Essays in Industrial Organization

by

Youngjun Jang

B.A.S. Economics and Mathematics

Stanford University, 2007

Submitted to the Department of Economics on May 15, 2015 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Economics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2015

Signature of Author: ... Signature redacted ...

Department of Economics

May 15, 2015

Certified by: ... Signature redacted ...

Glenn Ellison

Gregory K. Palm (1970) Professor of Economics

Thesis Supervisor

Accepted by: ... Signature redacted ...

Ricardo Caballero

Ford International Professor of Economics

Chairman, Departmental Committee on Graduate Studies

Essays in Industrial Organization

by

Youngjun Jang

Submitted to the Department of Economics on May 15, 2015 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Economics

Abstract

The first chapter investigates the effects of technological advances on the retail gasoline market. Since 2008, a Korean government website has posted daily prices of all gasoline stations. Combined with the rapid increase of smartphone and mobile technologies, this price information service may have changed the consumer search environment significantly. Using daily price data, sales data, and regional smartphone penetration rates, I find that price dispersion among gasoline stations and markups increase slightly when the smartphone penetration rate increases, even while additional descriptive evidence suggests that demand is becoming more price-sensitive. Structural estimation of a two-type consumer search model finds that the proportion of highly informed consumers increases as the smartphone penetration rate increases. A counterfactual analysis confirms that observed price changes are consistent with theoretical models of pricing, given the structurally estimated parameters.

The second chapter studies consumer decisions at gasoline pumps, using a detailed transaction-level dataset. About 36% of regular gasoline consumers chose to simply fill up, while the remaining 64% of consumers spent pre-selected dollar amounts. Descriptive analyses show that consumers become more active in quantity choices at gasoline pumps, and less likely to simply fill up, when retail prices are on an upward trend and when the current price level is unexpectedly high. Reduced-form results suggest that consumers expect that gasoline prices tend to move to the average price over time.

The third chapter analyzes the effects of ability grouping on the academic performance of high school graduating students in Korea. About half of the regions in Korea have adopted an equalization policy (EP), which means that students are randomly assigned. For the other non-EP regions, students are sorted among schools based on ability levels. I utilize a difference-in-differences strategy to exploit the adoption of the EP, an exogenous policy shift. I find that after the EP, performance of students above the median dropped 1.4% in terms of national percentiles, while that of students below the 30% percentile jumped 1.3%. In addition, there was an increasing trend in the achievement levels in the treatment regions, but after the introduction of the EP, this trend vanished.

Thesis Supervisor: Glenn Ellison
Title: Gregory K. Palm (1970) Professor of Economics

Acknowledgments

I would like to express my gratitude to Glenn Ellison and Sara Ellison for their invaluable support and guidance throughout the PhD program. I decided to study Industrial Organization after taking Glenn's Industrial Organization class, as course topics were intellectually stimulating and his efficient teaching style fascinated me. Through the course of my education at MIT, they provided countless insights and amazing comments, and helped me to find interesting topics and approach them critically using rigorous methods. In addition, they have been extremely friendly and accessible: I could seek their advice for almost everything, personal and professional. In short, they were like parents to me. I believe that choosing Glenn and Sara as my advisors was the best decision that I have made at MIT.

I would also like to thank the faculty at MIT who provided excellent comments. In particular, I would like to thank Paulo Somaini, Nikhil Agarwal, and the MIT Industrial Organization lunch seminar and weekly workshop participants.

I am grateful to my friends and colleagues at MIT. Graduate school is no bed of roses, and my MIT experience would have been less fulfilling without amazing friends. In particular, I thank Brandon Lee, who survived the two years of the coursework along with me, for being a wonderful roommate and classmate; Joonhwan Lee, my closest friend at MIT, for providing thoughtful comments and technical details; and Eunhee Sohn, for going through stressful job market processes with me.

After finishing the coursework, I served two years at the Korean Army Headquarters from 2009 to 2011. While it prevented me from focusing on economics for a long period, I believe that it was a worthwhile experience overall: I learned how to collaborate, how to endure, and how to focus under any circumstances. Such a rewarding experience and smooth transitions from academia to military (and back) would not have been possible without fantastic superiors and comrades at the Analysis and Assessment Group Division.

I gratefully acknowledge the financial support from the Kwanjeong Scholarship Foundation and MIT Presidential Fellowship. Having financial support not only helped me to concentrate on research, but also motivated me to constantly improve my research quality. I hope that I may contribute to the scholarship foundation and MIT in the future by supporting future students.

I was fortunate to have chances to meet industry experts and government officials for my study. I sincerely appreciate the Opinet staff and anonymous company officials who provided me wonderful data sets and answered countless questions. I would like to thank the Korea Institute for Curriculum and Evaluation for the opportunity to work on an educational evaluation research project, which was developed further to become the third chapter of the thesis.

Finally, I cannot thank my parents enough for their continuous support during the PhD program. They have always encouraged me to welcome challenges and pursue my dreams. I am delighted to present them with my thesis, and I look forward to opening the next chapter of my life.

Contents

Chapter 1

The Effect of the Internet and Mobile Search Technologies on Retail Markets: Evidence from the Korean Gasoline Market

Since 2008, a Korean government website has posted daily prices of all gasoline stations. Combined with the rapid increase of smartphone and mobile technologies, this price information service may have changed the consumer search environment significantly. This paper investigates the effects of these technological advances on the retail gasoline market, using daily price data, quantity data for select stations obtained from a credit card provider, and regional smartphone penetration rates. In daily price data for four regions from January 2010 to June 2012, price dispersion among gasoline stations and markups increase slightly when the smartphone penetration rate increases, even while additional descriptive evidence suggests that demand is becoming more price-sensitive. Structural estimation of a two-type consumer search model finds that the proportion of highly informed consumers increases as the smartphone penetration rate increases. A counterfactual analysis confirms that observed price changes are consistent with theoretical models of pricing, given the structurally estimated parameters.

1.1 Introduction

While the "law of one price" is an elegant and comprehensible economic theory, it is seldom true in practice. Since Stigler's seminal paper (1961) on search and price dispersion, many papers have noted that consumer search costs can lead to price dispersion. If there is a change in the search cost, consumer search behavior and the expected amount of price information consumers have will change. That will affect the supply side, and hence market outcomes such as price dispersion and markup will also change. Recent interest in the effects of consumer search costs has been fueled by the growth of e-commerce because the Internet can often supply settings with very low consumer search costs. However, there is no clear evidence that the Internet has made markets more efficient or decreased price dispersion. For example, Ellison and Ellison (2009) find that while consumers became extremely price-sensitive for some products, internet retailers developed obfuscation strategies to maintain their profit margins, and price dispersion has persisted. In fact, price dispersion seems to be increasing over time and price dispersion is an equilibrium phenomenon that depends on the market structure (Baye et al., 2004). A recent emergence of mobile technologies, in particular smartphones, reintroduces the topic of whether a new technology that seems to reduce consumer search costs may change the market structure and consumer behavior. Smartphone penetration rates have rapidly increased from 3% (2009) to 45% (2012) in the U.S., and similar growth rates are observed in other countries. In this paper, I study the effects of the increased usage of mobile technologies and the real time price information service on market outcomes.

The Korean gasoline retail market after 2008 is an excellent place to study the effects of search cost reduction on market outcomes for three reasons. First, gasoline is a fairly homogeneous good. Second, the Korean government has published free, real-time price information for all gasoline stations in the nation since 2008. As a consumer can access price information of all gas stations in the region of interest within 20 seconds via the Internet, his search cost is much lower than before. Furthermore, this price information service could be a paradigm shift in search instead of a mere search cost decrease, as most consumers previously had very limited price information and did not search at all before this price information service. Third, a rapid increase in mobile technologies helps to measure the effects of the price information service. While it is easy to check prices from the Internet, drivers usually choose which gas station to go to when they are on the road and running low on gas, instead of checking the prices before driving. Mobile devices, such as smartphones and in-car navigation systems that display gasoline prices, may significantly change the situation, as drivers can utilize the price information service whenever they want to. Moreover, there is enough variation in smartphone user population to estimate the effects. The smartphone penetration rate in Korea was below 5% before 2010 and reached 50% during 2012.

In addition, a combination of the unusually rich data environment created by the Korean government and private datasets that I could utilize provides an unique opportunity to examine market functioning. The Opinet information service offers daily retail prices of gas stations, station characteristics information, and weekly, national average distribution costs for all four major gasoline companies. Furthermore, I was able to obtain private daily sales data for select gasoline stations and one major telecommunication company. Since previous literature computes price dispersion and markup using only observed prices,[1] I compute more realistic measures that take into account quantity differences. Quantity-weighted price dispersion and markup measures ensure that included prices are real prices at which transactions occur, and prices with more frequent transactions are weighted more. Also, having a full set of station-level prices allows me to perfectly measure market-level price dispersion. Documenting stylized facts, I find an interesting phenomenon: price dispersion and markup levels did not decrease over time, even though consumer search costs decreased significantly. Graphs of all price dispersion measures and both unweighted and quantity-weighted markups show non-decreasing trends. To account for all factors which could drive price dispersion and markup measures, I run reduced-form regressions to examine these measures.

Regression results suggest that both price dispersion and markups slightly increase as smartphone penetration rates increase. This might be counter-intuitive, as it could be natural to assume that higher smartphone penetration rates would lead to more consumer search that would make stations compete more, and hence price dispersion and markup would decrease. The Opinet website and smartphone application usage trend confirms that more consumers uti-

[1]Mainly due to the lack of quantity data.

lized the price information service as smartphone penetration rates increased, which implies that the proportion of informed consumers increased. In addition, descriptive tests suggest that consumers became more price-sensitive, and the distribution of gas stations' markups moved toward a bimodal distribution. Intuitively, as the ratio of informed consumers grows, some stations set low prices to attract informed consumers, and some stations set high prices to serve less informed consumers. I find that search theory models could be consistent with these results. Many previous papers (see, e.g., Stigler, 1961; Diamond, 1971; Salop and Stiglitz, 1977; Varian, 1980; Morgan and Manning, 1985; Stahl, 1989 and 1996; Sorensen, 2000; Brown and Goolsbee, 2002; Baye et al., 2004; Clark and Houde, 2013) point out that the prices of homogeneous goods are quite dispersed and this price dispersion can be an equilibrium when there are some consumers who observe several prices while other consumers learn only one price. Search theory models with two types of consumers, where one type is more informed than the other, find that while there is no pure strategy Nash Equilibrium, there exists a symmetric mixed Nash Equilibrium where firms choose prices from an atomless distribution. In particular, it predicts a bimodal distribution of prices as firms either aim to be the lowest to attract informed consumers or to be the highest to get the maximum profit per customer. For instance, the Stahl model predicts that price dispersion increases when the ratio of informed consumers increases, as long as the ratio is below the critical value (Pennerstorfer et al., 2014).

Motivated by these descriptive results, I construct and estimate a discrete-choice demand model that includes both more and less informed consumers and a distance term that has different values for consumers at different locations. The structural model explains consumer choices and substitution patterns, and identifies a proportion of smartphone users who actually search for gasoline prices. These estimates are in line with the descriptive evidences: they indicate that the fraction of highly informed consumers has increased from 1.7% to 11.4% over the sample period, and that consumers have become more sensitive to prices as a result. A counterfactual analysis tells us what the new equilibrium prices would be if the informed consumer ratio changed. The structural model estimates that the increase in the fraction of informed consumers would be expected to result in an 0.57% increase in price dispersion, an 0.33% increase in the average markup, and an 0.09% increase in the quantity-weighted average markup. These findings are consistent with the descriptive analyses and the observed trends from the data. Lastly, the model interprets the effect of the distance between consumers and stations as how much more an average consumer is willing to pay to avoid driving an additional distance to visit a different gasoline station.

This paper attempts to measure the effects of new technological advances. While smartphones have become indispensable in our daily lives, no other research that I am aware of studies the impact of smartphones on the search environment. The Korean retail gasoline market is a quasi-experimental field where the effect of this new technology on search can be evaluated. Since smartphones combined with the real-time price information service directly change the consumer search situation and the subsequent decision, this paper helps us measure the effects of price information on retail market outcomes and how smartphones increase them. While there is no previous research that examines the effects of smartphone introduction on market outcomes, there are two empirical results about the introduction of mobile phones in developing countries. Jensen (2007) studies fish prices in several Indian towns and finds that the adoption of mobile phones by fishermen and wholesalers led to a significant reduction in price dispersion and an increase in social welfare. Aker (2008) also reports that the introduction of mobile phone service between 2001 and 2006 explains a 10 to 16 percent reduction in grain price dispersion in Niger. One difference, though, is that the effects of mobile phones and of smartphones are quite different. A mobile phone search, or making a phone call, provides one price quote at a time and it usually requires non-negligible time. On the other hand, a price search using a smartphone provides all the gasoline price information with a single search, and that single search only costs several seconds. Thus, the introduction of smartphones should not be interpreted as a mere decrease of search costs; rather, it changes the paradigm of searching and divides consumers into two groups: one with the full information, and the other with limited information.

I also contribute to the empirical literature on retail gasoline demand. Gasoline retail pricing has been studied in depth. Studies of price levels at individual stations have considered local market characteristics, region or time fixed effects, and individual station characteristics; see, for example, Barron et al. (2004), Eckert and West (2005), Hosken et al. (2008), and Lewis (2008) . In general, the findings have been mixed. In many cases, the impact of station characteristics on price is fairly small, and higher local station density implies lower price level and lower price dispersion. Also, there are some papers that study how search affects retail prices. Lewis and Marvel (2011) find that consumers search more as prices rise than they do when prices fall. As a result, when prices rise, margins are lower and there is less price dispersion. Similarly, Chandra and Tappata (2011) find that price dispersion increases with search costs. This paper provides an unusual case that station characteristics have significant effects on prices, and that price dispersion does not decrease with lower search costs.

Finally, this paper contributes to the empirical literature on search. While there are considerable theoretical results, relatively little empirical research has focused on measuring search costs and the effects of consumer search behavior changes in practice. As my empirical setting allows a two-type consumer search environment and two-type search costs, estimating search cost distribution is simplified to finding a proportion of informed consumers and this assists in understanding the change of search cost distribution easily. Recently, Hong and Shum (2006) presented structural methods to estimate search cost distributions using price data alone, and using similar methods, Wildenbeest (2011) focuses on the grocery markets in the United Kingdom and concludes that most of the observed price dispersion is explained by supermarket heterogeneity rather than search frictions. Yet I do not attempt to estimate the whole search cost distribution, nor to rationalize consumer search behavior. I focus on studying changes in market outcomes as a result of consumer search behavior changes.

The remainder of this article is organized as follows. In the next section I provide details about four data sets and discuss their merits and limitations. In section 1.3, I document stylized facts and present summary statistics. I construct regression models that estimate the effects of increasing smartphone penetration rates on price dispersion and gasoline station price levels. Section 1.4 explains consumer search behavior assumptions and the structural model setups. Structural estimation results and counterfactual analysis are reported in Section 1.5. Section 1.6 presents conclusions.

1.2 Data

In this paper, I combine four datasets: the Opinet price and station characteristics data; average wholesale prices; gasoline quantity data for select stations; and regional smartphone penetration rates. From the Opinet information service, I gathered daily retail prices and station characteristics of all gas stations in four regions. I utilized weekly national average distribution costs for all four major gasoline companies from the Opinet website to approximate the actual marginal costs. In addition, I obtained private daily sales data for select gasoline stations that are essential to perform daily level demand estimation and compute realistic price dispersion measures. Lastly, I use regional smartphone sales data to infer smartphone penetration rates.

1.2.1 Opinet Data

In 2008, the Korean government established an unusual resource for its citizens. Since April 15, 2008, every gas station in South Korea has been required to report its posted gasoline and diesel price at least once a day. Korea National Oil Corporation, a public institution, is in charge of collecting and publishing the real-time price information. I took advantage of this uncommon opportunity and scraped data from the Opinet website.[2] I obtained a complete set of historical daily prices for all gas stations in four regions from January 2010 to June 2012 (909 days) when mobile devices such as smartphones were diffusing rapidly. In particular, the smartphone penetration rate was nearly zero in January 2010, but went over 50% in June 2012.

Among the four regions, two regions are districts of Seoul, and the other two regions are small cities that are isolated by mountains.[3] The streets are laid out in a grid pattern for all regions, especially for the two districts of Seoul. On average, each region has 40 gasoline stations, and the four major gasoline companies have more than a 95% market share in total.

Each gas station owner can update the price information by calling the Opinet office, submitting the information on the Opinet website, or using an automated report system. According to an Opinet representative, most gas station owners use automated systems: for each credit or debit card transaction, price information for that transaction is electronically reported to

[2] I also contacted Opinet officials to supplement missing data.

[3] Except for two districts of Seoul, distances between any two regions are over 50 miles.

the Opinet server and price information is automatically updated.[4] Consumers can access the price data via the website www.opinet.co.kr or other methods, such as car navigation systems or Opinet smartphone applications.[5] In addition to the price information, Opinet also provides station characteristics such as location, self-service or full-service, car wash, repair shop, and convenience store availability. I collected these data as well for all of the stations in my sample.

1.2.2 Average Wholesale Price

Since most, if not all, of Korean oil imports come from Middle Eastern countries, the raw price Korean oil companies pay when they buy crude oil closely follows Dubai oil price futures. After importing crude oil, oil companies refine it and distribute gasoline to retailers. According to the gasoline station owners, each gasoline company sets a base distribution price once a week, and most transactions between the company and stations are made at that price during the week. The Opinet website publishes these weekly national average distribution costs for all four major gasoline companies. From now on, I call the distribution price AWP, or average wholesale price. These average wholesale prices are used as an approximation of the actual marginal costs.

1.2.3 Gasoline Quantity Data

Quantitative data such as the number of smartphones sold for each region and the quantity of gasoline sold for each station are very difficult to gather.[6] For gasoline quantity data, I have daily credit card transaction numbers of select individual gas stations of one major gas company that has about a 30% market share: for about 20% of the total gas stations, I have the quantity information. According to the company official, credit cards are used for most of the transactions, and the proportion of credit card transactions has been stable and similar for all four regions during the period 2009-2012. Moreover, since the average transaction amount has been stable during the period, either using daily transaction numbers or using daily transaction

[4] While the Opinet website advertises that it offers real-time data, it does not update new price information every time transaction information comes in. Instead, it updates six times a day, based on price information during each time period. However, as gasoline stations rarely change prices more than once a day (change once in a week on average), updating six times a day basically provides the real-time information.

[5] As the name of this service is usually called "Opinet", following the name of its website, I will use the term "Opinet" to represent this real-time price service and the name of institution that provides the price data.

[6] I thank anonymous company officials who provided quantity data of one major company in each sector.

amounts naturally delivers very similar results.[7] I present the results from using daily transaction numbers for the main analysis.[8]

Having quantity data provides two main advantages. First, quantity information is essential for any demand estimation. Second, more economically relevant statistics can be derived. In many markets, especially online markets, we cannot distinguish "real" prices at which transactions actually occur from unrealistic price listings that will never result in a sale. In addition, even after limiting the price listings to the ones at which trades are made, estimating price dispersion and average price without using the quantity information could be misleading.[9]

1.2.4 Smartphone Penetration Rates

I utilize data from a representative company in the telecommunication sector that has constant market share during the sample period to estimate total smartphone penetration rates. To compute daily smartphone penetration rates, I start from the quarterly number of regional smartphone users for one telecommunication company. This company is one of the three major telecommunication companies and its market share had been 30-35% during 2009-2012. Multiplying the number of users for this company by three, I estimate the total number of smartphone users of the region for each quarter. In regressions run at the daily level, I use a linear interpolation method to infer daily smartphone penetration rates.

Smartphones have become increasingly popular in the Korean market since late 2009. Before the introduction of iPhone 3G in November 2009, less than 1% of the population used a smartphone. Within 3 years, the smartphone penetration rate (the ratio of the number of smartphone users to the total population) went over 50%. Figure 1.1 shows smartphone penetration rates

[7] Transactions with more than one million Korean Won (about 1000 USD) are treated as outliers and excluded.

[8] All gasoline stations had 200 or more transactions per day on average. When a station had 50 or fewer transaction numbers in a day, that station-day observation is excluded from the sample, under the assumption that either a card reader system did not work properly or the station opened for a very short time on that date.

[9] For example, Amazon.com's book listing for Jean Tirole's "The Theory of Industrial Organization" on October 20, 2014, contains 30 prices, ranging from \$69.54 (plus \$3.99 shipping) to \$263.64 (plus \$3.99 shipping). Among the thirty listings, eight listings have prices \$120 or higher. Since it is unlikely that anyone would pay \$120 or more, including these prices in the consumer choice set would overestimate price dispersion and average price. Moreover, even if we can identify the set of "real" prices, having quantity data is essential. For simplicity, suppose that these are two prices, \$70 and \$120, and 99% of consumers pay \$70 and 1% of consumers end up paying \$120. Quantity-weighted price dispersion gives a correct picture of actual price dispersion, which is close to zero (as the quantity-weighted price is close to \$70), while unweighted price dispersion does not properly reflect the market situation.

of four regions during the data period. Note that regions 1 and 2 have the same rates, as these two regions are neighboring districts of Seoul and the company treated them as a single region when they collected sales data. While all regions have increasing trends, there are some regional differences.

1.3 Descriptive Analyses on Price Dispersion and Markup

In Section 1.3, I present reduced-form evidence on price dispersion and markups. The most basic observation is that both price dispersion and markups do not decrease over time. Regression results suggest that higher smartphone penetration rates lead to slightly higher price dispersion and average markups. To investigate whether these relationships are causal or just because consumers do not use price information, I perform additional analyses: usage statistics; a demand regression to look for price sensitivity changes; and a test to check whether stations' markup distributions are bimodal.

1.3.1 Definitions and Summary Statistics

Table 1.1 defines the variables used in the analysis. *RetailP, AdjustedP*, and *Mkup* are retail gasoline prices, adjusted gasoline prices, and markups of gasoline stations. *AvRetailP* (*AvMkup*) is unweighted daily regional average retail prices (markups), and *QwRetailP* (*QwMkup*) is quantity-weighted average retail prices (markups). *AWP* stands for average wholesale prices (see section 1.2.2), and *SmartPen* denotes the ratio of smartphone users to the region's population, or the smartphone penetration rate, as described in section 1.2.4. *Self, Carwash, Repair,* and *Store* are station characteristic dummy variables. Since station characteristics remained constant for the sample period, these variables do not have t subscript.[10]

For the initial descriptive analysis, I aggregate my station-level data up to the regional level and compute several measures of price dispersion which vary at the region-day level. The last four variables of Table 1.1 *(Range, Std, IDR,* and *IQR)* are price dispersion measures that have two indexes: r stands for region and t for time. In addition to the standard measures

[10] There are two cases that station characteristics changed for a short period of time. According to an Opinet official, it is likely that these are input errors.

of price dispersion, *Range* and *Std*, I also consider interdecile range (*IDR*) and interquartile range (*IQR*) to investigate the characteristics of price dispersion in detail. When calculating price dispersion, I use *AdjustedP*, prices adjusted for differences in station characteristics. I run a regression with day fixed effects, $RetailP_{jt} = \beta_0 + \Sigma\delta_t Day_t + \Sigma\gamma_j X_j + \epsilon_{jt}$, and define $AdjustedP_{jt} = RetailP_{jt} - \Sigma\hat{\gamma}_j X_j$, where X_j are station characteristics.

Table 1.2 presents summary statistics for all variables. I converted the data from Korean won per liter to U.S. dollars per gallon to aid interpretation.[11] While the mean and standard deviation of the quantity-weighted prices are very similar to the unweighted prices, the quantity-weighted markups are lower than the unweighted ones.

Summary statistics of the four price dispersion measures are shown in Table 1.3. The left side of Table 1.3 presents the unweighted statistics, and the right side presents descriptive statistics of the four dispersion measures that are computed with quantity-weighted data. For example, suppose that station 1 makes 2 sales at price p_1 and station 2 makes 6 sales at price p_2. For the quantity-weighted case, I consider that there are 2 entries of p_1 and 6 entries of p_2. The unweighted case assumes that two prices have equal proportions, and hence there are 4 entries of p_1 and p_2. For instance, if $p_1 = 2$ and $p_2 = 1$, then the unweighted standard deviation is 0.535 and the quantity-weighted standard deviation is 0.463. As expected, Table 1.3 confirms that the quantity-weighted dispersion measures have smaller values.

1.3.2 Average Markup and Price Dispersion Trends

I present average markup and price dispersion trend graphs and a brief interpretation of the graphs in this section.[12] Figure 1.2 shows trends of the four dispersion measures. None of the four measures decreases over time and they actually slightly increase in the second half of the sample period. Quantity-weighted dispersion measures also do not decrease over time; Figure 1.3 presents quantity-weighted and unweighted standard deviation measures. Note that the

[11] I used the following conversion rates: 1 gallon = 3.785 liter; 1 dollar = 1000 Korean won.

[12] Since there were nominal retail price differences between gasoline companies from 4/7/2011 to 7/7/2011, I treated that period separately (details in Appendix A). During this period, the Korean government asked the four major gasoline companies to cut their prices, and they agreed to reduce retail prices by 100 Korean won per liter. However, one gas company (SKE) chose to offer a rebate of 100 won per liter, instead of cutting the posted price directly. This policy caused artificial relative posted price differences between SKE stations and non-SKE stations.

quantity-weighted dispersion is lower than the unweighted one, as quantity-weighted dispersion reflects that stations with lower prices tend to make higher numbers of sales.

In addition, both quantity-weighted and unweighted markup levels remain steady (Figure 1.4). As stations with lower prices tend to have lower markups and sell more, it is not surprising that the quantity-weighted markups are lower than the unweighted markups. Also, the differences between two measures are fairly stable during the period.

There are two hikes: early 2012 and mid-2012, the latter being the end of the sample period. For both cases, while a rapid decline in the international oil prices caused a sharp drop of AWP, retail prices fell only slowly. For the second hike, I find that average retail prices decreased in the next month (right after the end of the sample period) and average markups went back to the usual levels. These cases are classic examples of asymmetric price adjustments, or "prices rise faster than they fall" (Peltzman, 2000).[13] For instance, AWP fell about 10% (about 64 cents per gallon) during the last four weeks of the sample period, but average retail prices fell only 3%.

1.3.3 Measures of Price Dispersion and Markup Changes

The graphs in the previous section were suggestive, but do not account for all factors which could drive price dispersion and markups. This section, therefore, will examine these quantities in a series of reduced-form regressions. These regressions are intended as descriptive in nature. I do, however, assume that smartphone penetration rates are exogenous with respect to gasoline price dispersion or markups of the stations. This assumption seems fairly innocuous – it is unlikely that people buy smartphones because gasoline price dispersion or markups are low or high, or are even aware of them for that matter.

As described in the section 1.3.1, I aggregate my station-level data up to the region level and compute several measures of price dispersion which vary at the region-day level. The base regression format is

$$PD_{rt} = \beta_0 + \beta_1 AWP_t + \beta_2 SmartPen_{rt} + \sum_{j=1}^{11} DM_j + \sum_{k=1}^{3} DR_k + \epsilon_{rt}$$

[13]Many previous papers reported asymmetric price adjustments in the gasoline market. See, e.g., Borenstein et al. (1997), Noel (2007), Verlinda (2008), and Lewis (2011).

The dependent variable, PD_{rt}, is one of the four dispersion measures. $SmartPen_{rt}$ is a smartphone penetration rate at time (date) t for region r, and AWP_t is an average wholesale price of gasoline at date t.[14] DM_j are month of the year dummies and DR_k are region dummy variables. In Table 1.4, I present regression results with the dependent variable *Std*, standard deviation. Results from the unweighted data are presented in columns 1 and 2, and results from the quantity-weighted case are presented in columns 3 and 4. Both month of the year dummies and region dummies are included in all regressions, and columns 2 and 4 have a time trend term included.

Positive *SmartPen* coefficients imply that dispersion is higher when there are more smartphone users. The regression results from the column 1 suggests that the standard deviation decreases by 0.252 cents per gallon when the smartphone penetration rate increases by 1%.[15] Including the time trend decreases the magnitude of *SmartPen* coefficients, but the significance levels do not change. Negative signs for the *AWP* coefficients mean that when marginal costs are higher (hence price levels are higher), price dispersion is smaller. For example, the results from column 1 suggest that the standard deviation decreases by 4.4 cents per gallon when the average wholesale cost goes up by one dollar per gallon.

As a robustness check, I present the regression results for all four price dispersion measures in Table 1.5. For the unweighted data, all coefficients are highly significant and have the same signs. For the quantity weighted case, coefficient estimates are less consistence across dispersion measures. This finding suggests that quantity-weighting might be important.

Since we are also interested in evidence on markups, not just price dispersion, so I also consider the following regression for completeness:

$$Mkup_{rjt} = \beta_0 + \beta_1 AWP_{jt} + \beta_2 SmartPen_{rt} + \sum_{i=1}^{4} DC_i + \sum_{j=1}^{11} DM_j + \sum_{k=1}^{3} DR_k + \sum_{k=1}^{4} DChar_{kj} + \epsilon_{rjt}$$

[14] Since it is possible that stations make purchases in advance and in fact pay *AWP* from a week or two weeks before, I also consider *AWP* a week before and *AWP* two weeks before as additional variables. Regression results show that there are no big differences, so I only report the regression with the *AWP* term.

[15] This result is contrary to the previous research on the effect of mobile phones (Jensen, 2000 and Aker, 2007). But as discussed before, the mechanism for potentially affecting price dispersion is much different here than in the previous cases.

The dependent variable, $Mkup_{rjt}$, is markup of station j at time t for region r. $DChar_{kj}$ is a set of dummy variables representing characteristics k of station j. For instance, if station j offers full service, $DSelf_j = 0$. Similarly, *Carwash* denotes whether a station has a car wash, *Repair* denotes whether a station has a repair shop, and *Store* denotes whether a convenience store is located at the station.

Column 1 of the Table 1.6 shows the results of the above regression, and column 2 includes the time trend term t. Results of the two columns are almost identical except for the *SmartPen* coefficients. Since *SmartPen* is increasing over time during the data period, including the time trend term decreases the *SmartPen* coefficient. *SmartPen* has a positive coefficient and implies that a markup is 0.344 (0.266 for column 2) cents per gallon higher when a smartphone penetration rate is 1% higher.[16] *AWP, Self,* and *Store* coefficients are significant and have expected signs: higher markups are expected when the marginal cost is lower, when a station offers full service, or has a convenience store.[17]

Combined with the previous dispersion regression results, this descriptive analysis suggests that higher smartphone penetration rates are associated with higher markup levels and higher price dispersion. These results seem counter-intuitive: it is natural to assume that higher smartphone penetration rates would lead to more consumer search, and more price information would intensify competition, hence lower prices and price dispersion. There are two possible explanations. First, consumers do not, or seldom, use smartphones to search for the price information. If this were the case, the increase in the smartphone penetration rate does not necessarily change the search intensity and the proportion of informed consumers. Second, consumers search more, but gas stations change their pricing strategies since they face different demand. As a result, we reach a new equilibrium, and price dispersion and markups do not decrease. In the next section, I present three additional descriptive analyses that suggest changes in consumer search behavior and gas stations' pricing strategies.

[16]Since *AWP* is the same for all stations belonging to the same gasoline company, and the differences of *AWP* between gasoline companies are small, we would expect the markup regression results to be very similar to those for *RetailP*, and they are. The *RetailP* regression results also suggest that *RetailP* is increasing in *SmartPen*.

[17]While *Carwash* and *Repair* terms are not significant, their signs can also be interpreted. A station with car wash is more desirable for consumers, and hence may charge higher price without losing consumers. For a station with a repair shop, it tries to attract consumers to the repair shop where profit margin is high. It is possible that this type of station uses gasoline prices as loss-leaders.

1.3.4 Additional Descriptive Evidence

Before presenting the structural model, I offer three types of suggestive evidence that consumers indeed searched more during the sample period, and gas stations changed their pricing strategies as a result. I document stylized facts about the Opinet price information service usage to find that a stable fraction of consumers and smartphone users have utilized the Opinet service. I also discuss a simple regression model to confirm that consumers became more sensitive about the minimum prices. Lastly, I adduce that gas stations tend to charge either high prices or low prices over time, using a test for unimodality.

Opinet Usage Has Increased

In the data section, I presented a graph that shows a dramatic increase of smartphone users and discussed how Opinet real-time price information is provided. However, it is not certain what the combined effect of these two is. I start with the basic question that is motivated from the regression results: do people actually use smartphones to access the Opinet service? An even more fundamental question is: how many people use the Opinet service?

Figure 1.5 shows trends of four variables. The variable labeled *Opinet Web/Car* shows the number of Opinet website visitors for each week as a fraction of the total number of registered cars in the nation. It is slightly above 2% before 2011 and about 5-6% during 2011 and 2012. If we assume consumers search for gas prices only when they fill up and it happens once per week on average (Byrne and Roos, 2014), then on average, 5-6% of the gas purchasers use Opinet to become highly informed without using a smartphone.

The *Total App/Car* variable, the ratio of the number of Opinet smartphone application (henceforth "app") downloads for each month (cumulative) over the number of cars, clearly shows an increasing trend. It has increased rapidly after 2010: it starts from less than 1% in December 2010, and reaches 12% at the end of the sample period. If we assume that all drivers who downloaded the Opinet app are informed and who did not download are not, this trend suggests that about 12% of drivers are highly informed at the end of the period.

iPhone App/iPhone represents the ratio of the number of Opinet iPhone app downloads to the number of iPhone users. The iPhone Opinet app was first introduced in May 2010, and the Android Opinet app was available from January 2011. After the initial release, about 5% of iPhone users downloaded the app, and this ratio remained fairly stable until January 2011. When Opinet also launched the Android version app and started advertising, it seems that more iPhone users became aware of the Opinet app: during 2011 and 2012, about 9.5% of iPhone users downloaded the app. The *Total App/Smartphone* variable, the proportion of the total Opinet app download numbers to the total number of smartphone users, shows a similar trend.[18]

In summary, these graphs imply that (i) the Opinet website received a steady stream of visitors, (ii) the proportion of smartphone users who downloaded the Opinet app was relatively stable, and (iii) with the rapid increase of smartphone penetration, the number of Opinet app downloads also increased.

While these results suggest that consumers actually utilized the Opinet service more frequently, and more consumers searched by using the service as the smartphone penetration rate increased, it should be noted that there are also other important channels through which the Opinet data have additional impacts. When the Opinet launched its Android version app in January 2011, it also started advertising the Opinet service and provided most of the Opinet information to other websites such as naver.com (Korean version of Google.com) and car navigation systems. While separate data for other methods that also provide the Opinet price information service are not available, Opinet officials suggested that the access rate trends for the other websites would be similar to that of the Opinet website. I construct estimates of the impact of Opinet as a function of smartphone penetration rate. Estimated effects should be thought of as reflecting the combined impact of the various channels, not just smartphones.

A Simple Demand Regression

In addition to the Opinet usage, I estimate a crude demand equation to suggest that consumers actually did search more in the latter periods when there were more smartphone users. If consumers search more, consumers will be more sensitive to the difference between its price and

[18]Since there was no Android or iPhone app available before May 2010, it is zero. From May 2010 to December 2010, there was only the iPhone app and the ratio remained stable at 1.5%. When the Android version was introduced, this ratio jumped to 8% and stayed 7-8% for the rest of the period. As the number of smartphone users was rapidly increasing, a stable *Total App/Smartphone* implies that the number of the app users was increasing also.

the minimum price. To examine this conjecture, I estimate a regression with station fixed effects on the 2010 quantity data and separately on the 2012 quantity data:

$$ln\left(Q_{jt}\right) = \alpha_1 p_{jt} + \alpha_2 \left(p_{jt} - p_{min,t}\right) + f_j + \epsilon_{jt}$$

α_1 captures the effect of own price increases on the quantity sold, α_2 represents the effect of the difference between its price and the minimum price of the region at time t, and f_j are station fixed effects. Three instruments are used for the price and price difference terms: average wholesale costs of gasoline; the difference of the costs between station j and the costs of the lowest price station(s) in the market; and the sum of characteristics of other rival stations within 1 mile radius.

I emphasize that the main purpose of this regression is not to identify the exact effect of own price changes or relative price changes, but to show how these coefficients change over time. In particular, I compare estimates from the first year (2010) and ones from the last year (2012) at Table 1.7. While α_1 does not change much, the magnitude of α_2 is higher for the latter period (α_2 is negative, so it means that α_2 decreases) and this difference is statistically significant. These results suggest that consumers became more sensitive to the price difference between station j's price and the minimum price during 2012, compared to 2010. In other words, these results suggest that there are more informed consumers for the latter period.

The Bifurcation of Markups

In the previous sections, I find that while consumers become more price-sensitive and search more, price dispersion and average markups do not change. One possible explanation is a polarization of stations as suggested by the Stahl model (1989, 1996). The first group of gasoline stations focus on the consumers who do not know the prices (no Opinet price information) and happen to visit these stations. They try to maximize profit given these uninformed consumers and concede consumers with price information. On the other hand, the second group of gasoline stations adopt a low-price high-volume policy and try to attract consumers with price information.

If this were true, we might be able to observe bimodality in a price or markup distribution graph. Since I use the same marginal cost (daily average wholesale cost) for all stations belonging to one gasoline company for a given day, and the differences between AWP of gasoline companies are fairly small, shapes of price graphs and markups graphs are similar. The first graph in the Figure 1.6 shows a kernel density of quantity-weighted markups for gasoline stations in region 3 at March 1st, 2010, and the second graph shows the one at March 1st, 2012. The 2010 graph shows little evidence of bimodality while the 2012 graph is suggestive of a double peaked distribution.

The figures shown above are just from a single day, but these features are robustly present across dates. To show this formally, I use Hartigan's Dip Test (1985), which tests for multi-modality in a sample by "the maximum difference, between the empirical distribution function, and the unimodal distribution function that minimizes that maximum difference". I compute proportions of days in a given period that a markup distribution of a given day is rejected to be unimodal, according to the test.[19] While the proportion of rejection rates of unimodality for unweighted markup distributions is low and does not vary significantly over time,[20] the proportion of rejection rates for quantity-weighted markup distributions shows an increasing trend. For example, average proportion of the rejection rates over the regions is 0.35 for 2010, 0.38 for 2011, and 0.56 for 2012: approximately 1/3 of the region-day observations reject the null hypothesis that a quantity-weighted markup distribution for a given region-day is unimodal at the beginning of the period, and more than half of the region-day observations reject the test at the end of the period. These results suggest that there were changes in the quantity-weighted markup distributions but not in the unweighted markup distributions, while average values for both markup measures remained stable during the period (the trend graphs in section 1.3.2).

1.4 Structural Model

While the descriptive analysis section suggests that consumers have become more price-sensitive and searched more, it also reports that both price dispersion and markup did not decrease. Structural model can help reconcile these descriptive results that seem counter to our naive intuition and provide more complete equilibrium pictures. Motivated by these results, I develop

[19] I calculate p-values for each region-day observation, under the null hypothesis that a given distribution is unimodal.

[20] These rejection rates are close to zero, as p-values for the test are higher than 0.2 for most of the time.

a structural model of consumer choice to estimate demand to answer the following questions: (i) Do mobile technologies and the price information service actually affect the proportion of informed consumers? If yes, then how much?, (ii) What are the effects of the factors related to the consumer choices in the retail gasoline market?, and (iii) How would gasoline stations adapt their pricing strategies in reaction to consumer search behavior changes?

In order to answer these questions, I study how the ratio of informed consumers changes, as smartphone penetration rates change, to measure the effect of the Opinet price information service on the market. As it is costly to gather prices by visiting gasoline stations, consumers who do not use the Opinet service typically have very limited price information. On the other hand, consumers who utilize the Opinet service may obtain all price information at once. Reflecting this price information distribution, I assume that there are two types of consumers (informed consumers and uninformed consumers). By estimating a modified version of a random-coefficients discrete-choice demand model (Berry et al., 1995, henceforth BLP) of the two types of consumers, I explain the effects of important factors such as prices and distances from consumers to stations. While I use the same taste parameters for both consumer types as there are no obvious differences between the two types except the amount of price information, consumers have different locations, and hence different distances to gasoline stations. Combining expected market shares from both types, I compare the total expected market shares with the observed market shares to find out the parameter values that minimize the objective function. In particular, I focus on the price sensitivity term, the distance sensitivity term, the proportion of informed consumers who search without smartphones, and the ratio of smartphone users who do search to get full price information.

1.4.1 Consumers on the Grid

To build a structural model of consumer search and purchase behavior, it is important to consider and incorporate factors that drive consumer choice in the retail gasoline market. There are three main factors: price, distance, and station amenities. Distance to a station is unique among these factors in horizontally differentiating gas stations: the value of a station's location depends on the consumers' location. Unlike other factors, such as the price at a station or whether that station offers a car wash, that are the same for all consumers, the distance to a certain station is different for consumers from different locations.

To consider how distances would affect consumer choices, I start with a simple model of location. Imagine a rectangle that covers the whole region. Divide it into $n-1$ by $m-1$ rectangles so that we have an n by m grid. I assume that consumers are located at each grid point according to a uniform distribution. For example, if $m = n = 10$ then each grid point has 1% of the consumers. Each of the four regions are approximate squares, so I employ square grids in each case.[21]

In contrast, Houde (2012) was able to use commuting patterns to assign consumer locations in a model of gasoline demand. He divides Quebec City into a grid and decomposes consumers into four components: workers, full-time students, unemployed, and outside commuters. Assuming non-congestion for travel paths, he computes the probability of commuting routes using road network and census data. However, the uniform distribution assumption without considering commuting patterns is realistic in my case, especially for two regions that are districts of Seoul. Figure 1.7 shows a map of one of the districts. As the streets are laid out in a grid pattern, the grid assumption reflects the actual road map well. Moreover, there are no clear commuting paths in this region as (i) residential and commercial areas are mixed, (ii) almost all roads are congested during rush-hours, and (iii) there are multiple ways to enter and exit the region.[22]

1.4.2 Two Types of Consumers: Informed Consumers

There are two types of consumers, informed consumers and uninformed consumers.[23] While informed consumers know all station information by using the Opinet service, uninformed consumers only have limited information. For example, if someone starts to look for a gas station when he runs low on gas and chooses to go to one of the first two stations he finds, he is an uninformed consumer. Since uninformed consumers do not have information for most gas stations,

[21]Region 1 and 2 are $9\times$ 9 miles and Region 3 and 4 are $10\times$ 10 and $7\times$ 7, respectively.

[22]While it would be ideal to have traffic volume information to estimate demand distribution more precisely by putting different weights on different grid points, it is almost impossible to measure the volume of traffic separately for the roads where stations are located.

[23]The two-type consumer assumption is applied in other empirical settings. One example is an online computer memory chips market where Moraga-Gonzalez and Wildenbeest (2008) find that the consumer population can be split into two groups which either have high search cost or low search cost.

or "products", their choice sets are limited. On the other hand, if someone uses her smartphone to check all the stations in the region and makes a decision to visit a certain station, she is an informed consumer. Since using the Opinet price service will provide all information (including gasoline prices) with a single search in several seconds, I assume that consumers who use the Opinet service gets all information without any cost.[24]

For informed consumers, following standard utility assumptions, the indirect utility of consumer at i for station (or product) j in market t is given by

$$u_{ijt} = d_{ij}\beta_d + X_j\beta_c - \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

where d_{ij} is a distance (in miles) from consumer at i (grid point i) to the station j.[25] X_j are observable station characteristics, β_c is a vector of consumer taste coefficients for station amenities, β_d is a distance coefficient, and α is a price coefficient (p_{jt} is price of station j in market t). Lastly, $\xi_{jt} = \xi_j + \Delta\xi_{jt}$ are unobserved station characteristics. As Nevo (2001) suggests, I include station-specific dummy variables as unobserved (by the econometrician) station fixed effects ξ_j. Market-specific unobserved components are included in $\Delta\xi_{jt}$ and are left as "error terms".

Market t is one day of a certain region. For example, Gangnam District, March 2nd, 2011 is one market. As the data from January 1, 2010 to June 27, 2012 are chosen, the number of markets is 909 times the number of regions. Note that this utility setup is a special case of the BLP model where consumers have different values (denoted by the subscript i) only for the $d_{ij}\beta$ term, except for the separable additive random shocks. As consumers are more likely to substitute toward stations that are close to each other, location of consumers with respect to stations is an important source of heterogeneity between consumers.

ϵ_{ijt} is an i.i.d random utility shock distributed according to a Type I extreme-value distribution. Then, the percentage of consumers at i who chooses station j in market t is[26]

[24]This single search assumption is used in previous literature, including the retail prescription drugs study of Sorensen (2001). He uses transactions data of prescription drugs to estimate a discrete-choice demand model that embeds a simple search decision. In particular, he assumes that search is "all or nothing": consumers either search exhaustively to learn all pharmacies' prices, or not at all.

[25]$d = 1$ means that a distance between a consumer and a station is one mile.

[26]In this setup, choosing the outside option at market t means that a consumer does not go to gas station at date t. As usual, I normalize the utility from the outside good to zero.

$$\frac{\exp\left(d_{ij}\beta_d + X_j\beta_c - \alpha p_{jt} + \xi_{jt}\right)}{1 + \sum_{j=1}^{J} \exp\left(d_{ij}\beta_d + X_j\beta_c - \alpha p_{jt} + \xi_{jt}\right)}$$

Note that in this setup, choosing the outside option in market t means that a consumer does not go to a gas station in market t. Let w_i denote the proportion of consumers at i in the population (for the base model, all $w_i = \frac{1}{n^2}$). The overall market share of station j is

$$s_{jt}^{informed} = \sum_i w_i \frac{\exp\left(d_{ij}\beta_d + X_j\beta_c - \alpha p_{jt} + \xi_{jt}\right)}{1 + \sum_{j=1}^{J} \exp\left(d_{ij}\beta_d + X_j\beta_c - \alpha p_{jt} + \xi_{jt}\right)}$$

1.4.3 Two Types of Consumers: Uninformed Consumers

To explain how consumers search without the price information service, a consumer search behavior assumption plays an important role in the model setup. The theoretical literature typically models consumer search in two ways: the fixed sample size search model, where consumers sample a fixed number of stores and choose to buy the highest utility one, and the sequential search model, where consumers decide to search one more if the expected benefit from the next search is higher than the search cost. Both types certainly could be applied in this case. One could argue that when a driver needs gas, he would try to look for several stations nearby and decide where to go. One example is a driver who observes several prices on the way to work and chooses one of them on the way back home. On the other hand, it is possible that the driver sees the first gas station and checks the price and both observed and unobserved station characteristics, and decides whether to continue search or just stop by the station and fill his car up, depending on his expectation of prices of other stations and his own search cost.

Since both models are plausible, I chose a fixed sample search approach which is more straightforward to implement.[27] Let m be the number of stations whose information is known to an uninformed consumer. Choosing $m = 1$ would cause a station to set very high price in equilibrium so that it can make huge profit from uninformed consumers who only know its information

[27] When there is a delay between the search decision and the search outcome, Morgan and Manning (1985) have shown that a fixed sample size search typically offers a better explanation of observed behavior than a sequential search. Santos, Hortacsu, and Wildenbeest (2012) argue that fixed sample size search models provide a better explanation of observed consumer search behavior in online book stores than sequential search models. While I do not access which model fits better in this paper, this result motivates me to apply the fixed sample size search model for the case when consumers do not conduct a smartphone price search and are not aware of the prices.

and concede all other consumers.[28] Thus, I choose $m = 2$: each uninformed consumer learns $m = 2$ station prices among J stations in the market.[29]

To determine which station information consumer at i gets, I assume that a probability of getting price information of the station j is proportional to the inverse of the distance, d_{ij}. In other words, Pr (consumer at i learns the price of the station j_1) / Pr (consumer at i learns the price of the station j_2) $= \left(\frac{d_{ij_2}}{d_{ij_1}}\right)$. Solving these equations, I get

$$r_{i,k} = \text{ Pr (consumer at } i \text{learns the price of the station } j_k) = \frac{\frac{1}{d_{ij_k}}}{\Sigma_s \frac{1}{d_{ij_s}}}$$

Using the previous result, I compute the probability of consumer at i learns the price of station j_{k_1} and j_{k_2}.

$$\begin{aligned} r_{i,k_1,k_2} &= \text{Pr (consumer at } i \text{learns the price of the station } j_{k_1} \text{and } j_{k_2}) \\ &= r_{i,k_1} r_{i,k_2} \left(\frac{1}{1 - r_{i,k_1}} + \frac{1}{1 - r_{i,k_2}} \right) \end{aligned}$$

I assume that after learning prices for two stations, an uninformed consumer would choose a station in this restricted set the same way an informed consumer would. Using the same distributional assumption, a market share for station j among uninformed consumers is

$$s_{jt}^{uninformed} = \sum_i w_i \left(\sum_{l \neq j} r_{i,j,l} \frac{\exp\left(d_{ij}\beta_d + X_j\beta_c - \alpha p_{jt} + \xi_{jt}\right)}{1 + \exp\left(d_{ij}\beta_d + X_j\beta_c - \alpha p_{jt} + \xi_{jt}\right) + \exp\left(d_{il}\beta_d + X_j\beta_c - \alpha p_{lt} + \xi_{lt}\right)} \right)$$

1.4.4 Demand Side

As I only observe total quantity sold for each station, I need to combine the results from the informed consumers and uninformed consumers so that I can compute the total market share to match the observed values and the expected ones. I assume that the size of market t is the number of cars in the region at that time. Let IR_t, the informed ratio, denote the proportion

[28] Moreover, interviews with several drivers confirmed that consumers tend to avoid making choices when they are given only one price information, as they are not certain that whether the given choice is a complete rip-off.

[29] The estimation results are fairly robust to the choice of m: for example, $m = 3$ does not change the results significantly. Moreover, it is unlikely that a driver would observe only one station during his trip.

of informed consumers in market t. Intuitively, IR_t should increase as time passes, since the proportion of informed consumers is bigger when there are more smartphone users. I allow for a linear growth in the informed ratio to change every day by assuming $IR_t = a_0 + a_1 SmartPen_t$. a_0 represents the fraction of drivers who are already informed without smartphones, and a_1 indicates the proportion of smartphone users who become informed consumers.

Then the total expected market share of station j is

$$\hat{s_{jt}} = IR_t \cdot s_{jt}^{informed} + (1 - IR_t) \cdot s_{jt}^{uninformed}$$

1.4.5 Supply Side

In addition to the demand side, I incorporate the profit-maximizing conditions for gas stations. As a profit for station j in market t is $(p_{jt} - mc_{jt})\, q_{jt} - F_{jt}$ where F_{jt} denotes fixed costs, the optimal behavior for gas stations is to follow the profit-maximizing condition: $(p_{jt} - mc_{jt})\, \frac{\partial q_{jt}}{\partial p_{jt}} + q_{jt} = 0$. This first-order condition reduces to $p_j = mc_j - \frac{s_j}{\partial s_j / \partial p_j}$, and this requires the derivatives of the market share function with respect to price. Thanks to the analytic market share formula, I can compute $\frac{\partial s_{jt}}{\partial p_{jt}}$ explicitly (details in Appendix B).

For simplicity, I begin by assuming that the marginal cost is both independent of output levels and linear in cost characteristics: $mc_j = \gamma_{0,j} + \gamma_{1,j} AWP_j + \eta_j$. $\gamma_{0,j}$ is a fixed component that mainly consists of the transportation cost and the variable labor cost.[30] AWP_j is an average wholesale price for station j, and η_j is an unobserved cost shock. As AWP_j is what station j needs to pay today to replenish,[31] I set $\gamma_{1,j} = 1$. We have $\eta_j = p_j - (\gamma_{0,j} + AWP_j) + \frac{s_j}{\partial s_j / \partial p_j} = p_j - b_j(p, X, \xi; \theta) - W_j \gamma$ where $W_j = (1\ AWP_j)$ and $\gamma = (\gamma_{0,j}\ 1)'$. I assume that while the unobservable cost term η_j might be correlated with ξ_j, it is mean independent of the average wholesale cost (and the observable station characteristics).

[30] In fact, this variable labor cost is fairly similar among gas stations, as most of the gas station workers are temporary part-time employees who are paid a national minimum hourly wage or slightly above the minimum.
[31] Also, this is the expected future average wholesale price given today's information.

1.4.6 Instruments

In this section, I explain which instruments are used and what their identifying assumptions are. Following previous literature, moment conditions are formed by interacting instruments with the unobservable error terms.

First, I exploit the panel structure of the data and use the average prices of the stations in different regions that belong to the same gasoline company as instruments.[32] The identifying assumption for these instruments is that the average prices of the stations that belong to the same gasoline company in two regions are correlated due to the common marginal cost changes, but they are independent of the region-specific unobserved valuation changes, $\Delta\xi_{jt}$. Since the regions in the data are far apart and no national demand shock is likely, the average prices of another region are independent of the unobserved valuation changes in the region.[33] Let these instruments be Z_1. Then, $E(\Delta\xi_{jt}|Z_1) = 0$ and we have the usual moment conditions $E(\Delta\xi_{jt} \cdot Z_1) = 0$. Similarly, $E(\eta \cdot Z_1) = 0$.

Another set of instruments Z_2 are cost shifters, average wholesale prices for each gasoline company. As average wholesale prices directly change marginal costs but only affect demand through prices, I have $E(\Delta\xi_{jt}|Z_2) = 0$. In addition, I assume that this exogenous element of the marginal cost shifter is not correlated with the unobservable cost η. This assumption gives additional moment conditions $E(\eta \cdot Z_2) = 0$.

While Z_1 and Z_2 have excellent time-series variation, they have the same values for stations in the same gas company. The next set of instruments, Z_3, the sum of characteristics of other rival stations within one mile radius, provides the other type of necessary variation: they vary substantially by station.[34] The identifying assumption is that characteristics of other stations are correlated with prices as the markup levels are affected by these measures of isolation in the product space, and "location" of gas stations in the characteristics space is exogenous, or predetermined. Also, I assume that these observable characteristics are mean independent of the unobserved cost shock, as the effects of observables are reflected in the $b_j(p, X, \xi; \theta) + W_j\gamma$ part

[32] This is similar to the instruments used by Hausman (1996).

[33] Since region 1 and 2 are neighboring districts, I use region 3 and 4 prices as instruments for region 1 and 2. For region 3 and 4, all other region prices are valid instruments. As distances between regions are more than 50 miles except for the region 1 and 2, it is unlikely that consumers can switch to other regions.

[34] However, they do not have good time-series variation.

of the marginal cost equation. Then, I have $E\left(\Delta\xi_{jt}\cdot Z_3\right) = 0$ and $E\left(\eta\cdot Z_3\right) = 0$.

I conclude with an informal discussion on identification of the model. The main parameters of interest are the price sensitivity, the distance sensitivity, and the ratio of informed consumers (α, β_d, and IR_t). In this model, variation in the market shares is due to (i) variation in the prices and attributes of gasoline stations, (ii) variation in the smartphone penetration rates, and (iii) variation in distances among consumers. Note that the distance effects are assumed to be stable over time ($d_{ij}\beta_d$ does not have a t subscript). The distance parameter β_d is identified from the correlation of distances and market shares that are observed in the data. The price sensitivity term α is identified from time-series variation of market shares and the high-frequency variation in prices charged by a station and its competitors.

Moreover, I know that price changes occur once a week on average, but retail prices for stations move up or down in tandem by the same amount on many different days. Consequently, market share differences during periods with the same prices (or identical price movements) can be used to identify the informed ratio parameters (a_0 and a_1 from the variation of IR_t). Another way to explain the informed ratio identification is that the changes in IR_t capture the changes in the consumer price elasticity. As more consumers have the full price information, consumers become more price-sensitive. Since I assume that α is the same for all periods, movements in IR_t reflect this consumer price sensitivity changes by putting different weights on the informed consumers. For example, a high IR_t value assigns a large proportion of more elastic consumers (informed consumers) so that the consumer price elasticity for the market level is high.

1.4.7 Estimation

I estimate the parameters of the model by following the BLP algorithm, except that I include station-specific dummy variables and I utilize a different approach for the contraction mapping part. For the inner loop, I compute the mean utility level δ_j and run the IV regression $\delta_j = X_j\beta_c - \alpha p_j + \xi_j + \Delta\xi_{jt}$ with a set of instruments $Z_d = \left[z_1, \ldots, z_{M_1}\right]$. Then, the moment conditions are $E\left[Z_d\cdot \Delta\xi_{jt}\left(\theta\right)\right] = 0$, where $\theta = \{\alpha, \beta_c, \gamma, \sigma\}$ and $\sigma = \{\beta_d, a_0, a_1\}$. Another set of the moment conditions is from the mean independence assumption of η_j and a set of instruments $Z_s = [z_{M_1+1}, \ldots, z_{M_1+M_2}]$: $E\left[Z_s\cdot \eta\left(\theta\right)\right] = 0$. Let $h\left(\theta\right)$ be these moment conditions such that

$E[h(\theta)] = 0$. For the outer loop, I calculate the empirical analogue of the moment conditions, $\hat{h}(\theta)$. Using the two-step generalized method of moments (GMM), I find the GMM estimate that minimizes the objective function $\hat{h}'(\theta)\Phi^{-1}\hat{h}(\theta)$, where Φ is a consistent estimate of $E[h(\theta)\, h'(\theta)]$.

The inclusion of the station fixed effects (in the inner loop IV regression) requires the minimum-distance procedure (Chamberlain, 1982) to estimate the taste parameters β_c. I follow Nevo (2001). Let f be the vector of station dummy coefficients (f_j captures both the quality of observed station amenities and the mean of the unobserved characteristics, $X_j\beta_c + \xi_j$). If we assume that $E[\xi|X] = 0$, the estimates of β_c and ξ are $\hat{\beta}_c = \left(X'V_f^{-1}X\right)^{-1} X'V_f^{-1}\hat{f}$ and $\hat{\xi} = \hat{f} - X\hat{\beta}_c$. $\hat{f}$ is the estimated coefficient vector in the regression $\delta_j = -\alpha p_j + f_j D_j + \Delta\xi_{jt}$, and V_f is the covariance matrix of these estimates.[35]

As α, β_c and γ enter the GMM objective function linearly, I only need to search for $\sigma = \{\beta_d, a_0, a_1\}$ in the outside loop. Formally, the error terms are calculated by $\Delta\xi_{jt} = \delta_j - X_j\beta_c + \alpha p_j - \xi_j$ and $\eta_j = p_j - b_j(p, X, \xi; \theta) - W_j\gamma$. Let $T = \begin{pmatrix} Y & 0 \\ 0 & W \end{pmatrix}$, where $Y = \{X_j, p_j\}$. Let Z denote instruments, and $P_Z = Z\left(Z'Z\right)^{-1} Z'$. There is an analytic solution for $\binom{\beta_c, \alpha}{\gamma}$ given $\delta(s, \sigma)$:

$$\begin{pmatrix} \beta_c, \alpha \\ \gamma \end{pmatrix} = \left(T'P_Z T\right)^{-1} T'P_Z \begin{pmatrix} \delta(s,\sigma) \\ p - b(p, X, \xi; \sigma) \end{pmatrix}$$

As a result, the GMM objective function is a function of σ only. This reduces the number of parameters I need to search in the outside loop. The outside loop only updates 3 parameters: β_d, a_0, and a_1.

The estimation procedure closely resembles the nonlinear GMM approach developed by BLP, except that I made a modification in the mean utility step. When I calculate the mean utility level δ_j, I need to use a different approach, as I do not have the quantity data for many stations. For example, I know market shares for $J_0 \approx 10$ stations out of $J \approx 45$ stations in the region. Since I have only 10 observations for each date, I cannot estimate 45 δ's for the date with the traditional BLP contraction mapping. I utilize the unusual high frequency of the data to construct weekly mean utility levels.

[35] D_j are station dummy variables.

I make an additional assumption that, for each week, the only variation of the mean utility level for station j comes from the variation in p_{jt}. In other words, $\delta_{j,Tue} = \delta_{j,Mon} + \alpha(p_{j,Mon} - p_{j,Tue}), \ldots, \delta_{j,Sun} = \delta_{j,Mon} + \alpha(p_{j,Mon} - p_{j,Sun})$. With this assumption, I can write $\delta_{j,Tue}, \ldots, \delta_{j,Sun}$ as a function of $\delta_{j,Mon}$. I have $J_0 \times 7 (\approx 70)$ equations, $s_{jt} = \hat{s_{jt}}(\delta_{j,Mon})$ for $J \approx 45$ parameters $\delta_{j,Mon}$ for each week:[36]

$$\begin{aligned} s_{jt} &= \hat{s_{jt}}(\delta_{j,Mon}) = IR_t \cdot s_{jt}^{informed} + (1 - IR_t) \cdot s_{jt}^{uninformed} = IR_t \cdot \sum_i w_i \frac{\exp(d_{ij}\beta_d + \delta_{j,t})}{1 + \sum_{k=1}^{J} \exp(d_{ij}\beta_d + \delta_{j,t})} \\ &+ (1 - IR_t) \cdot \sum_i w_i \left(\sum_{l \neq j} r_{i,j,l} \frac{\exp(d_{ij}\beta_d + \delta_{j,t})}{1 + \exp(d_{ij}\beta_d + \delta_{j,t}) + \exp(d_{il}\beta_d + \delta_{l,t})} \right) \end{aligned}$$

for $j = 1, \ldots, J_0$ and $t = Mon, Tue, \ldots, Sun$. I replace the BLP contraction mapping part by finding J parameters ($\delta_{j,Mon}$) that minimize squared distances between the actual market shares and the model expected ones. Note that the IV regression in the inner loop becomes $\delta_{j,Mon} = X_j\beta_c - \alpha p_{j,Mon} + \xi_j + \Delta\xi_{j,Mon}$, or $\delta_{j,Mon} = -\alpha p_{j,Mon} + f_j D_j + \Delta\xi_{j,Mon}$, where D_j are station dummy variables and f_j are station fixed effects. While this approach reduces the number of observations for the regression to the number of weeks in the dataset, thanks to the high frequency of the data set, I have enough observations for each j.[37]

Note that as long as at least one p_{jt} changes, the right-hand-side value changes as the denominator $1 + \sum_{k=1}^{J} \exp(d_{ij}\beta_d + \delta_{j,t})$ changes. In the case when all prices remain stable, say during t_1 to t_2, then this equation becomes $E(\hat{s_{jt}}(\delta_{j,Mon})) = \frac{\Sigma s_{jt}}{t_2 - t_1}$. While this would decrease the number of equations by $t_2 - t_1 - 1$, it does not happen often and I have at least J equations in most cases. In rare cases when I do not have J equations, I assume that for two weeks (instead of one week), the only variation of the mean utility level for station j comes from the variation in p_{jt}.

1.5 Structural Estimates

This section presents structural estimation results and provides interpretations of the estimated parameter values. Structural estimation results suggest that the fraction of informed consumers

[36]As a reminder, w_i is a weight on grid point i ($\frac{1}{n^2}$ for the uniform distribution n by n grid case). $r_{i,j,l}$ is a probability of that an uninformed consumer located at i learns station j and l prices.

[37]For the whole period, the number of observations for each j is 129 (909 days, so 129 weeks).

actually increases from 1.7% to 11.4% during the sample period, as the smartphone penetration rate increases from 1% to 53%. For the effect of distances, I find that an average consumer is indifferent between driving 0.25~0.4 miles more and saving ten cents per gallon. The counterfactual analysis section derives a new set of equilibrium prices under a different informed consumer ratio, and confirms that observed price dispersion and markup changes are consistent with theoretical models of pricing, given the structurally estimated parameters.

1.5.1 Results and Interpretation

I report parameter estimates from the structural model and connect them to the descriptive findings for a better interpretation. Estimation results are presented in Table 1.8. Column 1 shows estimates for January 2010 to December 2010 when no Android version of the Opinet app was available.[38] Column 2 provides estimates for 2011 to 2012, when both iPhone and Android Opinet apps were available and the Opinet service was advertised nationally. Column 3 presents estimates for all four regions for the whole sample period, and column 4 is for regions 1 and 2, two districts of Seoul, where average income level is higher, traffic congestion is severe, and smartphone penetration rates are higher.

First of all, a_1, the proportion of informed consumers to the number of smartphone users, is positive and significant. The estimated value of 0.093 in column 3 can be interpreted as indicating that 9.3% of smartphone users are highly informed consumers. The results in columns 1 and 2 indicate that this has changed over time: 3.5% in 2010 and 16.5% in 2011-2012. As it is likely that the ratio of informed consumers to smartphone users is proportional to the ratio of the Opinet app download numbers to the total number of smartphone users, this change is not surprising: in section 1.3.4 we show that *Total App/Smartphone* is about 1.1% in 2010 and 7.0% in 2011-2012. The model results suggest that even before the Android version Opinet app and national advertising of the Opinet service, 3.5% of smartphone users visited Opinet website using smartphones to get price information, and 16.5% of smartphone users were informed after 2010. The region 1 and 2 estimate of a_1 is slightly higher than the all-four-region estimate, as consumers in the two districts of Seoul are more information-sensitive than consumers in the two

[38] From January 2010 to April 2010, no Opinet app was available, and only the iPhone app was available from May 2010 to December 2010.

rural cities.

The baseline proportion of informed consumers, a_0, is estimated to be 1.8% in column 3 (full sample). Again, the estimates from the earlier and later subsamples are different: 1.4% in 2010 and 3.1% in 2011-2012. This pattern can be thought of as similar to what is captured by the *Opinet Web/Car* variable, the proportion of weekly Opinet website visitors to the number of cars. In fact, as performing price search without the Opinet service is too costly, these two variables should have a similar trend, and the results confirm this: *Opinet Web/Car* is 2.3% in 2010 and 5.1% in 2011-2012 (see section 1.3.4).

The increase of a_0 over time is likely due to the higher consumer awareness of the Opinet service, thanks to the Opinet advertising campaign and word-of-mouth information diffusion. Like in the a_1 case, the region 1 and 2 estimate of a_1 is slightly higher than the all four region estimate of a_1. Knowing a_0 and a_1, and the smartphone penetration rate, I can estimate the ratio of informed consumers (IR_t) in the market: the informed ratio started from 1.7% in January 2010 and reached 11.4% at the end of the sample period (June 2012).

Now let us look at two other key parameters, the price sensitivity term α and the distance sensitivity term β_d. For the full sample case (column 3), $\alpha = 26.131$ and $\beta_d = -6.715$. Both estimates are highly significant and have expected signs: consumer utility decreases as prices increase or traveling distances increase. Since α is a price sensitivity term and β_d is a distance sensitivity term, it is better to interpret magnitudes of these two terms together. As a distance between neighboring grid points is 1 mile, an average consumer is willing to travel $-\frac{0.1\alpha}{\beta_d}$ miles to save ten cents per gallon. For example, $\alpha = 10$ and $\beta_d = -2.5$ means that an average consumer is indifferent between traveling 0.4 miles more and saving ten cents per gallon. A distance an average consumer is willing to travel to save ten cents per gallon varies depending on the specifications, with the minimum of 0.254 miles (column 4) and the maximum of 0.407 miles (column 1). In terms of the willingness of a typical consumer to travel an additional mile, an average consumer asks for \$0.25~\$0.39 lower prices per gallon.[39] If we assume an average

[39]Compared to the results of Manuszak and Moul (2009), \$0.065~0.084 per gallon, these estimates have higher values. However, as average price level in my data is about six times higher (about \$7 per gallon) than the average price level of \$1.2 per gallon during their sample period (Chicago and northern Indiana, 2001), their estimates would become \$0.39~0.504 per gallon, considering these differences. Moreover, they focus on tax differences in the area and assume that taxes were stable and consumers were aware of the resulting price differences. In my

driving speed in these regions to be 20 miles/hour and consumers buy five gallons of gas per each purchase, then the implied opportunity cost of time is $24.55~39.43 per hour, which is fairly reasonable.

As column 4 is for two district of Seoul where average income level is higher and traffic congestion is severe, it is not surprising that column 3 estimates of α and β_d are higher than column 4 estimates. Also, magnitudes of the α and β_d estimates of the first period (column 1) are smaller than those of the latter period (column 2); however, these differences are not statistically significant.

Finally, I discuss station characteristic coefficients, β_{Self}, $\beta_{Carwash}$, β_{Repair}, and β_{Store}. As consumers are likely to prefer stations with full-service, car wash, and convenience stores, negative β_{Self} and positive $\beta_{Carwash}$ and β_{Store} are expected results. Two coefficients, β_{Self} and $\beta_{Carwash}$, are significant at the 5% level for all specifications. For the full sample case (column 3), estimates imply that an average consumer would travel 0.58 miles more for full-service and 0.52 miles more for a car wash. On the other hand, β_{Store} is less significant under some specifications, and β_{Repair} does not even have a constant sign.

1.5.2 Counterfactual

When I presented the reduced-form results, I noted that they raised a puzzle: why have price dispersion and markups not decreased as more consumers have become informed? In this section, I provide a resolution of this puzzle by presenting counterfactual simulations examining how equilibrium prices would be expected to change as more consumers became informed, given the degree of consumer substitution and the magnitudes of the proportion of informed consumers that seems to be present. Changing a_0 or a_1 leads to different informed ratios and hence results in different equilibrium prices. For example, I study what the equilibrium prices would have been if the informed ratio were 0.05 or 0.2, instead of the original value 0.1.

To do this, I first estimate implied marginal costs that are consistent with the stations' pricing decisions. Note that without daily, individual station shocks, it is impossible to establish that stations follow the profit-maximizing condition every day. For the previous section, I used the

case, since consumers are not aware of the amount of potential savings unless they know the prices by searching, it is likely that the willingness to pay amount is different.

sum of average wholesale prices and a constant as approximations of marginal costs. If I denote daily differences between the real, unobserved marginal cost and the average wholesale cost as e_{jt}, I can find e_{jt} from the first order condition:

$$(p_{jt} - [AWP_{jt} + e_{jt}]) \frac{\partial q_{jt}}{\partial p_{jt}} + q_{jt} = 0$$

$$(p_{jt} - [AWP_{jt} + e_{jt}]) \frac{\partial s_{jt}}{\partial p_{jt}} + \frac{q_{jt}}{Q_t} = 0$$

For the $\frac{q_{jt}}{Q_t}$ term, I use the actual q_{jt} for the stations that I have quantity data, and model computed $s_{jt} = \frac{q_{jt}}{Q_t}$ for the stations without quantity data. Then I get

$$e_{jt} = p_{jt} - AWP_{jt} + \frac{q_{jt}}{Q_t} / \frac{\partial s_{jt}}{\partial p_{jt}}$$

For simplicity, I define $ex_t(i,j) = \exp(d_{ij}\beta_d + X_j\beta_c - \alpha p_{jt} + \xi_{jt})$. Based on the analytic market share formula, I can compute $\frac{\partial s_{jt}}{\partial p_{jt}}$ (details in Appendix B).

$$\frac{\partial s_{jt}}{\partial p_{jt}} = \frac{1}{n^2}\left\{ IR_t \sum_i \left(\frac{-\alpha \cdot ex_t(i,j)\left(1 + \Sigma_{k \neq j} ex_t(i,k)\right)}{\left(1 + \sum_{k=1}^{J} ex_t(i,k)\right)^2} \right) + (1 - IR_t) \sum_i \left(\sum_{l \neq j} r_{i,j,l} \frac{-\alpha \cdot ex_t(i,j)\left(1 + ex_t(i,l)\right)}{\left(1 + ex_t(i,j) + ex_t(i,l)\right)^2} \right) \right\}$$

Empirical distribution of e_{jt} suggests that $e_{jt} = f_j + \epsilon_{jt}$, where f_j is an unobserved station fixed cost and $\epsilon_{jt} \sim N\left(0, \sigma_e^2\right)$. This fixed component refers to characteristics of the station j that are invariant during the period. These include, for instance, rent for the space, and labor cost for managers and core workers. The transitory component ϵ_{jt} reflects daily temporary changes. Empirical distributions of ϵ_{jt} are similar for both stations that I observe quantities and stations that I do not, and the estimate of σ_e is 0.057. As an average retail price during the period is about 7 dollars per gallon, a daily cost shock is about 0.8% of the average retail price.

Using the real marginal cost $c_{jt} = AWP_{jt} + e_{jt}$, I derive a new equilibrium prices when informed ratio IR_t is a different value. For each market t, I have n parameters $(p_1, p_2, \ldots, p_n)$and

n profit maximizing conditions:

$$(p_j - c_j)\frac{\partial q_j}{\partial p_j} + q_j\left(p_j, p_{-j}\right) = 0$$

Note that q_j is a function of $p_1, p_2, \ldots, p_n$ and for each equation, I calculate a best response: fix p_{-j} and find optimal p_j. For each step, we get new $p_1', p_2', \ldots, p_n'$ as best responses from $p_1, p_2, \ldots, p_n$. Using the $\frac{\partial s_{jt}}{\partial p_{jt}}$ formula (Appendix B), the profit-maximizing condition becomes

$$p_{jt}' = c_{jt} + \frac{1}{\alpha}\frac{IR_t s_{jt}^{informed} + (1 - IR_t)\, s_{jt}^{uninformed}}{IR_t s_{1t} + (1 - IR_t)\, s_{2t}}$$

where $s_{1t} = \sum_i \left(\frac{ext(i,j)\left(1+\Sigma_{k\neq j} ext(i,k)\right)}{\left(1+\sum_{k=1}^{J} ext(i,k)\right)^2}\right)$ and $s_{2t} = \sum_i \left(\sum_{l\neq j} r_{i,j,l}\frac{ext(i,j)(1+ext(i,l))}{(1+ext(i,j)+ext(i,l))^2}\right)$. Repeating these steps, I can find a fixed point (converging point) as new equilibrium prices.[40] Examining price dispersion (standard deviation) and markup levels at the new equilibrium prices, I find that both price dispersion and markup levels are slightly increasing as the informed ratio goes up, which is consistent with the results form the descriptive analysis section.

Table 1.9 shows how price dispersion (as measured by standard deviation) and markup levels change when the fraction of informed consumers, IR, changes from the baseline case, 10%. As the structural model estimates show that IR was close to 1% at the beginning of the sample period and about 10% at the end of the period, I start with comparing statistics for the 1% and 10% case. The counterfactual results suggest that there are moderate changes in price dispersion and average markups. The change in equilibrium prices, due to changes in demand, would increase the standard deviation of prices by 0.57%, the average markups by 0.33%, and the average quantity-weighted markups by 0.09%. Hence, this model could explain the "puzzle", the slight increase in both measures that we observe in the actual data. One way to interpret these moderate changes, in terms of magnitude of changes as the informed ratio increases, is that consumers have strong location preferences. As consumers prefer certain gasoline stations or stations close to them, their choices would not change much even if they learned the prices of stations that are far away.

[40] I use $\Sigma\left(p_j - p_j'\right)^2 < 10^{-12}$ as a convergence criterion: in this empirical setting, convergence took less than 20 steps, or one minute.

Since mobile technologies continue to be developed and could reduce consumer search costs even further, the fraction of informed consumers is likely to grow. Thus, it would be interesting to consider what would happen if IR goes beyond 10%. Table 1.9 shows the results up to 35%. While price dispersion increases for all IR levels, both unweighted and quantity-weighted markups start to decrease at some point. Unweighted markups stop increasing around 30%, and quantity-weighted markups show a reversal when the informed ratio reaches 20%. As the proportion of informed consumers goes up, setting low prices to draw informed consumers becomes more attractive to stations. Since stations with lower prices make more sales, these reversals in markup trends happen earlier for the quantity-weighted case.

The counterfactual analysis also allows me to compute consumer welfare changes. By introducing the Opinet price information service, the Korean government intended to enhance consumer welfare. However, my results show the opposite. While informed consumers earn about $200 per year, uninformed consumers lose about $50 per year since they are paying higher prices.[41] As 90% of consumers are uninformed, consumers as a whole lose $25 per year on average.[42]

1.6 Conclusion

In this paper, I investigate how real-time gasoline price information and the spread of mobile technologies affect market outcomes, e.g., consumer search behavior and price dispersion among gasoline stations. I combine daily, individual gasoline station price and quantity data with regional smartphone penetration data for the analysis. The universe of daily station-level prices for each region allows me to perfectly measure daily market-level price dispersion. In addition, I utilize daily station-level quantity information to compute more realistic measures of price dispersion and markup levels. Having both price and quantity data enables me to estimate daily demand in the gasoline market, which was impossible for the previous literature due to the lack of data.

[41] Allen et al. (2014) also note that effects are not spread equally among consumers and consumers with lower search costs tend to benefit more.

[42] This is consistent with the finding that markups actually increased.

This paper is motivated by two technological advances: a free, real-time gasoline price information service provided by the Korean government, and the introduction and rapid growth of mobile technologies, in particular, smartphones. The price information service reduces consumer search costs, and the existence of smartphones facilitates the use of this price information service by allowing consumers to search for prices while driving. In particular, from the standpoint of the gasoline retail market, the introduction of smartphones is an exogenous technology shock.

To measure the impact of these changes, I analyze the data and find interesting stylized facts: both price dispersion and markup did not decrease, even though smartphone penetration rates increased significantly. Similar trends are found in quantity-weighted measures. This observation is contrary to many previous research results that a search cost reduction leads to higher competition, and hence lower price dispersion and markup. There are two potential reasons: first, it is possible that consumers may not use smartphones to search; second, consumers indeed search more, but this change in consumer search behavior affects gas stations' pricing strategies and changes market outcomes.

Additional descriptive evidence suggests that consumers searched more and became more price-sensitive. I analyze the Opinet website visitor numbers and the Opinet smartphone application download numbers to find that 2-5% of consumers visited the website throughout the period and a constant fraction[43] of smartphone users downloaded the Opinet app. Since the smartphone penetration rate has increased rapidly during the sample period, the Opinet service usage has increased, and it is likely that the ratio of informed consumers has increased. Moreover, a simple regression test suggests that consumers became more sensitive to price differences, in particular, the difference between a price of a station and the minimum price of the region.

I find that search theory models can explain these surprising trends. As indicated by the counterfactual analysis, a model with differentiated products and two consumer types predicts that under a two-type consumer environment (one type more informed than the other) where the ratio of informed consumers is low, price dispersion can increase until the ratio reaches a critical point. The fact that quantity-weighted markup distributions moved toward a bimodal distribution coincides with theoretical search model results: intuitively, it is due to a bifurcation

[43] About 1.5% until Dec 2010, and 7% after 2010.

of firm strategies with some setting low prices to attract informed consumers and some setting high prices to serve less informed consumers.

I develop a discrete choice model of consumer demand for spatially differentiated gasoline stations to estimate how consumer choices and the informed consumer ratio change as smartphone penetration rates increase. The structural model results suggest that consumers became more price sensitive as the informed ratio increased from 1.7% to 11.4% during 2010-2012. The counterfactual analysis studies what the new equilibrium prices would be if the proportion of informed consumers were different. In particular, when the informed ratio moves from 1% to 10%, price dispersion, unweighted markups, and quantity weighted markup levels are expected to increase 0.57%, 0.33%, and 0.09%, respectively. These estimates are consistent with the reduced-form section results and the actual data: both price dispersion and markup levels increase slightly as the informed ratio increases, and the magnitude of the increases fit the observed patterns.

References

[1] Jenny Aker. Does digital divide or provide? the impact of cell phones on grain markets in niger. *Center for Global Development Working Paper*, (154), 2008.

[2] Jason Allen, Robert Clark, and Jean-François Houde. Search frictions and market power in negotiated price markets. Technical report, National Bureau of Economic Research, 2014.

[3] John M Barron, Beck A Taylor, and John R Umbeck. Number of sellers, average prices, and price dispersion. *International Journal of Industrial Organization*, 22(8):1041–1066, 2004.

[4] Michael R Baye, John Morgan, and Patrick Scholten. Price dispersion in the small and in the large: Evidence from an internet price comparison site. *The Journal of Industrial Economics*, 52(4):463–496, 2004.

[5] Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890, 1995.

[6] Severin Borenstein, A Colin Cameron, and Richard Gilbert. Do gasoline prices respond asymmetrically to crude oil price changes? *The Quarterly Journal of Economics*, 112(1):305–339, 1997.

[7] Jeffrey Robert Brown and Austan Goolsbee. Does the internet make markets more competitive? evidence from the life insurance industry. *Journal of Political Economy*, 110(3):481–507, 2002.

[8] David P Byrne and Nicolas De Roos. Search and stockpiling in retail gasoline markets. *Available at SSRN 2427556*, 2014.

[9] Ambarish Chandra and Mariano Tappata. Consumer search and dynamic price dispersion: an application to gasoline markets. *The RAND Journal of Economics*, 42(4):681–704, 2011.

[10] Robert Clark and Jean-François Houde. Collusion with asymmetric retailers: Evidence from a gasoline price-fixing case. *American Economic Journal: Microeconomics*, 5(3):97–123, 2013.

[11] Babur De los Santos, Ali Hortacsu, and Matthijs R Wildenbeest. Testing models of consumer search using data on web browsing and purchasing behavior. *The American Economic Review*, 102(6):2955–2980, 2012.

[12] Peter Diamond. A model of price adjustment. *Journal of Economic Theory*, 3:pp. 156–68, 1971.

[13] Andrew Eckert and Douglas S West. Price uniformity and competition in a retail gasoline market. *Journal of Economic Behavior & Organization*, 56(2):219–237, 2005.

[14] Glenn Ellison and Sara Fisher Ellison. Search, obfuscation, and price elasticities on the internet. *Econometrica*, 77(2):427–452, 2009.

[15] John A Hartigan and PM Hartigan. The dip test of unimodality. *The Annals of Statistics*, pages 70–84, 1985.

[16] Jerry A Hausman. Valuation of new goods under perfect and imperfect competition. In *The economics of new goods*, pages 207–248. University of Chicago Press, 1996.

[17] Han Hong and Matthew Shum. Using price distributions to estimate search costs. *The RAND Journal of Economics*, 37(2):257–275, 2006.

[18] Daniel S Hosken, Robert S McMillan, and Christopher T Taylor. Retail gasoline pricing: What do we know? *International Journal of Industrial Organization*, 26(6):1425–1436, 2008.

[19] Jean-François Houde. Spatial differentiation and vertical mergers in retail markets for gasoline. *The American Economic Review*, 102(5):2147–2182, 2012.

[20] Robert Jensen. The digital provide: Information (technology), market performance, and welfare in the south indian fisheries sector. *The Quarterly Journal of Economics*, 122(3):879–924, 2007.

[21] Matthew Lewis. Price dispersion and competition with differentiated sellers. *The Journal of Industrial Economics*, 56(3):654–678, 2008.

[22] Matthew S. Lewis. Asymmetric price adjustment and consumer search: An examination of the retail gasoline market. *Journal of Economics & Management Strategy*, 20(2):409–449, 2011.

[23] Matthew S. Lewis and Howard P. Marvel. When do consumers search? *The Journal of Industrial Economics*, 59(3):457–483, 2011.

[24] Mark D Manuszak and Charles C Moul. How far for a buck? tax differences and the location of retail gasoline activity in southeast chicagoland. *The Review of Economics and Statistics*, 91(4):744–765, 2009.

[25] José Luis Moraga-González and Matthijs R Wildenbeest. Maximum likelihood estimation of search costs. *European Economic Review*, 52(5):820–848, 2008.

[26] Peter Morgan and Richard Manning. Optimal search. *Econometrica: Journal of the Econometric Society*, pages 923–944, 1985.

[27] Aviv Nevo. Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342, 2001.

[28] Michael D. Noel. Edgeworth price cycles: Evidence from the toronto retail gasoline market. *The Journal of Industrial Economics*, 55(1):69–92, 2007.

[29] Sam Peltzman. Prices rise faster than they fall. *Journal of Political Economy*, 108(3):466–502, 2000.

[30] Dieter Pennerstorfer, Philipp Schmidt-Dengler, Nicolas Schutz, Christoph Weiss, and Biliana Yontcheva. Information and price dispersion: Evidence from retail gasoline. Technical report, 2014.

[31] Steven Salop and Joseph Stiglitz. Bargains and ripoffs: A model of monopolistically competitive price dispersion. *The Review of Economic Studies*, 44(3):pp. 493–510, 1977.

[32] Alan T. Sorensen. Equilibrium price dispersion in retail markets for prescription drugs. *Journal of Political Economy*, 108(4):pp. 833–850, 2000.

[33] Alan T Sorensen. An empirical model of heterogeneous consumer search for retail prescription drugs. Technical report, National Bureau of Economic Research, 2001.

[34] Dale O. Stahl. Oligopolistic pricing with sequential consumer search. *The American Economic Review*, 79(4):pp. 700–712, 1989.

[35] Dale O. Stahl. Oligopolistic pricing with heterogeneous consumer search. *International Journal of Industrial Organization*, 14(2):243–268, 1996.

[36] George J. Stigler. The economics of information. *Journal of Political Economy*, 69(3):pp. 213–225, 1961.

[37] Hal R Varian. A model of sales. *The American Economic Review*, pages 651–659, 1980.

[38] Jeremy A. Verlinda. Do rockets rise faster and feathers fall slower in an atmosphere of local market power? evidence from the retail gasoline market. *The Journal of Industrial Economics*, 56(3):581–612, 2008.

[39] Matthijs R Wildenbeest. An empirical model of search with vertically differentiated products. *The RAND Journal of Economics*, 42(4):729–757, 2011.

Appendix A. Government Intervention Period

There were two international issues that caused a big jump in Dubai oil prices in March, 2011. The series of protests and demonstrations across Middle East and North Africa caused social unrest. Also, trade sanctions against Iran directly affected oil supply. Following the international price spike, the Korean domestic gasoline prices went up by more than 150 Korean won per liter (about 57 cents per gallon) within a month. Since the price of gasoline plays a important role in a retail price index that people are interested in, the Korean government chose to intervene in the retail gasoline market to stabilize prices, and asked four major gasoline companies to cut gasoline distribution prices (average wholesale prices, or AWP. SKE, a leading gasoline company, announced a price cut of 100 Korean won per liter from 4/7/11 to 7/7/11 and other companies followed. According to an Opinet representative, discounts by three companies (GSC, HDO, and SOL) were reflected in the price data, as they cut the distribution price directly and posted prices went down. However, SKE offered refund bonus points that were equivalent to 100 Korean won per liter discount to customers after their purchases. Thus, posted prices for SKE stations did not reflect the discount. Note that I have not used dollar per gallon metric in this section to emphasize the impact of 100 Korean won per liter differences. Figure 1.8 presents that there are big gaps between national average prices among gas companies from April 2011 to July 2011.

To test this information, I compared mean prices for SKE gas stations and those of three other major companies. If the different discount methods were the only reason of the spike, we should be able to observe that compared to other periods, SKE average prices are around 100 won (per liter) higher than average prices of other companies from April 2011 to July 2011.[44] For the main analysis, I used to modified prices for SKE stations (retail prices were subtracted by 100 Korean won) from 4/7/11 to 7/7/11. As a robustness check, I also tried using the data without this period, and using the data without SKE stations from 4/7/11 to 7/7/11. The results were similar in all cases.

[44] It is difficult to understand what exactly happened during this period, as many issues such as supply chain networks and political considerations are involved.

Appendix B. $\frac{\partial s_{jt}}{\partial p_{jt}}$ term

I derive an analytic form of $\frac{\partial s_{jt}}{\partial p_{jt}}$. For simplicity, I define $ex_t(i,j) = \exp(d_{ij}\beta_d + X_j\beta_c - \alpha p_{jt} + \xi_{jt})$. Then $\frac{d(ex_t(i,j))}{dp_{jt}} = -\alpha \cdot ex_t(i,j)$. Then

$$s_{jt}^{informed} = \sum_i w_i \frac{ex_t(i,j)}{1+\sum_{k=1}^{J} ex_t(i,k)},\ s_{jt}^{uninformed} = \sum_i w_i \left(\sum_{l\neq j} r_{i,j,l} \frac{ex_t(i,j)}{1+ex_t(i,j)+ex_t(i,l)}\right)$$

$$\begin{aligned}
\frac{\partial s_{jt}^{informed}}{\partial p_{jt}} &= \sum_i w_i \left(\frac{-\alpha \cdot ex_t(i,j)\left(1+\Sigma_{k\neq j} ex_t(i,k)\right)}{\left(1+\sum_{k=1}^{J} ex_t(i,k)\right)^2}\right) \\
\frac{\partial s_{jt}^{uninformed}}{\partial p_{jt}} &= \sum_i w_i \left(\sum_{l\neq j} r_{i,j,l} \frac{-\alpha \cdot ex_t(i,j)\left(1+ex_t(i,l)\right)}{\left(1+ex_t(i,j)+ex_t(i,l)\right)^2}\right) \\
\frac{\partial s_{jt}}{\partial p_{jt}} &= IR_t \frac{\partial s_{jt}^{informed}}{\partial p_{jt}} + (1-IR_t)\frac{\partial s_{jt}^{uninformed}}{\partial p_{jt}}
\end{aligned}$$

Using $w_i = \frac{1}{n^2}$ (the uniform distribution over the n by n grid), we can simplify further:

$$\frac{\partial s_{jt}}{\partial p_{jt}} = \frac{1}{n^2}\left\{ IR_t \sum_i \left(\frac{-\alpha \cdot ex_t(i,j)\left(1+\Sigma_{k\neq j} ex_t(i,k)\right)}{\left(1+\sum_{k=1}^{J} ex_t(i,k)\right)^2}\right) + (1-IR_t)\sum_i \left(\sum_{l\neq j} r_{i,j,l} \frac{-\alpha \cdot ex_t(i,j)\left(1+ex_t(i,l)\right)}{\left(1+ex_t(i,j)+ex_t(i,l)\right)^2}\right)\right\}$$

$$s_{jt} = \frac{1}{n^2}\left\{ IR_t \sum_i \frac{ex_t(i,j)}{1+\sum_{k=1}^{J} ex_t(i,k)} + (1-IR_t)\sum_i \left(\sum_{l\neq j} r_{i,j,l} \frac{-\alpha \cdot ex_t(i,j)\left(1+ex_t(i,l)\right)}{\left(1+ex_t(i,j)+ex_t(i,l)\right)^2}\right)\right\}$$

Then the profit-maximizing condition becomes:

$$\begin{aligned}
&\alpha\left(p_{jt} - mc_{jt}\right)\left\{ IR \sum_i \left(\frac{ex_t(i,j)\left(1+\Sigma_{k\neq j} ex_t(i,k)\right)}{\left(1+\sum_{k=1}^{J} ex_t(i,k)\right)^2}\right) + (1-IR_t)\sum_i \left(\sum_{l\neq j} r_{i,j,l} \frac{ex_t(i,j)\left(1+ex_t(i,l)\right)}{\left(1+ex_t(i,j)+ex_t(i,l)\right)^2}\right)\right\} \\
&= \left\{ IR_t \sum_i \frac{ex_t(i,j)}{1+\sum_{k=1}^{J} ex_t(i,k)} + (1-IR_t)\sum_i \left(\sum_{l\neq j} r_{i,j,l} \frac{ex_t(i,j)}{1+ex_t(i,j)+ex_t(i,l)}\right)\right\}
\end{aligned}$$

where d_{ij} are given and $r_{i,j,l}$ are probabilities from section 1.4.3. α, β_c, β_d, and IR_t (or a_o and a_1) are the model estimates.

Appendix C. Figures and Tables

Figure 1.1: Smartphone Penetration Rates

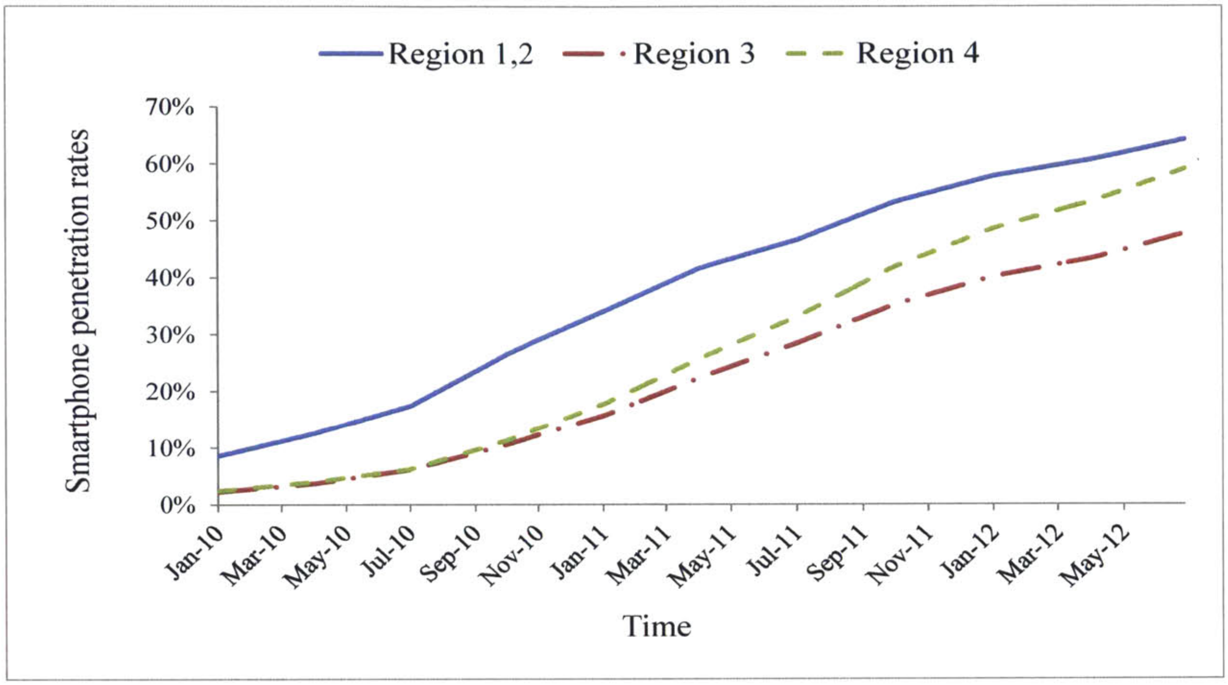


Figure 1.2: Four Price Dispersion Measures

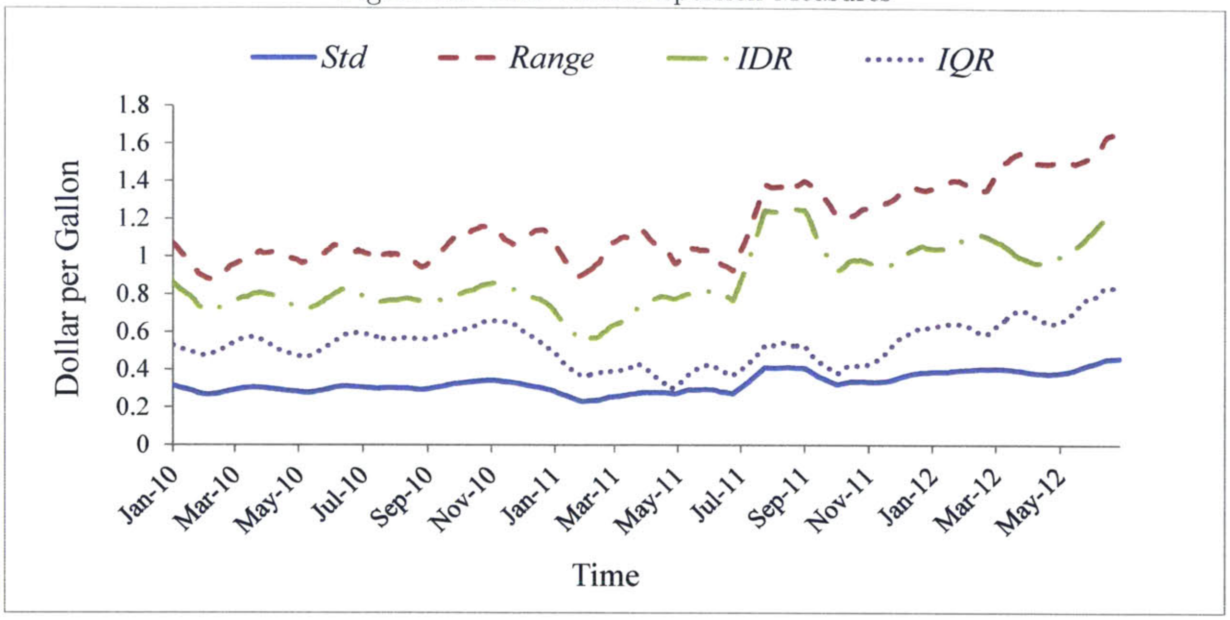


Figure 1.3: Quantity-weighted and Unweighted *Std*

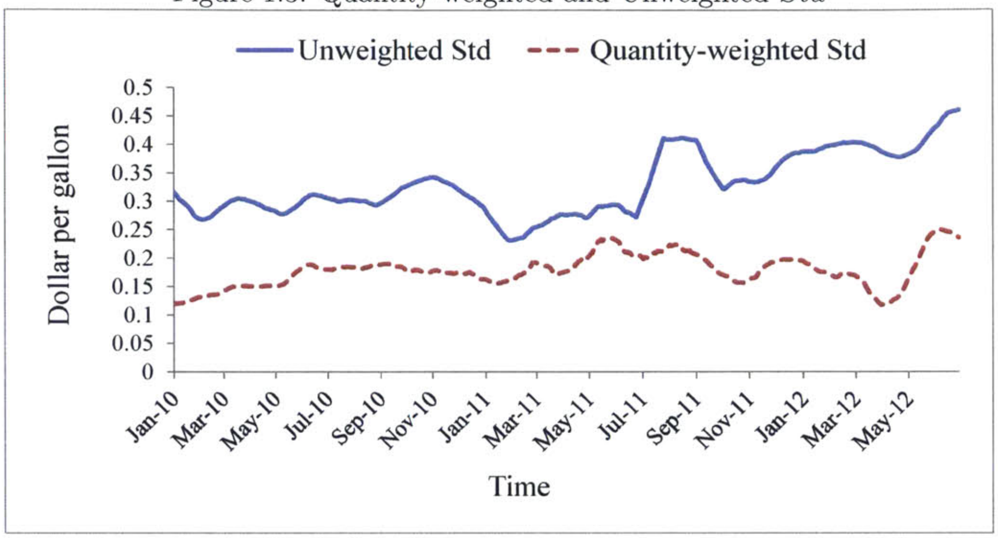


Figure 1.4: Quantity-weighted and Unweighted Markup Trend

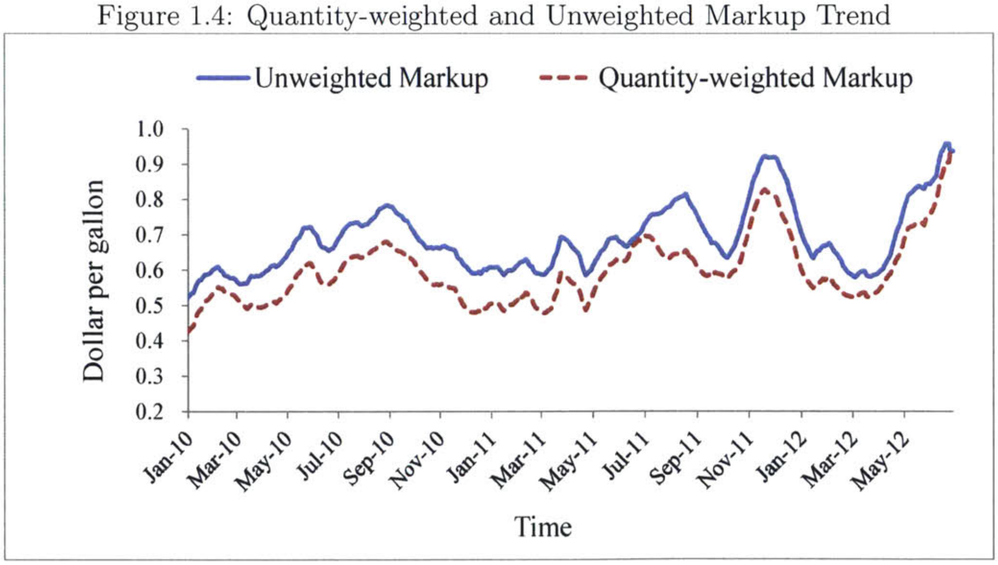


Figure 1.5: Opinet Website and Application Usage Trends

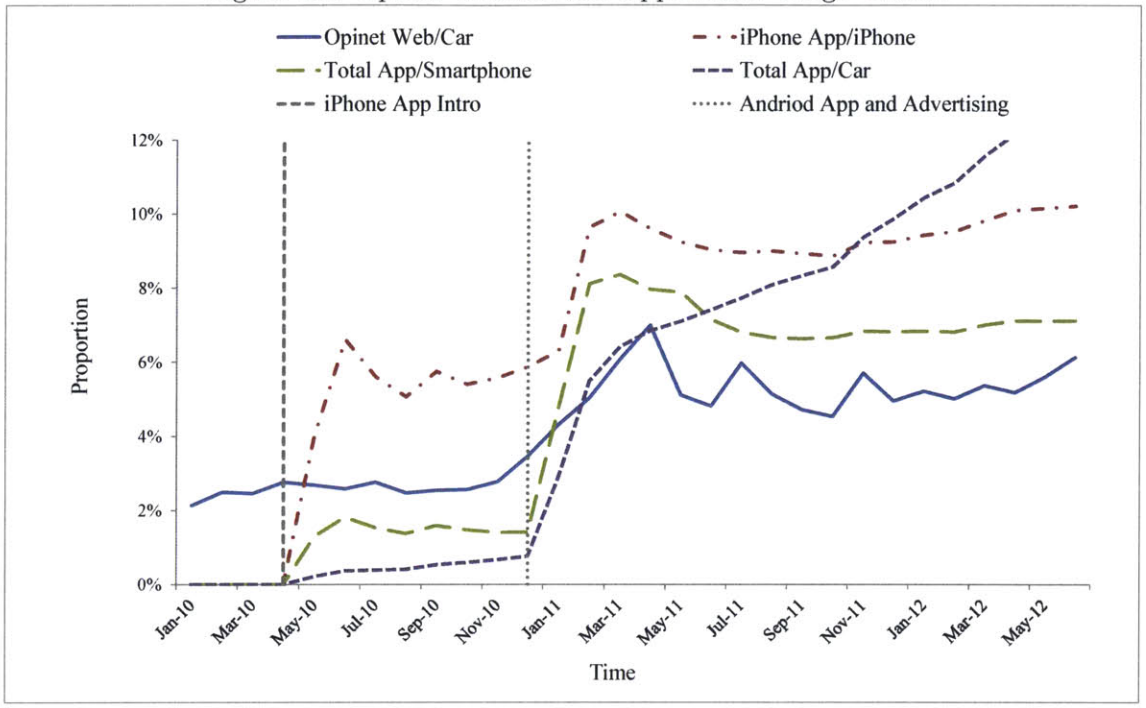


Figure 1.6: Quantity-weighted Markup Kernel Density (2010 March 1st, 2012 March 1st)

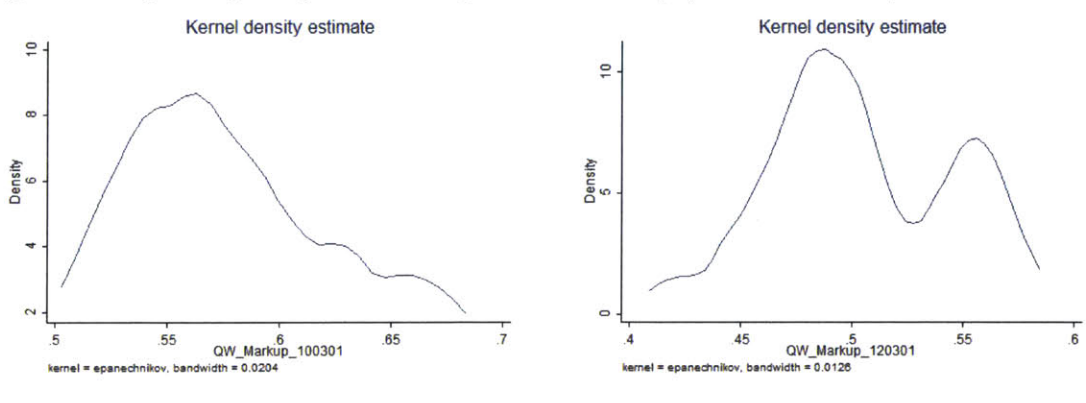
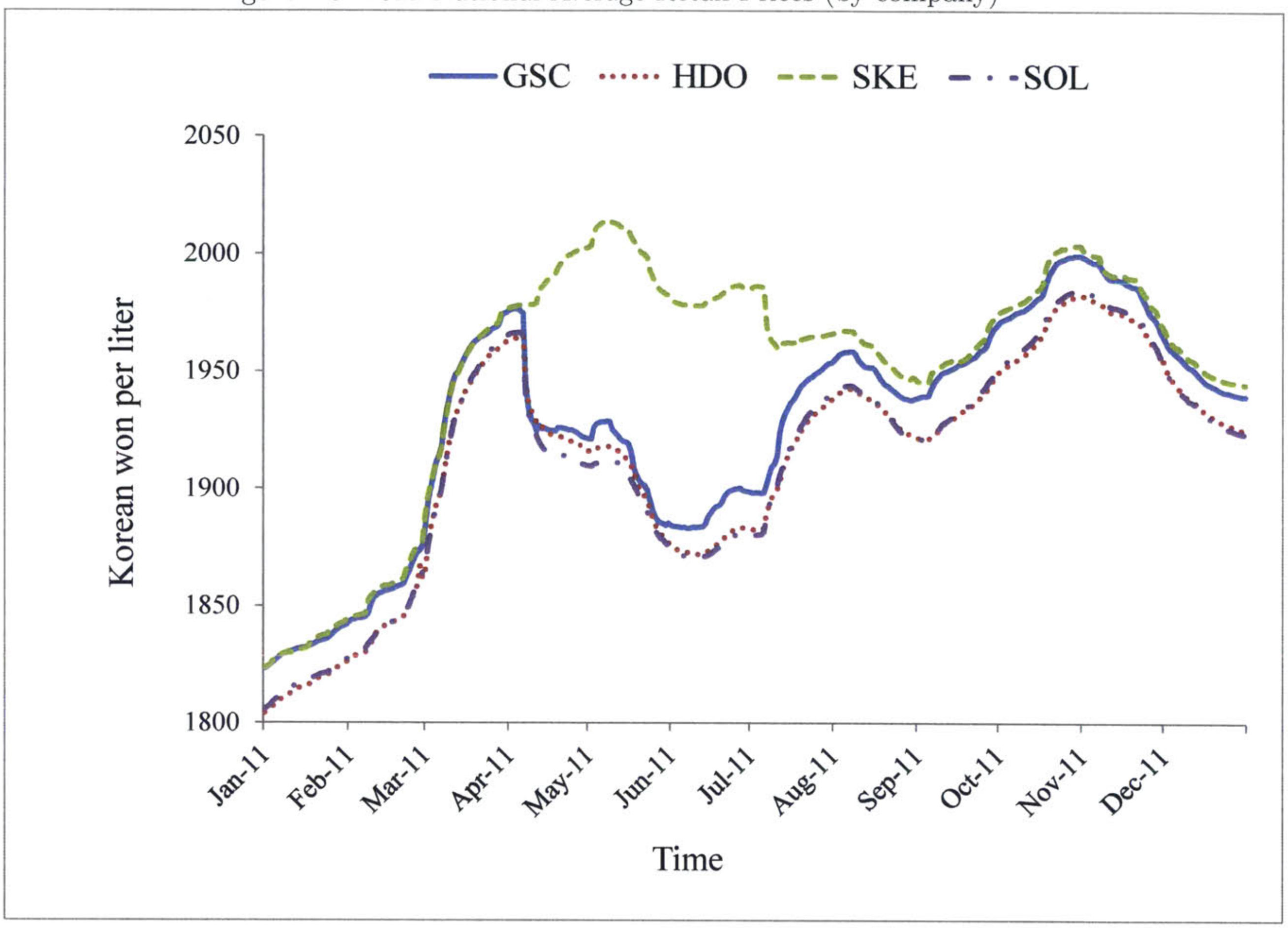


Figure 1.7: Region 1 (a district of Seoul) Map

Figure 1.8: 2011 National Average Retail Prices (by company)



Notes: Korean won per liter metric is used to present the effects of the 100 Korean won price cut.

Table 1.1: Definition of Variables

Variable	Indexes vary over	Definition
RetailP	j, t	Retail gasoline price (posted price)
AvRetailP	r, t	Daily average retail gasoline price of each region
QwRetailP	r, t	Quantity-weighted daily retail prices of each region
AdjustedP	j, t	Price net of characteristics fixed effects
Mkup	j, t	Markup of station j (*RetailP* - *AWP*)
AvMkup	r, t	Daily average markups of each region
QwMkup	r, t	Quantity-weighted daily markups of each region
AWP	j, t	Average wholesale price
SmartPen	r, t	The ratio of smartphone users to the total population
Self	j	1 if a station offers self-service, 0 otherwise
Carwash	j	1 if a station has a car wash, 0 otherwise
Repair	j	1 if a station has a repair shop, 0 otherwise
Store	j	1 if a station has a convenience store, 0 otherwise
Range	r, t	Maximum price - minimum price
Std	r, t	Standard Deviation
IDR	r, t	Interdecile Range (90% percentile price - 10% percentile price)
IQR	r, t	Interquartile Range (75% percentile price - 25% percentile price)

Notes: r: region, t: market, j: station.

Table 1.2: Summary Statistics (Variables)

Variable	Mean	Std. Dev.	Min.	Max.	Obs.
RetailP	7.279	0.626	6.014	9.076	150634
AvRetailP	7.279	0.565	6.213	8.473	3636
QwRetailP	7.223	0.580	6.176	8.587	3636
AdjustedP	7.342	0.571	5.963	8.666	150634
Mkup	0.667	0.394	-0.086	2.510	150634
AvMkup	0.667	0.281	0.155	1.574	3636
QwMkup	0.587	0.306	0.120	1.508	3636
AWP	5.943	0.435	5.226	6.780	150634
SmartPen	0.386	0.177	0.086	0.642	3636
Self	0.101	0.302	0	1	150634
Carwash	0.602	0.489	0	1	150634
Repair	0.205	0.404	0	1	150634
Store	0.133	0.339	0	1	150634

Notes: Dollar per gallon metric is used. Region-level variables have 3636 (909 times 4, or the number of days times the number of regions) observations, and station-level variables have 150634 observations.

Table 1.3: Summary Statistics (Price Dispersion Measures)

Unweighted					Quantity-weighted				
Variable	Mean	Std. Dev.	Min.	Max.	Variable	Mean	Std. Dev.	Min.	Max.
Range	1.170	0.341	0.837	1.623	*Range*	1.170	0.341	0.837	1.623
Std	0.329	0.057	0.226	0.496	*Std*	0.178	0.034	0.098	0.383
IDR	0.882	0.182	0.506	1.283	*IDR*	0.447	0.106	0.193	0.992
IQR	0.536	0.123	0.240	0.935	*IQR*	0.245	0.096	0.053	0.799

Table 1.4: Price Dispersion Regression Result (*Std*)

	Unweighted		Quantity-weighted	
	(1)	(2)	(3)	(4)
Dependent var.	*Std*	*Std*	*Std*	*Std*
AWP	-0.044**	-0.043**	-0.005	-0.005
	(-8.78)	(-8.76)	(-1.09)	(-1.21)
SmartPen	0.252**	0.213**	0.195*	0.167*
	(19.47)	(7.93)	(1.76)	(1.54)
time trend	No	Yes	No	Yes
R^2	0.917	0.921	0.701	0.832

Notes: the number of observations is 3636. t-statistics are in parentheses. Month of the Year dummies and region dummies are included. $*$: Significant at the 10% level. $**$: Significant at the 5% level.

Table 1.5: Four Price Dispersion Measures Regression Result

	Unweighted				Quantity-weighted			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent var.	*Std*	*Range*	*IDR*	*IQR*	*Std*	*Range*	*IDR*	*IQR*
AWP	-0.044**	-0.122**	-0.135**	-0.161**	-0.005	0.085**	0.043**	-0.065**
	(-8.78)	(-5.96)	(-7.60)	(-11.10)	(-1.09)	(2.99)	(3.15)	(-5.35)
SmartPen	0.252**	1.006**	0.689**	0.454**	0.195*	0.017	-0.167**	0.156**
	(19.47)	(19.00)	(15.30)	(12.00)	(1.76)	(0.23)	(-4.74)	(4.79)
R^2	0.917	0.905	0.829	0.816	0.701	0.798	0.491	0.387

Notes: the number of observations is 3636. t-statistics are in parentheses. Month of the Year dummies and region dummies are included. $*$: Significant at the 10% level. $**$: Significant at the 5% level.

Table 1.6: Markup Regression Result

	(1)	(2)
Dependent var.	*Mkup*	*Mkup*
AWP	-0.392**	-0.397**
	(-29.54)	(-35.34)
$SmartPen$	0.344**	0.266**
	(22.21)	(6.00)
Self	-0.038**	-0.038**
	(-3.14)	(-3.13)
$Carwash$	0.005	0.005
	(0.41)	(0.41)
Repair	-0.023	-0.023
	(-1.44)	(-1.44)
Store	0.036**	0.036**
	(3.02)	(3.02)
time trend	No	Yes
R^2	0.561	0.647

Notes: the number of observations is 150634. t-statistics are in parentheses. Month of the Year dummies and region dummies are included. $**$: Significant at the 5% level.

Table 1.7: A Simple Demand Regression Result

	2010		2012	
	(1)	(2)	(3)	(4)
Dependent var.	$ln\,(Q)$	$ln\,(Q)$	$ln\,(Q)$	$ln\,(Q)$
p_{jt}	-0.179**	-0.173*	-0.206**	-0.159**
	(-2.482)	(-1.734)	(-3.301)	(-2.615)
$p_{jt} - p_{min,t}$		-0.277		-1.246**
		(-0.875)		(-8.468)
R^2	0.542	0.681	0.562	0.762

Notes: the number of observations is 13476 for columns 1 and 2, and 7464 for columns 3 and 4. t-statistics are in parentheses. Month of the Year dummies and region dummies are included. $*$: Significant at the 10% level. $**$: Significant at the 5% level.

Table 1.8: Structural Model Estimates

	(1) No Android App Jan 2010 - Dec 2010	(2) Both Apps Jan 2011- Jun 2012	(3) All Regions 2010-2012	(4) Region 1 and 2 only 2010-2012
α	25.251**	29.238**	26.131**	22.709**
	(3.922)	(2.053)	(2.298)	(2.474)
β_d	-6.199**	-8.481**	-6.715**	-8.955**
	(1.252)	(0.528)	(0.534)	(1.201)
$-\frac{0.1\alpha}{\beta_d}$	0.407	0.345	0.389	0.254
a_0	0.014*	0.031**	0.018**	0.023**
	(0.008)	(0.003)	(0.005)	(0.004)
a_1	0.035**	0.165**	0.093**	0.121**
	(0.011)	(0.014)	(0.006)	(0.013)
β_{Self}	-4.196**	-4.297**	-3.900**	-4.696**
	(0.365)	(0.464)	(0.397)	(0.303)
$\beta_{Carwash}$	3.695**	3.394**	3.503**	4.101**
	(1.182)	(0.401)	(0.615)	(0.825)
β_{Repair}	-0.123	0.438*	0.516	-0.468
	(0.180)	(0.261)	(0.655)	(0.474)
β_{Store}	3.704**	2.300*	2.504**	1.905
	(0.758)	(1.197)	(0.923)	(1.187)

Notes: Station characteristics coefficients are estimates from a minimum-distance procedure. Other parameters are GMM estimates. Standard errors are given in parentheses. $*$: Significant at the 10% level. $**$: Significant at the 5% level.

Table 1.9: Counterfactual Results

IR	1%	5%	10%	15%	20%	25%	30%	35%
Std	-0.57%	-0.23%	0	0.13%	0.25%	0.34%	0.41%	0.43%
Mkup	-0.33%	-0.15%	0	0.10%	0.17%	0.21%	0.22%	0.20%
QwMkup	-0.09%	-0.02%	0	0.04%	0.06%	0.05%	0.03%	0.01%

Chapter 2

How Do Consumers React to Price Movements? Evidence from Consumers Filling Up Their Cars

This paper studies consumer decisions at gasoline pumps, using a detailed transaction-level gasoline sales dataset for select stations in Korea. There are two unique market properties: first, almost all gasoline stations in Korea offer only full service; second, gas pumps can be set to fuel up to pre-set integer dollar amounts. From February 2011 to December 2012, about 36% of regular gasoline consumers chose to simply fill up, while the remaining 64% of consumers spent pre-selected dollar amounts. Descriptive analyses show that the fraction of consumers who fill up increases by 0.11% when the current price becomes 10 Korean Won (1 cent) higher, decreases by 0.94% on days when the average retail price is 1% higher than the average of the previous three days, and decreases by 0.22% when there is a positive random price shock of 10 Korean Won. These results suggest that consumers become more active in quantity choices at gasoline pumps, and less likely to simply fill up, when retail prices are on an upward trend and when the current price level is unexpectedly high. Among the four proposed models, the model that offers predictions that are consistent with the reduced-form results is the one in which consumers expect that gasoline prices will tend to move to the average price over time.

2.1 Introduction

Consumer decision making in response to price changes have been studied in various aspects: in financial markets, researchers have found that consumers may expect asset prices to follow major trends or to converge to the "average" level in the long run (Poterba and Summers, 1988; Grinblatt et al., 1995; Balvers et al., 2000; Moskowitz et al., 2012); Tversky and Kahneman (1991) suggest that losses and disadvantages have greater impact on preferences than gains and advantages on consumer choices, and Kőszegi and Rabin (2006) extend this loss-aversion theory with a reference-dependent preferences; Barsky et al. (2007) approach this issue in terms of sticky price models in durable goods; Thaler (1985, 1999) and Hasting and Shapiro (2013) focus on mental accounting (narrow budgets) and find that consumers may maintain separate budgets for different categories and exhibit excess sensitivity to category-level price shocks. However, previous literature finds little to no evidence regarding quantity decisions of consumers: will consumers change their purchasing patterns or amounts when prices change? Understanding why consumers select certain amounts is crucial to estimate consumer demand and firms' pricing strategies. A retail gasoline market is an excellent place to investigate this issue, as gasoline is a homogeneous good and consumers make quantity decisions frequently. In this paper, I study the effects of the current price level and the past price movements on consumer decisions at gasoline pumps.

The Korean gasoline retail market has two unique features that allow me to analyze consumer quantity decisions at the gas pump. First, almost all gasoline stations offer only full service: consumers simply needed to say the amount that they desire. Second, gasoline pumps can be set to fuel up to a pre-specified amount, unlike in the U.S: this feature allows consumers to purchase the exact amount they want. Using private credit card transaction records for select gasoline stations, I observe total transaction amounts both in dollars and liters and the transaction date and time of each purchase, for individual station level. Documenting stylized facts, I find interesting phenomena: (i) the proportion of the integer dollar amount transactions is very high (90%) while the proportion of the integer liter amount transactions is low (2%), (ii) the proportions of certain dollar amounts are high (for example, 13% of total transactions are 30,000 KRW (Korean Won) and 27% are 50,000 KRW),[1] and (iii) these proportions do not change much even if retail prices change.

Why do consumers choose various amounts? And what are the factors that affect consumer decisions? In this paper, I focus on a special version of this question: when do consumers tend to fill up and what are the governing factors? Note that a transaction is considered as "full" if the transaction amount is not a multiple of 10,000 KRW or if the purchased quantity is over 40 liters (10.57 gallons). Economic models and interviews with consumers who visit gasoline stations suggest that there are two main potential reasons to fill up: first, consumers try to minimize the number of gasoline stations visits to save time; second, consumers speculate that they will face higher price at the next visit. On the other hand, not filling up and choosing pre-set amounts allows consumers to track their gasoline expenditures easily. Also, if consumers expect gasoline prices to go down, they may choose to buy a smaller amount today. Considering these features, I propose two models of rational consumers who have certain speculations and two models of consumers with bounded rationality.

Consumer speculations of future asset prices are studied in depth in financial markets, and previous research papers find that both trend following (or momentum) and mean reversion may be applicable depending on the empirical settings (for trend following, see, e.g., Grinblatt et al., 1995, Fung and Hsieh, 2001; Szakmary et al., 2010; Moskowitz et al., 2012. For mean reversion, see, e.g., Poterba and Summers, 1988; Bessembinder et al., 1995; Balvers et al., 2000). Specifically, the Trend Following Hypothesis assumes that consumers are rational and use all available information at the time of purchase. In addition, they speculate that gasoline prices tend to follow major trends: if gasoline prices were rising in the previous week, it is likely that prices will continue to go up. The other rational model, the Mean Reversion Hypothesis, differs in the direction of gasoline price speculation: consumers believe that gasoline prices tend to move to the average price level over time. Two behavioral models assume that consumers have limited cognitive capacity and may not remember, or would like to not remember, past prices and transactions.

On the other hand, behavioral models assume that consumers have limited information-

[1] 1,000 KRW (Korean Won) is about 1 US dollar.

processing abilities due to cognitive costs. The Attention Hypothesis proposes that consumers pay more attention to their quantity choices when the current price level is high: consumers simply make decisions based on the current price levels. Going one step further, the Narrow Budget Hypothesis assumes that consumers also consider dynamic effects: when a consumers faces higher prices, he is more likely to start managing his gasoline expenditures, fearing that he may spend too much. Compared to simply filling up, choosing pre-set amounts helps him to track his total gasoline spending as he only need to count the number of gasoline purchases, not the amounts for all transactions.

The four hypotheses offer different predictions regarding three factors that may affect consumer decisions at the gas pump: the current price level, the past price movement trends, and daily random price shocks.[2] I run reduced-form regressions to examine the effects of the three factors on the fraction of consumers who fill up. Regression results suggest that this fraction increases by 0.11% when the current price level increases by 10 KRW (1 cent), decreases by 0.22% when there is a positive daily price shock of 10 KRW, and decreases by 0.94% when today's gasoline price is 1% higher than the average price of the past three days. Among the four hypothesis, the Mean Reversion Hypothesis is consistent with the regression results: the reduced-form analysis suggests that consumers think that gasoline prices follow mean-reverting processes and expect price levels to decline when price levels rose recently. In addition, the regression results also suggest that consumers are more likely to fill up if they live in districts of Seoul (higher average income regions) and if they visit gasoline stations during working days.

This paper contributes to the consumer behavior literature by studying gasoline quantity choices that most people frequently make. While there are considerable results on the effectiveness of investing strategies based on trend following or mean reversion on financial markets, relatively little empirical research has focused on other retail markets. Compared to the financial markets, the range of participating consumers is much wider and the frequency of transactions are higher in retail gasoline markets: thus, consumer behavior on gasoline choices may represent more general consumer decision preferences. In addition to adding one more piece of evidence that consumers expect mean reverting processes to occur, this paper also tests psychological

[2]Daily price shocks are measured as residuals from a fixed effect regression with day and region fixed effect dummies.

models of decision-making. In particular, I study consumer inattention due to cognitive costs (see, e.g., Abel et al., 2007; De Clippel et al., 2014; Grubb, 2014) and mental accounting (Thaler, 1985; Moon et al., 1999). While most evidence on mental accounting comes from hypothetical choices or incentivized laboratory settings, this paper illustrates what happens in the real world (Thaler, 1999; Anderson et al., 2010; Hastings and Shapiro, 2013).

I also contribute to the empirical literature on retail gasoline demand. Going to a gasoline station is one of the most frequent purchasing activities for consumers, and it is also economically important for consumers, as motor fuels alone account for more than 5% of all consumer expenditures in the U.S. (Anderson et al., 2011). Naturally, consumer reactions to gasoline prices have been investigated in various aspects: for example, Busse et al. (2009) examines whether the increase in gasoline prices changed automobile choices and prices; Lewis and Marvel (2011) analyzes when consumers search; and Hastings and Shapiro (2013) studies magnitude of substitution from premium gasoline to regular gasoline when prices increase. Gasoline retail pricing has been studied in depth, but Eckert (2013) points out that most papers focused on price levels at individual stations and consumer decisions on station choices (Eckert and West, 2005; Hosken et al., 2008; Houde, 2012). This paper focuses on consumer choices about how much to buy after choosing which stations to visit. While not only are consumer decisions at the extensive margin (station choices) but also consumer choices at the intensive margin (purchasing quantity choices) are crucial to understand overall demand, no other research I am aware of examines consumer decisions at the pump. Note that instead of justifying consumer behaviors at the pump, I analyze how certain factors affect consumer decisions and report what consumers actually do.

The remainder of this article is organized as follows. In the next section I provide details about two data sets and empirical settings. In section 2.3, I document stylized facts and present an method to measure the fraction of consumers who fill up. Section 2.4 presents four models that may explain consumer behavior at the pump. In section 2.5, I construct regression models that estimate the effects of the current price level, the past price movement trends, and random price shocks and discuss whether these results are consistent with the model predictions. Section 2.6 presents conclusions.

2.2 Data

I utilize a unique transaction-level sales records and daily price information to examine consumer quantity choices at gasoline pumps. I also discuss two main advantages of this empirical setting: full-service and gasoline pump setups.

2.2.1 Sales Transaction Records

I obtained credit card transaction records of select individual gas stations in four regions from February 2011 to December 2012. Among the four regions, two regions are districts of Seoul, and the other two regions are small cities that are isolated by mountains.[3] On average, each region has 40 gasoline stations, and the four major gasoline companies have more than a 95% market share in total. Gasoline stations in my dataset are from one major gasoline company that has about 30% market share: for about 20% of the total gas stations, I have the sales information.

This sales data is a set of transaction receipts, without the credit card number or cardholder information: each data entry consists of transaction date, time, quantity sold, unit price, oil type, total amount paid, and a station code. According to the company official, credit cards are used for most of the transactions, and the proportion of credit card transactions has been stable and similar for all four regions during 2011-2012. Prices are represented in terms of KRW (Korean Won. Exchange rate is roughly 1 USD = 1,000 KRW) per liter (1 gallon = 3.785 liters). Transactions with more than 200,000 KRW (about 200 USD) or more than 100 liters (about 26.4 gallons) are excluded, as they are not usual gasoline purchases.[4]

2.2.2 Opinet Data

In 2008, the Korean government established an unusual resource for its citizens. Since April 15, 2008, every gas station in South Korea has been required to report its posted gasoline and diesel price at least once a day. Korea National Oil Corporation, a public institution, is in charge of collecting and publishing the real-time price information. I scraped data from the Opinet website

[3]Except for two districts of Seoul, distances between any two regions are over 50 miles.

[4]According to the gasoline station workers, such transactions are mostly gift-card purchases. These transactions are less than 1% of the total transactions.

and contacted Opinet officials to supplement missing data. I obtained a complete set of historical daily prices for all gas stations in four regions from February 2011 to December 2012 (699 days).[5] Each gas station owner can update the price information by calling the Opinet office, submitting the information on the Opinet website, or using an automated report system. According to an Opinet representative, most gas station owners use automated systems: for each credit or debit card transaction, price information for that transaction is electronically reported to the Opinet server and price information is automatically updated.

2.2.3 Gasoline Stations and Pumps

In addition to the unusually detailed receipt-level dataset, gasoline stations in South Korea during 2011-2012 have two main advantages to study how consumers choose each transaction amount. First, most gasoline stations only offered full-service: consumers simply needed to say the amount they want to a gas station worker. It eliminates one potential factor of consumer decision, hassle. For instance, consumers may want to stop fueling earlier than usual if the weather is bad (too hot, heavy snow, etc) if they visit self-service stations.

Second, gasoline pumps can be set to fuel up to a pre-specified amount, unlike in the U.S. This feature allows consumers to purchase the exact amount they want, either by putting the number themselves (for a self-service case) or telling a station worker the amount they want (for a full-service case). Figure 2.1 (top picture) shows a sample screen of a gas pump in Korea. On the left side, one may choose pre-set dollar amount, multiple of 10,000 KRW[6] from 10,000 KRW to 90,000 KRW, or simple choose "full" option at the left-bottom. On the right side, one may type any specific dollar amount or liter amount, and select an option to "round" up to the nearest multiple of 1,000 KRW (nearest integer USD). If the full option is chosen, usually the round up option is automatically chosen as well. The bottom picture of Figure 2.1 presents an example of a gas pump screen after choosing the full option and automatically fueled up to 70,000 KRW.

[5]In fact, there are 700 days from February 2011 to December 2012. Since my dataset is missing one day (May 31, 2012), I have 699 days.

[6]In terms of the U.S. dollars, this is equal to multiples of 10 USD.

2.3 Descriptive Analyses

In Section 2.3, I present consumer quantity decisions and propose two hypotheses to explain the effect of past price movements. I start with documenting stylized facts and find evidence that consumers make how much to buy based on dollar amounts, not liter (volume). Also, most transaction amounts are multiples of 10,000 KRW, and consumers prefer certain amounts such as 50,000 KRW. I define the fraction of consumers who fill up from the purchased quantity to find that less than 40% of transactions are "full" (filling up) and this proportion seems to be unrelated with retail price levels.

2.3.1 Average Price and Transaction Number Trends

I present average retail prices and daily transaction number trend graphs and a brief interpretation of the graphs.[7] Figure 2.2 shows trends of the average regular gasoline retail prices of the four regions. All graphs (Figures) use weekly moving averages unless noted otherwise. Districts of Seoul (regions 1 and 2) have higher gasoline prices, as average income levels and rents are higher. Regions 3 and 4 are small, isolated cities, and they have very similar average prices that are lower than regions 1 and 2. Overall, average price trends for all regions follow similar trends, as they are basically dependent on international oil prices. Also, consistent regional gaps are observed: for instance, average prices of region 1 are usually 100 KRW higher than those of regions 3 and 4.

The next graph, Figure 2.3, presents daily average prices and transaction numbers of the four regions. In general, fewer transactions occur when average price levels are higher: are the magnitudes consistent with reduced use of gasoline? Previous literature reports a wide range of price elasticity estimates depending on regions and period. Dahl (2012) summarize historical studies of 70 countries to find that the mean price elasticity is -0.34 and the price elasticity estimate of Korea is -0.60, which is one of the highest in the sample.[8] If the quantity purchased

[7]Since there were nominal retail price differences between gasoline companies from 4/7/2011 to 7/7/2011, I treated that period separately (details in Appendix A). During this period, the Korean government asked the four major gasoline companies to cut their prices, and they agreed to reduce retail prices by 100 Korean won per liter. However, one gas company (SKE) chose to offer a rebate of 100 KRW per liter, instead of cutting the posted price directly. This policy caused artificial relative posted price differences between SKE stations and non-SKE stations.

[8]For instance, the price elasticity of the United States is -0.30 in this study. Levin et al. (2009) also obtained

is proportional to the number of transactions, the price elasticity from using the transaction numbers, instead of the purchased quantity, should give us a similar estimate. The estimated price elasticity using the transaction numbers is -0.74, which is slightly more elastic. This suggests that the price movement alone may not fully explain the changes in the number of transactions: consumers may also change purchase amounts when price levels change.

2.3.2 Consumer Decision: Price or Quantity?

I study how much gasoline consumers buy per each transaction to find consumer decisions at gasoline pumps. Two histograms in Figure 2.4 suggest that consumers make purchase decisions based on dollar amounts, not liter amounts. The left histogram shows that transaction dollar amounts are clustered around multiples of 10,000 KRW. In particular, about 11% of total transactions are 30,000 KRW, 27% are 50,000 KRW, 9% are 70,000 KRW, and 7% are 100,000 KRW.

If consumers (i) cannot expect future gasoline prices and (ii) are not budget-constrained, not filling up is clearly sub-optimal, as going to a gasoline station is costly (mainly in time, but also in terms of additional driving to stations.) On the other hand, the right histogram of transaction liter amounts does not have a clear pattern. While there is a cluster around 25 liter, it is unlikely that consumers choose these liter amounts: this clustering is a simple consequence of consumers choosing 50,000 KRW often (50,000 KRW is about 25 liters, as average price during this period is about 2,000 KRW per liter.)

Graphs in Figure 2.5 show that the proportion of transactions that are round dollar amounts is in sharp contrast with that of liter amounts. The left graph shows the proportion of transactions that are multiples of 10,000 KRW or 1,000 KRW. These proportions are fairly stable regardless of the price level, and very high: 70% of total transactions are multiples of 10,000 KRW and more than 90% are multiples of 1,000 KRW.[9] On the other hand, the proportion of transactions that are integer liter amounts are mostly 2-3%, except for a few points. These points are when retail prices $\times$ integer liter amounts = multiples of 10,000 KRW. For example, when a retail price is

price elasticity estimates ranging from -0.29 to -0.61.

[9] In terms of USD, 70% of total transactions are multiples of 10 USD and more than 90% are integer dollar amounts.

2,000 KRW, any liter amounts that are multiples of five would give total dollar amounts that are multiples of 10,000 KRW, and the fraction of integer liter amount transactions is extremely high (87%). Also, when a retail price is 2,083 KRW, this fraction is 36%, since buying 50,000 KRW is equal to buying 24.00 liter. Note that when a retail price is 2,082 or 2,084 KRW, the fraction of integer liter transactions is below 3%.[10] Thus, these patterns indicate that consumers tend to choose certain dollar amounts, not liter amounts, while they can theoretically choose any liter amounts equally easily: one can either say 50,000 KRW or 25 liter to a gasoline station worker or type these numbers in a gas pump keypad (recall Figure 2.1), for instance.

Lastly, Figure 2.6 shows the proportion of transactions that are 30,000 KRW, 50,000 KRW, 70,000 KRW, and 100,000 KRW for difference price levels. As retail prices go up, the proportion of consumers who chooses 100,000 KRW slightly increases and the ratio of consumers who choose other amounts slightly decrease. However, these proportions are fairly stable for all price levels, implying that consumers who choose these dollar amounts are not sensitive to the retail prices.

2.3.3 Do Consumers Fill up?

In the previous section, I observed that consumers tend to purchase amounts that are multiples of 10,000 KRW very often, and certain amounts are chosen frequently: for example, 27% of total transactions are 50,000 KRW and 13% are 30,000 KRW. Since 50,000 KRW is about 25 liters or 6.6 gallons (assuming 2,000 KRW per liter), choosing 50,000 KRW or 30,000 KRW is not filling gas tanks completely, as consumers tend to visit gasoline stations when they are running low on gas and tank capacity is usually 50 liters or more. If we focus on the opportunity time cost of visiting (and finding) gasoline stations, consumers should always choose to fill up so that they can minimize the number of visits. However, simply considering consumers who buy 30,000 KRW and 50,000 KRW shows that we are missing something.

Why do consumers choose various amounts? And what are the factors that affect consumer decisions? In this paper, I focus on a more specific question: when do consumers tend to fill up and what are the governing factors? To start, I define filling up, or choosing the "full" option. First, transactions that are not multiples of 10,000 KRW are likely to be "full", as

[10] To emphasize contrasts between 2,083 KRW and 2,082 and 2,084 KRW, I do not use moving average for this graph.

consumers seldom choose specific amounts that are not multiples of 10,000 KRW in advance: these non-multiples are results of choosing the full button at gas pumps. Also, there are some full transactions among transactions that are multiples of 10,000 KRW, as choosing the full option usually includes the "round" up option automatically. Thus, high volume transactions should be treated as filling up, even though they are multiples of 10,000 KRW. I choose 40 liters as a cutoff: any transaction that is over 40 liters is considered as a "full" transaction.[11] Note that this counts 100,000 KRW purchases as fill-ups. Third, there could be "full" cases that happen to be multiples of 10,000 KRW for transactions that are less than 40 liters: suppose a consumer simply chose the full option, but it ended up at a multiple of 10,000 KRW, say 70,000 KRW. To account for such possibilities, I assume that the proportion of consumers who choose to fill up and end up at 70,000 KRW are an average of consumers who end up at 69,000 KRW or 71,000 KRW: if 0.35% of transactions are 69,000 KRW and 0.33% are 71,000 KRW, then I assume that 0.34% of transactions that are 70,000 KRW are fill-ups and not results of setting 70,000 KRW in advance. Under these assumptions, I compute $FullFrac$, the fraction of consumers who fill their cars up, simply by looking at the transaction amounts and counting three types of fill-ups.

Figures 2.7 and 2.8 show the fraction of consumers who choose the "full" option. Figure 2.7 tells us that while trends of the two districts of Seoul (regions 1 and 2) are almost identical, there are significant gaps between other regions: region 3 is about 5% below regions 1 and 2, and region 4 is about 10% below. Another notable fact is that the fraction of consumers who fill up is low: even for regions 1 and 2, $FullFrac$ is below 50% during the whole period.

In addition to regional variation, Figure 2.7 also shows that there is time-series variation. Where does this time-series variation come from? An average retail price level is an obvious candidate: however, Figure 2.8 presents that the effect of the average price level is not clear. It presents how the $FullFrac$ variable changes as average retail price increases: there is no clear tendency, and on average less than 40% of transactions are "full." Thus, the average price level cannot fully explain consumer choices. In the next section, I discuss four rational and behavioral models that may explain variation of the fraction of "full" transactions.

[11]The reduced-form analyses are not sensitive to this assumption. For instance, choosing 45 liters as a cutoff does not qualitatively change the results.

2.4 Rational and Behavioral Models

In this section, I present four hypotheses, two rational speculation models and two bounded rationality models, to answer these questions: why do some consumers fill up and some do not? And what are the factors that affect consumer decisions? One potential reason for not filling up is that consumers may face liquidity concerns. For instance, if a consumer needs to wait until he receives his paycheck at the end of the week, he may choose to purchase the minimum amount of gasoline before his payday. However, such a liquidity issue would not have a big impact on consumers in this empirical setting, as I only consider credit card transactions. Since a consumer may pay for the gasoline she buys today more than a month later, when her credit card statement is due, she is basically free to choose any amount, unlike cash transactions that she needs to pay immediately. Thus, I concentrate on consumer reactions to gasoline prices. In particular, I focus on these three factors: the current price level, the past price movement trend, and random price shocks: I analyze predictions of the four hypothesis regarding the effects of the three factors on consumer fill-ups (Table 2.1).

2.4.1 The Trend Following Hypothesis

Suppose that a driver is running low on gas today and he expects gasoline prices to rise in the near future. Then he is more likely to fill his car up, as it will be more costly to buy gas later. On the other hand, if he expects gasoline prices to decline, then he would be hesitant to fill up, as he might buy gasoline at a lower price later.

While there is no guaranteed theoretical method that can predict future gasoline price movements,[12] retail gasoline prices tend to move in the same direction, continuously decreasing or increasing for several days, or even weeks. Figure 2.3 shows that there were six major uptrends (and downtrends) during this two-year sample period. Knowing this, sophisticated consumers might make an "educated guess" based on recent price movements: if prices are on an upward trend recently, consumers might hoard gasoline, expecting that this increasing trend will continue (see, e.g., Grinblatt et al., 1995, Fung and Hsieh, 2001; Szakmary et al., 2010; Moskowitz et al., 2012). Thus, this hypothesis predicts that the fraction of consumers who fill up is higher when

[12] According to the efficient market theory, all relevant information is already reflected to the current price.

the past price movement trend is going up.

Similarly, this theory predicts that this fraction would be lower if there is a positive random price shock. Since the expected value of random shock at the next visit is zero, the price at the next visit is likely to be lower than the current price. Thus, consumers are less likely to fill up today. On the other hand, the Trend Following Hypothesis implies that the fraction of fill-ups does not depend on the current price level.

2.4.2 The Mean Reversion Hypothesis

As in the previous section, I assume that consumers make rational speculations of the future gasoline prices. In this setup, consumers are making a different "educated guess" based on the past prices. Instead of following recent movements, consumers assume that gasoline prices will revert to the mean level. This mean reversion theory is widely studied in other assets, especially financial markets (see, e.g., Poterba and Summers, 1988; Bessembinder et al., 1995; Balvers et al., 2000). According to the Mean Reversion Hypothesis, consumers are less likely to fill up their cars when prices are on an upward trend, as they expect that prices are likely to go back to the mean level.

While this hypothesis provides a different expectation regarding the effect of the past price movement trends, it offers qualitatively the same predictions about the effect of a positive random price shock and the effect of the current price level. There is no reason for consumers to modify their behavior solely based on the current price level itself. Also, observing a positive random price shock implies that the price level is more likely to drop to the mean level, which reduces consumers' incentive to fill up today.

2.4.3 The Attention Hypothesis

The Attention Hypothesis provides a behavioral explanation of why the current price level might matter. When gasoline prices are low, most consumers simply do not care much about their gasoline purchases, as they are not spending more than they expect (see, e.g., Abel et al., 2007; De Clippel et al., 2014; Grubb, 2014). While choosing the full option is economically efficient if we only consider total travel time to gasoline stations, many consumers report that they choose the full option for a different reason: it is hassle-free and they do not need to think about how much to purchase.

However, when gasoline prices are high, more consumers pay attention and decide to actively manage the purchases by choosing certain amounts before starting fueling. For example, a consumer tends to simply fill up when prices are in the 1,900-1,999 KRW range, but chooses 70,000 KRW when retail prices are 2,000 KRW or higher. Note that this hypothesis implicitly assumes that consumers are myopic – that they only care about the current prices, not price movement trends. In fact, if consumers believe that future gasoline prices are unpredictable, according to the efficient market theory, it is rational to only focus on the current price. Therefore, this hypothesis predicts that the fraction of consumers who fill up is higher when the current price level is higher, but does not depend on price movement trends. Moreover, consumers do not necessarily react to random shocks, as they only pay attention to the current price level.

2.4.4 The Narrow Budget Hypothesis

The Narrow Budget Hypothesis extends the Attention Hypothesis and assumes that consumers also consider dynamic aspects. Under this setup, consumers are concerned with both the current price level and price movement trends. In addition to the reasons described in the previous section, if a gasoline price level is high, an incentive to actively manage purchases is also high. If a consumer wants to keep track of how much he spends on gasoline, choosing pre-set amounts is helpful: paying $50 each time for gasoline allows one to easily calculate how much he has spent on gasoline, as he only needs to remember how many times he has visited gasoline stations during a given period of time (Hastings and Shapiro, 2013). On the other hand, if he fills up whenever he visits, it is difficult to remember all transaction amounts. This is actually in line with the fact that the proportion of 50,000 KRW transactions is very high, as the calculation is easy. Thus, the fraction of consumers who fill up should be low when the current price level is high. Similarly, consumers are also less likely to fill up if there is a positive random price shock.

Moreover, the Narrow Budget Hypothesis implies that consumers become more anxious about gasoline spending when prices are on an upward trend, and as a result, consumers are more inclined to choose specific amounts, instead of simply filling up. Suppose that gasoline prices are going up. Now a consumer worries that he may spend too much on gasoline, if he continues to fill up. For instance, he may spend \$2,000 per month, and most of his expenditures are not elastic: he has to pay \$1,200 for rent, \$200 for utility bills, \$500 for groceries, and so on. Even if there are no concrete divisions, mental accounting theory suggests that consumers tend to allocate separate budgets for different categories (Thaler, 1999; Hastings and Shapiro, 2013). For example, a consumer might plan to spend \$100 for gasoline for one week, \$500 for groceries, and so on, instead of spending \$2,000 per month in total. Then, his residual disposable income for gasoline has a narrow range, and the fact that he is likely to face higher prices in the near future (if he believes that the increasing trend is likely to continue) strengthens his incentive to control his gasoline purchases, as his budget will be even narrower in this case. Therefore, the Narrow Budget Hypothesis predicts that consumers are less likely to fill up when the price movement trend is going up.

2.5 Reduced-Form

The graphs and histograms in the previous sections were suggestive, but do not account for all factors which could affect the fraction of consumers who fill up. Therefore, this section analyzes how this fraction changes as the past price movement trend changes, using reduced-form regressions, to test four proposed hypotheses of interest. Starting with the definitions and summary statistics of key variables, I present regression models and discuss results.

2.5.1 Definitions and Summary Statistics

Table 2.2 defines the variables used in the analysis. As defined before, $FullFrac$ is the fraction of consumers who choose to fill up. $CurrentP$ is daily, regional average retail gasoline prices and $AvgRetailP_{r,t-period,t-1}$ is the average price of past $period$ days (from time $t-1$ to $t - period$) for region r. Using these two variables, I define a past price movement trend variable, $TrendP_{r,t-period} = CurrentP_{rt}/AvgRetailP_{r,t-period,t-1} - 1$. If the current price is 1% higher than the average price of past days, the trend variable is 1 (values are in percent terms). $DailyShock$ denotes daily, regional variation of prices that is not explained by day and region fixed effects. In other words, this variable is a residual term of the fixed-effect regression $CurrentP_{rt} = \alpha_0 + \Sigma\gamma_r Region_r + \Sigma\delta_t Day_t + \epsilon_{rt}$. $Seoul$ and $Workday$ are dummy variables that denote whether region r is a district of Seoul or time t is a working day.[13]

Summary statistics of these variables are shown in Table 2.3. Note that $FullFrac$ and $TrendP$ are in percent terms, while $Seoul$ and $Workday$ are ratios (that are between zero and one). All of $CurrentP$, $AvgRetailP$, and $DailyShock$ are in terms of Korean Won (KRW) per liter. $FullFrac$ has a mean of 36.11(percent) and a standard deviation of 7.37, which implies that more than half of transactions are not "full" for more than 90% of the cases. $TrendP$ has a mean very close to zero (0.015%), which makes sense as there should not be large changes for 3 days. Since two regions are districts of Seoul and two are not, $Seoul$ has a mean of 0.50 and a standard deviation of 0.50. $Workday$ has a mean of 0.68 that is less than $\frac{5}{7} \fallingdotseq 0.714$, as both weekend and national holidays are not counted as working days.

2.5.2 Regression

I proceed to estimate the effect of past price trends on the fraction of "full" consumers as follows:

$$FullFrac_{rt} = \beta_0 + \beta_1 TrendP_{r,t-period} + \beta_2 CurrentP_{rt} + \beta_3 DailyShock + \beta_s Seoul_r + \beta_w Workday_t + \epsilon_{rt}$$

where r and t stand for region and time(date), respectively. The main variable of interest is $TrendP_{r,t-period}$, the ratio of the current price ($CurrentP_{rt}$) and the mean of past prices during a certain period of time ($AvgRetailP_{r,t-period,t-1}$). In other words, $TrendP_{r,t-period} = CurrentP_{rt}/AvgRetailP_{r,t-period,t-1} - 1$. $DailyShock$ measures the effects of temporary price shocks on the fraction of "full" consumers. Lastly, to control for region and time fixed effects, I use $Seoul_r$ to indicate whether a certain region is a district of Seoul, and $Workday_t$ to indicate whether time t is a working day.

Table 2.4 shows the regression results. Column 1 provides estimates for the regression with $period = 3$ (mean prices of three previous days are compared to the current prices): the fraction

[13] A date is a working day if it is not a national holiday or weekend.

of "full" consumers ($FullFrac$) is 0.11% higher when the current average price among stations in the region is 10 KRW higher (10 Korean Won is about 1 cent in USD), 0.22% lower when there is a price shock of 10 KRW, and 0.94% lower when the current price is 1% higher than the mean price of the previous three days. I interpret coefficient estimates of the main regression, column 1, below.

While the $CurrentP$ estimate is statistically significant, the magnitude is quite small. For simplicity, let us assume that the average retail price is 2000 KRW (per liter). If a retail price becomes 1% higher, or 20 KRW higher, then the fraction of "full" consumers increases by 0.22%. As prices seldom move more than 2% or 40 KRW within a week, the effect of the current price level *per se* is less than 0.44% for most cases. This is consistent with Figure 2.8 that the fraction of "full" consumers does not vary much on price levels. Also, this coefficient estimate suggests that the effect of the current price *per se* is a 4.68% difference in the fraction of "full" consumers, even if I compare the minimum and the maximum price during the period, since the range of average retail prices is from 1825 KRW to 2250 KRW during the sample period. This result is close to the predictions of the Trend Following Hypothesis and the Mean Reversion Hypothesis, as these models imply that the current price level should have no meaningful effect on the fraction of fill-ups.

The negative $DailyShock$ coefficient estimate suggests that consumers are less likely to fill up during "unusually" high price days. The magnitude of this effect is twice that of the current price level: other things being equal, the total effect of having a 10 KRW price shock is -0.11%, the sum of the higher current price level effect (0.11%) and the positive shock effect (-0.22%). Except for the Attention Hypothesis, all proposed hypotheses predict that the effect of a random price shock to be negative, and these predictions are consistent with what regressions report.

On the other hand, relative price movement trends ($TrendP$) have a much bigger impact. When the oil price is on an upward trend, consumers are less likely to fill up their cars. Suppose that the average price level is 1% higher today, compared to that of the past 3 days. There are two effects: the fact that the current price level is higher accounts for a 0.11% increase, but that the price trend is going up (by 1%) leads to a 0.94% decrease of $FullFrac$. Thus, the combined effect is a 0.83% decrease. Note that the magnitude of the price movement trends is almost nine times that of the current price level. These results are consistent with an explanation that consumers do not care much about the price level *per se*, but consider a general price trend when they make quantity decisions at the gas pumps. Note that this negative direction is consistent with the Narrow Budget Hypothesis and the Mean Reversion Hypothesis.

The estimates of the three explanatory variables suggest that the regression results support the Mean Reversion Hypothesis. While the Trend Following Hypothesis predicts that consumers will fill up more when the past price movement trends are going up, the $TrendP$ term is significantly negative. Also, the current price level does not decrease the fraction of "full" consumers, unlike what the Narrow Budget Hypothesis predicts. Lastly, predictions of the Attention Hypothesis is only consistent with the $CurrentP$ term, as this model predicts no effect for both of the past price movement trends and random price shocks.

Next, I discuss region and time fixed effects. Whether this region is a district of Seoul (or a small rural city) has a large effect, 6.67%, on the fraction of consumers who choose to fill up. Given that districts of Seoul (regions 1 and 2) are quite different from other small cities (regions 3 and 4) in terms of average income and traffic congestion levels, it is not surprising that regional differences are significant. Since Seoul consumers are more likely to have higher incomes, they would care less about small gasoline price differences. It is also possible that the cost of going to a gas station is higher in Seoul, in terms of driving time and opportunity costs.

For the time fixed effects, whether it is a working day or not is found to be important: consumers are 2.93% more likely to fill up during working days. Since drivers are more likely to be time-constrained during working days, it is puzzling that more consumers choose the "full" option as filling tanks takes slightly more time than not filling. One possible explanation is that the income distribution of consumers during working days and the distribution of consumers during weekends (including holidays) are different: consumers who visit gasoline stations during working days are more likely to be higher income. This sample selection story is similar to the region fixed effect interpretation in the previous paragraph.

Column 2 is for the $period = 7$ case and column 3 is for the $period = 14$ case. Estimates are similar with the baseline case of $period = 3$: $CurrentP$ is positively significant but small; $TrendP$ and $DailyShock$ are negatively significant. Both fixed effect terms, $Seoul$ and

$Workday$, are positively significant and magnitudes are consistent for all specifications. Note that the magnitude of the relative price movement trend term ($TrendP$) is also fairly stable across different time periods. This suggests that these estimation results are robust to the length of the period. Lastly, column 4 shows that estimates are quite similar to the ones from column 1, even without the relative price trend term.

2.6 Conclusion

This paper analyzes consumers' quantity decisions at gasoline pumps, using a detailed transaction-level dataset from four regions in Korea. The Korean retail gasoline market from 2011 to 2012 has two properties that help to reveal consumer choices: almost all gasoline stations offer full service and gasoline pumps have an option to fuel up to pre-set amounts. These features eliminate other factors that may affect consumer decisions, such as weather or temperature.

While consumer decisions on station choices are discussed in the previous literature, consumer decisions after choosing stations have not yet been studied in depth. To analyze consumer choices and firms' strategies, not only is where to go important, but also how much to fuel is critical: average quantity purchased is related to both how consumers choose which station to visit and how often they visit. In particular, I study the fraction of consumers who choose to fill their cars up as a measure of consumer quantity decision.

This paper provides empirical evidence of certain factors that affect consumer decisions to fill up or not. Descriptive analyses find that three main factors – the current price level, the past price movement trends, and random price shocks – influence how often consumers choose to fill up: consumers are slightly more likely to fill up when the current price level is higher; and consumers tend to fill up less often when prices are going up recently and when there is a positive price shock. In particular, the proportion of filling up increases by 0.11% when the current price level becomes 10 KRW higher, decreases by 0.94% when retail price levels increase by 1% compared to the past three days, and decreases by 0.22% when there is a price shock of 10 KRW.[14] Among the four proposed hypotheses, the Mean Reversion Hypothesis offers the best

[14] 10 Korean Won (KRW) is about 1 cent in USD

predictions that are consistent with the reduced-form results: consumers expect that gasoline prices will go back to "mean" levels when there are changes in price levels.

However, the overall magnitude of difference is small: about 35% of consumers always fill up and more than 50% of consumers always choose some integer dollar amount. Less than 15% of consumers change their decisions from time to time, and about two-thirds of these movements can be attributed to whether it is a working day or not and whether a region is a district of Seoul. This surprisingly persistent propensity emphasizes the importance of personal habits in consumer purchase decisions.

References

[1] Andrew B Abel, Janice C Eberly, and Stavros Panageas. Optimal inattention to the stock market. *The American economic review*, pages 244–249, 2007.

[2] Soren T Anderson, Ryan Kellogg, and James M Sallee. What do consumers believe about future gasoline prices? Technical report, National Bureau of Economic Research, 2011.

[3] Ronald Balvers, Yangru Wu, and Erik Gilliland. Mean reversion across national stock markets and parametric contrarian investment strategies. *The Journal of Finance*, 55(2):745–772, 2000.

[4] Robert B Barsky, Christopher L House, and Miles S Kimball. Sticky-price models and durable goods. *The American economic review*, pages 984–998, 2007.

[5] Hendrik Bessembinder, Jay F Coughenour, Paul J Seguin, and Margaret Monroe Smoller. Mean reversion in equilibrium asset prices: evidence from the futures term structure. *The Journal of Finance*, 50(1):361–375, 1995.

[6] Meghan R Busse, Christopher R Knittel, and Florian Zettelmeyer. Pain at the pump: the differential effect of gasoline prices on new and used automobile markets. Technical report, National Bureau of Economic Research, 2009.

[7] Carol A Dahl. Measuring global gasoline and diesel price and income elasticities. *Energy Policy*, 41:2–13, 2012.

[8] Geoffroy De Clippel, Kfir Eliaz, and Kareen Rozen. Competing for consumer inattention. *Journal of Political Economy*, 122(6):1203–1234, 2014.

[9] Andrew Eckert. Empirical studies of gasoline retailing: A guide to the literature. *Journal of Economic Surveys*, 27(1):140–166, 2013.

[10] Andrew Eckert and Douglas S West. Retail gasoline price cycles across spatially dispersed gasoline stations. *Journal of Law and Economics*, 47(1):245–273, 2004.

[11] Glenn Ellison. Bounded rationality in industrial organization. *Econometric Society Monographs*, 42:142, 2006.

[12] William Fung and David A Hsieh. The risk in hedge fund strategies: Theory and evidence from trend followers. *Review of Financial studies*, 14(2):313–341, 2001.

[13] Mark Grinblatt, Sheridan Titman, and Russ Wermers. Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *The American economic review*, pages 1088–1105, 1995.

[14] Michael D Grubb. Consumer inattention and bill-shock regulation. *The Review of Economic Studies*, page rdu024, 2014.

[15] Justine S Hastings and Jesse M Shapiro. Fungibility and consumer choice: Evidence from commodity price shocks. *The Quarterly Journal of Economics*, page qjt018, 2013.

[16] Daniel S Hosken, Robert S McMillan, and Christopher T Taylor. Retail gasoline pricing: What do we know? *International Journal of Industrial Organization*, 26(6):1425–1436, 2008.

[17] Jean-François Houde. Spatial differentiation and vertical mergers in retail markets for gasoline. *The American Economic Review*, 102(5):2147–2182, 2012.

[18] Botond Kőszegi and Matthew Rabin. A model of reference-dependent preferences. *The Quarterly Journal of Economics*, pages 1133–1165, 2006.

[19] Laurence Levin, Matthew S Lewis, and Frank A Wolak. High frequency evidence on the demand for gasoline. Technical report.

[20] Matthew S Lewis and Howard P Marvel. When do consumers search? *The Journal of Industrial Economics*, 59(3):457–483, 2011.

[21] Philip Moon, Kevin Keasey, and Darren Duxbury. Mental accounting and decision making:: The relationship between relative and absolute savings. *Journal of Economic Behavior & Organization*, 38(2):145–153, 1999.

[22] Tobias J Moskowitz, Yao Hua Ooi, and Lasse Heje Pedersen. Time series momentum. *Journal of Financial Economics*, 104(2):228–250, 2012.

[23] James M Poterba and Lawrence H Summers. Mean reversion in stock prices: Evidence and implications. *Journal of financial economics*, 22(1):27–59, 1988.

[24] Andrew C Szakmary, Qian Shen, and Subhash C Sharma. Trend-following trading strategies in commodity futures: A re-examination. *Journal of Banking & Finance*, 34(2):409–426, 2010.

[25] Richard Thaler. Mental accounting and consumer choice. *Marketing science*, 4(3):199–214, 1985.

[26] Richard H Thaler. Mental accounting matters. *Journal of Behavioral decision making*, 12(3):183–206, 1999.

[27] Amos Tversky and Daniel Kahneman. Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics*, pages 1039–1061, 1991.

Appendix A. Government Intervention Period

There were two international issues that caused a big jump in Dubai oil prices in March, 2011. The series of protests and demonstrations across Middle East and North Africa caused social unrest. Also, trade sanctions against Iran directly affected oil supply. Following the international price spike, the Korean domestic gasoline prices went up by more than 150 Korean won per liter (about 57 cents per gallon) within a month. Since the price of gasoline plays a important role in a retail price index that people are interested in, the Korean government chose to intervene in the retail gasoline market to stabilize prices, and asked four major gasoline companies to cut gasoline distribution prices (average wholesale prices). SKE, a leading gasoline company, announced a price cut of 100 Korean won per liter from 4/7/11 to 7/7/11 and other companies followed. According to an Opinet representative, discounts by three companies (GSC, HDO, and SOL) were reflected in the price data, as they cut the distribution price directly and posted prices went down. However, SKE offered refund bonus points that were equivalent to 100 Korean won per liter discount to customers after their purchases. Thus, posted prices for SKE stations did not reflect the discount. Figure 2.9 presents that there are big gaps between national average prices among gas companies from April 2011 to July 2011.

To test this information, I compared mean prices for SKE gas stations and those of three other major companies. If the different discount methods were the only reason of the spike, we should be able to observe that compared to other periods, SKE average prices are around 100 won (per liter) higher than average prices of other companies from April 2011 to July 2011.[15] For the main analysis, I used to modified prices for SKE stations (retail prices were subtracted by 100 Korean won) from 4/7/11 to 7/7/11. As a robustness check, I also tried using the data without this period, and using the data without SKE stations from 4/7/11 to 7/7/11. The results were similar in all cases.

[15] It is difficult to understand what exactly happened during this period, as many issues such as supply chain networks and political considerations are involved.

Appendix B. Figures and Tables

Figure 2.1: Sample Gas Pump Screens

Figure 2.2: Average Regular Gasoline Prices of Four Regions

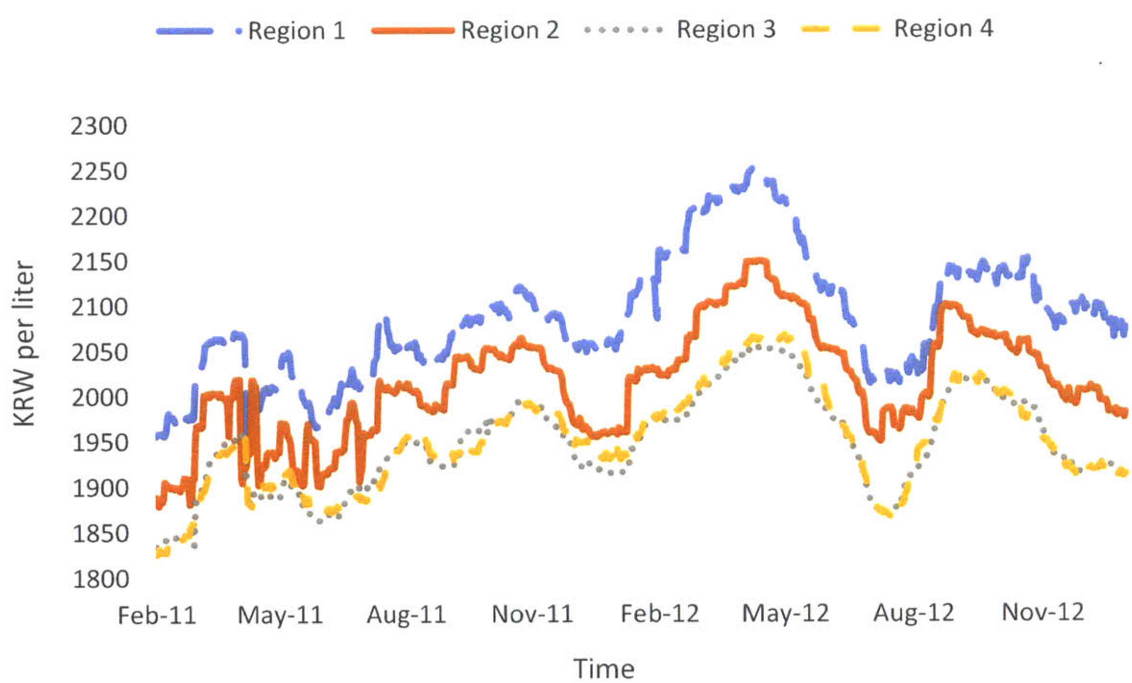


Figure 2.3: Average Price and Daily Transaction Numbers

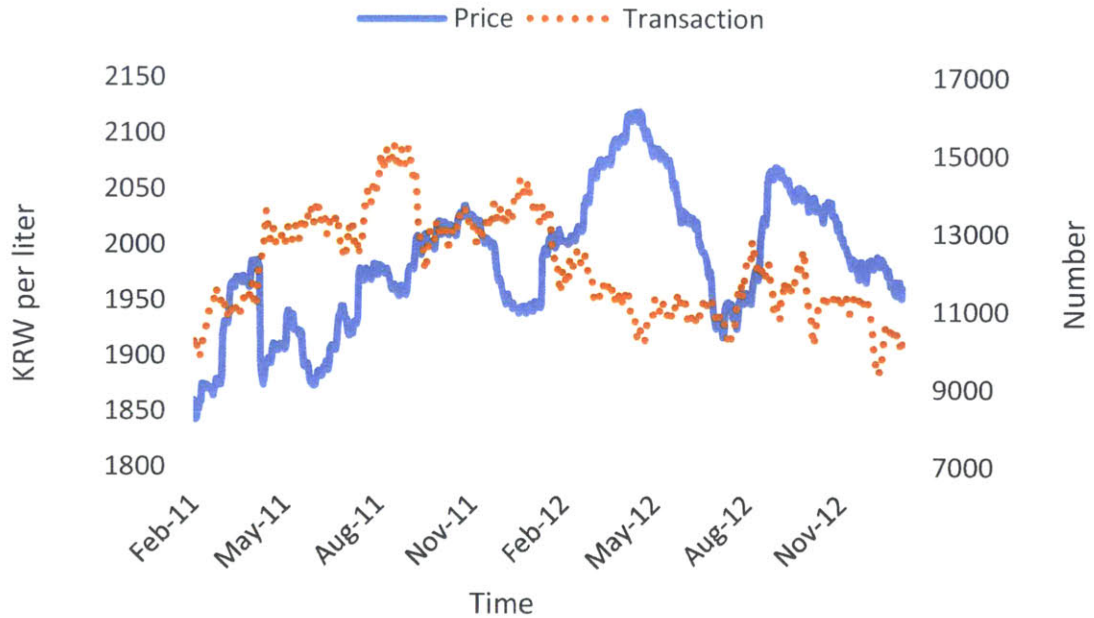


Figure 2.4: Purchased Amounts Frequency, Dollar and Liter

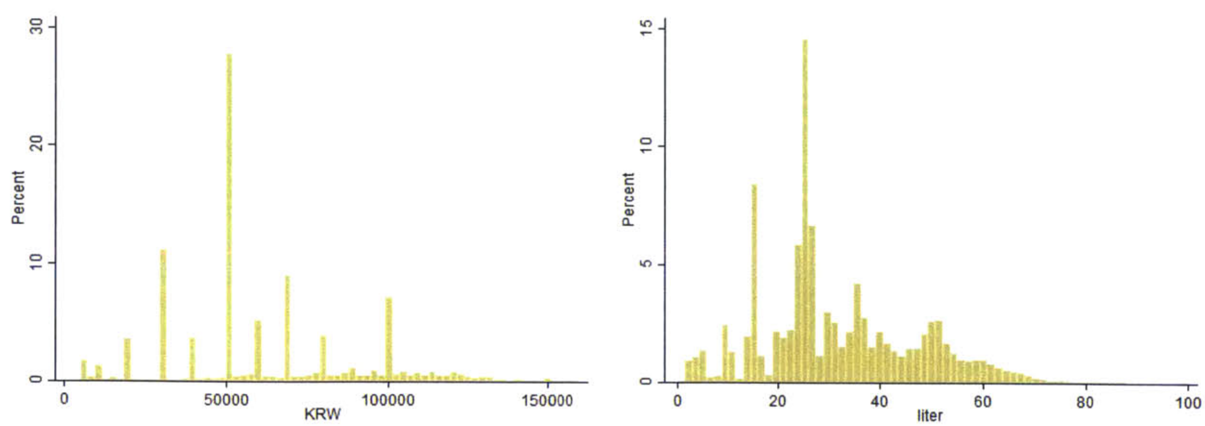


Figure 2.5: Round Numbers Frequency, Dollar and Liter

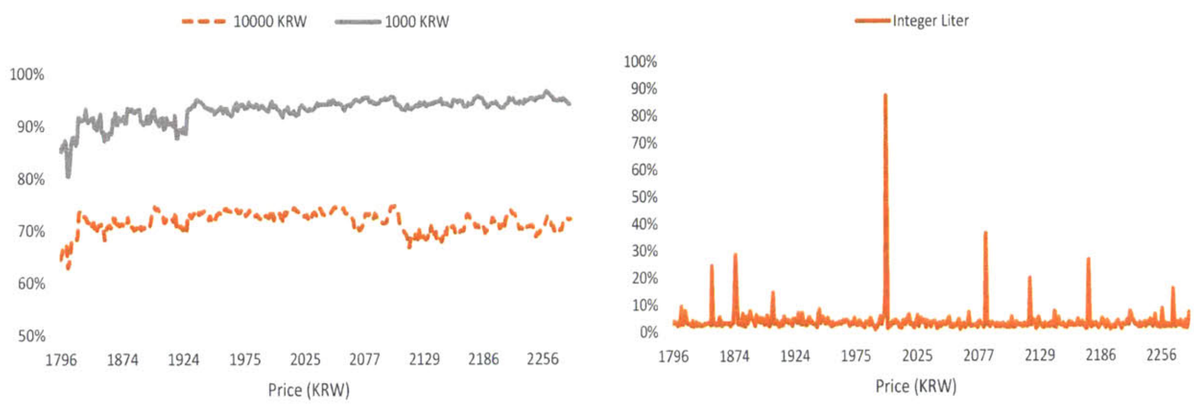


Figure 2.6: Proportion of Transactions

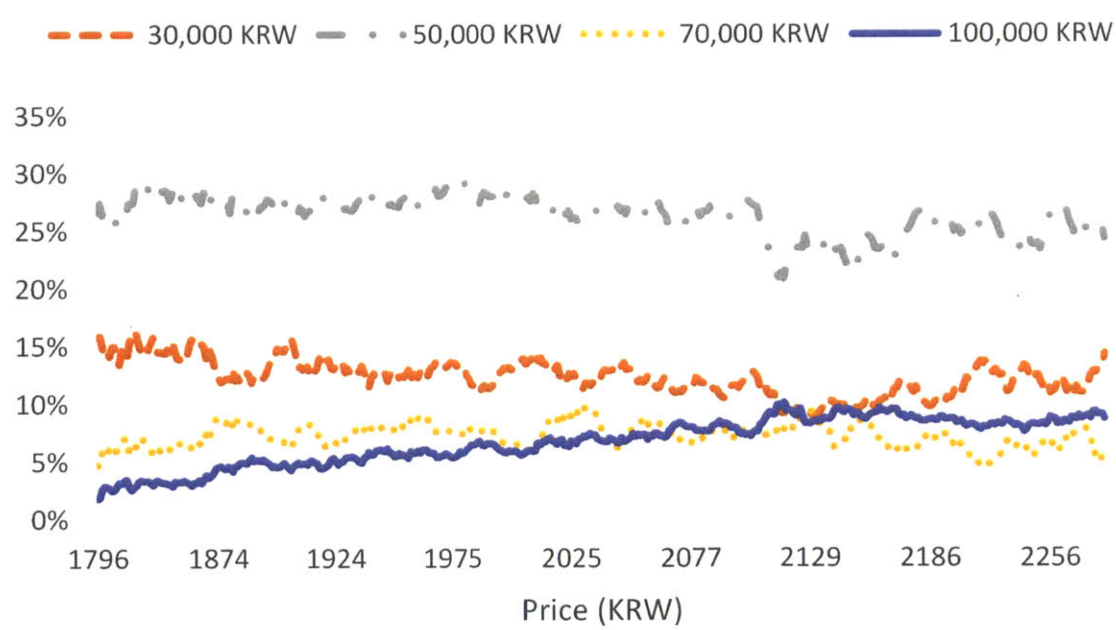


Figure 2.7: The Fraction of Consumers Who Fill Up (Region)

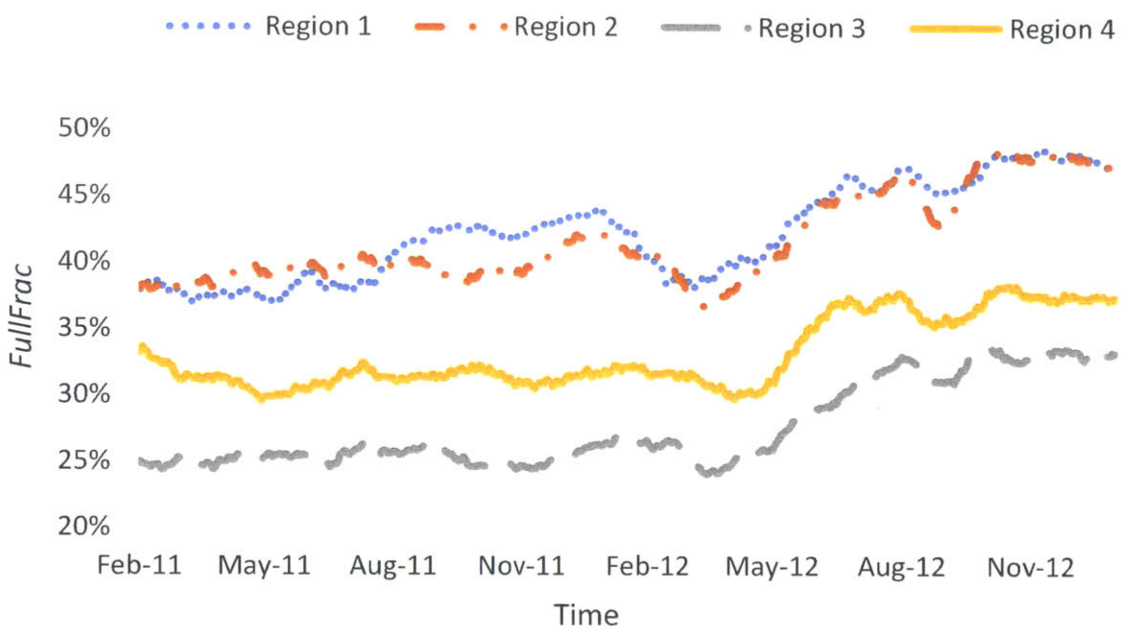


Figure 2.8: The Fraction of Consumers Who Fill Up (Price)

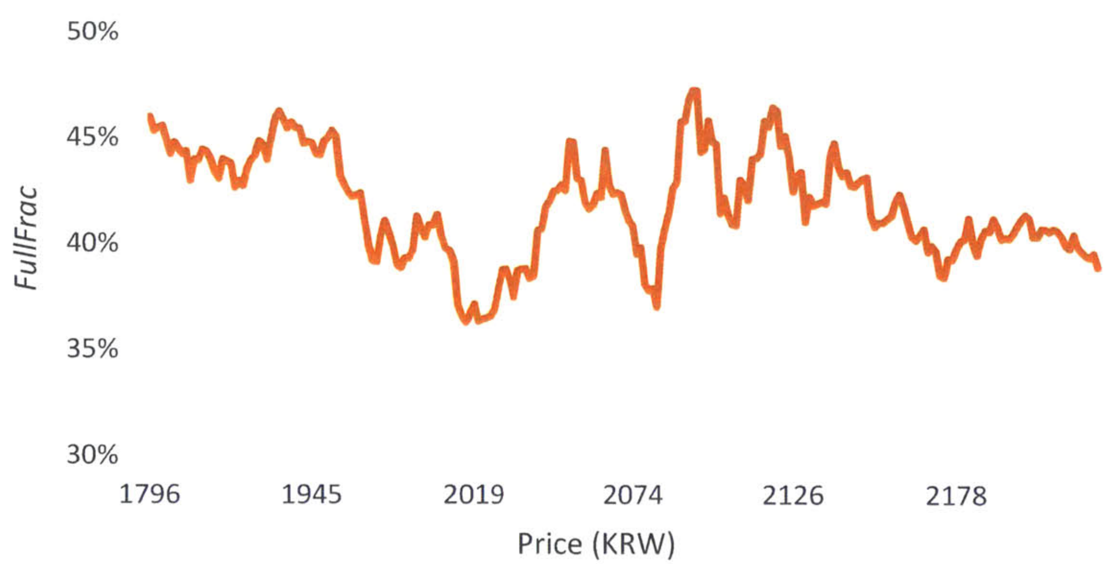


Figure 2.9: 2011 National Average Retail Prices (by company)

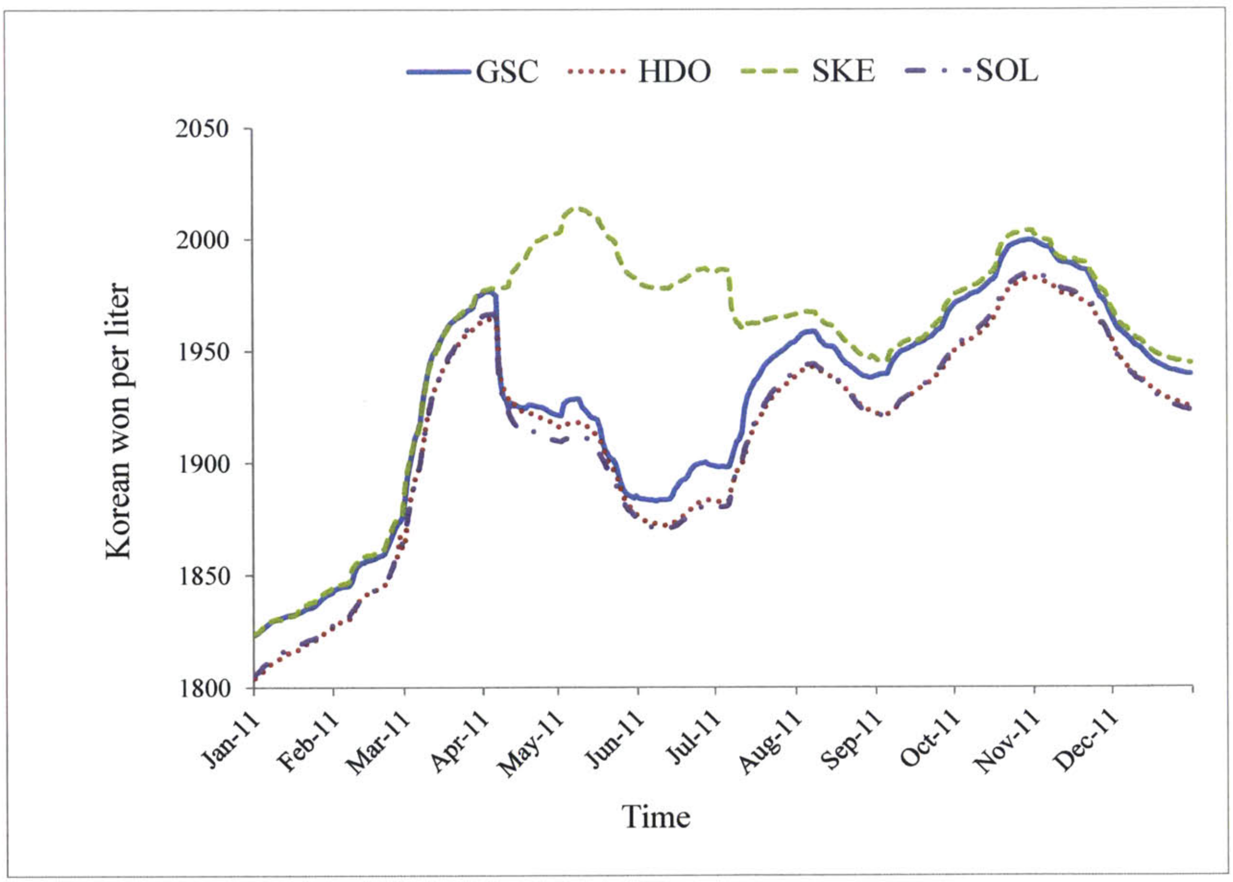


Table 2.1: Predictions of the Four Hypotheses

	Trend Following	Mean Reversion	Attention	Narrow Budget
Current Price Level	No Effect	No Effect	Negative	Negative
Past Price Movement Trend	Positive	Negative	No Effect	Negative
Random Shocks (Residuals)	Negative	Negative	No Effect	Negative

Table 2.2: Definition of Variables

Variable	Indexes vary over	Definition
$FullFrac$	$r,\ t$	The fraction of consumers who choose to fill up
$CurrentP$	$r,\ t$	Daily average retail gasoline prices of each region
$AvgRetailP$	$r,\ t,\ t-p$	Average retail gasoline prices of the past p days of each region
$TrendP$	$r,\ t-p$	The ratio of the current price and the average price of the past p days
$DailyShock$	$r,\ t$	Residuals after taking out day and region fixed effects
$Seoul$	r	1 if a region is a district of Seoul, 0 otherwise
$Workday$	t	1 if time t is a working day, 0 otherwise

Notes: r: region, t: time, p: a period of time (3 days, a week, and two weeks)

Table 2.3: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	Obs.
$FullFrac$	36.11	7.37	20.71	54.30	2796
$CurrentP$	1987.87	83.07	1825.70	2251.90	2796
$AvgRetailP$	1987.61	83.18	1825.70	2250.37	2796
$TrendP$	0.015	0.42	-4.85	2.95	2796
$DailyShock$	0.00	13.43	-40.25	52.21	2796
$Seoul$	0.50	0.50	0	1	2796
$Workday$	0.68	0.47	0	1	2796

Notes: Korean Won (KRW) per liter metric is used for prices. Ratios are presented in percent terms. $AvgRetailP$ and $TrendP$ are $period = 3$ case. The number of observation is 2796, 699 days times 4 regions. $FullFrac$ and $TrendP$ are in percent terms, while $Seoul$ and $Workday$ are ratios (that are between zero and one).

Table 2.4: Regular Gasoline Regression Results

Dependent Var.	$FullFrac$			
	(1)	(2)	(3)	(4)
	$period = 3$	$period = 7$	$period = 14$	No $period$
$CurrentP$	0.011**	0.012**	0.013**	0.011**
	(5.85)	(6.06)	(6.58)	(5.80)
$TrendP_{r,t-3}$	-0.94**			
	(-3.40)			
$TrendP_{r,t-7}$		-1.04**		
		(-5.49)		
$TrendP_{r,t-14}$			-0.98**	
			(-7.60)	
$DailyShock$	-0.022**	-0.020**	-0.017*	-0.026**
	(-2.44)	(-2.15)	(-1.86)	(-2.87)
$Seoul$	6.67**	6.62**	6.49**	6.68**
	(19.23)	(19.20)	(18.96)	(19.15)
$Workday$	2.93**	2.92**	2.90**	2.87**
	(12.19)	(12.21)	(18.96)	(11.90)

Notes: the number of observation is 2796. t-statistics are in parentheses.
$*$: Significant at the 10% level. $**$: Significant at the 5% level.

Chapter 3

Ability Grouping and Student Achievement: Effects of the Equalization Policy in Korea

This paper analyzes the effects of ability grouping on the academic performance of high school graduating students in Korea. About half of the regions in Korea have adopted the equalization policy (EP), which means that high school students are randomly assigned. For the other non-EP regions, students are sorted among schools based on ability levels. I utilize a difference-in-differences strategy to exploit the adoption of the EP during 2000-2005, which resulted from the exogenous policy shifts of several regions. Studying a dataset of Korean SAT scores for the entire pool of exam takers from 1997 to 2012, I find that after the EP, performance of students above the median dropped 1.4% in terms of national percentiles, while that of students below the 30% percentile jumped 1.3%. In addition, there was an increasing trend in the achievement levels in the treatment regions in the pre-treatment period, but after the introduction of the EP, this trend vanished.

3.1 Introduction

The effects of peer groups and social interactions have played a prominent role in various policy debates: in particular, the potential for school or classroom peers to affect academic performance is undeniably crucial for education policy decisions. The impact of peers has been studied in depth, but previous literature reports that findings vary widely across studies, depending on the dataset and empirical settings. For instance, while grouping students by academic ability may encourage students to put in more effort and hence improve achievement levels, it may disadvantage low-achieving students by not providing a chance to interact with high-achieving peers. Due to such lack of consensus on the peer effects, potentially from the difficulty in identification of the causal effects by separating peer effects from other confounding factors, it is difficult to derive consistent policy implications (see, e.g., Evans et al., 1992; Hanushek et al., 2003; Burke and Sass, 2013; Angrist, 2014). In this paper, I use policy shifts from ability grouping to random assignment in Korea to examine whether a change to random assignment affects student achievement levels and whether the impact of the policy change is different for different ability levels.

Changes in education policies in Korea provide an excellent empirical setting to analyze the effects of ability grouping and random assignment on peers. Starting from 1974, the Korean government has adopted the equalization policy (hereafter EP) in several areas. Roughly two-thirds of middle school students enter general high schools, and students are either assigned by a lottery-based enrollment system or by high school entrance examinations, depending on whether a region is under the EP or not. If a region is under the EP, students from this region are randomly assigned to general high schools in the region. If a region is not under the EP, students are assigned by entrance exam results: each high school offers an entrance exam and admits students based on the test scores (Kim et al., 2008). As there is a widely recognized hierarchy of high schools based on their seniors' performance on national college entrance exams, students are sorted among general high schools based on their academic abilities in the non-EP region (Kang et al., 2007). Moreover, general high schools in Korea are fairly homogeneous in terms of the education environment (except for student compositions) that might be considered as endogenous in other settings, as all schools follow national standards and policies: the same centralized curriculum, teaching guidelines, government financial support, and teacher qualifications (Lee et al., 2014). In addition, classrooms are homogeneous in terms of race and ethnicity, unlike in the U.S.

In this paper, I use Korean, mathematics, and English scores of the yearly national college entrance exams in South Korea, from 1997 to 2012,[1] to measure student achievement levels. As this national college entrance exam, The College Scholarly Aptitude Test, is the Korean version of SAT, I call it "the Korean SAT" hereafter. Since Korean SAT scores are valid for only one year, I develop a measure to compare performance of a certain region to the whole nation for different years so that I can evaluate a time-trend performance of the region. Suppose a median student in a region scored 70 in 2001, and this score was the 48th percentile for the whole nation. The next year, 2002, a median student in this region scored 68, which was the 49th percentile for the nation. In this case, I find that a median student in the region performed 1% better in 2002, in terms of national percentiles (by comparing 48 and 49). Note that the scores themselves do not mean anything, as each Korean SAT is treated separately and cannot be directly compared.

I exploit the EP adoptions of six regions from 2000 to 2005 to measure the impact of ability grouping on academic performance: comparing before and after the policy shift allows me to utilize a difference-in-differences strategy. Since I have the Korean SAT scores of all exam takers, I can analyze the effect of the EP adoption in different achievement levels: for instance, I may focus on the 10% and 20% percentile scores of a region to check the effects on low-achievement groups. The reduced-form analysis suggests that there are both immediate and persisting effects, and the directions and magnitudes of these effects are different for different percentiles. For low-achieving students, the immediate effect of the EP adoption is positively significant: the positions of students in the national score distributions move up by 1.3% and 1.4%, respectively, for students at the 10% and 20% percentiles in treatment regions. For the median and above students, I observe negatively significant immediate effects of the EP introduction. For example, the national percentiles dropped 1.80% with the EP introduction for the students at the 80%

[1]Year 2008 is excluded, as no score data were available. Only "grades", a range of percentiles, were reported. For example, If you are told that your "grade" is 2, it means that your score is between top 4-11% percentile among all exam takers.

percentiles in treatment regions.

The regression results also suggest that there exist trend-changing effects in the long run. For students in the 30-50% percentiles, I find that there was a positive trend of about 0.3% increase per year (in terms of the national percentiles) before the EP adoption, but this trend vanished after the introduction of the EP: in other words, there was a persisting negative effect of the EP that offsets the pre-existing positive trend. For the median and above students, I also observe persisting negative effects of the EP. For instance, there was an increasing trend of 0.3% per year before the EP, but a decreasing trend of 0.2% per year is found after the EP for the students at the 80% percentiles. Thus, the total persisting effect of the EP on the 80% percentile may be a change from 0.3% to -0.2%, or 0.5% decrease. Note that these persisting effects are less clearly identified, since there could be other confounding factors, such as changes in demographics, that I cannot fully control for.

This paper contributes to the literature that investigates peer effects. In general, much of the previous literature finds that modest peer effects exist for higher education, but there is no consensus for elementary and secondary education (Epple and Romano, 2011). Studying peer effects, researchers have faced the endogeneity problem: it is difficult to separate the effects that the peer group has on the individual from the effect the individual has on the group. Also, since individuals may self-select into certain peer groups by their choices, the peer effects are intertwined with the selection effects (Manski, 1993; Moffitt, 2001; Sacerdote, 2001). One method to overcome these issues is to use previous peer achievement as an instrument for current achievement (Hanushek et al., 2003; Burke and Sass, 2008). Another method is to exploit empirical settings in which individuals are randomly assigned to peer groups. Zimmerman (2003) uses data from Williams College to find that the middle of the SAT distribution does worse when they share a room with the bottom 15%, while neither weak nor strong students are affected significantly by peers. Carrell et al. (2009) use the random assignment of students at the United States Air Force Academy and suggest that peer effects are significant in math and science, but insignificant in physical education and foreign language courses. Imberman et al. (2012) study the impact of an exogeneous inflow of students (due to Hurricanes Katrina and Rita) on academic performance and find evidence in support of monotonic peer effects: all students benefit from high-achieving peers and are harmed by low-achieving peers. Using within-school variation in the proportion of low-ability students in Israeli schools, Lavy et al. (2012) find that having more low-achieving students in the same class results in a negative peer effect on the scores of other students.

I also contribute to the literature on the impacts of the ability grouping and random assignment policy on student performances. The effects of ability grouping (or tracking) on academic outcomes have been a hot policy issue: advocates of the ability grouping system emphasize that sorting would enhance performance of students of all achievement levels by offering them more efficient, focused classes, and by encouraging competition; on the other hand, proponents of the random assignment system argue that low-achieving students are helped by high-achieving students under the ability mixing system and hence it improves equity, while high-achieving students are unharmed (Argys et al., 1996; Betts and Shkolnik, 2000; Figlio and Page, 2002; Epple and Romano, 2011). Previous literature finds mixed evidence: Glewwe (1997) studies the secondary school students in Philippines data and suggest that ability grouping raises average math scores; Figlio and Page (2002) use the National Education Longitudinal Study and suggest that ability grouping might actually help low-achieving students; Hanushek et al. (2003) find no clear evidence of the effects of the ability grouping policy; Kang (2007) finds that weak students are likely to benefit from ability mixing, while strong students benefit from ability grouping in Korea; and Duflo et al. (2011) use experimental data on tracking from Kenya to find that ability grouping of students by prior achievement levels raised scores for all students, even those assigned to low-achieving classes.

The remainder of this paper proceeds as follows. In section 3.2, I describe the Korean SAT and the equalization policy. Section 3.3 explains the Korean SAT score dataset from 1997 to 2012. The empirical strategy is discussed in section 3.4, and the reduced-form results are presented in section 3.5. Section 3.6 discusses identification issues and performs robustness checks. Section 3.7 provides conclusions.

3.2 Empirical Setting

Since the Korean education system is quite different from that of the United States, I devote this section to illustrating what the equalization policy and the Korean SAT are and how they could be used to measure the effects of ability grouping on academic performance. I start with providing institutional background and the history of the equalization policy. Next, I explain three unique features of the Korean college entrance exam, or the Korean SAT, that make the Korean SAT scores appropriate measures for students' performance. Lastly, I discuss why I consider only graduating students from general high schools and why these general high schools may be considered as homogeneous.

3.2.1 Equalization Policy (EP)

The equalization policy (hereafter, EP) is a random lottery system that assigns students to schools in a certain region. If a region is under the EP, all schools, both public and private, follow a random lottery system: students are allocated randomly among schools in the relevant residence-based school districts. If a region is not under the EP, students are assigned by their test scores.[2] In other words, students are grouped by their academic abilities for non-EP regions and are randomly assigned for EP regions. For the high school level (school years 10-12), about half of the regions have adopted the EP: at the beginning of the sample period (1997), 45.2% of general high schools (55.8% of total students) were subject to the EP, and it reached 50.4% (61.1% of total students) at the end of the period (2012). For the middle school level (school year 7-9), all schools in the nation were under the EP during the sample period.[3]

Before 1974, all general and vocational high schools in Korea offered their own entrance exams and admitted new students based on the exam results. Since there were clear and widely established rankings among general high schools, the middle school graduates (finishing grade 9 in the U.S. system) were sorted by their academic performance: as a result, ability of students within each general high school were fairly homogeneous. This system was regarded as effective in

[2] If a region is not under the EP, the general high school admission process in this region is similar to the U.S. college application process.

[3] The equalization policy for middle schools started in 1969 and all middle schools in the nation were under the EP by 1975.

training elite groups, but was also blamed for exacerbating inequality. Moreover, financial burden of parents to hire private tutors for the entrance exams and psychological stress level of middle school students during the exam preparation formed public opinion against the competition-based admission system for general high schools.

In 1974, the Korean government launched a new assignment system, the equalization policy, to discourage private tutoring and dismantle elite magnet schools. Since the main goal of the equalization policy is to prevent excessive early competition among young students and repress private cram schools, the Korean government started with major cities. The program began in the two biggest cities, Seoul and Pusan, in 1974, and the national government adopted the EP in most major cities by 1981. However, in 1988, the government allowed each region to choose whether to adopt the EP or not, due to political issues: some people argued that the equalization policy resulted in downward standardization that harmed all students. For a long time after 1981, the only changes were reversals that happened in 1990 and 1991 that several regions chose to go back to the ability grouping system: residents of these regions worried that students performed worse under the EP.

The next switches to the EP happened in 2000, almost 20 years after the last major change in 1981. Following an announcement from the Korean government about a new national education curriculum and revisions to college admission processes that would take effect from 2002, several cities (three in 2000, seven satellite cities near Seoul in 2002, and three more in 2005) decided to adopt the EP, based on local public opinion and political pressure: it was argued that the EP system would suit the new system better and increase the probability of entering good colleges.

3.2.2 Korean SAT

In this paper, I use Korean, mathematics, and English scores of the national college entrance exams in South Korea, from 1997 to 2012,[4] to measure student achievement levels. Since this national college entrance exam, The College Scholastic Aptitude Test, is the Korean version of SAT, I call it the Korean SAT, or the KSAT hereafter. The Korean SAT is a comprehensive one-

[4]Year 2008 is excluded, as no score data were available. Only "grades", a range of percentiles, were reported. For example, if you are told that your "grade" is 2, it means that your score is between top 4-11% percentile among all exam takers.

day exam that tests various subjects including Korean, mathematics, English, natural science, and social science.

There are three main advantages of using the KSAT to measure student performance. First, it is safe to assume that students take the KSAT very seriously, as KSAT scores are given great weight in college admissions. In some cases, a good KSAT score is the only requirement to enter desired universities. Second, the KSAT score distribution is close to the whole population distribution, as almost all graduating students from general high schools take the KSAT. Third, the KSAT is offered only once every year and only valid for that year. Thus, I can treat KSAT scores in different years separately, unlike American SAT scores.

Since KSAT is an annual test that carries a significant weight on college admissions, it is common to take KSAT multiple times, especially for high percentile students seeking to enter top universities. Figure 3.1 shows the proportion of graduating students among KSAT takers: the ratio of exam re-takers is about 30% at the beginning of the period and tends to decrease over time. Since the ability or score distribution of KSAT re-takers is very different from that of the whole population, I exclude re-takers from the sample and only consider graduating students (high school seniors) for the analysis.

3.2.3 High School Types

The Korean government groups high schools into three types: general, vocational, and special-purpose. About two-thirds of high school students attend general high schools: these are humanities high schools that prepare students for college. About one-third of high school students go to vocational schools where they learn practical skills to seek jobs right after graduation. Lastly, there is a small fraction of students (less than 5%) who attend special-purpose high schools that are designed to teach talented students in various areas, including arts, science, and foreign languages.

I drop both special-purpose and vocational high schools from the sample for the main analysis, as they are fundamentally different from general high schools, especially in terms of scholastic ability. In general, students from vocational schools and arts or sports special-purpose schools tend to underperform, and students from other special-purpose schools tend to outperform. In addition, special-purpose schools and vocational schools have not been under the EP for the whole period: all special-purpose and vocational schools have selected their students by their own entrance exam results, not by random lottery results. Thus, including them would not help me to study the effects of the EP.

3.2.4 Classroom Settings

In South Korea, all general high schools are homogeneous, and this fact helps me to separate the effect of the policy change without detailed micro-level data. All schools follow national standards and policies, in almost every aspect: lectures are given according to the centralized curriculum, and teachers are supposed to follow national teaching guidelines. Moreover, public and private general high schools are essentially identical in terms of education environment, as they are under the same regulations regarding finances and teacher qualifications. All public and private school finances are integrated nationwide, and teachers in both types of schools must meet the same qualifications set by the central government, and are subject to the same salary schedules. The only major difference between public and private high schools is that private schools were founded by non-governmental institutions, and this is not relevant in studying the effect of an education policy on student achievement levels. In addition, classrooms are homogeneous in terms of race and ethnicity, unlike in the U.S.

3.3 Data

In this paper, I combine two datasets from 1997 to 2012: the KSAT scores of all exam takers, and high school characteristics information. To measure student achievement levels, I focus on three main subjects, Korean (verbal), mathematics, and English, as these subjects are required for all students. I study scores of all exam takers from 1997 to 2012, except 2008, to estimate the effects of ability grouping and random assignment (caused by the adoption of the equalization policy). As students from vocational schools or special-purpose schools are very different in terms of academic performance distributions and these schools have not been under the EP, I limit my dataset to general high school students. Similarly, I only consider graduating students (school year 12) who are taking the test for the first time.[5]

Contrasting academic performance of students from the EP regions and students from non-EP regions allows me to estimate the effects of the EP (during high school years). However, simply comparing EP regions and non-EP regions would not capture the true effect of the EP, as there exist fundamental differences between the two groups. For example, EP regions have higher income and test scores on average, as they include most major cities. Instead of directly comparing EP and non-EP regions, I exploit policy shifts in the six regions. Note that these policy adoptions are exogenous, as these are reactions to the national education system changes.[6] In particular, I find that trends of population size and average income in these regions are stable: thus, it is unlikely that these two factors caused policy shifts. Figure 3.2 shows the trends of regional population sizes as the ratio of regional population to the national population, and Figure 3.3 shows the trends of regional average income as the ratio of regional average income to the national average income: for both graphs, regional variation exists, but time-trend variation is limited.

During 2000-2005, there were three years in which several areas switched from ability grouping to random assignment (EP): three in 2000, seven satellite cities near Seoul in 2002, and three in 2005. Multiple neighboring cities were grouped into a single region if these cities were close that there were frequent movements between them and pre-EP school assignments are mixed within these neighboring areas.[7] For instance, seven small cities near Seoul are grouped into a single region. With this definition, I may assume that almost all movements of students are within regions, not across regions. After this adjustment, I have six regions that newly adopted the EP.[8] These six regions are treatment regions, and "treatment" is the introduction of the EP, or random assignment system. In this setup, the control group is the rest of the nation.

Figure 3.4 shows the locations of the six regions and when they switched: red (2000, regions 1 and 2), blue (2002, region 3), and green (2005, regions 4, 5, and 6). Note that if a region adopted the EP in 2000, it means that students entering high schools in 2000 were the first

[5]See Section 3.2.2 for details.

[6]It was argued that the EP system would suit the new system better and increase the probability of entering good colleges. Depending on public opinion and political pressure levels, several regions decided to adopt the EP.

[7]For example, suppose areas 1 and 2 had mixed assignments before the EP. It means that it is common for students in area 1 to apply for high schools in area 2.

[8]1: Gunsan and Iksan, 2: Ulsan, 3: Seoul satellite cities, 4: Mokpo, 5: Sunchun, 6: Yeosu

students who were under the EP in this region and that they took the 2003 KSAT. As the new year starts in March in South Korea, the 2003 KSAT is for students who enter universities in March 2003 and the entrance exam, the 2003 KSAT, was taken in November 2002. Thus, while the EP was adopted in 2000, the effect of the EP introduction should first appear in the 2003 KSAT. Similarly, for regions that switched in 2002 and 2005, the first students under the EP in these regions took the 2005 and 2008 KSATs, respectively.

I use national percentiles to compare performance of a certain region to the whole nation for different years.[9] Let me start with a simple example. Suppose a median student in region 1 scored 70 in 2001, and this score was the 48th percentile for the whole population (nation). The next year, 2002, a median student in region 1 scored 68, and this score was the 49th percentile for the nation. In this case, I find that a median student in the region performed better in 2002, by comparing 48 and 49.[10] Formally, I define this measure as the following: $Y_r^{p_{region}} = p_{nation}$ denotes that the p_{region}th percentile of region r is the p_{nation} th percentile for the whole nation. The main advantage of using this measure is that I can easily compare performance across the regions for multiple years.

Figure 3.5 presents a bar graph of the average national percentiles of the six treatment regions of year 2000 and 2010. Graphically, for the regional percentile where the blue bar is higher than the red bar, students of treatment regions performed better in 2000, as the corresponding national percentile is higher. For example, the blue bar is 51% and the red bar is 49% for the 50% regional percentile: the average national percentiles of the 50% percentile students in the six treatment regions is 51% and 49% for year 2000 and 2010, respectively. The bar graph suggests that while students below the 20% percentile in treatment regions performed slightly better in 2010, students above the 50% percentile did worse.

Figure 3.6 shows average trends of 20th, 40th, 60th, and 80th percentiles of the six regions, before and after the EP. The first vertical line indicates $t = 0$, the last year for students under the ability grouping system, and the second vertical line indicates $t = 1$, the first year for students under the random assignment system. There are substantial decreases between the two periods

[9]Since the exam scores are standardized at the national level for each exam year, it is difficult to compare scores directly to measure performance changes.

[10]Note that the scores themselves are not important and actually do not mean anything in this setup, as each exam is treated separately and cannot be directly compared.

for all percentiles except the 20th percentile case.

3.4 Empirical Strategy

Since this project exploits the policy shift, it is natural to utilize a difference-in-differences method. In particular, using a difference-in-differences model, I study how the locations of regional percentile levels move within the national score distribution, before and after the equalization policy (EP).

As different effects on different ability groups are expected, I run a separate regression for each percentile p (p =10, 20, ... , 90, 95, 99):

$$Y^p_{r,s,t} = \lambda_{rs} + \beta_1 D_{1,r,t} + \beta_2 D_{2,r,t} + \beta_3 D_{3,r,t} + \epsilon_{rst}$$

where r stands for each region, s for subject, and t for year. The dependent variable ($Y^p_{r,s,t}$) denotes where each regional (r) percentile level (p) is located in the national distribution. For example, $Y^{50}_1 = 48$ means that the median (50%) of region 1 is the 48th percentile for the whole nation. Thus, higher dependent variable values imply that students from this region are doing better, compared to the rest of the nation. λ_{rs} are region-subject fixed effects that control existing differences between regions and subjects.

The main variable of interest, D_1, is the usual difference-in-differences indicator that denotes the effect of the EP policy adoption: D_1 is zero until year $t-1$ and 1 for year t and after. In addition, I also measure before and after the policy shift trends, using a separate indicator variable for each. It is possible that there is a gradual change or movement in student achievement levels as a result of the policy change. For instance, teachers could gain more experience under the new system as time goes on, and it may take some time for the new system to successfully settle in. Such post-trend changes are supposed to be captured by $D_2 = max$ (year-t, 0): D_2 is zero until year $t-1$, and linearly increase for year t and after. Similarly, $D_3 = min$(year-t, 0) is designed to represent pre-existing trends before the policy shift: D_3 linearly decreases until year t, and zero after t.

3.5 Results and Interpretation

Table 3.1 shows the main difference-in-difference model results. There are two kinds of effects: immediate effects that are applied at the first year of the equalization policy (EP) adoption, and persisting effects that change the pre-existing trends or create new trends. Furthermore, the effects are different for different achievement levels.

I start with the immediate effects that are represented by the main variable of interest, D_1. For low-achieving students, there are positively significant immediate effects that move up the students at the 10% and 20% percentiles in treatment regions by 1.27% and 1.39%, respectively, in terms of the national percentiles. On the other hand, for students who are above the median, there are negatively significant results: for example, the position of students in the national score distribution fell by 1.80% for students at the 80% percentile in treatment regions. Lastly, for students in 30%-50% percentiles, I do not find significant results, indicating that the effects of the EP on these achievement levels are not clear. Thus, the D_1 coefficient estimate imply that under-performing students were positively impacted by the EP introduction and high-performing students were negatively impacted, while the effects on the middle group is ambiguous.

In addition to the instantaneous effects, the D_2 and D_3 estimates indicate that there are also persisting effects of the EP adoption that affect time-trends: note that D_2 represents the post-EP trends of student achievement levels and D_3 denotes the pre-EP trends. However, note that interpretations of these estimates come with a caveat: these trend effects are less clearly identified, since there are more potential confounding factors. For instance, there could be demographic trend changes that the model cannot fully control using year and region fixed effects.

Similar to the D_1 case, the estimated trends effects are heterogeneous for different achievement levels. For the bottom (10%) percentile, both of the D_2 and D_3 estimates are not significant: there are no significant pre or after-EP trends. For high-achieving students who are 70% of above, I find that there exist positive pre-EP trends, while negative trends are found after the EP introduction: for example, the relative location of students at the 80% percentile in treatment regions within the national distribution was going up by 0.40% per year, but it changed the direction and started to decrease by 0.19% per year after the EP adoption. Thus, the EP created the negative trend of 0.59% (a difference of -0.19% and 0.40%) per year for the 80% percentile students. For students between 10% and 70% percentiles, I find that there are positively significant pre-existing trends, but these trends disappeared after the introduction of the EP: for example, the national percentiles of the median students in the treatment regions were on an upward trend of 0.43% per year, but I do not find a significant after-EP trend for the median students.

These results suggest that (i) there are both instantaneous and persisting effects of the EP adoption, (ii) students in the lower percentiles immediately performed better with the EP adoption but lost positive pre-EP trends, (iii) students in the middle percentiles do not have clear immediate performance changes but lost positive pre-EP trends, and (iv) high-achieving students (above 60%) performed worse with the EP instantaneously and suffer negative post-EP trends. For instance, let me interpret the model suggested effects on the 20% and the 80%. For the 20% percentile, there was a gradual upward trend that the position of students in the national score distribution went up by 0.12% per year for students at the 20% percentile in treatment regions. When the EP was introduced, the relative position of the 20% percentile students in treatment regions jumped up by 1.39% (instantaneous effect), and there was no statistically significant after-EP-trend (trend effect). For the 80% percentile, the position of the 80% percentile students (in treatment regions) in the national score distribution dropped by 1.80% with the EP introduction (instantaneous effect). Furthermore, there was an increasing trend of 0.40% per year before the EP, but a decreasing trend of 0.19% per year is found after the EP. Thus, the total effect of the EP is a sum of the instantaneous effect of 1.80% drop and the persisting effect of -0.59% per year (a change of trend from 0.40% to -0.19%).

In summary, both immediate and persisting effects of the EP are negative for high-achieving students. On the other hand, the immediate effect of the EP is positive for low-achieving students and the persisting effect is slightly negative, since the pre-existing positive trends disappeared after the EP adoption. These results suggest that moving from ability grouping to random assignment by the introduction of EP benefits low-achieving students but hurts high-achieving students, at the first year of the EP introduction. Furthermore, the negative persisting effects for all percentiles except the bottom groups (the 10% percentiles) imply that random assignment might deteriorate academic performance of most students in the long run. In terms of the peer effects on academic outcomes, these findings suggest that all students benefit from high-achieving peers and are hurt by low-achieving peers.

3.6 Discussion

In this section, I discuss identification issues by checking two model setup assumptions. I also check robustness of the estimation results by including special-purpose schools or vocational schools to the sample.

3.6.1 Identification

The model setup in the previous section requires the following two assumptions to hold: 1) the EP introduction was exogenous, and 2) there were no endogenous movements of students across different regions. In the empirical setting section, I document that most of the EP regions are big cities, while most of the non-EP regions are small cities or rural areas. If the decisions to adopt the EP during the sample period depend on other factors, such as future trends in the population size, average income, or test results, the identifying assumption is not valid. I checked that the first two factors had been fairly stable during the period (recall graphs in Figure 3.2 and 3.3) and unlikely to cause the policy shifts. Here, I consider another potential factor, the test results: while it is plausible that test scores of a certain region might lead the region to change education policies, these regression results show that it is unlikely. Note that the coefficients of D_3 are positively significant for all percentiles except the 10% percentile, which implies that student achievement levels in these treatment regions had been improving before the EP introduction: it is unlikely to adopt a new policy due to the student performance, if students were already doing better.

For the second assumption, if there were significant cross-region movements to avoid random assignments, there should be higher effects for students in the higher percentiles. As these regions are fairly large (contain several cities), moving to another region is costly: unless students are aiming for top-tier schools, it is not worthwhile to move to non-EP regions.[11] Yet, all of D_1, D_2, and D_3 coefficients show similar t statistics and the same trends for percentiles above median, which implies that movements of students across regions due to the EP introduction are quite limited.

3.6.2 Special-purpose Schools

As discussed in the empirical settings, there are three types of high schools in South Korea: general, vocational, and special-purpose high schools. For the main specification, I considered graduating students from general schools. special-purpose high schools are account for about 2-3% of the whole population, and less than 2% for the treatment regions. They have different goals and completely different student compositions compared to general high schools, by definition. In addition, special-purpose schools have not been under the EP for the whole period: all special-purpose schools have selected their students by own entrance exam results, not by random lottery results. Thus, it is natural to exclude them from the sample.

Moreover, there have been very limited changes of the proportion of students attending special-purpose schools within treatment regions. Figure 3.7 shows the proportions of the KSAT takers who attended special-purpose high schools for the six treatment regions: these proportions are stable during the sample period. Therefore, including special-purpose schools would not affect the results significantly. Furthermore, if there exist effects from the special-purpose schools, we should observe differences among the high-performing students, say 70% to 95%, and the top students (95% or above), as students who attend the foreign language schools or science high schools tend to be at the top percentiles. Table 3.2 shows that the results for students above median have the same trends, and the magnitudes and significance of coefficients are almost identical to the main results.

3.6.3 Vocational Schools

After 2000, the Korean government has highly encouraged vocational schools to concentrate on getting jobs right after the graduation by learning latest skills, instead of applying to colleges: new

[11]The only way to avoid random assignment is that a student's entire family members move to a different district.

types of vocational schools such as "Internet" high schools were founded, and existing vocational schools also changed curriculum. Also, the government asked colleges to offer more admission choices that allow students to apply without KSAT scores, which affected vocational high school students' decision to prepare for KSAT exams. While the proportion of students attending vocational schools has been relatively stable during the period, the proportion of vocational school students who choose to take KSAT has dropped significantly. As a result, the proportions of vocational school students among KSAT takers for all regions have decreased, and these trends are shown in Figure 3.8. Except for region 1, all regions (and the whole nation) show constantly decreasing trends, and region 1 joins this downward trend after 2002.

If the composition of vocational high school students changes over time in terms of the achievement level, there may be a selection issue. For instance, suppose that achievement levels of vocational high school students were uniformly distributed from 0% to 40% before 2002, but uniformed distributed from 0% to 50% after 2002, as students slightly below median became more interested in direct employment opportunities from "Internet" high schools. If such hypothetical trends are different for regions, say I can observe such changes only in regions 2 and 3, I compare different student population distributions for different regions.

As a robustness check, I also present the regression results with vocational school KSAT takers in Table 3.3. I expect two changes, compared to the Table 3.1 results. Since students from vocational schools tend to locate below median, the effects of including vocational school exam takers would mainly impact below-median coefficients. Moreover, vocational schools are not affected by the EP policy. Hence, including them would basically dilute the treatment effect of the EP adoption.

Table 3.3 confirms these two predictions: only results for the below-median cases change, and they change in a way that the magnitudes of coefficients decrease so that I lose significance levels but have the same signs. Note that for students around and above median, the results are basically the same: instant negative effect, slightly negative after-the-EP trend, and positive before-the-EP trend.

3.7 Conclusion

In this paper, I investigate how a policy shift from ability grouping to random assignment of general high school students affected academic performance. Using Korean college entrance exam (Korean SAT) scores of all test takers from 1997 to 2012 as indicators of student achievement levels, I exploit exogenous policy shifts of six regions to apply a difference-in-differences strategy. Before adopting the equalization policy (EP), students in these regions were assigned to high schools by their academic abilities based on test scores. After the policy introduction, students are now randomly assigned to general high schools within the region by lottery results. Thus, this empirical setting provides an excellent opportunity to study the effects of ability grouping on student achievement levels.

Descriptive analyses suggest that there are both immediate and persisting effects of the equalization policy and that the effects vary for different percentiles. While the positions of students in the national score distribution went up by 1.3% for low-achieving students who are at the 20% percentile or below in treatment regions, as a result of the EP adoption, the positions of students above the median fell by from 0.3% to 1.8% after the first year of the EP introduction. Moreover, there are trend-changing impacts. There were gradual upward trends (in terms of the positions in the national score distribution) for all achievement levels except the 10% percentile in treatment regions before the EP, but these trends disappeared after the policy shifts: in fact, decreasing trends from 0.1% to 1.6% per year are found after the introduction of the EP for students above the median in treatment regions. Thus, the total effect of EP is significantly negative for students above the median and positive for students below the 30% percentile.

While both immediate and persisting effects of the EP are negative for high-achieving students, the immediate effect of the EP is positive for low-achieving students and the persisting effect is slightly negative (as the positive pre-EP trend vanished). These results suggest that moving from ability grouping to random assignment (ability mixing) helps low-achieving students but harms high-achieving students initially. Furthermore, the persisting negative effects for all percentiles except the bottom groups (the 10% percentiles) imply that random assignment might hurt most students, potentially by hurting teaching quality or academic atmosphere. In terms of the peer effects on academic outcomes, these findings provide evidence for monotonicity in peer effects: all students benefit from high-achieving peers and are hurt by low-achieving peers.

References

[1] Joshua D Angrist. The perils of peer effects. *Labour Economics*, 30:98–108, 2014.

[2] Joshua D Angrist and Kevin Lang. Does school integration generate peer effects? evidence from boston's metco program. *American Economic Review*, pages 1613–1634, 2004.

[3] Laura M Argys, Daniel I Rees, and Dominic J Brewer. Detracking america's schools: Equity at zero cost? *Journal of Policy analysis and Management*, 15(4):623–645, 1996.

[4] Julian R Betts and Jamie L Shkolnik. The effects of ability grouping on student achievement and resource allocation in secondary schools. *Economics of Education Review*, 19(1):1–15, 2000.

[5] Julian R Betts and Jamie L Shkolnik. Key difficulties in identifying the effects of ability grouping on student achievement. *Economics of Education Review*, 19(1):21–26, 2000.

[6] Julian R Betts and Andrew Zau. Peer groups and academic achievement: Panel evidence from administrative data. *Unpublished manuscript*, 2004.

[7] Mary A. Burke and Tim R. Sass. Classroom peer effects and student achievement. *Journal of Labor Economics*, 31(1):pp. 51–82, 2013.

[8] Scott E Carrell, Richard L Fullerton, and James E West. Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics*, 27(3), 2009.

[9] Esther Duflo, Pascaline Dupas, and Michael Kremer. Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. *American Economic Review*, 101(5):1739–74, 2011.

[10] Dennis Epple, Elizabeth Newlon, and Richard Romano. Ability tracking, school competition, and the distribution of educational benefits. *Journal of Public Economics*, 83(1):1–48, 2002.

[11] Dennis Epple and Richard Romano. Peer effects in education: A survey of the theory and evidence. *Handbook of social economics*, 1(11):1053–1163, 2011.

[12] William N Evans, Wallace E Oates, and Robert M Schwab. Measuring peer group effects: A study of teenage behavior. *Journal of Political Economy*, pages 966–991, 1992.

[13] David N Figlio and Marianne E Page. School choice and the distributional effects of ability tracking: does separation increase inequality? *Journal of Urban Economics*, 51(3):497–514, 2002.

[14] Paul Glewwe. Estimating the impact of peer group effects on socioeconomic outcomes: Does the distribution of peer group characteristics matter? *Economics of Education Review*, 16(1):39–43, 1997.

[15] Eric A Hanushek, John F Kain, Jacob M Markman, and Steven G Rivkin. Does peer ability affect student achievement? *Journal of applied econometrics*, 18(5):527–544, 2003.

[16] Changhui Kang. Classroom peer effects and academic achievement: Quasi-randomization evidence from south korea. *Journal of Urban Economics*, 61(3):458–495, 2007.

[17] Changhui Kang, Cheolsung Park, and Myoung-Jae Lee. Effects of ability mixing in high school on adulthood earnings: quasiexperimental evidence from south korea. *Journal of Population Economics*, 20(2):pp. 269–297, 2007.

[18] Taejong Kim, Ju-Ho Lee, and Young Lee. Mixing versus sorting in schooling: Evidence from the equalization policy in south korea. *Economics of Education Review*, 27(6):697–711, 2008.

[19] Victor Lavy, M Daniele Paserman, and Analia Schlosser. Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom*. *The Economic Journal*, 122(559):208–237, 2012.

[20] Soohyung Lee, Lesley J. Turner, Seokjin Woo, and Kyunghee Kim. All or nothing? the impact of school and classroom gender composition on effort and academic achievement. Working Paper 20722, National Bureau of Economic Research, December 2014.

[21] Robert A Moffitt et al. Policy interventions, low-level equilibria, and social interactions. *Social dynamics*, 4:45–82, 2001.

[22] Bruce Sacerdote. Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly Journal of Economics*, page 681, 2001.

[23] Ron W Zimmer and Eugenia F Toma. Peer effects in private and public schools across countries. *Journal of Policy Analysis and Management*, 19(1):75–92, 2000.

[24] David J Zimmerman. Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics*, 85(1):9–23, 2003.

Appendix A. Figures and Tables

Figure 3.1: The Proportion of Graduating Students Among KSAT Takers

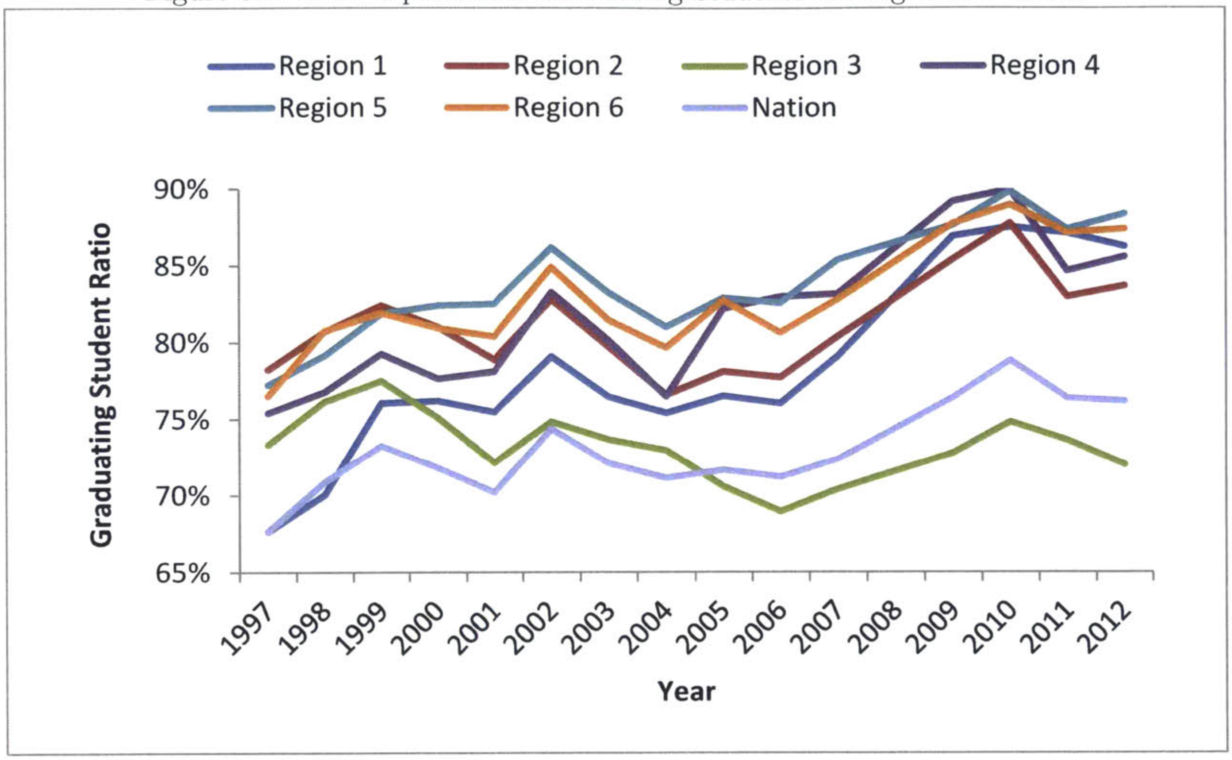
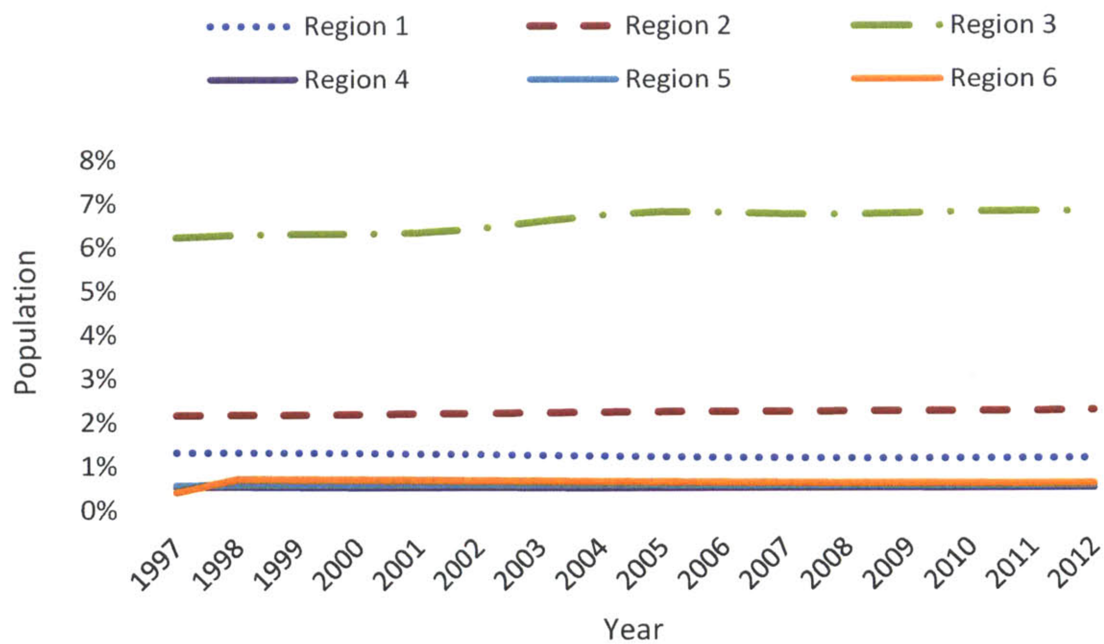
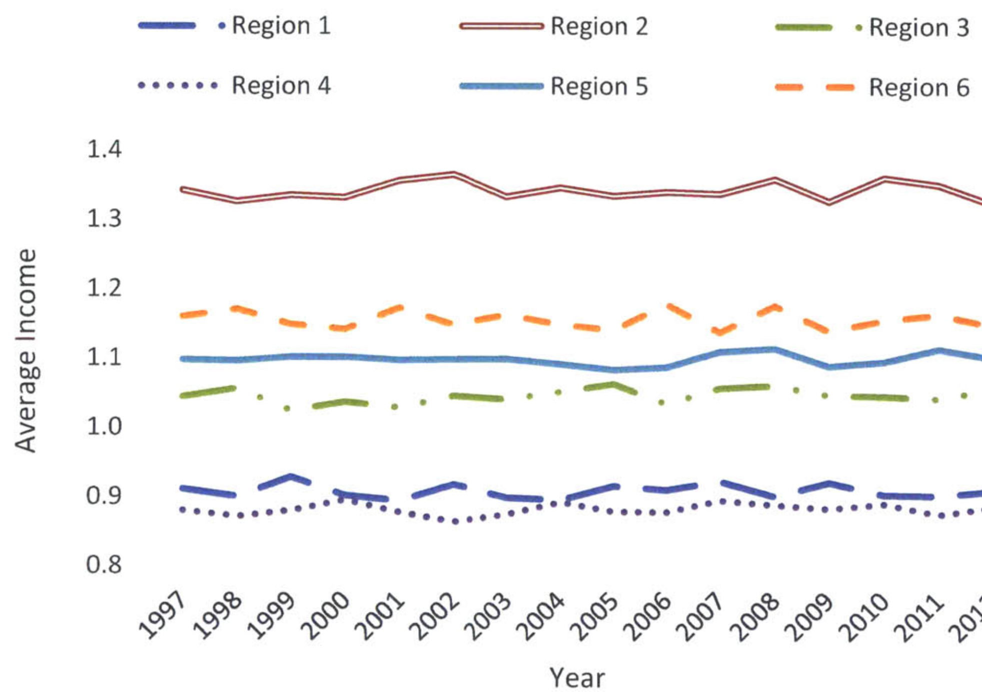


Figure 3.2: Population Trends of the Six Treatment Regions



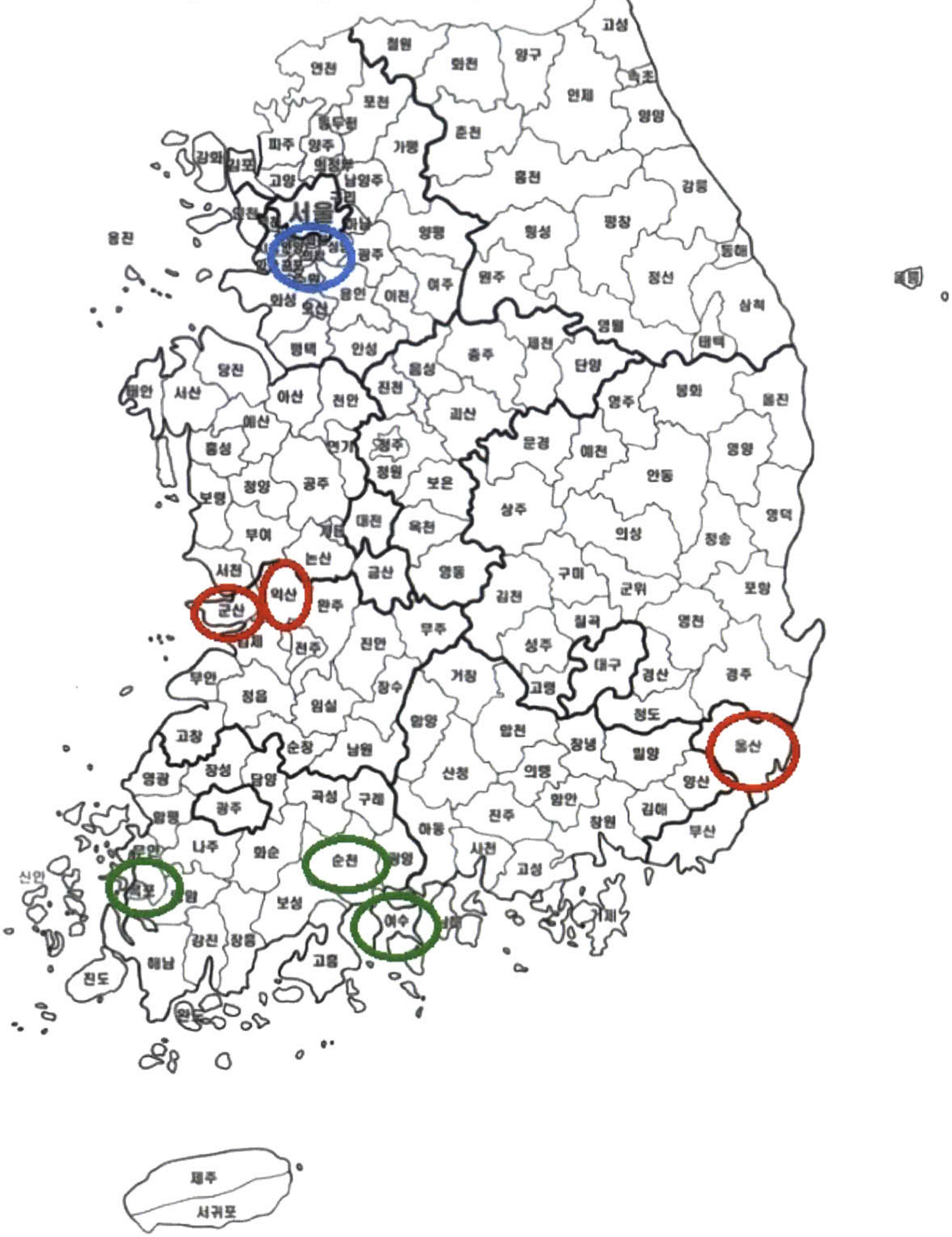
Notes: The vertical axis represents the ratio of regional population to the national population. Since region 4, 5, and 6 have very similar population sizes, three trends lines overlap.

Figure 3.3: Average Income Trends of the Six Treatment Regions



Notes: The vertical axis represents the ratio of average regional income to the national average income.

Figure 3.4: Regions that adopted the Equalization Policy



Note: the year of the EP adoption is 2000 (red), 2002 (blue), and 2005 (green), respectively.

Figure 3.5: Average Trends of the Six Treatment Regions

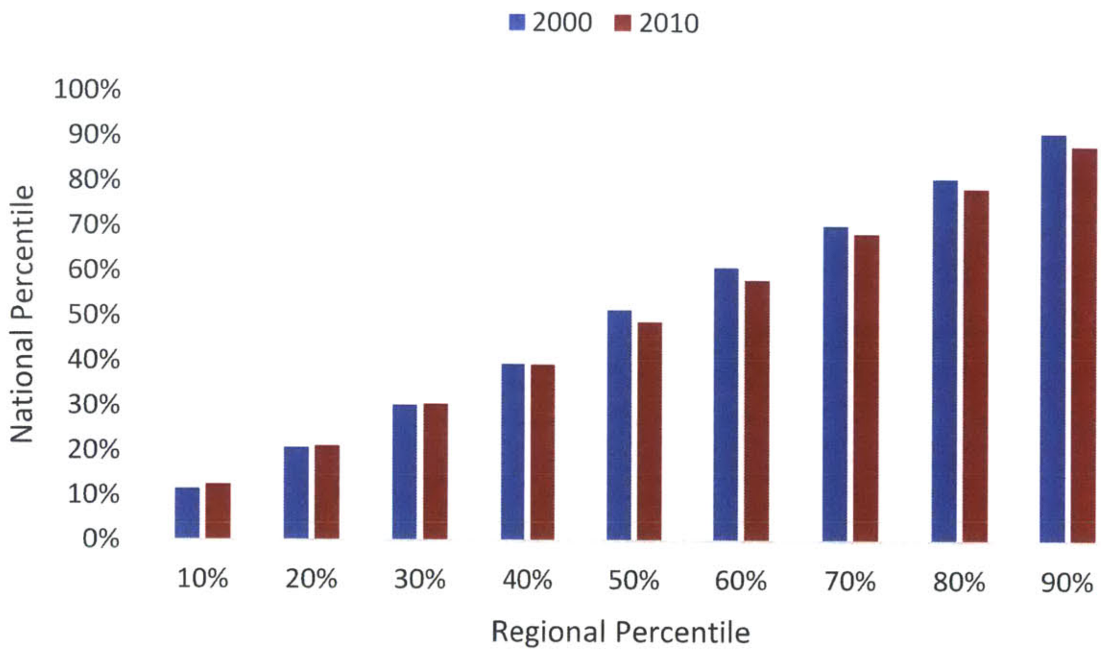


Figure 3.6: Average Trends of the Six Treatment Regions

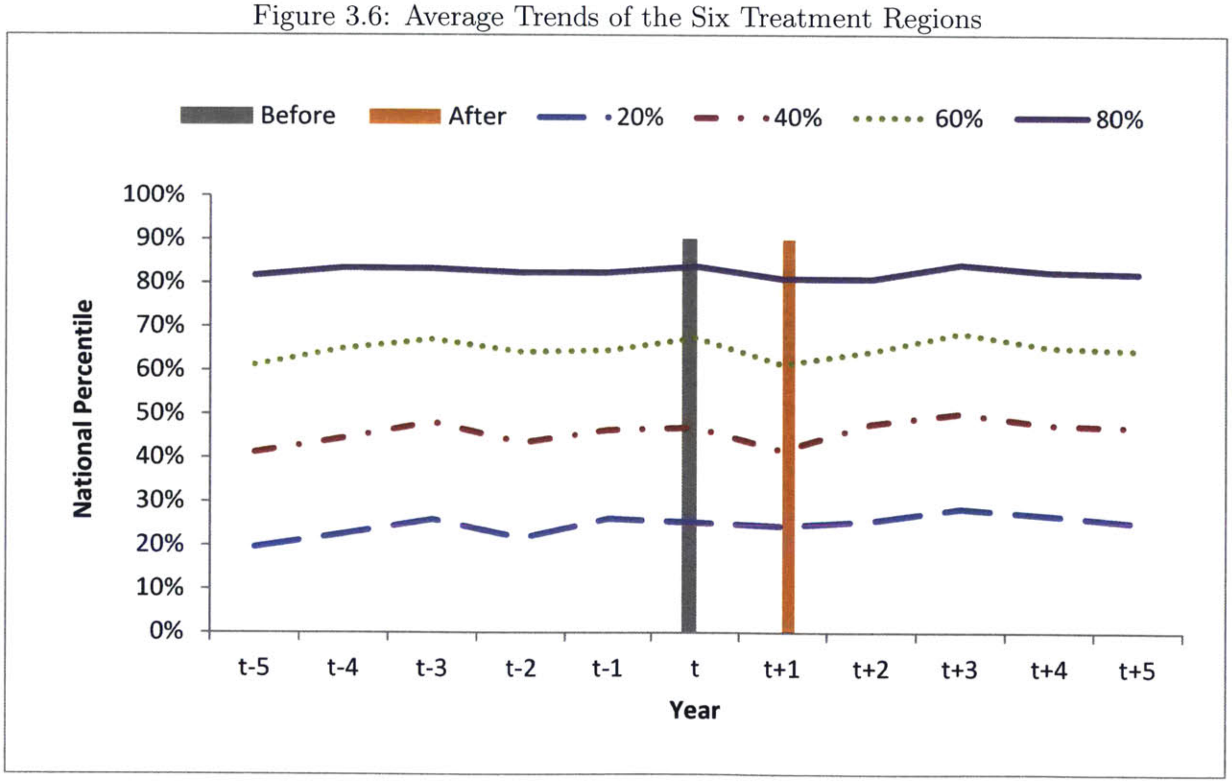
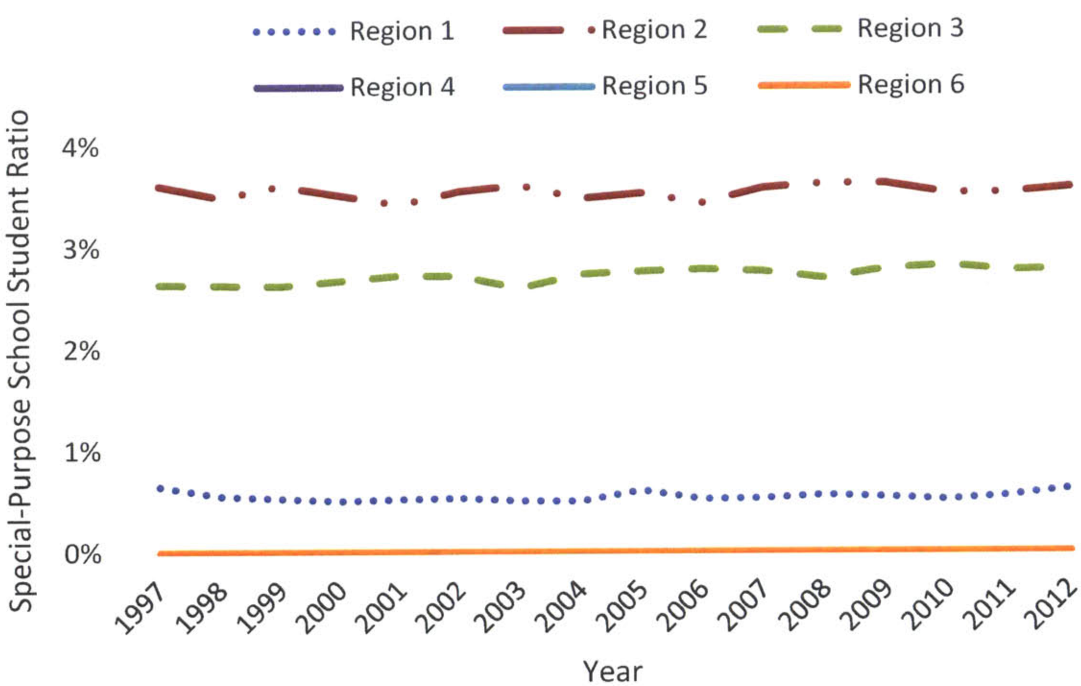


Figure 3.7: The Ratio of the KSAT Takers from Special-Purpose Schools



Notes: For region 4, 5, and 6, the ratio is zero (or very close to zero).

Figure 3.8: The Ratio of the KSAT Takers from Vocational Schools

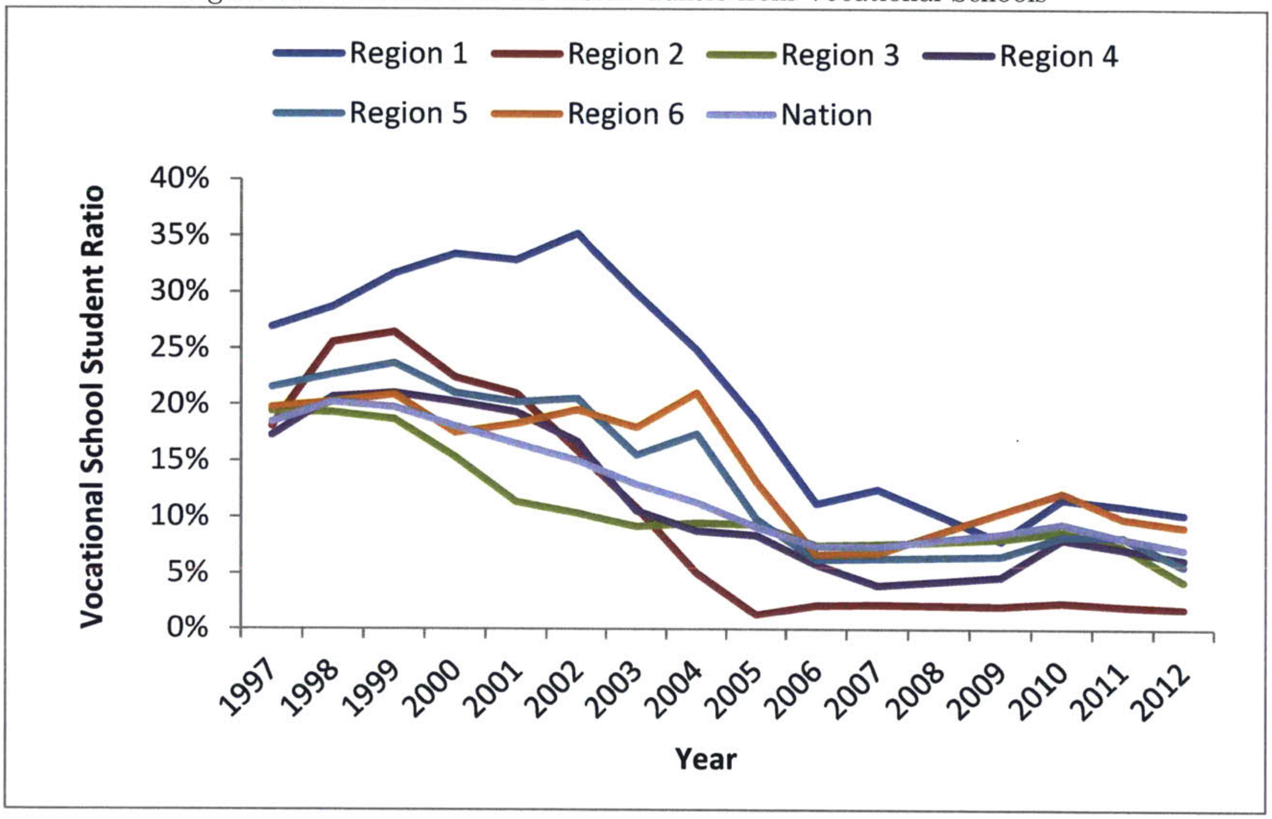


Table 3.1: Differences-in-Differences Result

Dependent Var.	$Y^p_{r,s,t}$: where each regional (r) percentile level (p) is located in the national distribution										
Percentile	10	20	30	40	50	60	70	80	90	95	99
D_1	1.27**	1.39**	1.03*	2.00	-0.76	-1.38**	-1.48**	-1.80**	-1.35**	-0.80**	-0.26**
	(3.38)	(2.94)	(1.87)	(0.35)	(-1.39)	(-2.47)	(-2.85)	(-4.14)	(-4.52)	(-4.16)	(-3.26)
D_2	0.10	0.08	-0.04	-0.07	-0.12	-0.17*	-0.23**	-0.19**	-0.44**	-0.09**	-0.03**
	(1.44)	(0.93)	(-0.43)	(-0.84)	(-1.47)	(-1.88)	(-2.78)	(-2.95)	(-3.01)	(-2.67)	(-2.15)
D_3	0.05	0.12**	0.23**	0.34**	0.43**	0.50**	0.49**	0.40**	0.27**	0.15**	0.05**
	(1.07)	(2.04)	(3.38)	(4.61)	(6.06)	(7.66)	(7.79)	(7.22)	(7.25)	(6.15)	(5.01)

Notes: the number of observation is 480. t-statistics are in parentheses.
$*$: Significant at the 10% level. $**$: Significant at the 5% level.

Table 3.2: Including Special-Purpose School KSAT Takers

Dependent Var.	$Y^p_{r,s,t}$: where each regional (r) percentile level (p) is located in the national distribution										
Percentile	10	20	30	40	50	60	70	80	90	95	99
D_1	1.21**	1.33**	1.03*	2.03	-0.79	-1.36**	-1.48**	-1.84**	-1.44**	-0.72**	-0.29**
	(3.47)	(2.97)	(1.82)	(0.37)	(-1.45)	(-2.49)	(-2.74)	(-4.03)	(-4.40)	(-4.06)	(-3.85)
D_2	0.10	0.08	-0.04	-0.07	-0.12	-0.17*	-0.23**	-0.18**	-0.45**	-0.08**	-0.04**
	(1.49)	(0.95)	(-0.44)	(-0.87)	(-1.40)	(-1.82)	(-2.71)	(-2.85)	(-3.06)	(-2.35)	(-2.27)
D_3	0.05	0.11*	0.22**	0.32**	0.41**	0.48**	0.47**	0.38**	0.25**	0.13**	0.04**
	(1.02)	(1.94)	(3.21)	(4.38)	(5.76)	(7.28)	(7.40)	(6.86)	(6.71)	(5.38)	(4.01)

Notes: the number of observation is 480. t-statistics are in parentheses.
$*$: Significant at the 10% level. $**$: Significant at the 5% level.

Table 3.3: Including Vocational School KSAT Takers

Dependent Var.	$Y^p_{r,s,t}$: where each regional (r) percentile level (p) is located in the national distribution										
Percentile	10	20	30	40	50	60	70	80	90	95	99
D_1	0.23	0.58	-0.13	-1.10	-1.95**	-2.23**	-2.04**	-1.93**	-1.22**	-0.63**	-1.62*
	(0.40)	(0.84)	(-0.17)	(-1.51)	(-2.85)	(-3.02)	(-3.42)	(-3.86)	(-3.39)	(-2.63)	(-1.68)
D_2	0.10	0.06	-0.02	-0.00	-0.06	-0.10	-0.16*	-0.15**	-0.11**	-0.08**	-0.03
	(1.25)	(0.58)	(-0.21)	(-0.04)	(0.61)	(-1.00)	(-1.72)	(-2.02)	(-2.03)	(-2.15)	(-1.59)
D_3	0.29**	0.34**	0.43**	0.46**	0.57**	0.59**	0.54**	0.44**	0.25**	0.17**	0.05**
	(3.35)	(3.22)	(3.56)	(4.04)	(5.40)	(5.04)	(5.55)	(4.73)	(3.90)	(3.98)	(3.53)

Notes: the number of observation is 480. t-statistics are in parentheses.
$*$: Significant at the 10% level. $**$: Significant at the 5% level.