station brand: S-Oil	0.11	0.31	0	1
HD	Station brand: Hyundai Oilbank	0.12	0.33	0	1
AD	Station brand: Alddle	0.02	0.13	0	1
Unbranded	Station brand: Unbranded	0.03	0.17	0	1
Brand Share	Share of same-brand stations within 1.5 miles	0.27	0.16	0	1
Multi	Station selling regular and premium gasoline	0.33	0.47	0	1
Store[a]	Station having a convenience store	0.09	0.28	0	1
Carwash[a]	Station having an automatic carwash equipment	0.67	0.46	0	1
Repair[a]	Station having a auto-repair facility	0.26	0.44	0	1
Income[b]	Household monthly income in district (unit: KRW million)	3.69	0.43	2.72	4.49
Car[b]	Number of vehicles in district (unit: thousand)	104.0	41.8	37.2	193.65
Stations	Number of gasoline stations in Seoul	602.8	25.4	558	658

Note: Average across all stations in all time periods, except for Income and Car

[a] Observations at the station level in two days; the first Wednesday of May 2010 and 2011

[b] Observations at the district level in two months; May 2010 and May 2011

Lastly, I have cross-sectional information on each station's bundling and price, collected on May 17, 2017. At no charge, stations can publicize their special promotions in OPINET and consumers can see it through the OPINET website or smartphone

[18]For stations selling premium gasoline, there is another transaction price reported as in "premium gasoline." Following the construction of stations' entry and exit, I also assume that a station that does not report premium gasoline prices for four weeks has stopped selling premium gasoline.

[19]The city of Seoul has 25 districts, each with 30 gas stations on average.

application. Of the 539 stations operating in Seoul in May 2017, 117 stations advertised a special promotion, including free coffee or carwash. I assume that stations with no advertisements did not offer any promotions. Unfortunately, OPINET does not provide information on promotions in the past.

2.4 Results

2.4.1 Full-Service Stations' Decision

When self-service technology is introduced in the gasoline market, full-service stations must decide to either close, convert to self service, or continue to sell full-serve gasoline. In this section, I examine full-service stations' decisions in response to their local competitive environments. Based on entry and exit information, I identify "permanent exit" for full-service stations when they stopped reporting prices and did not return by the end of the sample period; "conversion to SS" as those that stopped reporting full-service prices but later reported self-service prices in the same location; and "FS continuation" when they neither closed nor converted during my study period.[20] The top panel of Table 2.3 provides the number of stations corresponding to each situation during my study period.

Table 2.3: Full-Service Stations' Decisions during the Sample Period.

	Permanent Exit	Conversion to SS	FS continuation
# of stations in the market[a]		$593 \rightarrow 426$	
# of instances	122	65	406
Avg. Num^{SS1} on the first day	0.93	1.02	0.72
Avg. Num^{FS1} on the first day	9.20	8.77	8.63

[a] The total number is from one on the first day to one on the last day of the sample. $593-122-65<426$; the difference is the number of full-service stations newly entered during the period.

I specify a multinomial logistic model of full-service station i's decision j to be

[20]In general, full-service stations take time to convert to self service. The average number of weeks during which a station is closed for conversion is 12; the median is 4; the minimum is zero; the maximum is 88.

correlated with the number of local competitors and some other covariates. The multinomial logistic model allows me to simultaneously consider three options: continuing as full service, switching to self service, and station exit. The coefficients of interest are β and γ in Equation (2.1):

$$log(\frac{\pi_{ij}}{\pi_{iJ}}) = \alpha_j + \beta_j Num_i^{SS1} + \gamma_j Num_i^{FS1} + \zeta_j \boldsymbol{Z}_i + \varepsilon_{ij} \tag{2.1}$$

where $j = \{$"permanent exit" or "conversion to SS"$\}$ and $J =$ "FS continuation"

where $\frac{\pi_{ij}}{\pi_{iJ}}$ is the odds that full-service station i falls in category j as opposed to the baseline outcome (J="FS continuation"). Num_i^{SS1} and Num_i^{FS1} are the number of self- and full-service competitors within a one-mile radius of full-service station i, respectively.[21] $\boldsymbol{Z}_i$ is a vector of all covariates described in the data section (Table 2.2). I allow the standard errors to be correlated within district in that the decisions of stations are possibly correlated with their nearby competitors.

Equation (2.1) is a cross-sectional specification that uses the first day of the sample. A panel regression or hazard model of self-service adoption that uses information on the timing of switching would potentially use the available information more efficiently. In my particular setting, however, the endogeneity of the number of local competitors and their service levels worsens if I add the time domain in the specification; today's exit and switch decisions of stations are interdependent of their local competitive conditions in the past. That is, the error structure of a panel regression or hazard model becomes hard to specify in later time periods. Because of that concern, I select a cross-sectional regression as my main specification.[22]

Table 2.4 presents relative-probability ratios for each outcome relative to "FS continuation", with two different days. I start by showing results with the first day of

[21] The total number of competitors does not need to be included in the specification because every station is either full service or self service in this market.

[22] A multinomial logistic regression with panel data does not make a big different to the results of the cross-sectional specification. I do not show results of panel data in this paper.

Table 2.4: Explaining the Relative Probability to FS Continuation

	(1) May 05, 2010	(2) May 04, 2011
Outcome: Permanent Exit		
Num^{SS1}	1.293**	1.286***
	(0.145)	(0.125)
Num^{FS1}	1.047**	1.073***
	(0.023)	(0.026)
Outcome: Conversion to SS		
Num^{SS1}	1.413*	1.345**
	(0.267)	(0.178)
Num^{FS1}	1.021	1.048
	(0.032)	(0.033)
Control variables	Y	Y
Obs	591	546

The coefficients are the relative probability of one outcome to the base outcome, FS continuation.
The coefficients on control variables are generally insignificant.
Robust standard errors in parentheses, clustered by district.
* p<10%; ** p<5%; *** p<1%.

my sample (May 5, 2010) in Column (1). For each additional self-service competitor nearby, the relative probability of exiting is 29% higher and relative probability of converting is 41% higher, respectively, than the probability of remaining full service. Consistent with Figure 1, this finding implies the presence of strategic complementarities in service in this market. Two possible explanations for this phenomenon are learning by consumers, and learning by sellers. In the first case, in the early period of the market transformation, consumers need to learn how to use self-service pumps. In the second case, stations may choose to learn about demand by observing their competitors' success or failure before making the transition. Switching carries a huge financial burden to sellers, so a rash decision to switch leads to serious damages to their future business.[23]

[23] The average total cost of switching is approximately 500,000 USD, including the replacement of gas pumps and the installation of a new payment system at pumps (Kang, 2010). In this context, the selection of a self-service format may be also correlated with financial constraints. One official in the retail gasoline industry reported that lessee-dealer stations had initially received financial support from their brand contractor to encourage them to switch to self service (Kwon, 2010).

The coefficient on Num^{FS1} shows how competition among full-service stations is correlated with their decisions. Having one more full-service station nearby increases the probability of closing by 5% compared to the probability of remaining full service (at a significance level of 0.05), although increased full-service competition does not statistically predict their decision to switch to self service.

In Column (2), I estimate Equation (2.1) again, this time using data from May 4, 2011 (one year after the first day of my sample). The results generally mirror those for the first day of the sample and the statistical significance of estimates becomes stronger.[24]

Overall, my results support the notion that entry of low-price competitors drives high-cost marginal stations out of a market. Full-service stations that neither exited nor converted had faced less intense competition (bottom panel of Table 2.3). Nevertheless, the less competitive environments may not be sufficient for them to successfully survive against low-price competitors for long periods. The next section focuses on the strategic choices of each type of station in terms of pricing.

2.4.2 Price Competition

The emergence of self-service stations can affect market price through two channels: self-service stations can charge a low price due to their lower labor costs (direct effect) and drive down competitors' price if the market is imperfectly competitive (indirect effect). In this section, I quantitatively examine both channels using a difference-in-

[24]The magnitude of estimates or even their signs can differ by degree of market transformation because the progress of the transition is not constant. For example, when the number of self-service stations is in equilibrium, entry of self-service stations should no longer predict conversion from full service to self service.

difference technique as follows:

$$\ln P_{it} = \beta_1 \ln Dist_{it}^{SS1} + \beta_2(\ln Dist_{it}^{SS1} * Full_{it}) + \beta_3 \ln Dist_{it}^{FS1} + \beta_4(\ln Dist_{it}^{FS1} * Full_{it})$$
$$+ \theta Full_{it} + \phi(Full_{it} * Trend_t) + \zeta \boldsymbol{X}_{it} + \delta_i + \mu_t + \varepsilon_{it}. \quad (2.2)$$

P_{it} is the price of station i at time t. $Dist_{it}^{SS1}$ and $Dist_{it}^{FS1}$ are the great-circle distance to the nearest self- and full-service competitors of station i in time t, respectively. $Full_{it}$ is an indicator of whether station i sells full-serve gasoline at time t. $Trend_t$ is an incremental number that increases by 1/294 in each week, starting from 1/294 on the first day to 294/294 on the last day of my sample; so the coefficient on $Full_{it} * Trend_t$ captures the increase in the full-service premium, with appears to grow (Figure 2.3) over the sample period. $\boldsymbol{X}_{it}$ is a vector of control variables described in Table 3.1 which primarily helps isolate competitive effects on price. δ_i is a station fixed effect that captures time-invariant characteristics correlated with gasoline prices, such as location and traffic. μ_t is a time fixed effect that captures changes in average prices, mostly driven by wholesale gasoline prices. The error term ε is clustered by station and is robust to interdependent pricing of sellers.

I show estimation results of Equation (2.2) in Columns (1)-(2) in Table 2.5. Starting with Column (1), which omits the trend of the full-service premium, the coefficient on Full shows that *ceteris paribus*, full-service stations charge 5% more per gallon than self-service stations on average during my sample period. Including the trend in the specification in Column (2), the full-service premium more than triples during the period; it increases from 1.7% in the first day to 6.7% by the end of my sample (i.e., when **Trend**=1).

Next, I show indirect effects of self-service stations on price (i.e., competitive effects) in Column (1). When the distance to the nearest self-service station doubles, station-level prices increase by 0.7%, regardless of station's service level.[25] The esti-

[25] 0.7% is calculated by 0.0096*ln(2)*100. The coefficient on the interaction terms shows differential

Table 2.5: Estimation Results, Price Competitive Effects

	(1) 1st-nearest	(2) 1st-nearest +Full Trend	(3) 2nd-nearest +Full Trend	(4) within 0.5-mi	(5) within 0.5-mi +Full Trend	(6) within 1-mi +Full Trend
Full	0.0521***	0.0168	0.0175**	0.0609***	0.0231**	0.0327**
	(0.0103)	(0.0104)	(0.0089)	(0.0094)	(0.0100)	(0.0135)
Full*Trend		0.0495***	0.0491***		0.0486***	0.0487***
		(0.0062)	(0.0065)		(0.0063)	(0.0012)
$Dist^{SS}$	0.0096***	0.0027	0.0130***			
	(0.0034)	(0.0032)	(0.0048)			
$Dist^{SS}$*Full	-0.0048	0.0035	0.0032			
	(0.0042)	(0.0042)	(0.0060)			
$Dist^{FS}$	-0.0049	0.0008	0.0026			
	(0.0043)	(0.0041)	(0.0059)			
$Dist^{FS}$*Full	0.0039	-0.0025	-0.0004			
	(0.0043)	(0.0040)	(0.0063)			
Num^{SS}				-0.0087***	-0.0042	-0.0048**
				(0.0034)	(0.0033)	(0.0021)
Num^{SS}*Full				-0.0011	-0.0064	-0.0037
				(0.0041)	(0.0042)	(0.0025)
Num^{FS}				0.0040*	-0.0002	-0.0002
				(0.0025)	(0.0026)	(0.0012)
Num^{FS}*Full				-0.0051*	-0.0007	-0.0010
				(0.0025)	(0.0026)	(0.0012)
SK	0.0308***	0.0332***	0.0324***	0.0300***	0.0322***	0.0326***
	(0.0086)	(0.0088)	(0.0086)	(0.0085)	(0.0087)	(0.0088)
GS	0.0220***	0.0246***	0.238***	0.0222***	0.0246***	0.0240***
	(0.0080)	(0.0080)	(0.0078)	(0.0078)	(0.0079)	(0.0078)
SO	0.0106	0.0117	0.0108	0.0096	0.0109	0.0116
	(0.0118)	(0.0121)	(0.0115)	(0.0117)	(0.0120)	(0.0119)
HD	0.0105	0.0133	0.0126	0.0117	0.0141	0.0126
	(0.0112)	(0.0114)	(0.0115)	(0.0110)	(0.0119)	(0.0116)
AD	-0.0186	-0.0211	-0.0214	-0.0221*	-0.0225*	-0.0230*
	(0.0138)	(0.0137)	(0.0139)	(0.0127)	(0.0135)	(0.0138)
Brand Share	0.0361**	0.0286**	0.0287**	0.0349**	0.0289**	0.0270*
	(0.0148)	(0.0150)	(0.0149)	(0.0149)	(0.0148)	(0.0148)
Multi	0.0213***	0.0207***	0.0205***	0.0206***	0.0204***	0.0208***
	(0.0067)	(0.0067)	(0.0067)	(0.0068)	(0.0067)	(0.0066)
Obs	175940	175940	175940	175940	175940	175940
Adj R^2	0.900	0.900	0.901	0.900	0.900	0.901

All specifications also include station and time fixed effects.

LHS variable is station log price.

RHS variables are the distances to nearby stations in Columns (1)-(3); the numbers of nearby stations in Columns (4)-(6).

Unbranded is omitted.

Robust standard errors in parentheses, clustered by station.

* p<10%; ** p<5%; *** p<1%.

mate loses its significance after controlling for the trend of the full-service premium in Column (2). This may be because the trend variable of the full-service premium would absorb most of the variation in the distances, resulting in lack of estimation power of the distances on price level in my data.

In Column (3), I replace the distances to the nearest stations by the distance to the second-nearest station. The results are generally similar and the coefficient on Dist^{SS} becomes significant at the 1% level.[26] Overall, stations' prices are more elastic with respect to competition when their local competitors are self service than when they are full service. One possible explanation for this is that self-service customers who purchase only gasoline are more likely to be more price sensitive than full-service customers who buy a bundle of both gasoline and service.

The coefficients in Table 2.5 should be interpreted with caution as service level and location may be endogenous. For example, self-service stations may prefer to locate in markets with many full-service stations to attract more consumers through a relatively cheap price. In this case, the coefficient on Dist^{SS} would be biased downward in absolute value. It is also possible that full-service stations may prefer to locate in high-income areas, where they can charge a high premium to less-price-sensitive consumers, causing the coefficient on Dist^{FS} to be biased downward in absolute value. Although station fixed effects do not eliminate these concerns, they greatly mitigate them since seller characteristics are generally fixed for long periods in the retail gasoline industry.

To check the robustness of the results above, I estimate competitive effects this time using the alternative measure of competition: the number of stations within half a mile. This specification allows for gasoline stations to engage in competition with

sensitivity to the degree of competition when a station is full sevice.

[26]This result is also robust to including both distances to the nearest and the second-nearest stations together in the specification

more than just one or two stations:

$$\ln P_{it} = \beta_1 Num_{it}^{SS0.5} + \beta_2(Num_{it}^{SS0.5} * Full_{it}) + \beta_3 Num_{it}^{FS0.5} + \beta_4(Num_{it}^{FS0.5} * Full_{it}) + \theta Full_{it} + \phi(Full_{it} * Trend_t) + \zeta \boldsymbol{X}_{it} + \delta_i + \mu_t + \varepsilon_{it} \tag{2.3}$$

where $Num_{it}^{SS0.5}$ and $Num_{it}^{FS0.5}$ are the number of self- and full-service stations within a half-mile radius of station i in time t, respectively. Here, I use the natural numbers of stations rather than taking logs because some gasoline stations do not have competitors within half a mile.[27]

The estimation results are shown in Columns (4)-(6) of Table 2.5, consistent with the distance-specification results. Starting with Column (4), a one-station increase in the number of self-service stations within half a mile is associated with a 0.9% reduction in station prices, and similarly, the estimate becomes insignificant after controlling for the trend of the full-service premium in Column (5). Replacing the numbers by those of stations within a 1-mile radius, I find in Column (6) that self-service stations drive down their local competitors' prices by 0.5%, which is statistically significant. These results support my earlier finding that price competition is most intense in local markets with many self-service stations.

Lastly, the coefficients on control variables are shown in the bottom of Table 2.5; the estimates are consistent across all specifications. The coefficient on brand share is positive, implying less intense competition within brands. Other things equal, stations that sell both premium and regular gasoline charge 2% more for their regular gasoline than those selling only regular gasoline. This positive sign was a focus of Lee et al. (2015). More interestingly, they also observe an increasing price gap between "premium" and "regular" stations.[28] Although it is beyond the scope of my paper

[27] I estimate a log-log specification using a five-mile radius; all stations have at least one competitor within five miles. The results are qualitatively similar.

[28] They define "premium" stations as ones selling both premium and regular gasoline and "regular" stations as ones selling only regular gasoline.

to identify a link between a full-service premium and a product-grade premium, the similarity in results suggests that the different pricing strategies of full- vs. self-service stations might have a common basis with those of premium vs. regular stations.[29]

To summarize, I find that self-service stations affect the average price of gasoline both directly and indirectly, whereas full-service stations charge higher prices and have little indirect impact on price. Yet, this finding does not explain how full-service stations have retained their business while they do not seem to be competing on price. I attempt to shed light on this question in the following section.

2.4.3 Product Differentiation

Recall, from Figure 2.4, that the distribution of full-service prices displays an increasing right-tail over time, meaning a growing share of stations charge a very high premium. In this section, I mainly focus on the striking increases in the variability of full-serivce prices, and apply insights from models of product differentiation to interpret the unique feature of full-service stations' pricing.[30]

A broad literature has developed on the relationship between price and competition along with product differentiation, and many of the literature have demonstrated that competition is less intense if products are differentiated among sellers.[31] Consistent with existing studies, I find that full-service stations that charge an extra premium tend to differentiate their product on one or more dimensions.[32]

[29]I estimate the relationship between the level of services and the type of stations using a logit regression where other covariates are controlled. The estimate is insignificant in my setting.

[30]The story of price competition alone is not enough to explain the trend of the full-service premium in my setting. See Appendix 2.B for discussion in detail.

[31]Particularly in the context of retail gasoline, Netz and Taylor (2002), Lewis (2008), and Kim (2017) examine pricing along with market segmentation – e.g., Netz and Taylor on a spatial dimension, Lewis for a brand dimension, and Kim for a service dimension.

[32]Matsa (2011) examines the relationship between competition and product quality in the U.S. supermarket industry, and finds that competition increases sellers' incentive to provide quality and the largest improvement of quality arises in lower income areas of the markets affected by Wal-Mart's entry.

Descriptive Evidence

I start by providing some descriptive evidence that product differentiation increases largely in full-service stations. Moreover, product differentiation takes far more subtle and varied forms than simply location and brand often considered in most retail-gas-pricing studies.

Some gasoline stations have added another type of business, such as a fast-food restaurant, dry cleaner, or nail salon to their station. A few gas stations were completely rebuilt as a complex-mall structure with multiple floors, rented to other businesses.[33] I display four examples of these stations in Figure 2.6. Station (a), located in Yeoeuido where many businesses and government agencies have offices, operates a dry cleaner within the station. Station (b) includes a coffee shop along with the gasoline pumps. Station (c) advertises itself as "a female-friendly station" equipped with a powder room and nail-care shop and decorated in purple. Station (d) has a multiplex structure; the entrance to the gas station is the right wing of the building, on which the SK logo is shown.

The average price at these four stations is 12% higher than the market average for full-service gasoline on the last day of my sample. With the exception of Station (b) whose price is 4% lower than its nearby stations, in alphabetical order, they also charge 2%, 17%, and 25% more than their neighborhood within one mile, respectively. This evidence is consistent with earlier studies showing that firms enjoy less competitive pressure when their product is differentiated. To the best of my knowledge, these business formats did not exist in the Seoul gasoline market before the self-service model took off in 2008.[34]

I observe further evidence of product differentiation from advertisements of each

[33] Only selected types of businesses are allowed to operate within gas stations, according to the enforcement rule under Safety Control of Dangerous Substances Act (Chapter 37 of Title 3: Gas-Station Location, Structure, and Facility).

[34] The four stations reopened on (a) 12/2012; (b) 11/2011; (c) 11/2008; (d) 12/2012.

(a) Dry Cleaner

(b) Coffee Shop

(c) Female-Friendly; Powder Room & Nail Care

(d) Complex Mall Station

Figure 2.6: Examples of Product Differentiation within Gas Stations in Seoul
Source: (a) http://www.korea-news.com/news/articleView.html?idxno=46986; (b) http://cafe.daum.net/oilproject; (c) http://blog.skenergy.com/4; (d) http://blog.skenergy.com/176

Table 2.6: Price Comparisons of Stations by Bundling Status

	Full Service		Self Service	
May 17, 2017	# Stations	Avg. Log Price	# Stations	Avg. Log Price
Bundled[a]	55	7.444	11	7.321
Not Bundled[b]	337	7.360	136	7.298
Difference		0.085		0.023
P-value		(0.000)		(0.073)
All	392	7.372	147	7.300
Difference		0.072		
P-value		(0.000)		

[a] stations that dispense "free" bundled products when the purchase amount of gasoline is 50,000 KRW or more.

[b] stations that do not dispense "free" bundled products.

station, collected on May 17, 2017, approximately 17 months after the end of my panel dataset. Some gasoline stations provide additional products along with gasoline, such as coffee, tissue, or even vacation packages. Some of these stations dispense "free" bundled products, as long as consumers purchase gasoline (50,000 KRW or more).[35] Table 2.6 presents the number of unique stations and their average price by service level, and the detailed description on bundled ("free") products and their locations are shown in Table 2.10 and Figure 2.8 and 2.9 in Appendix 2.A. The total number of stations operating on May 2017 was 539, and 55 full-service stations and 11 self-service stations provide bundled products, respectively.

I draw two conclusions from Table 2.6. First, full-service stations are more than twice as likely to provide bundled offers as self-service stations, suggesting that full-service stations are more likely to differentiate their product on this subtle dimension. Second, prices at stations that offer extra products are higher than prices elsewhere; this pattern is much stronger for full-service stations than for self-service stations.[36] Taken together, the results support the argument that differentiation – offering a va-

[35] This is the amount that average consumers pay at one time fuel in Korea. See the survey from Korean Transportation Database (KTDB, 2013). 50,000 KRW, or approximately 50 USD, is equivalent to the cost of eight gallons of a gas.

[36] I find a qualitatively similar pattern when I estimate a cross-sectional regression of station prices on a dummy of whether stations have bundled offers, controlling for often observed station characteristics.

riety of bundled products and services – enables stations to charge an extra premium, mostly to less-price-sensitive customers by engaging in less-intense competition on gasoline price.[37]

Evidence from Price Data

One may think that the descriptive evidence is selective and the bundled products are evidence observed after the end of my study period. In this section, I use price data to strengthen my argument above. I first identify stations in my main dataset that offer bundled products in May 2017. After that, I compare prices of stations that bundle products in May 2017 and those that do not bundle (this is because the timing of the bundling is unknown).

The intuition of this practice is that if gasoline stations were differentiated in 2010 or 2015 as much as they were in May 2017, they would have charged a higher price than others. Among 539 stations in Seoul on May 17, 2017, I confirm that 536 stations are matched with my main dataset (i.e., only three stations entered in a new location after my study period), and restrict the matched stations by removing 22 stations that had changed service level as of May 2017. For each station, I take prices on the first and the last days of my sample, and then compare the prices of stations that later bundle products and those that do not.

At the top panel of Table 2.7, I show the log prices on the first day of my sample (May 5, 2010), for stations that later offered bundled products and those that did not. Focusing on full-service stations, the price gap between stations that later bundled products and those that did not was 1.5%, which is much smaller than the difference on May 17, 2017 (8.5%). In addition, there was essentially no difference in self-service prices between stations that later bundled than those that did not. The results support

[37] Selling bundled product may not be profitable if an extra premium at bundling stations is merely a cost of additional services. Regardless of that, it is still valid that some stations stay in business by doing differentiation their product on additional dimension.

Table 2.7: Price Comparisons of Stations by Later Bundling Status

	Full Service		Self Service	
May 05, 2010	# Stations	Avg. Log Price	# Stations	Avg. Log Price
Bundled in 2017[a]	48	7.514	2	7.503
Not bundled in 2017[b]	290	7.499	50	7.475
Difference		0.015		0.028
P-value		(0.011)		(0.268)
All	338	7.501	52	7.476
Difference			0.025	
P-value			(0.000)	

	Full Service		Self Service	
Dec 16, 2015	# Stations	Avg. Log Price	# Stations	Avg. Log Price
Bundled in 2017[a]	55	7.407	7	7.282
Not bundled in 2017[b]	332	7.326	115	7.258
Difference		0.081		0.023
P-value		(0.000)		(0.156)
All	387	7.338	122	7.259
Difference			0.078	
P-value			(0.000)	

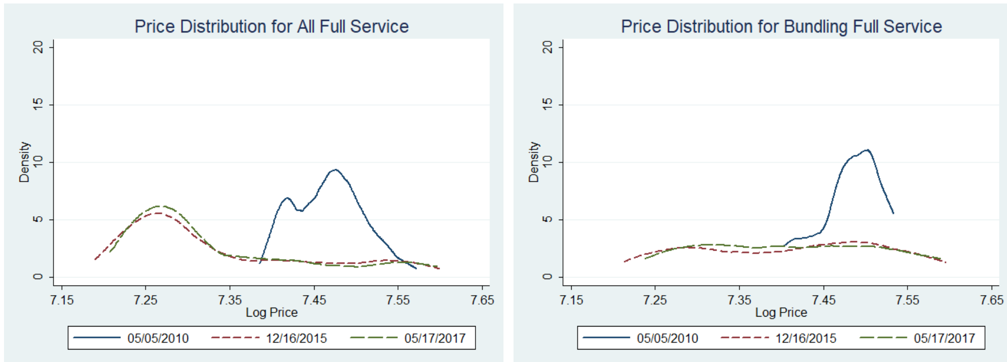
[a] stations that provide bundled their products in May 2017

[b] stations that do not provide bundled products in May 2017

the notion that the stations were much less differentiated on this dimension in May 2010 than in May 2017, and this result is strongest for full-service stations.

In the bottom panel, I show the result of the same analysis above, this time using the last day of my sample (December 16, 2015). The premium at the bundling (or bundling-anticipated) stations was 8.1%, very similar to the actual premium on May 17, 2017 (8.5%). Figure 2.7 shows the price distributions of all full-service stations (panel a) and the full-service stations that later bundle (panel b) at three points in time.[38] From this analysis, I infer that some full-service stations that charge high premiums for their gasoline maintain market share by providing bundled products, and that bundling is closely related to the increasing variability of full-service prices in this market.

[38]These prices are adjusted for inflation to real values for December 16, 2015.



(a) All Full-Service Stations (b) Full-Service Stations that Later Bundle

Figure 2.7: The Price Distributions of Full-Service Stations on the Selected Days
Note: Bundling information is as of on May 17, 2017

2.4.4 Price Search and Relative Price's Change

I continue to discuss product re-positioning, but in this section, I rely on insights from models of search to supplement the shortcoming of explicit information on subtle differentiation. Models of search assume a certain fraction of consumers who are unaware of which sellers charge high or low prices, which allows for price dispersion in a market for one homogeneous product and variation in prices of sellers from one period to the next.[39]

Given the predictions of the models, I hypothesize a testable implication that if a product is homogeneous across stations, their relative prices should not move systematically; otherwise, consumers will be informed about stations' prices. To test for this, for each week, I create price rankings of stations and group them by octile, and then calculate the transition probabilities to determine relative price stability.[40]

Table 2.8 shows results of the transition probabilities using all stations in my sample. Each row represents an initial octile and each column is a final octile. The probabilities along the diagonal line show that the stability of price rankings is higher for higher-end stations than for lower-end stations with the exception of the first

[39] Baye et al. (2016) provide a detailed overview of both theoretical and empirical research on search and price dispersion.

[40] My results are qualitatively similar when I group rankings by quartile (1/4) or hexadecile (1/16).

Table 2.8: Price Octile Transition Matrix [174,648 Obs.]

		T+1								
		1	2	3	4	5	6	7	8	Total
Lowest 12.5%	1	77.35	20.03	2.14	0.35	0.09	0.04	0.00	0.00	100.0
	2	18.32	59.12	20.58	1.69	0.21	0.07	0.00	0.02	100.0
	3	2.73	18.19	59.71	18.25	0.97	0.11	0.02	0.01	100.0
T	4	0.67	2.38	15.30	66.23	14.92	0.45	0.04	0.01	100.0
	5	0.28	0.43	1.74	12.83	73.93	10.56	0.20	0.03	100.0
	6	0.15	0.13	0.20	0.68	9.37	81.04	8.37	0.07	100.0
	7	0.04	0.06	0.08	0.17	0.37	7.43	86.53	5.32	100.0
Highest 12.5%	8	0.06	0.05	0.05	0.05	0.07	0.28	4.87	94.58	100.0
Total		12.37	12.56	12.48	12.56	12.46	12.45	12.52	12.60	100.0

Rows show an initial octile and columns are a final octile. This matrix represents one-week transition probabilities of price rankings of stations.

octile. This result suggests that higher-end stations keep price positions much more stable than other stations. Specifically, the probability that the eighth-octile stations stay in the same octile in the next week is 95%, whereas the average probability over all octiles is 75%. The finding is not consistent with the prediction of search models I described above. Instead, I infer that a product of higher-end stations is more likely to be differentiated from a product of other stations.

The high stability of higher-priced stations could occur if their consumers were unaware of other sellers' prices. In my setting, however, there are many examples this would appear unlikely. As one example, there were seven stations in the Gangnam downtown area, all within a 1-mile area, where the highest-priced station charged 20% more than the lowest-priced station. The large price gap remained for multiple years despite their similar (observed) amenities. It is not reasonable to believe that higher-end stations' consumers are entirely ignorant of large price differences across stations in this market.

To strengthen my inference above, I generalize the results of the transition probabilities using a regression framework that allows me to control for station characteristics. I specify a linear probability model of stability, including a full-service indicator, all control variables described in the data section, and station and time fixed effects:

$$Stable_{it} = \beta Octile_{it-1} + \theta Full_{it} + \gamma \boldsymbol{X}_{it} + \delta_i + \mu_t + \varepsilon_{it} \tag{2.4}$$

$$\text{where} \quad Stable_{it} = \begin{cases} 1 & \text{if } Octile_{it-1} = Octile_{it}; \\ 0 & \text{if } Octile_{it-1} \neq Octile_{it}. \end{cases}$$

where $Stable_{it}$ is a dichotomous variable that equals 1 if station's octile at week t-1 is the same as in week t, and 0 if it is not.[41] $Octile_{it-1}$ is a categorical variable based on station i's price ranking from 1 to 8 (the lowest 12.5% to the highest 12.5%) in t-1. The coefficient β picks up how relative price stability depends on a stations' initial position.

Table 2.9 shows estimation results of Equation (2.4).[42] Column (1) shows results controlling for station and time fixed effects. I find that a one-octile increase is associated with a three percentage point increase in the probability that the station's price remains in the same octile one week to the next, which confirms my observation that the price rankings of higher-priced stations are more stable than those of lower-priced stations. As discussed earlier, if the product were homogeneous across stations, this finding would show a breakdown of search models as the high-priced stations settle into fairly fixed prices during my study period relative to the lower-priced stations. This finding supports the notion that "gasoline" sold at high-priced stations is systematically different from "gasoline" sold at lower-priced stations.

The results are robust to adding more covariates. In Column (2), I add controls for a full-service indicator and the number of competitors within one mile. The estimated coefficient on Octile remains approximately 0.03. Full-service stations are four percentage points more likely to have prices in the same octile from one period to the next. Also, the positive coefficient on Num^{FS1} means that greater competition with

[41] Stability is undefined in the first week a station appears in the data and after a temporary exit.

[42] The results of a linear probability model are qualitatively similar to those of a logit regression. Here, I show the results a linear probability model for ease of interpretation, and present logit results in Table 2.11 in Appendix 2.A.

Table 2.9: Estimation Results: Relative Stability of Price Rankings

	(1) Baseline	(2) Controls	(3) By Service
Octile	0.0303***	0.0296***	0.0092*
	(0.0023)	(0.0022)	(0.0053)
Octile*Full			0.0251***
			(0.0055)
Full		0.0433***	-0.0390***
		(0.0164)	(0.0237)
Num^{SS1}		-0.0000	0.0003
		(0.0038)	(0.0037)
Num^{FS1}		0.0095***	0.0104***
		(0.0023)	(0.0023)
Control variables	N	Y	Y
Station and Time FE	Y	Y	Y
Obs	174648	174648	174648
Adj R^2	0.092	0.092	0.093
% predicted outside [0, 1]	1%	2%	3%

LHS variable is a dummy of whether station's octile at week *t*-1 is the same as in week *t*.
In Column (3), the coefficient on Full is positive whenever Octile >=2.
Robust standard errors in parentheses, clustered by station.
* p<10%; ** p<5%; *** p<1%.

full-service stations increases the stability of a station's pricing position, consistent with the earlier result that price competition is relatively less intense in local markets with many full-service stations. In Column (3), I add an interaction of octile and a full-service dummy. I find that the pattern – the higher priced stations display more stable relative price rankings – is much stronger for full-service stations than for self-service stations, as long as the price octile is greater than one.[43]

Full-service stations are unlikely to compete on price with self-service stations given that their marginal costs are higher than those of self-service stations. Together with the finding in Section 2.4.2 that self-service stations compete primarily for price-sensitive consumers, my results show that the strategic choices of stations evolve in different ways when the market is undergoing a massive restructuring, each type of station using its unique position to its advantage. Furthermore, it may be necessary that full-service stations compete through product differentiation because they represent a market where services differ in value across customers.

[43] The estimates in Column (3) of Table (2.9) imply that stability is greater for full-service stations when Octile is greater than one.

2.5 Concluding Remarks

This paper contributes to a literature on revisiting the evolution of price strategies in a market undergoing a significant transition, and the literature on the topic is growing thanks to increased accessibility to rich, micro-level data. Using a case study of the retail gasoline industry in Korea, I document a new stylized fact: the price gap between a full-service format and a self-service format has increased during the transformation of gasoline stations from full service to self service.

I show that the strategic options for incumbents differ by seller heterogeneity in service type: some stations close down their business while others adopt the self-service model. Intense competition on price is positively correlated with these choices. A third option is that stations differentiate their product on one or more dimensions, which enables sellers to raise their prices. This approach is an alternative that allows sellers to maintain their business by engaging in less-intense competition on price.

Grabowski and Vernon (1992) document empirical evidence similar to my paper: they study price competition in the pharmaceutical industry and find that the entry of generic drugs causes the price of major branded drugs to *increase*. Although Grabowski and Vernon's setting is very different from mine with regard to market structure, the pattern of sellers' pricing in the two papers appears similarly and is able to be well understood through insights from models of price competition with differentiated products. I think of it as the beauty of economic theory which helps us understand data.

2.6 Appendices to Chapter 2

2.6.1 Appendix 2.A: Supplementary Results

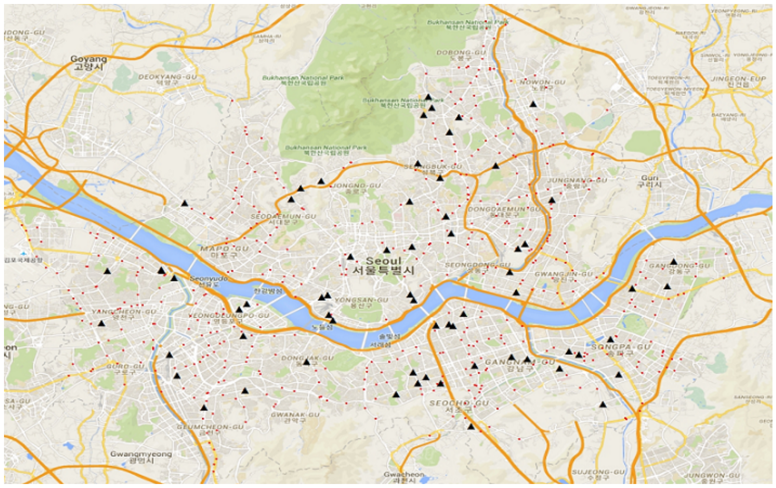


Figure 2.8: Locations of Stations with "Free" Offers on May 17, 2017

Table 2.10: Description of Bundled Goods and the Number of Stations on May 17, 2017

	# Full Service	# Self Service	Total #
Reward points[a]	29	4	33
Carwash	9	3	12
Carwash, if buying premium gas	1	0	1
Carwash & Coffee	1	0	1
Carwash & Coffee & Washer fluid	1	0	1
Coffee or Tea	3	1	4
Coffee & Washer fluid	1	0	1
Coffee & Facial tissue	2	0	2
Water	3	0	3
Facial tissue	1	1	2
Water or Facial Tissue	1	1	3
Car Inspection	2	0	2
Service for Diplomatic vehicles	1	0	1
Coffee, Soda, Noodle, Copy/Fax, TV, Lounge[b]	0	1	1
Total #: 539 (Full Service 392; Self Service 147)	55/392 (14%)	11/147(7%)	66/539(11%)

[a] Reward points can be redeemed as gasoline or station's own bundled products: the bundled products are various ranging from wiper blades to vacation packages.

[b] This station seems to serve largely tractor-trailer drivers.

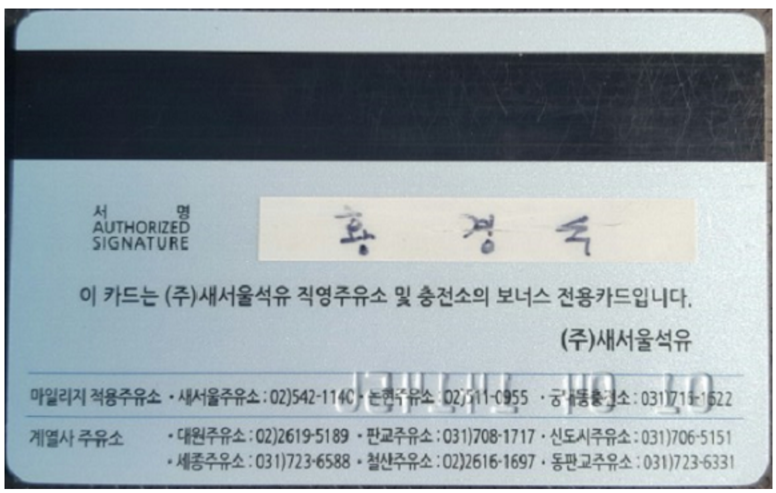
Table 2.11: Logistic Regressions: Stability of Price Rankings

	(1) Baseline	(2) Controls	(3) By Service
Octile	0.170***	0.164***	0.039***
	(0.005)	(0.005)	(0.011)
Octile*Full			0.157***
			(0.012)
Full		0.267***	-0.242***
		(0.015)	(0.057)
Num^{SS1}		-0.015	-0.012
		(0.012)	(0.012)
Num^{FS1}		0.051***	0.058***
		(0.008)	(0.008)
Control variables	N	Y	Y
Station and Time FE	Y	Y	Y
Obs	174648	174648	174648

Standard errors in parentheses.

* p<10%; ** p<5%; *** p<1%.

(a) Reward Point Card, Front



(b) Reward Point Card, Back

Figure 2.9: centering The Example of a Station-Specific Point Card
Note: It is not same as a brand-membership card.

2.6.2 Appendix 2.B: Robustness Checks

Table 2.12: Robustness Check on the Main Specification

	(1) Baseline	(2) Squared[a]	(3) Trimmed[b]	(4) Trimmed[b]
Full	0.0168***	0.0195***	0.0201***	0.0153*
	(0.0044)	(0.0059)	(0.0051)	(0.0093)
Full*Trend	0.0550***	0.0466***	0.0477***	0.0421***
	(0.0066)	(0.0057)	(0.0045)	(0.0049)
Competition[c]	N	Y	N	Y
Control variables	N	Y	N	Y
Station FE	N	Y	N	Y
Time FE	Y	Y	Y	Y
Obs	175940	175940	158289	158289
Adj R^2	0.687	0.900	0.748	0.915

[a] I add squared terms of the (log) distances in the main specification.
[b] Specification (4) and (5) drops 5% of top and bottom stations ranked by price, in each time by service level.
[c] The measure of competition is the distance to the nearest station.
Robust standard errors in parentheses, clustered by station.
*** p<1%.

Models of competition, including a monopolistic competitive model, generally predict that increased competition lowers prices.[44] If the gasoline market is segmented by service level, the increased number of self-service stations over my sample period, and the falling number of full-service stations, along would predict an increase in the full-service premium. Interestingly, the changes in the service composition of stations cannot explain the increasing full-service premium in my setting.

Column (1) of Table 2.12 shows that, without any controls, the coefficient on Full$*$Trend in Equation (2.2) is 0.055, consistent with Figure (2.3-b). Adding controls – including station fixed effect and distances or number of nearby stations – barely changes the coefficient (See Column (2) of Table 2.5). The trend of the price gap between the two services is largely unexplained by station brands, location, and more importantly, competitive conditions. This finding implies that there is another underlying principle behind the increasing premium of full service that monopolistic-competition models cannot explain.

[44]Barron et al. (2004) argue that the U.S. retail gasoline market resembles a setting of monopolistic competitive market, and Kim and Kim (2011) find similar evidence in the Korean market.

My main specification assumed a monotonic relationship between competitive conditions and price, but this assumption may not be justified in that the market has been in a process of the diffusion of self-service technologies.[45] For example, the initial entry of self-service stations may not be a threat to full-service stations if only few consumers know how to handle self-service pumps. When the number of self-service stations gets large enough, full-service stations may be perfectly differentiated from self-service stations; no competition across services happens. Adding squared terms of my measure of seller concentration, I allow for a non-monotonic relationship in price and competitive conditions of sellers. In Column (2) of Table 2.12, I confirm that the presence of nonmonotonicity does not explain the pattern of the full-service premium; the coefficient on Full∗Trend is still around 0.05.[46]

A Least Square procedure estimates a conditional *mean* function. One concern for this procedure is that if the increasing full-service premium resulted from only a few stations that charge extremely high (or low) prices, average effects of competitive conditions may not capture variation in the full-service premium. To check on this issue, I estimate the specification using a trimmed sample in which I drop observations that rank at the top and bottom 5% stations (about 60 stations) in each week, by service level.[47] The results are shown in Column (3)-(4).[48] The significance of the premium remains unchanged, although the fitted trend of the full-service premium is reduced in the restricted sample when compared to the full sample. This practice supports that the increasing full-service premium in which "average" stations involved cannot be well explained by changes in competitive environments.

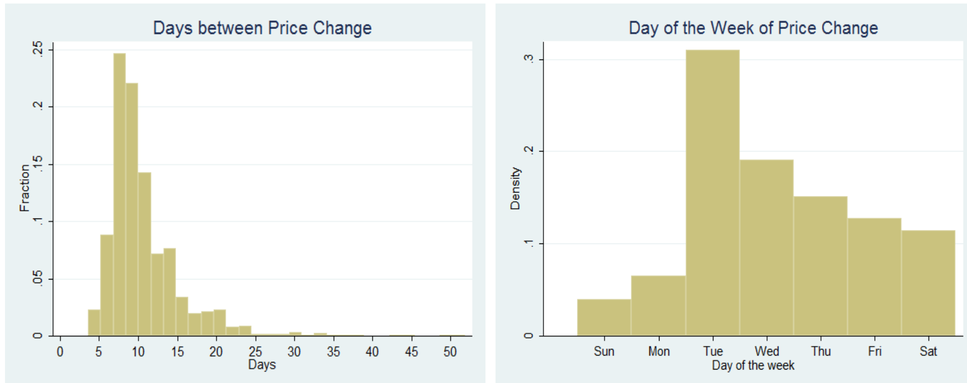
[45] It is a well-known phenomenon that the number of firms is nonmonotonic over the evolution of new industries (Graddy, 1990 and Jovanovic and MacDonald, 1994).

[46] I also estimate a regression by replacing the squared distances to the nearest station by those to the further located station, as well as the squared numbers, but these changes do not make a big difference to the coefficient on Full∗Trend.

[47] A quantile regression may be an alternative method to check this issue, but the quantile regression is not trivial to apply to longitudinal data with fixed effects, although some treatments have been developed.

[48] I also try to drop 1% and 3% outliers from my sample. The results are qualitatively simiar.

2.6.3 Appendix 2.C: Price Information



(a) Timing (b) Frequency

Figure 2.10: Price Change: Timing and Frequency

More information on Figure 2.10: The left-hand figure presents that stations change price once average 11 days on average (median=11; sd=5.22; min=3.50; max=45.1). The right-hand figure shows that stations' price changes are more common on Tuesdays and least common on Sudays.

(a) Missing Days (b) Compromise

Figure 2.11: Missing and Compromise

More information on Figure 2.11: The left-hand figure shows the frequency of missing observations in the full dataset. The right-hand figure shows my compromise for instances which do not report price for up to four weeks, but report no changes of observed station characteristics. This compromise applies to 80 instances of a total of 175,940 observations.

Chapter 3

Price Discrimination in Retail Gasoline: Evidence from a Natural Experiment

3.1 Introduction

This paper uses a natural experiment to examine the pricing effects of changes in demand due to the introduction of government contracts in retail gasoline in Korea. Specifically, since 2013 official vehicles operated by the Korean government have been required to refuel at contracted gasoline station to get a fixed discount relative to the posted price. According to Kim (2012), the gasoline purchase of public vehicles accounts for about 7.7% of market demand in Korea. The initial contract terminated in November 2015, and another group of sellers took over the contract.

To empirically test the effect of changes in the contract on prices, I specify a difference-in-difference specification and use price data on all stations in Seoul for four weeks between 2012 and 2017. My estimates demonstrate that contractors' posted prices are 1.2% and 2.2% higher than non-contractors' prices in the first contract and the second contract periods, respectively. The increased price is consistent with the prediction of price discrimination models that contracted stations supply gasoline to price-inelastic government workers.

The result varies by stations' service levels and intensity of competition. Full-service contracted stations display a larger increment than do self-service contracted stations, and larger increases are found at contracted stations with fewer nearby competitors. Furthermore, prices of non-contracted stations very close to contracted stations are about 1.7% smaller than prices that non-contracted stations otherwise charge. This result implies that consumers can find a lower price across the streets of contracted stations.

I describe the institution of the government procurement in Section 2, and explain the data used in my paper in Section 3. I show my main results on the effect of the contract on price in Section 4. Section 5 concludes.

3.2 Institutions and Theoretical Framework

3.2.1 Institutions

The Korean government implemented the policy that all government officials' vehicles must refuel at gas stations under contract with Korea Public Procurement Service (KPPS) in December 2013.[1] KPPS is a governmental organization in charge of government procurement in Korea. This policy was introduced to increase the buying power of the public sector and save government expenditures on retail gasoline (Park, 2012).[2]

To bid the contract, firms must have over 1200 participating gas stations in the country and more than one participating station in each of 25 districts. There are four oil firms that satisfy these conditions in the Korean market – SK Energy, GS Caltex, Hyundai Oilbank, S-Oil. Before biding, firms are required to identify the participating stations.

Figure 3.1 shows the detailed timing of the contracts. In November 2013, GS Caltex won a two-year contract to supply gasoline to public vehicles at a discount of 3.99% from

[1]This rule applies all the time, except for emergencies – e.g., empty tank or natural disaster.

[2]Duggan and Scott Morton (2006) study a similar natural experiment: they examine the effects of government procurement on prices for drugs covered by US Medicaid health insurance. To determine the price the federal-state Medicaid program pays, Medicaid uses the average private sector price. The authors claim that such procurement rules create an incentive for pharmaceutical firms to increase prices, and they find results consistent with the claim.

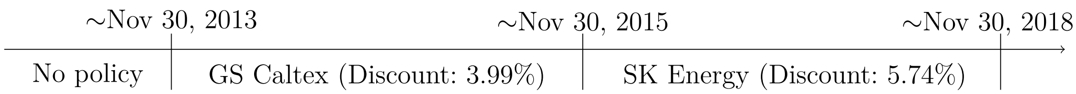


Figure 3.1: The timeline of contracts and fixed discounts for procurement of gasoline

the posted price. Two years later, SK Energy won a three-year contract to provide gasoline until the end of November, 2018 at a discount of 5.74%. These discounts do not include a 1.1% cash back that it provides all credit-card consumers. To get discounts, buyers must pay for gasoline with credit cards that KPPS issues.[3] Contracted stations first charge the posted price in full and reimburse discounts through the credit card companies within two weeks.

3.2.2 Theoretical Framework

In this section, I provide a theoretical framework to understand how stations change their prices when the contract takes effect.

Assume that each contracted station is a local monopolist at constant MC and *two* distinct demand curves for retail gasoline. The private market has a downward-sloping demand curve, while the public market consumes a fixed amount of gasoline with perfectly inelastic demand. The actual price charged to government workers is lower than the posted price due to the contracted discount. Each contracted station solves

$$\max_{p} \pi(p) = pq_{\text{pvt}}(p) + \alpha p\bar{q}_{\text{govt}} - c(q_{\text{pvt}}(p) + \bar{q}_{\text{govt}}), \quad 0 < \alpha < 1 \text{ and } MC = c. \tag{3.1}$$

where q_{pvt} is quantity demanded by private market and a function of posted price; q_{govt} is quantity demanded by government workers who do not care about price level; α is a discount parameter; and c is a constant marginal cost equal to the wholesale price and possibly some labor costs.

[3]KPSS issues these cards jointly with two financial companies – Sinhan and NH Card Inc.

The solution to this maximization problem is:

$$p^{*} = c - \frac{\alpha\overline{q}_{\text{govt}}}{q^{'}_{\text{pvt}}(p)} - \frac{q_{\text{pvt}}(p)}{q^{'}_{\text{pvt}}(p)} \quad where \quad q^{'}_{\text{pvt}}(p) < 0. \tag{3.2}$$

Equation (3.3) shows the incentive for a contractor to raise its price. That is, the larger the share of station's customers who are price inelastic, the higher the profit-maximization price:[4]

$$\frac{\Delta p^{*}}{\Delta\overline{q}_{\text{govt}}} = -\frac{\alpha}{q^{'}_{\text{pvt}}(p)} > 0 \quad where \quad q^{'}_{\text{pvt}}(p) < 0. \tag{3.3}$$

In the meantime, each non-contracted station solves the similar problem and gets the optimal price (denoted p**) when $\overline{q}_{govt}$ is zero.

$$p^{**} = c - \frac{q_{\text{pvt}}(p)}{q^{'}_{\text{pvt}}(p)} \tag{3.4}$$

$$p^{*} - p^{**} = -\frac{\alpha\overline{q}_{\text{govt}}}{q^{'}_{\text{pvt}}(p)} > 0 \quad where \quad q^{'}_{\text{pvt}}(p) < 0. \tag{3.5}$$

It is clear that the optimal price of a contracted station is larger than that of a non-contracted station in Equation (3.4). In Section 4, I compare prices of contracted stations with those of non-contracted stations, and empirically test how the government contract changes the profit-maximizing price of contractors and document the magnitude of this effect.

3.3 Data

The station-level data come from two sources. The first source is the Korea Public Procurement Service (KPPS) and contains the name and address of contracted stations in Seoul, a city of approximately 10 million people, at three different dates (November 11, 2015; May 3, 2017; and September 27, 2017). The second source is the Oil Price Information Network (OPINET). OPINET is a website, *www.opinet.co.kr*, operated by Korea National Oil Cor-

[4] $\overline{q}_{govt}$ is assumed to be a fixed positive number, not necessarily differentiable.

poration, where gasoline prices of every station in the market are posted daily. OPINET also provides stations' address, brand type, and service level.[5] I merge the status of procurement contracts with station prices based on address.

Table 3.1: Summary Statistics [3437 Obs.]

Variable	Description	Mean	SD	Min	Max
P	Price of gasoline (unit: KRW/liter)	1710.6	238.3	1371	2385
lnP	Log price of gasoline	7.43	0.13	7.22	7.77
SK	Station brand: SK Energy	0.40	0.49	0	1
GS	Station brand: GS Caltex	0.27	0.44	0	1
HD	Station brand: Hyundai Oilbank	0.14	0.35	0	1
SO	Station brand: S-Oil	0.13	0.34	0	1
AD	Station brand: Alddle	0.02	0.15	0	1
Unbranded	Station brand: Unbranded	0.03	0.15	0	1
Num^{FS15}	Number of full-service competitors within 1.5 miles	12.90	5.82	0	29
Num^{SS15}	Number of self-service competitors within 1.5 miles	3.70	2.38	0	11
Full	Station offering full-serve gasoline	0.75	0.13	0	1
ShortDist	Dist. of non-contractors from closest contractors	0.32	0.19	0.01	0.94
Stations	Num. of stations in Seoul [a]	560	27.40	537	606

Average across all stations in all time periods.

[a] Stations' entry and exit are active. The number is: 606 in Aug 2012; 559 in Nov 2015; 541 in May 2017; and 537 in Sept 2017.

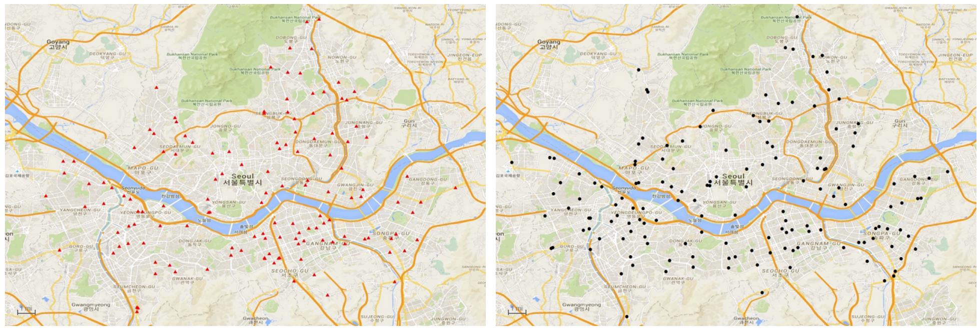
The procurement contracts are systematically correlated with station brands, so I add brands as control variables in estimation.[6] There are five brands operating in the market. In descending order by market share, these are: SK Energy, GS Caltex, Hyundai Oilbank, S-Oil, and Alddle. After geocoding all station addresses, I count the numbers of full- and self-service competitors within a 1.5-mile radius for each station, and use them as my measure of intensity of competition in my regression analysis. Lastly, stations are perfectly partitioned by service level at each point in time in this market. I also include service level in my estimation in that it has a large affect on station pricing. Table 3.1 presents the summary statistics on these variables for the full sample of 643 stations over all time periods.

The locations of these stations are shown in Figure 3.2, and they seem well dispersed in the market. The closest distance of a non-contracted station from a contracted station is 0.32 miles on average. Including the three selected dates provided by KPPS, I use every Wednesday of a week in August 2012 as a reference date of "no policy."[7] In a personal

[5] See the earlier chapters for a detailed description of OPINET.

[6] All branded stations are not necessarily contractors. 92% of GS stations and 65% of SK stations were contractors in the first and second contracts, respectively.

[7] It is not clear whether such government contracts were in effect between October 2012 and



(a) GS contractors, November 2015 (b) SK contractors, September 2017

Figure 3.2: Locations of Contracted Stations, by period

communication, one official staff at KPPS told me that once stations enter into a contract, the contract does not change very often. However, in my analysis, I use only selected dates because of concerns about measurement error. I discuss this issue in more detail in the next section.

Table 3.2: Summary Statistics: contracted vs. non-contracted stations

	Log price[a]			Means of control variables[a,b]			# of Stations
	Mean	Min	Max	Full	Num^{FS15}	Num^{SS15}	
no policy [2012.08]							
non-contractors	7.576	7.499	7.776	0.854	15.49	2.417	606
	(0.069)			(0.352)	(6.107)	(1.671)	
GS Contract [2015.11]							
contractors	7.360	7.242	7.598	0.786	12.80	3.360	136
	(0.102)			(0.411)	(5.673)	(2.226)	
non-contractors	7.341	7.223	7.647	0.751	12.66	3.841	423
	(0.092)			(0.432)	(5.613)	(2.339)	
SK Contract [2017.05; 2017.09]							
contractors	7.426	7.268	7.652	0.702	11.40	4.104	141
	(0.105)			(0.457)	(5.821)	(2.507)	
non-contractors	7.354	7.268	7.654	0.708	11.52	4.167	394
	(0.087)			(0.454)	(5.393)	(2.554)	

Average across all stations given contract periods.
Variables defined in Table 3.1.
[a] Standard deviation in parentheses.
[b] Differences of means between contractors and non-contractors are not statistically different at 5% level.

Table 3.2 provides summary statistics on the log price of contracted and non-contracted stations and a few comparisons between those stations on two dimensions: service and competitive environment.[8] First, contracted stations' prices are higher than non-contracted

November 2013.

[8]The Korean retail gasoline market has transitioned from full service to self service since early 2009. This explains why the number of self-service stations nearby increases over time. See Kim (2018) for details.

stations, but it is unclear whether it is a result of the contract or something else. In addition, the means of the control variables seem different between contractors and non-contractors, but these differences are not statistically significant. In the next section, using a regression framework, I formally analyze the effect of the government contracts on prices, and examine how the effect varies with station characteristics.

3.4 Results

3.4.1 Effects of Contract on Prices

To investigate the effect of the government contract on prices, I specify a reduced-form model of station prices as follows:

$$\ln P_{it} = \beta GovtContract_{it}^{GS} + \gamma GovtContract_{it}^{SK} + \zeta \boldsymbol{X}_{it} + \delta_i + \mu_t + \varepsilon_{it} \tag{3.6}$$

where P_{it} is the price of station i at period t. $GovtContract^{\boldsymbol{GS}}{}_{it}$ is a dichotomous variable that equals 1 if station i is a procurement supplier with a brand of GS Caltex at time t, and 0 if it is not. Similarly, $GovtContract^{\boldsymbol{SK}}{}_{it}$ is 1 if station i is a procurement supplier with SK Energy at time t, and 0 if it is not. $\boldsymbol{X}_{it}$ is a vector of control variables, such as service level and intensity of competition, described in Table 3.1. δ is a station fixed effect that controls for time-invariant characteristics, such as location and traffic, that affect pricing. μ is a time fixed effect that captures changes in average prices, mostly driven by wholesale gasoline prices. ε is an error term clustered by station level.

Estimation results are shown in Table 3.3. Column (1) shows the effect of the contract on prices when only controlling for station brands. Column (2) shows the effect after adding other variables that potentially affect gasoline prices. Finally, I include a station fixed effect in Column (3).

The coefficients on GovtContract are positive and significant at 10% level across all spec-

Table 3.3: Effects of government procurement on price

	(1)	(2)	(3)
GovtContract^{GS}	0.0140**	0.0096*	0.0125**
	(0.006)	(0.005)	(0.005)
GovtContract^{SK}	0.0341***	0.0331***	0.0219***
	(0.009)	(0.008)	(0.006)
SK	0.0568***	0.0574***	0.0493***
	(0.008)	(0.009)	(0.016)
GS	0.0399***	0.0380***	0.0505***
	(0.008)	(0.008)	(0.015)
HD	0.0069	0.0078	0.0253
	(0.008)	(0.007)	(0.017)
SO	0.0169**	0.0202**	0.0213
	(0.009)	(0.008)	(0.017)
AD	-0.0280***	-0.027***	-0.0045
	(0.006)	(0.008)	(0.022)
Full		0.0589***	0.0544***
		(0.004)	(0.008)
Num^{FS15}		-0.0005	-0.000
		(0.000)	(0.000)
Num^{SS15}		-0.0110***	-0.0055***
		(0.001)	(0.002)
F-statistic[a]	2.41	4.14	1.05
[p-value]	[0.12]	[0.04]	[0.30]
Time FE	Yes	Yes	Yes
Station FE	No	No	Yes
Obs.	3437	3437	3437
# of stations	643	643	643
R^2	0.777	0.822	0.932

Unbranded is omitted.

Robust standard errors in parentheses, clustered by station.

[a] F-statistics for equal effects of GS and SK contracts.

* p<10%; ** p<5%; *** p<1%.

ifications.[9] Starting with Column (1), all else equal, GS Caltex contractors charge 1.4% more than non-contractors, and SK contracted stations charge 3.4% more than non-contracted stations. These results are consistent with the prediction from Equation (3.3) that quantity demanded by the government increases the optimal price of a contracted station. It is also confirmed that the optimal posted price gets larger when the discount parameter is larger. These estimates generally decrease with more controls in Columns (2)-(3).

Considering the amount of discounts offered to government purchases, the point estimates on GovtContract in Column (3) indicate that the discounts actually offered are 2.74% (3.99-1.25) under the GS Contract and 3.80% (5.74-2.19) under the SK Contract. The government policy affects the price for everyone in the market: not only firms and government but also consumers.[10]

Even if the price increase leads to the loss of consumers in the private market, the contract may, in fact, increase station profits.[11] For one thing, the loss of private consumers is (partially or fully) offset by government demand. Moreover, if procurement suppliers are perceived by private consumers as trustworthy sellers, consumers' willingness to pay may increase.[12]

For the estimates above to be unbiased, the policy must be implemented or changed independent of the pricing of contracted stations. This condition is reasonable in my setting because stations have no incentive to increase their prices in advance. They would otherwise lose consumers. However, my estimates are not entirely free from selection bias in that stations close to public offices may already charge a high price and prefer being a contracted

[9] The F-test confirms that the coefficient on $\text{GovtContract}^{\text{GS}}$ is not statistically different from that on $\text{GovtContract}^{\text{GS}}$ at the 5% level.

[10] Assuming no changes in the contracts during the intervals in Figure 3.1, I re-estimate the specification with the extended sample that includes every Wednesday of a week between 2012 and 2017. I suspect that measurement error is so dramatic that the coefficients of interest are attenuated considerably. See Table 3.6 for details on Appendix 3.A.

[11] Courty and Pagliero (2012) show that price discrimination generates about greater revenues, using data from concern tickets which are collected by Billboard magazine.

[12] Koo (2017) documents one interview of a gas-station owner talking about the selection of the contract: "stations with low price do not like a making contract. Why do we cost 70 or 80 KRW? (it is equivalent to 30 cents per gallon). Then there is nothing left..." This story sounds that the contract is voluntary. Nevertheless, I find some contracted stations with lower price than their nearby competitors, implying that oil companies may enforce their branded stations for participating in the contract.

station. Although station fixed effects do not completely eliminate these concerns, they greatly remove station-specific characteristics that are fixed.

The middle-panel of Table 3.3 shows how control variables affect prices. Focusing on Column (2), brand coefficients pick up a branded price relative to that of unbranded stations; branded gasoline is usually expensive than unbranded gasoline, except for AD. The coefficient on Full shows that full-service stations charge 6% more than self-service stations. Regarding the effect of intensity of competition on prices, having one more self-service station within a 1.5-mile radius decreases gasoline price by 1.1%, while full-service stations do not significantly affect competitors' price. Lastly, controlling for station fixed effects in Column (3), most of the coefficients become smaller but signs are robust to this specification.[13]

Next, I examine whether the effect of the contract on prices is uniform across station characteristics. Starting with the point estimates in Column (1) of Table 3.4, GS contracted stations charge 1.87% (0.0253-0.0065) more when they are full service than when they are self service. Similarly, SK contracted stations charge 3.77% (0.0573-0.0196) more when they are full service than when they are self service.[14] Interestingly, these increases are a result of much larger increases in full-service stations' prices and SK contracted self-service stations even charge *lower* prices than other stations. One possible explanation for this is that government workers constitute a very small share of total demand, although about 30% of contracted stations are self serve (Table 3.2).

Column (2) shows heterogeneous effects of intensity of competition. GS contracted stations that do not have competitors nearby charge 2.73% more than those that do have at least one self-service station nearby. The effect is similarly observed in the period of the SK Contract. This finding suggests that contracted stations still compete with non-contracted stations on price, as conventional models of competition predict: price falls when competition is high.

I include all interactions together in Column (3). The coefficient on the SK contract

[13]Recall that the conversion of gasoline stations from full service to self service is in progress in this market. This would make the coefficients on control variables generally significant. There were 69 instances of service changes and 84 instances of brand changes during the sample period.

[14]The calculation of the discounted price charged to government workers is consistent with earlier results.

Table 3.4: Heterogeneous effects of government procurement on price

	(1)	(2)	(3)
GovtContract^{GS}	-0.0065	0.0273**	0.0097
	(0.004)	(0.012)	(0.011)
GovtContract^{GS}*Full	0.0253***		0.0246***
	(0.006)		(0.006)
GovtContract^{GS}*Num^{FS15}		-0.0001	-0.0003
		(0.000)	(0.000)
GovtContract^{GS}*Num^{SS15}		-0.0038**	-0.0032*
		(0.002)	(0.002)
GovtContract^{SK}	-0.0196***	0.0290**	-0.0059
	(0.007)	(0.014)	(0.013)
GovtContract^{SK}*Full	0.0573***		0.0557***
	(0.009)		(0.010)
GovtContract^{SK}*Num^{FS15}		0.0001	0.0003
		(0.000)	(0.000)
GovtContract^{SK}*Num^{SS15}		-0.0045**	-0.0040**
		(0.002)	(0.002)
F-statistic[a]	7.92	0.01	7.09
[p-value]	[0.00]	[0.93]	[0.00]
Time FE	Yes	Yes	Yes
Station FE	Yes	Yes	Yes
Obs.	3437	3437	3437
# of stations	643	643	643
R^2	0.933	0.932	0.934

All specifications (1)-(3) also include control variables such as station brand, service level, and the numbers of competitors by service level.

Robust standard errors in parentheses, clustered by station.

[a] F-statistics for equal heterogeneous effects of GS and SK contracts.

* p<10%; ** p<5%; *** p<1%.

becomes insignificant. This is possibly because the variation in the number of self-service stations nearby is strongly correlated with the presence of self-service pumps at contracted stations.[15] To identify an effect, I would need to have more data on the contracts at different dates.

The increase in price of contracted stations possibly affects their nearby competitors. If it does, the magnitude and sign of the impact depends on the elasticity of residual demand each competitor faces. To examine the effect on nearby competitors, I create a dummy variable, **Neighborhood**, that equal to 1 if non-contracted station i is located within a specified distance of contracted station j at time t. I use distances 0.25-miles, 0.2-miles, and 0.15-miles. Otherwise, it is 0.

Table 3.5: Indirect effects of government procurement on price

	(1) 0.25-mile	(2) 0.2-mile	(3) 0.15-mile
Neighborhood	-0.0089	-0.0107*	-0.0156**
	(0.006)	(0.006)	(0.007)
GovtContract^{GS}	0.0117**	0.0119**	0.0118**
	(0.005)	(0.005)	(0.005)
GovtContract^{SK}	0.0209***	0.0209***	0.0210***
	(0.006)	(0.006)	(0.006)
F-statistic[a]	1.02	0.98	1.02
[p-value]	[0.31]	[0.32]	[0.31]
Time FE	Yes	Yes	Yes
Station FE	Yes	yes	Yes
Obs.	3437	3437	3437
# of stations	643	643	643
R^2	0.932	0.932	0.932

All specifications also include all other control variables.
Robust standard errors in parentheses, clustered by station.
[a] F-statistics for equal effects of GS and SK contracts.
* p<10%; ** p<5%; *** p<1%.

The results are shows in Columns (1)-(3) Table 3.5. The coefficients are negative across specifications, and the statistical significance of these coefficients becomes stronger when I narrow the geographic dimension of price competition. Specifically in Column (3), non-contracted stations that compete with contracted stations within 0.15 miles post a 1.5% lower price than non-contracted stations that do not face such competition. This finding im-

[15] These variables move together while the market is transitioning from full service to self service. See Kim (2018)

plies that non-contracted stations competing with contracted stations have elastic residual demand, so they attract their price-elastic local customers by offering a lower price.

Overall, my results support the notion of models of price discrimination that firms charge more in markets with the lower elasticity of demand. My estimates implies that consumers would pay about 2-3% *less* at contracted stations if the regulatory contract had not been implemented. This means that the discounted price charged to the government is partially funded by private consumers who are price-inelastic and/or less-informed of price dispersion in the market.

3.4.2 Welfare Implication

The discount offered to the government is associated with saving government expenditure on retail gasoline purchases. The total gasoline expenditure of the country is approximately $21 billion in 2014 dollar and $18 billion in 2015 dollar.[16] Therefore, an increase in government surplus equivalent to 7.7% of gasoline expenditure (Kim, 2012) is on the order of $1.6 billion and $1.4 billion for each year, respectively. Although the contract saves money on retail gasoline purchases, it does not necessarily account for an increase in total welfare.

3.5 Concluding Remarks

This paper studies how the pricing behavior of firms changes due to changes in the elasticity of demand. I find that that contracted stations post higher prices than otherwise similar stations. Because contracted stations supply their gasoline at a discounted price to the government, they have a strong incentive to charge private consumers *more* for their gasoline than they otherwise would. This pattern is much stronger in full-service contracted stations, which price-inelastic consumers prefer, than at self-service stations. This empirical finding supports the prediction of models of price discrimination.

[16] This is my calculation; I use the average price of gasoline in 2014 and 2015 and the total consumption of gasoline for the corresponding year, as well as the exchange rate of 1 USD for 1000 KRW. The Korea Energy Statistical Information System (KESIS) provides the raw data.

3.6 Appendix to Chapter 3

3.6.1 Appendix 3.A: Extended Sample

Table 3.6: Effects of government procurement on price, every Wednesday of a week

	(1) original	(2) one-week	(3) one-month	(4) three-month	(5) one-year
GovtContract^{GS}	0.0125***	0.0125***	0.0122***	0.008**	0.006*
	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)
GovtContract^{SK}	0.0219***	0.0222***	0.0231***	0.0229***	0.0100**
	(0.006)	(0.005)	(0.005)	(0.005)	0.004)
Time FE	Yes	Yes	Yes	Yes	Yes
Station FE	Yes	Yes	Yes	Yes	Yes
Obs.	3437	5073	8351	33317	80722
# of stations	643	643	643	650	651
R^2	0.932	0.923	0.906	0.900	0.875

Station brands are included.

Robust standard errors in parentheses, clustered by station.

* p<10%; ** p<5%; *** p<1%.

I estimate the main specification with the extended sample, and the results are shown in Table 3.6. From Column (2), I assume no changes in the contracts for one week, so the sample used in the estimation contains three dates provided by KPPS and also dates of one week ahead based the KPPS sample for each time. The estimates of the contracts are very robust in this case, and also robust to the assumption that the contracts do not change for one month in Column (3).

However, in Column (4), the estimate on GS Contract is significantly attenuated, while the estimate on SK Contract is still consistent. In fact, I confirm that there were only three stations newly entered into the contract between May 2017 and September 2017. As discussed in the Data section, one staff at KPPS told me in November 2017 that stations tend to keep their contract once they made. These results, however, indicate that he might talk only about the current SK contract, not about the GS contract.

Lastly in Column (5), attenuation bias seems dramatic in both coefficients with the assumption of no changes in the contracts for one year. I think that this setting may be not reasonable. Since the Korean retail gas market has been experiencing a massive restructuring during my study period, the ownership of stations are likely to change significantly, which I did not control for in my analysis. The precision of the estimates can be improved with more data on the contracts in other dates.

Bibliography

Barron, J. M., Taylor, B. A., and Umbeck, J. R. (2004). Number of Sellers, Average Prices, and Price Dispersion. *International Journal of Industrial Organization*, 22(8):1041–1066.

Basker, E., Foster, L., and Klimek, S. D. (2017). Customer-Employee Substitution: Evidence from Gasoline Stations. *Journal of Economics & Management Strategy*, 26(4):876–896.

Basker, E. and Noel, M. (2009). The Evolving Food Chain: Competitive Effects of Wal-Mart's Entry into the Supermarket Industry. *Journal of Economics & Management Strategy*, 18(4):977–1009.

Baye, M. R., Morgan, J., and Scholten, P. (2016). Information, Search, and Price Dispersion. In Hendershott, T., editor, *Handbook on the Economics and Information Systems*. Elsevier (pp. 323–371).

Borenstein, S. (1991). Selling Costs and Switching Costs: Explaining Retail Gasoline Margins. *RAND Journal of Economics*, 22(3):354–369.

Borenstein, S. and Rose, N. L. (1994). Competition and Price Dispersion in the US Airline Industry. *Journal of Political Economy*, 102(4):653–683.

Bronnenberg, B. J. and Ellickson, P. B. (2015). Adolescence and the Path to Maturity in Global Retail. *Journal of Economic Perspectives*, 29(4):113–134.

Chandra, A. and Lederman, M. (2018). Revisiting the Relationship Between Competition and Price Discrimination. *American Economic Journal: Microeconomics*, 10(2):190–224.

Chandra, A. and Tappata, M. (2011). Consumer Search and Dynamic Price Dispersion: An Application to Gasoline Markets. *RAND Journal of Economics*, 42(4):681–704.

Chen, Y. and Riordan, M. H. (2013). Profitability of Product Bundling. *International Economic Review*, 54(1):35–57.

Ching, A. T. (2010). Consumer Learning and Heterogeneity: Dynamics of Demand for Prescription Drugs after Patent Expiration. *International Journal of Industrial Organization*, 28(6):619–638.

Courty, P. and Pagliero, M. (2012). The Impact of Price Discrimination on Revenue: Evidence from the Concert Industry. *Review of Economics and Statistics*, 94(1):359–369.

Duggan, M. and Scott Morton, F. M. (2006). The Distortionary Effects of Government Procurement: Evidence from Medicaid Prescription Drug Purchasing. *Quarterly Journal of Economics*, 121(1):1–30.

Eckert, A. and West, D. S. (2005). Rationalization of Retail Gasoline Station Networks in Canada. *Review of Industrial Organization*, 26(1):1–25.

Ellison, G. and Ellison, S. F. (2009). Search, Obfuscation, and Price Elasticities on the Internet. *Econometrica*, 77(2):427–452.

Foster, L., Haltiwanger, J., and Krizan, C. J. (2006). Market Selection, Reallocation, and Restructuring in the US Retail Trade Sector in the 1990s. *Review of Economics and Statistics*, 88(4):748–758.

Genesove, D. (1995). Search at wholesale Auto Auctions. *Quarterly Journal of Economics*, 110(1):23–49.

Gerardi, K. S. and Shapiro, A. H. (2009). Does Competition Reduce Price Dispersion? New Evidence from the Airline Industry. *Journal of Political Economy*, 117(1):1–37.

Grabowski, H. G. and Vernon, J. M. (1992). Brand Loyalty, Entry, and Price Competition in Pharmaceuticals after the 1984 Drug Act. *Journal of Law and Economics*, 35(2):331–350.

Graddy, E. (1990). The Evolution of New Industries and the Determinants of Market Structure. *RAND Journal of Economics*, 21(1):22–44.

Harvey, A. C. (1976). Estimating Regression Models with Multiplicative Heteroscedasticity. *Econometrica*, 43(3):461–465.

Hastings, J. S. (2004). Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California. *American Economic Review*, 94(1):317–328.

Hortaçsu, A. and Syverson, C. (2015). The Ongoing Evolution of US Retail: A Format Tug-of-War. *Journal of Economic Perspectives*, 29(4):89–111.

Houde, J.-F. (2012). Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline. *American Economic Review*, 102(5):2147–2182.

Jovanovic, B. and MacDonald, G. M. (1994). The Life Cycle of a Competitive Industry. *Journal of Political Economy*, 102(2):322–347.

Kang, H. (2010). 저렴한 셀프주유소 찾기가 힘들다 왜? (Why Is It Hard to Find a Cheap Self-Service Station?). *Nocut News*, 18 March, 2010, available at http://www.nocutnews.co.kr/news/694134 (accessed 29 April 2018).

Kim, D.-W. and Kim, J.-H. (2011). The Impact of the Entry of Self-Service Stations in the Korean Retail Gasoline Market: Evidence from the Difference-in-Differences Methods. *Korean Journal of Economic Studies*, 59(2):77–99.

Kim, S. (2012). 공공기관 차량용 유류 1조원 규모 공동구매 (1 Trilion KRW Worth of Public Procurement for Gasoline for Public Cars). *Public Procurement Service*, 29 August, 2012, available at https://www.pps.go.kr/bbs/selectBoardList.do?boardId=PPS085 (accessed 12 March, 2018).

Kim, T. (2017). Price Competition and Market Segmentation in Retail Gasoline: New Evidence from South Korea. unpublished paper, University of Missouri.

Kim, T. (2018). Changing Market Structure and Evolving Ways to Compete: Evidence from Retail Gasoline. unpublished paper, University of Missouri.

Koo, H. (2017). 비싼 기름값: 공공기관 지정 주유소의 함정 (High Price: a Trap of Contracted Gas Stations). *JTBC News*, 28 September, 2017, available at http://news.jtbc.joins.com/article/article.aspx?news_id=NB11528545 (accessed 10 April, 2018).

KTDB (2013). 자동차이용실태조사 (Survery on Vehicle Usage). Korea Transport Database, available at https://www.ktdb.go.kr/www/index.do (accessed 29 April 2018).

Kwon, Y. (2010). 셀프주유소 확장 산넘어 산 (The Difficulty of Expanding Self-Service Stations). *Energy & Environment News*, 24 May, 2010, available at http://e2news.com /news/articleView.html?idxno=38093 (accessed 29 April 2018).

Lach, S., . M.-G. J. L. (2015). Asymmetric Price Effects of Competition. Centre for Economic Policy Research (CEPR) Discussion Papers 10456.

Lach, S. (2002). Existence and Persistence of Price Dispersion: An Empirical Analysis. *Review of Economics and Statistics*, 84(3):433–444.

Lee, S., Kim, H., and Park, M. (2015). Pricing Puzzle in a Retail Gasoline Market. unpublished paper, Sungkyunkwan University.

Lewis, M. (2008). Price Dispersion and Competition with Differentiated Sellers. *Journal of Industrial Economics*, 56(3):654–678.

Lewis, M. S. (2011). Asymmetric Price Adjustment and Consumer Search: An Examination of the Retail Gasoline Market. *Journal of Economics & Management Strategy*, 20(2):409–449.

MacMinn, R. D. (1980). Search and Market Equilibrium. *Journal of Political Economy*, 88(2):308–327.

Marvel, H. P. (1976). The Economics of Information and Retail Gasoline Price Behavior: an Empirical Analysis. *Journal of Political Economy*, 84(5):1033–1060.

Matsa, D. A. (2011). Competition and Product Quality in the Supermarket Industry. *Quarterly Journal of Economics*, 126(3):1539–1591.

Mazzeo, M. J. (2002). Competitive Outcomes in Product-Differentiated Oligopoly. *Review of Economics and Statistics*, 84(4):716–728.

Ministry of Government Legislation (2017). 위험물 안전관리법 시행규칙 (Safety Control of Dangerous Substances Act). National Law Information Center, available at http://www.law.go.kr/main.html (accessed 29 April 2018).

Netz, J. S. and Taylor, B. A. (2002). Maximum or Minimum Differentiation? Location Patterns of Retail Outlets. *Review of Economics and Statistics*, 84(1):162–175.

Noel, M. D. (2016). Retail Gasoline Markets. In Basker, E., editor, *Handbook on the Economics of Retailing and Distribution.* Edward Edgar (pp. 392–412).

Park, H. (2012). 조달청, 공공부문 유류 공동구매 서비스 개시 (Start Up Public Procurement for Gasoline for Public Cars). *Newsis*, 29 October, 2012, available at http://news.joins.com/article/9725465 (accessed 12 March, 2018).

Png, I. P. and Reitman, D. (1994). Service Time Competition. *RAND Journal of Economics*, 25(4):619–634.

Reinganum, J. F. (1979). A Simple Model of Equilibrium Price Dispersion. *Journal of Political Economy*, 87(4):851–858.

Remer, M. (2015). An Empirical Investigation of the Determinants of Asymmetric Pricing. *International Journal of Industrial Organization*, 42(1):46–56.

Salop, S. and Stiglitz, J. (1977). Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion. *Review of Economic Studies*, 44(3):493–510.

Saxonhouse, G. R. (1976). Estimated Parameters as Dependent Variables. *American Economic Review*, 66(1):178–183.

Shepard, A. (1991). Price Discrimination and Retail Configuration. *Journal of Political Economy*, 99(1):30–53.

Soetevent, R. A. and Bružikas, T. (2017). The Impact of Process Innovation on Price: Evidence from Automated Fuel Retailing in the Netherlands. Tinbergen Institute Discussion Paper No.045/VII.

Sorensen, A. T. (2000). Equilibrium Price Dispersion in Retail Markets for Prescription Drugs. *Journal of Political Economy*, 108(4):833–850.

Stahl, D. O. (1996). Oligopolistic Pricing with Heterogeneous Consumer Search. *International Journal of Industrial Organization*, 14(2):243–268.

Tappata, M. (2009). Rockets and Feathers: Understanding Asymmetric Pricing. *RAND Journal of Economics*, 40(4):673–687.

Varian, H. R. (1980). A Model of Sales. *American Economic Review*, 70(4):651–659.

Wildenbeest, M. R. (2011). An Empirical Model of Search with Vertically Differentiated Products. *RAND Journal of Economics*, 42(4):729–757.

Yang, H. and Ye, L. (2008). Search with Learning: Understanding Asymmetric Price Adjustments. *RAND Journal of Economics*, 39(2):547–564.

VITA

Taehwan Kim was born in Seoul of South Korea. He lived in the city until he entered the U.S. in July 2012 for his challenge to get a Ph.D. degree in Economics at Mizzou.

Taehwan Kim passed a qualifying exam in August 2013 and a (written) comprehensive exam in August 2014. When choosing his research field, he had no hesitation; he wanted to be a student of Emek Basker in November 2014. This decision was potentially risky, however, because he had had no chance to officially meet her in class or somewhere else on campus. But he was very serious regarding his decision so he tried very hard to show her his earnestness and diligence. She had multiple interviews with him and finally accepted his request in March 2015 – it was the first step that enabled his challenge to be done with a happy ending. Since then, he has spent another three and half years on writing a good-quality dissertation under the supervision of Emek Basker. One of his dissertation chapters is already published in a peer-reviewed journal, and he had tried to publish the rest of them in a high-quality journal in a few years.

After 24 years(!) of schooling, Taehwan Kim is going to start up his professional career with a position of Associate Research Fellow at Korea Energy Economics Institute (KEEI) from the fall of 2018.