TI 2014-123/VII
Tinbergen Institute Discussion Paper

Detailed Data and Changes in Market Structure: The Move to Unmanned Gasoline Service Stations

Tadas Bruzikas[1]
Adriaan R. Soetevent[2]

[1] *University of Groningen, the Netherlands;*
[2] *University of Groningen, and Tinbergen Institute, the Netherlands.*

Detailed Data and Changes in Market Structure:

The Move to Unmanned Gasoline Service Stations*

Tadas Bružikas† University of Groningen

Adriaan R. Soetevent‡ University of Groningen Tinbergen Institute

September 16, 2014

Abstract

We illustrate the impact of detailed data in empirical economic research by considering how the increased data availability has changed the scope and focus of studies on retail gasoline pricing. We show how high-volume, high-frequency price data help to identify and explain long-term trends using original data for the Dutch retail gasoline market.

We find that 22% of the observed increase in the highway/off-highway price gap can be explained by the trend towards more unmanned stations; another 13% can be explained by major-to-non-major re-brandings. In one of the first applications of event study analysis to non-financial price data, we show that the adjustment to the new, lower price level is almost immediate in case of manned-to-unmanned conversions but takes one to two months in case of major-to-non-major re-brandings. The impact of both events is asymmetric with no measurable price impact of changes in the opposite direction.

Keywords: retail gasoline pricing, big data, competitive spillovers, event study analysis
JEL classification: L1, L13, L81

*We are particularly grateful to Pim Heijnen for his assistance and for providing insightful and useful comments during numerous discussions. We thank Irina Stanga, Marco Haan, Ruud Koning, and seminar participants at EARIE 2013, ASSET 2013 and the University of East Anglia for their valuable comments. Views and opinions expressed in this paper as well as all remaining errors are solely those of authors.

†University of Groningen, Faculty of Economics and Business, Nettelbosje 2, 9747 AE Groningen, The Netherlands, t.bruzikas@rug.nl.

‡Corresponding author: University of Groningen, Faculty of Economics and Business, Nettelbosje 2, 9747 AE Groningen, The Netherlands, a.r.soetevent@rug.nl.

1 Introduction

To understand how markets work, one needs an intimate knowledge of how firms set prices and compete. The “big data” revolution has made micro-level data on the economic decisions of consumers and firms available at an unprecedented scale. The world’s general capacity to store information digitally has grown with an average rate of 23% per capita per year in the years between 1986 and 2007 (Hilbert et al., 2011). Next to having a higher volume of recorded data, data nowadays arrive at a much higher frequency than ten to twenty years ago and cover activities previously unobserved.

One prime market for which the nature of the available data has changed tremendously is the retail gasoline market. Data sets used in empirical work published before the year 2000 all contain at most 10,000 price quotes, with the interval between observations in the majority of cases being a fortnight or longer. Studies published after 2010 instead use data sets with more than 10,000 observations in over 90 percent of cases and in the majority of studies, the frequency of recording price quotes is once every two days or higher (daily or even hourly). In numerous countries private parties have taken initiatives to publicize price information on the Internet,[1] in others price comparison websites have been established by law.[2] This high public interest in information on retail gasoline prices is explained by the significant budget share of fuel in the expenditures of households in most countries.[3]

The purpose of this paper is twofold. First, we survey how data availability of retail gasoline prices has changed and how this in turn has influenced the type of questions addressed in empirical research. Traditionally, price information has been relatively easier to obtain for retail gasoline markets than for most other markets because firms post them on large advertising signs next to the station. For this reason, an established body of empirical research of pricing studies is available that together span several decades and cover various countries. As a result, developments in data collection and analysis have left their traces in the form of shifts in research focus of published studies in this field, such as shifts in the methods used and the type of questions addressed.

Second, based on own data, we will show the possibilities offered by a long panel with detailed price quotes to identify trends and causality between trends. Although the volume-characteristic is the aspect of big data usually picked up by the media, we believe that the changes in the frequency will have at least as large an impact on empirical research. Importantly, the high frequency characteristic

[1] For example, Athlon Car Lease in the Netherlands `http://www.athloncarlease.com/speciallease/Producten-diensten/Online-diensten/Goedkoper-tanken/`

[2] For example in Australia (`http://www.fuelwatch.wa.gov.au`)) and, more recently, Germany (`http://www.clever-tanken.de/`).

[3] For example, the average share of fuel in the consumption expenditures of the 7.5 million Dutch households in 2012 was 4.3%, amounting to total expenditures of over 12 billion euro (Statistics Netherlands, 2014).

allows for the application of event study analysis (Brown and Warner, 1985; MacKinlay, 1997). The application of this technique has so far been limited to the field of finance, where it has been used to measure e.g. the impact of earning announcements on a firm's stock price by estimating abnormal ex post returns in the stock's performance. In the field of industrial economics, this event study analysis has been applied to estimate the impact of EU merger control decisions on consumer surplus (Duso et al., 2007), but again by considering the stock market prices of the firms involved in the decision.[4] The second part of this paper extends the use of event study analysis to non-financial price data using detailed data on the Dutch retail gasoline market. The central event in our analysis is the move in recent years of a high number of gasoline stations to unmanned retailing, i.e. stations without an attendant where customers pay by debit or credit card. This development is not limited to the Netherlands, but exemplary of a trend observed throughout Europe.[5]

We explore to which extent the direct and competitive effects of this trend towards unmanned stations can explain the increasing highway/off-highway price differential we observe and to which extent other changes such as re-brandings, entries and exits have had an impact. Not surprisingly because of their lower cost structure, a regression based on within-station variation reveals that stations that convert from manned to unmanned reduce prices with 4.5 and 3.1 eurocent per liter (cpl) on- and off-highways, respectively. Moreover, for off-highway sites we find that these conversions lead to significant competitive spillovers: A doubling of the number of unmanned stations in one's direct neighborhood leads to a price decrease of 0.22 cpl. Together, the direct and competitive effects of the increase in the number of unmanned stations explain 22% of the increase in the highway/off-highway price gap, another 13% can be explained by major-to-non-major re-brandings.

The event study analysis considers more in detail how prices responded in the time period immediately before and after the events of a manned/unmanned conversion or a major/non-major re-branding. Consistent with the regression estimates, we find price drops at the event day for both events. However, these effects are asymmetric in the sense that changes in the opposite direction do not have a long-term price impact. Notably, the price adjustment patterns of the two events are very different, with the adaptation to the lower price level being almost immediate in case of manned-to-unmanned conversions but taking one to two months in case of major-to-non-major re-brandings.

We are not the first to consider the question how the availability of larger and more complex

[4] We have found one study using an event study for non-financial data: McKenzie and Thomsen (2001) studying the impact of recalls (due to contamination) on wholesale beef prices.

[5] See for example `http://www.datamonitor.com/store/product/unmanned_service_station_retailing_in_europe?productid=CM00032-005`.

data sets will influence empirical research in economics. In a thoughtful piece, Einav and Levin (forthcoming) describe the key aspects of the transformation, the need for novel ways to summarize and analyze the information in these data and the possible application in economics of tools developed in statistics and computer science.[6] The purpose of the current paper is to illustrate these developments for the particular field of applied economics that studies retail gasoline prices.

Although the empirical focus in this paper is on the retail gasoline market, the outcomes have a wider relevance as the research techniques employed will be applicable to other markets as well. In particular, we have in mind other retail markets that are characterized by frequent price changes and consumer search, such as the markets for groceries, financial products and online markets.

Section 2 will describe how the availability of more and more detailed data has influenced empirical research on gasoline retailing. Section 3 introduces the data on the Dutch retail gasoline market that is the basis for the analysis in the subsequent sections. In particular, we will pinpoint changes in the market structure that have taken place since the start of this data collection effort in 2005, such as re-brandings, entry and exits and the conversions to unmanned retailing mentioned above. Section 4 looks at trends in the price data and identifies a steadily increasing differential between highway and off-highway prices. This trend is linked to the underlying trends in market structure, which culminates in a number of research hypotheses that are tested in Sections 5 and 6.

The number and frequency of price quotes enables us to use a regression approach with a very rich set of station-level and day fixed effects in Section 5. The impact of the changes in market structure on price is identified using within station variation only. This illustrates how detailed data assist in establishing causality. Counterfactual analysis allows us to quantify the effect of the number of unmanned sites on the highway/off-highway price differential. Section 6 studies with an event study approach in detail what happens to prices at stations (close to stations) experiencing a re-branding or conversion to unmanned retailing in the final days prior to the event and the first days after the event. Section 7 concludes.

2 Empirical studies on retail gasoline pricing

The empirical literature on retail gasoline markets is vast. In his survey, Eckert (2013) counts more than 100 empirical studies. Whereas Eckert (2013) reviews and summarizes the main findings of this

[6]Varian (2014) describes new statistical and econometric techniques needed to manipulate and analyze these data that can be used to obtain good out-of-sample predictions when conventional techniques do not suffice. The Economist (2013) paints a picture on how big data induce a revolution in infographics, the way we present information. `http://www.economist.com/news/books-and-arts/21580446-revolution-taking-place-how-visualise-information-winds-change`.

stream of research and briefly mentions the main data sources that researchers have used, our paper will focus on the the way different data characteristics have shaped the literature on retail gasoline, thereby complementing his survey. We refer the reader to Eckert (2013) for an extensive overview of the research questions analyzed and the main insights that have resulted.

This section broadly describes how greater data availability and the digital revolution has shaped the nature of empirical studies on retail gasoline pricing. Section 2.1 discusses which research questions typically have been addressed in the last two decades and what data have been employed to answer them. Section 2.2 zooms in on how the three aspects usually associated with big data – high volume, high frequency and high variety – have impacted the research orientation. Our assessment is based on a selection of empirical studies listed in Table 1. The selection procedure can be briefly described as follows. First, we searched JSTOR for keywords "retail gasoline". Then we selected empirical studies which use retail-price data. Finally, to keep the table compact we shortlisted the papers published in journals ranked 3 and 4 according to the Keele list.[7] At various points, we will refer to this table which also succinctly summarizes the main characteristics of the individual studies relevant for our purposes.

2.1 Scope

Table 1 shows that the majority of studies uses price quotes from the US or Canada. The share of high quality economic research being done in these countries has undoubtedly influenced this observation. However, readily-available station-level price data sources in North America, such as Oil Price Information Service (OPIS), have also played a major role in making these studies possible. Retail gasoline price data on North American markets are available from OPIS since 1999.[8] 13 out of 62 data sets in Table 1 are obtained from OPIS. The Energy Information Administration (EIA) and MJ Ervin & Associates Inc., providing gasoline price data for the US and Canadian markets respectively, have served as the source for a significant number of data sets too. Remarkably, 11 out of 19 daily-price data sets in Table 1 were provided by OPIS. All of these 11 data sets covered US markets. Evidently, accessibility and quality of gasoline price data in the US and Canada have significantly contributed to the literature on retail gasoline pricing.

In terms of topics, the studies can be roughly categorized into four research areas as Eckert (2013) suggests. The first and by far the largest group studies the asymmetry of retail price responses to increases and decreases in wholesale prices. The first two econometric analyses of this problem were

[7] Accessible at: http://www.uncp.edu/home/ashraf/JournalRankings/KeeleEconJournalRanks.pdf.
[8] Source: http://www.opisnet.com/about/opis.aspx.

those by Karrenbrock (1991) and Bacon (1991). The former reflects a large interest in this topic from politicians, industry specialists, and the media. The latter is inspired by the Monopolies and Mergers Commission inquiries in the UK.[9] The common finding is that retail prices adapt faster to increases in wholesale gasoline price than to decreases. The explanations for this phenomenon vary from consumer search (Tappata, 2009) to tacit collusion (Borenstein *et al.*, 1997). In a large cross-market comparison, Peltzman (2000) finds that more than two thirds of the roughly 200 investigated product prices respond to positive input price shocks faster than to negative ones. The great public interest in the issue, the occurrence of the same phenomenon in a wide array of other markets combined with easily accessible price information explain why studying response-asymmetry using retail gasoline price data has been a popular and worthwhile research topic. The second group analyzes whether and why the data exhibit Edgeworth cycles, the asymmetric cycles where prices fall till profit margins become very small or even negative, at which point prices jump back to the original price in a sudden and sharp movement (see Maskin and Tirole, 1988; Noel, 2007a,b, or Atkinson, 2009).

One third of all papers in Table 1 (21 articles) analyze either the asymmetric passthrough of upstream cost shocks or Edgeworth cycles. A few papers try to establish a connection between both phenomena although there is not much consensus in the findings (see Section 2.2.2). Edgeworth cycle theory assigns an important role to firm's best responses to their competitors prices. The availability of high frequency price data has played a major role in the development of the empirical literature on Edgeworth cycles by enabling researchers to observe these price responses in great detail. Of the papers using higher-than-daily frequency data, two-thirds focus on analyzing asymmetric price cycles.

The third research area concentrates on the impact of vertical relations and regulation on retail gasoline prices (see Barron and Umbeck, 1984; Hastings, 2004; Taylor and Hosken, 2007). Hastings (2004) for example shows that the presence of independent retailers helps to decrease local retail prices; Barron and Umbeck (1984) provide empirical evidence indicating that prices rise when a refiner-controlled station changes to a franchise operation.

[9] The Monopolies and Mergers Commission was replaced by the Competition Commission which was recently superseded by the Competition and Markets Authority.

Table 1: Empirical studies on retail gasoline markets

Author(s)	Length	Frequency	Market	Level of aggregation	Missing data	Reduced form	Category
Livingston & Levitt (1959)	cross section	–	6 metro. areas (Midwest US)	7234 stations	–	–	
Maurizi (1972)	1 year	monthly	US	10 city avgs.	N	Y	
Marvel (1976)	7 yrs.	monthly	23 metro. areas (US)	23 city max and min	N	Y	D
Marvel (1978)	7 yrs.	monthly	22 metro. areas (US)	22 city avgs.	N	Y	D
Barron & Umbeck (1984)	5 yrs.	*unknown*	Maryland (US)	99 stations	N	Y	V
Borenstein (1991)	8 yrs.	yearly	US	63 city avgs.	Y	Y	
Shepard (1991)	cross section	–	Massachusetts (US)	1527 stations	–	Y	
Slade (1992)	3 mths.	daily	Vancouver, BC (Canada)	10 stations	N	N	
Castanias & Johnson (1993)	8 yrs.	weekly	Los Angeles, CA (US)	city average	N	Y	E
Shepard (1993)	cross section	–	Massachusetts (US)	1527 stations	–	Y	D
Png & Reitman (1994)	cross section	–	Massachusetts (US)	1501 stations	–	Y	D
Borenstein & Shepard (1996)	7 yrs.	monthly	US	43 city avgs.	N	Y	
Borenstein et al. (1997)	7 yrs.	(bi-)weekly	US	national avg.	N	Y	A
Slade (1998)	3 mths.	daily	Vancouver, BC (Canada)	14 stations	N	Y	V
Asplund et al. (2000)	17 yrs.	daily	Sweden	7 firms (advice p.)	N	Y	A
Barron et al. (2000)	4 years	bi-monthly	LA Basin area, CA (US)	600 stations	N	Y	D
Peltzman (2000)	19 yrs.	monthly	US	national avg.	N	Y	A
Eckert (2002)	5 yrs.	weekly	Windsor, ON (Canada)	city avg.	N	Y	A
Bachmeier & Griffin (2003)	14 yrs.	daily	US	national avg.	N	Y	A
Eckert (2003)	6 yrs.	weekly	Canada	19 city avgs.	N	Y	E
Barron et al. (2004)	cross section	–	4 areas (US)	3197 stations	–	Y	D
Eckert & West (2004)	5 mths.	daily	Vancouver, BC (Canada)	404 stations	Y	Y	E
Hastings (2004)	4 mths.	monthly	LA and San Diego, CA (US)	669 stations	N	Y	V
Chen et al. (2005)	12 yrs.	weekly	US	national avg.	N	Y	A
Eckert & West (2005)	6 mths.	daily	Vancouver, BC (Canada)	391 stations	Y	Y	D
Skidmore et al. (2005)	20 yrs.	monthly	US	50 state avgs.	N	Y	
Abrantes-Metz et al. (2006)	7 yrs.	daily	Louisville, KY (US)	279 stations	Y	Y	D
Radchenko & Tsurumi (2006)	25 yrs.	monthly	US	national avg.	N	Y	
Cooper & Jones (2007)	4 mths.	monthly	Lexington, KY (US)	13 stations	N	Y	
Noel (2007a)	5 mths.	twice-daily	Toronto, ON (Canada)	22 stations	N	Y	E
Noel (2007b)	10 yrs.	weekly	Canada	19 city avgs.	N	Y	E
Taylor & Hosken (2007)	3 yrs.	daily	US	7 city avgs.	Y	Y	V

Table 1: (continued)

Author(s)	Length	Frequency	Market	Level of aggregation	Missing data	Reduced form	Category
Balmaceda & Soruco (2008)	4 yrs.	weekly	Santiago (Chile)	50 stations	N	Y	A
Barron et al. (2008)	3 mths.	daily	3 areas in California (US)	> 54 stations	Y	N	
Deltas (2008)	15 yrs.	monthly	US	48 states avgs.	N	Y	A
Doyle & Samphantharak (2008)	1 year	daily	Northeast US	6000 stations	Y	Y	
Hosken et al. (2008)	3 yrs.	weekly	Washington, DC (US)	272 stations	Y	Y	
Lewis (2008)	2 yrs.	weekly	San Diego area, CA (US)	327 stations	Y	Y	D
Simpson & Taylor (2008)	5.5 yrs.	daily	Michigan (US)	6 city avgs.	Y	Y	V
Verlinda (2008)	9 mths.	weekly	South California (US)	about 100 stations	N	Y	A
Alm et al. (2009)	16 yrs.	monthly	US	50 state avgs.	N	Y	
Atkinson et al. (2009)	3 mths.	bi-hourly	Guelf, ON (Canada)	27 stations	N	–	D
Lewis (2009)	2 yrs.	daily	US	85 city avgs.	N	Y	A, E
Manuszak & Moul (2009)	3 weekends	daily	Illinois and Indiana (US)	485 stations	Y	Y	
Neilson (2009)	2 mths.	twice-daily	Bryan-College Station, TX (US)	28 stations	N	Y	
Noel (2009)	5 mths.	twice-daily	Toronto, ON (Canada)	22 stations	N	Y	A, E
Wang (2009)	3.5 yrs.	hourly	Perth, WA (Australia)	286 stations	N	Y	E
Erutku & Hildebrand (2010)	2 yrs.	weekly	Quebec (Canada)	3 city avgs.	N	Y	D
Manuszak (2010)	5 yrs.	monthly	Hawaii (US)	1350 stations	Y	N	V
Sen & Townley (2010)	6 yrs.	monthly	Canada	10 city avgs.	N	Y	
Taylor et al. (2010)	3 yrs.	monthly	LA and San Diego, CA (US)	1001 stations	N	Y	V
Chandra & Tappata (2011)	18 mths.	daily	CA, FL, TX, NJ (US)	> 25000 stations	Y	Y	D
Hosken et al. (2011)	6 yrs.	daily	California (US)	3623 stations	Y	Y	V
Lewis (2011)	2 yrs.	weekly	San Diego, CA (US)	369 stations	Y	Y	A
Lewis & Marvel (2011)	4 mths.	daily	US	103 city avgs.	N	Y	A
Lewis & Noel (2011)	20 mths.	daily	US	90 city avgs.	N	Y	A, E
Marion & Muehlegger (2011)	20 yrs.	monthly	US	50 state avgs.	N	Y	
Myers et al. (2011)	1 year	daily	3 metro. areas (US)	5203 stations	Y	Y	
Erutku (2012)	5.5 yrs.	weekly	Quebec (Canada)	2 city avgs.	N	Y	D
Houde (2012)	11 yrs.	bi-monthly	Quebec City, QC (Canada)	429 stations	Y	N	V
Lewis (2012)	2 yrs.	daily	32 states (US)	30755 stations	Y	Y	E
Lewis (2012)	2 yrs.	tri-hourly	32 states (US)	1613 Speedway stations	N		

Note: Papers included are empirical studies on retail gasoline pricing which are published in journals ranked 3 and 4 in Keele's list. The category-column shows to which line of research a study contributes: **A:** Asymmetric cost shock pass-through; **E:** Edgeworth cycles; **V:** Vertical relations; **D:** Price dispersion. As also noted by Eckert (2013), some studies contribute to multiple categories whereas some others do not fit into any of the four main research areas.

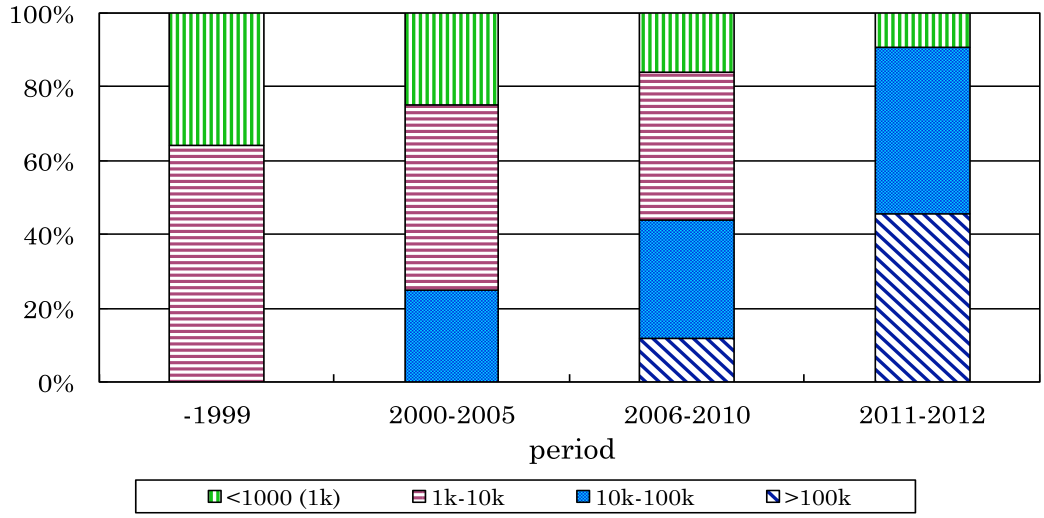
The final category of articles analyzes station-level price dispersion and examines to which extent price differences can be explained by differences in station characteristics or by differences in the level of local competition (see Barron *et al.*, 2004, or Clemenz and Gugler, 2006). By their nature, the completeness and variety of data (individual characteristics, location etc.) is more important for studies in this category than the number and frequency of price quotes. For example, besides retail price of gasoline, Barron *et al.* (2004) use an extensive set of controls (such as the stations' brand, location, number of pumps, availability of additional services, hours of operation, etc.) in their investigation of the relation between station density and price dispersion. Data on individual station characteristics are nowadays relatively easy to obtain because various market research companies routinely collect them and offer them for sale. In the following section, we discuss the impact that the main characteristics of big data have had on the retail gasoline literature.

2.2 Impact of better data availability

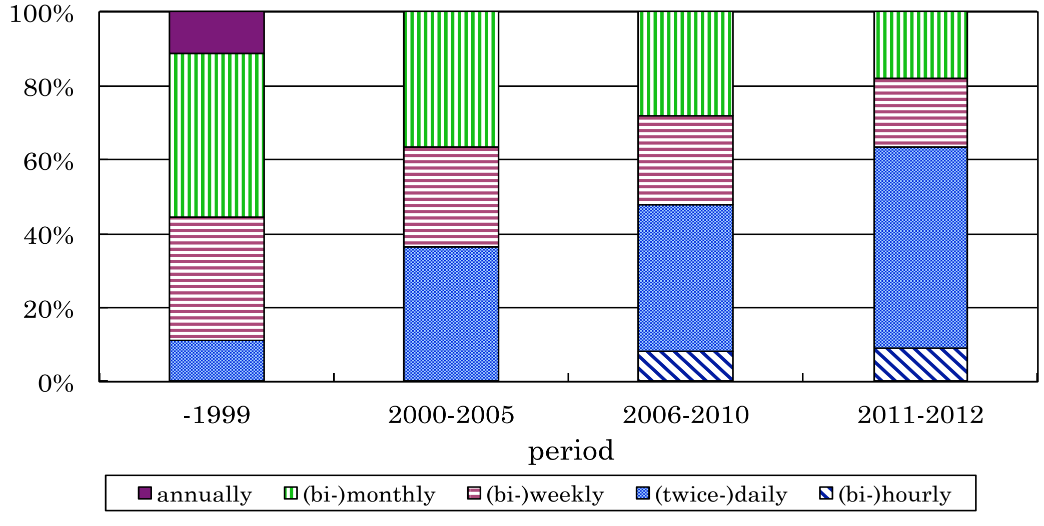
Research on gasoline retailing is largely driven by data availability. A greater level of detail in the data often allows new topics to be studied empirically. For example, the advent of weekly and daily price quotes has enabled researchers to study whether the competitive strategies firms apply in empirical practice fit the theoretical mechanism Maskin and Tirole (1988) have put forward to explain the occurrence of Edgeworth price cycles. Without price data arriving at a frequency exceeding or at least matching the frequency with which firms update their prices, we would not be able to tell the adequacy of the dynamic reaction functions that feature prominently in Maskin and Tirole's theory. Indeed, there are no empirical studies on Edgeworth cycles using bi-weekly or monthly data. In this subsection we discuss the trends in the characteristics of data used in studies on retail gasoline pricing as well as the implications of these trends.

2.2.1 Volume

The term "big data" is most often associated with the volume of the data. However, the empirical literature on retail gasoline cannot be solely characterized by the use of massive data sets counting millions or billions of observations. Despite a few recent studies which actually use very large data sets (see Chandra and Tappata, 2011; Hosken *et al.*, 2011; Lewis, 2012), the majority of recent papers does not rely on the volume but instead on the level of detail of the data. Nevertheless, Figure 1a shows that data sets used in research on retail gasoline are getting larger. There are several reasons underlying this trend. For example, the frequency of price quotes is also increasing over time as Figure 1b shows. Hence the data sets of the same length are larger than in the past. The volume of data sets has also increased due to changes in data collection techniques. The Internet had played a vital role in this development. Whereas in the past, aggregate price indices were constructed by statistical agencies or branch organizations, price quotes of individual stations are nowadays mostly available on firms' websites. Moreover, some companies also collect credit card data with fuel price records which are then sold to oil companies, distributors, traders, government, and academia. Many large data sets listed in Table 1 have been obtained from OPIS which provides the data on North American gasoline markets. These trends illustrate that research on retail gasoline and microeconomics in general are moving towards more intensive use of "big data".



(a) Volume



(b) Frequency

Figure 1: Developments of empirical literature by data-set characteristics.
Note: Number of papers per period from earliest to latest: 14; 12; 25; 11.

2.2.2 Frequency

Table 1 shows that before the 21st century, only two empirical studies using daily price quotes were published (see Slade, 1992, 1998). But even Slade's (1992, 1998) data set has a sample of only 14 stations in Vancouver which can represent a small sub-market at best. After 1994, cross-section data were used only once by Barron *et al.* (2004). These trends towards using less cross-section (more longitudinal) and more daily and hourly data are evident in Figure 1b.

Researchers use the high frequency property in different ways to answer specific questions. For example, Doyle and Samphantharak (2008) use daily prices to determine short-term pass-through of taxes by comparing two days before and two days after the tax change. Atkinson *et al.* (2009) use bi-hourly prices to study timing of responses to price increases and decreases. In both cases, the high frequency characteristic is crucial for identification. Atkinson (2008) actually shows that while (twice-) daily price data are suitable to identify the Edgeworth cycles, they are not sufficient to identify which stations initiate the cycles and which ones drive them down.

We find that six out of seven most recent studies on asymmetric price patterns used price data with daily to hourly frequencies.[10] Using daily data, Lewis and Marvel (2011) provide evidence that consumer search can explain asymmetric responses to cost shocks.[11] Daily price data have also been used to relate asymmetric response and Edgeworth cycles (see Noel, 2009; Lewis and Noel, 2011). While Noel (2009) argues that Edgeworth cycles might be a cause of or amplify asymmetric-response pricing, Lewis and Noel (2011) find that cost changes are passed on two to three times faster in markets exhibiting Edgeworth cycles. Lewis (2012) uses daily and tri-hourly price data to examine the price leadership and coordination mechanism in markets with Edgeworth cycles. He finds that the market leader signals the beginning of the cycle by raising a price to the same level at all its stations. The competitors then follow within 24 hours. These studies demonstrate how our knowledge about price asymmetries and price cycles has been informed by high frequency data.[12]

[10] These papers are Lewis (2009, 2011, 2012), Noel (2009), Wang (2009), Lewis and Marvel (2011), and Lewis and Noel (2011). An exception is Lewis (2011) who uses weekly prices. However, his sample covers over 50% of all sites in San Diego market.

[11] A theoretical explanation was proposed earlier by Tappata (2009).

[12] We would like to note that 3 out of 5 data sets with higher-than-daily frequency in Table 1 have been hand-collected. Other two used either credit card transactions data (Wang, 2009) or prices were downloaded directly from the company's website (Lewis, 2012). Hence, hourly retail prices are still rather difficult to obtain. Although OPIS now offers real-time price data too, see `http://www.opisnet.com/products/retail-fuel-station-prices.aspx`.

2.2.3 High variety and increased coverage

The availability of station-level price quotes has greatly enhanced the possibilities to study the impact of station characteristics and market structure on pricing. To accomplish the former, price data are merged with information on site characteristics including details about vertical contracts, brands, service hours, or other services that are available on-site (e.g. a grocery shop, a car wash, automotive services, etc). For the later, price data have been frequently merged with demographic information to construct proxies for station-specific demand which is usually unobservable to researchers (Doyle and Samphantharak, 2008; Verlinda, 2008). Advances in geocoding have given researchers more accurate information about the location of sites. By calculating distances between individual gasoline stations, most researchers nowadays have very precise measures of the intensity of local competition at their disposal. The approaches of gathering these data have differed substantially: Some use street addresses or postal codes, e.g. Barron *et al.* (2000), others use exact geographical coordinates using Google Earth (Soetevent *et al.*, 2014) or manually using a GPS unit (Verlinda, 2008). Given the developments in geocoding technology, we expect a strong growth of market power studies that use geocoded information.[13]

The exact market coverage of each data set listed in Table 1 is known only if the authors mention this explicitly. For example, Manuszak and Moul (2009), Wang (2009), and Chandra and Tappata (2011) note that their data cover (almost) 100% of the markets they study. This is exemplary of the trend to use less aggregated and more complete data: Note that in Table 1 the use of city- or state-averaged data is decreasing over time; also, until the year 2000, none of the papers had more than 100 stations in their sample (unless they used a cross section).

In this paper we exploit the fact that we have a sample of more than 85% of all gasoline stations in the Netherlands. This gives us detailed information about the structure of local markets and to identify changes in these markets. It also enables us to estimate various spillover effects consistently. The completeness of our market coverage combined with the length of the panel allows us to distinguish a number of trends in the Dutch retail gasoline market and to study possible causal relations between these trends.

[13] Our own experience is that third-party-collected street addresses are not always error-proof. For the Dutch market, we have come across cases where the address provided by the car-lease company is the one of the headquarters of the firm rather than a particular station. Fortunately, these cases seem rare.

3 Data description

3.1 Data collection

In August 2005, Athlon Car Lease (Athlon, hereafter), the leading car leasing company in the Netherlands with a fleet of over 125,000 cars, launched an online website where it has published the gasoline prices paid by its fleet-card owners ever since. Immediately after the launch, we have started a procedure to download these data on a daily basis and by the end of September 2005, this procedure was fully automated. We collected price data for all grades of gasoline, diesel, and liquefied petroleum gas (LPG), but in this paper we limit attention to the prices for regular unleaded 95 octane gasoline (known as Euro 95) which is the most commonly used type of fuel in the Netherlands.[14]

The price at a particular station at a given day is recorded if and only if at least one fleet-card owner frequented the site that day. This may bias our sample of observed prices towards the lower-end of the price distribution when drivers structurally avoid visiting the higher-priced stations. Fortunately, because Athlon's customers are employees of large companies or institutions who do not pay for the fuel themselves, they are likely to be price-insensitive.[15] Another bias would arise when site characteristics unobserved by the researcher would lead lessees to structurally avoid visiting some outlets. This however does not seem to be an issue given that our data contains more than 80% of all sites. The price data is appended with other explanatory variables such as the exact geographic coordinates and the (Euclidean) distances between all pairs of stations.[16] We also know the brand of station and whether it is an unmanned site.[17]

Our final sample includes 4,799 individual gasoline stations (at 3,828 sites) and runs from October 1, 2005 to April 25, 2011. According to TankPro.nl there were 4,206 gasoline stations in the Netherlands in June 2011, this means that our data cover 85 percent of all outlets in the country.[18] Together with all the individual station characteristics and other appended data, our data set is one of the largest and most detailed samples used in the literature on retail gasoline pricing.

[14] The same data set has been used in Soetevent, Haan and Heijnen (2014) and Heijnen, Haan and Soetevent (2014).

[15] Also note that only one lessee needs to visit a site for the price of that site to be recorded. So the sample will be unbiased as long as not all of the 125,000 drivers avoid higher-priced stations.

[16] These data were obtained using Google Earth.

[17] We know whether a particular station is unmanned either from its brand in the Athlon price data (TinQ, Tango, Shell Express, Q8 Easy etc.) or from the website `http://www.onbemandetankstations.nl` which lists all automated sites in the Netherlands. We have downloaded the list from this website twice, in the middle and at the end of the sample period.

[18] `http://www.tankpro.nl/brandstof/2011/11/30/aantal-tankstations-in-nederland-blijft-stabiel/`. Original source: PetrolView. We count 3,562 active sites in February 2011.

3.2 Within-station differences and changes in local market structure

Due to the long period of observation, our panel contains a large number of changes at the station level. Part A of Table 2 reports these changes. The table shows that the most pronounced trend has been the conversion of manned to unmanned sites, with many more such conversions off-highway (291) than on-highway (7). We also observe a drop in the number of off-highway sites carrying one of the six largest brands. In total, there are 247 off-highway (12 highway) sites that have been re-branded from major (Shell, Esso, BP, Texaco)[19], TOTAL or Q8 to another brand, 68 (6) stations have made a change in the opposite direction.

These changes at the station level have induced significant changes in the local market context in which highway and off-highway stations operate. Part B of Table 2 shows that 951 (135) off-highway (highway) stations experienced an increase in the number of off-highway competitors within a 2 km (5 km) radius and 713 (102) a decrease.[20] In other words, both for highway and off-highway stations competition by off-highway stations intensified, at least when measured as the number of neighboring off-highway stations.[21] Using the same 2 km (5 km) radius, there were 1,207 off-highway (159 highway) sites that experienced an increase in the number of off-highway unmanned sites in their neighborhood, resulting from either conversion or entry. Some sites experienced multiple changes in their local environment: The numbers in brackets reflect these double counts. Even though we observe only two sites which converted from unmanned to manned, there is a significant number of highway and off-highway sites that saw the number of unmanned off-highway sites in their neighborhood decrease at some point. This is because a number of unmanned off-highway stations ceased operating altogether. A total of 1,070 off-highway (146 highway) stations saw at some point the number of off-highway majors, TOTAL or Q8 within a 2 km (5 km) radius decrease, while 647 (109) stations saw an increase.[22] Taking entries and exits of stations into account, the number of stations in the colors of one of six most important brands decreased by roughly 170 (cf. Table A.1). The next section describes how these market trends have developed over time and relates them to changes in pricing.

[19]This definition of "major" is also used by the Netherlands Competition Authority (Nederlandse Mededingingsautoriteit, NMa).

[20]As in Soetevent *et al.* (2014) we define the relevant market for off-highway (highway) sites as all other sites within a 2 km (5 km) radius.

[21]Table 2 also shows the corresponding numbers for changes in neighboring highway stations. These numbers are much smaller and for this reason not discussed in the text.

[22]Note that both sets are not mutually exclusive.

Table 2: Number of sites that experienced changes in site characteristics and local market characteristics

A. Site characteristics				
Unmanned	*2005*	NO → YES	YES → NO	*2011*
Highway	*5*	7	0	*12*
Off-highway	*424*	291	2	*713*
Major		NO → YES	YES → NO	
Highway	*176*	4	8	*172*
Off-highway	*1469*	60	130	*1399*
TOTAL		NO → YES	YES → NO	
Highway	*24*	2	2	*24*
Off-highway	*419*	5	91	*333*
Q8		NO → YES	YES → NO	
Highway	*15*	0	2	*13*
Off-highway	*109*	3	26	*86*

B. Local market characteristics		
# off-highway sites	Increase	Decrease
≤5 km from highway site	135 [224]	102 [156]
≤2 km from off-highway site	951 [1221]	713 [837]
# unmanned off-highway sites	Increase	Decrease
≤5 km from highway site	159 [320]	46 [51]
≤2 km from off-highway site	1207 [1742]	270 [295]
# major+TOTAL+Q8 off-highway sites	Increase	Decrease
≤5 km from highway site	109 [171]	146 [278]
≤2 km from off-highway site	647 [802]	1070 [1466]
# highway sites	Increase	Decrease
≤5 km from highway site	3 [3]	0 [0]
≤2 km from off-highway site	4 [4]	2 [2]
# unmanned highway sites	Increase	Decrease
≤5 km from highway site	5 [5]	0 [0]
≤2 km from off-highway site	7 [10]	0 [0]
# major+TOTAL+Q8 highway sites	Increase	Decrease
≤5 km from highway site	14 [14]	16 [16]
≤2 km from off-highway site	20 [20]	21 [21]

Notes: Figures in brackets indicate the total number of events (including changes affecting the same site). Figures in *italics* indicate the total number of sites in certain category on 31-12-2005 and 31-03-2011, respectively.

4 Market description and trends

Whereas the previous section reported for a number of variables the total number of observed changes, this section discusses whether the Dutch retail gasoline market shows any structural trends in market concentration, the conversion to unmanned stations and in pricing.

4.1 Prices

Figure 2: Average retail gasoline price, ARA spot price, and crude oil prices (Oct. 2005 – April 2011).

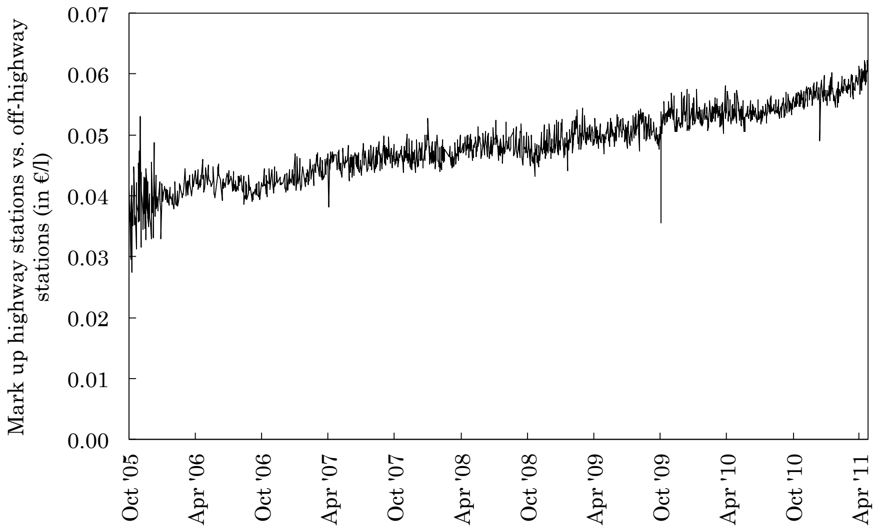


Figure 3: Absolute mark up highway vs. off-highway sites (01/10/2005–25/04/2011).
Note: Dates with less than 20 price quotes for highway sites and/or less than 30 price quotes in total have been excluded.

Gasoline prices in the Netherlands have varied widely over the last few years. Figure 2 shows that fluctuations in aggregate retail prices reflect the dynamics of the crude oil spot price: They move closely in line with the Amsterdam-Rotterdam-Antwerp (ARA) premium unleaded gasoline spot price. The ARA price gradually increased until the onset of the Great Recession in August 2008 initiated a sharp decline. In the two years following, prices recovered and reached their previous peaks in Spring 2011. In Figure 3, prices are disaggregated by comparing the average daily prices charged by highway stations with those charged off-highway. The figure shows that highway stations consistently charge higher prices. This difference does not come as a surprise; in two important merger cases the European Commission has also judged that highway stations constitute a separate product market (Exxon/Mobil 29/9/1999 and TotalFinaElf 9/2/2000).[23] What is surprising is not the gap *per se*, which may just reflect a premium paid for filling your tank on highways, but that the gap between average highway and off-highway prices has been increasing from roughly 4 cpl in October 2005 to 6 cpl in April 2011 and that this increase has been almost linear. This is the first trend we distill from the descriptive statistics.

Trend 1 *The price differential between highway and off-highway stations in the Dutch retail gasoline market has increased very gradually from 4 to 6 eurocents per liter from October 2005 to April 2011.*

4.2 Market concentration

Figure 4 shows the development of the market shares of the seven most important firms in the Dutch retail gasoline market: the four traditional majors plus TOTAL, Q8 and Gulf. Due to the absence of revenue data, the market share of a firm is defined as the percentage of all gasoline stations operating under one of its brand names.[24,25] For any given day, we calculate the number of active sites on that day by counting the number of stations for which the first price quote is recorded before, and the final price quote after that particular day. In this way, the calculated market shares are insensitive to fluctuations caused by some sites not being sampled on a particular day.[26]

With a market share over 14 percent, Shell is the market leader in 2011 despite having lost 3 percentage points of its market share since January 2006. Gulf increased its market share significantly during the sample period. Together, these firms own more than two thirds of all stations. TOTAL has experienced the largest decrease in market share with a fall of nearly 4 percentage points.

As Figure 5 attests, the development at highways is different. The four major brands plus TOTAL and Q8 capture almost the entire market. Whereas the off-highway market has a significant share of

[23]NMa has also mentioned differences in regulations and policies on approval for highway and off-highway stations (NMa, 2004).

[24]For Texaco, the market share is calculated including the sites it operations under the Firezone-brand; for Q8 including the sites it operates under the Tango-brand (automated stations); for Gulf including the sites it operates under the TinQ-brand (automated stations).

[25]Table A.1 in Appendix A compares the market shares as calculated from our data with the statistics from Catalist. Our data is accurate in terms of sampled stations and covers more than 80 percent of almost all six major-branded sites (the exception is Q8 with 69 percent). Figure A.1 shows that the number of sampled stations is growing over time, improving the representativeness of our sample.

[26]We drop the first and the last month of observation (October 2005 and April 2011) from our analysis to avoid that stations entering and exiting the database in these months, respectively, are misclassified as entering or exiting the market.

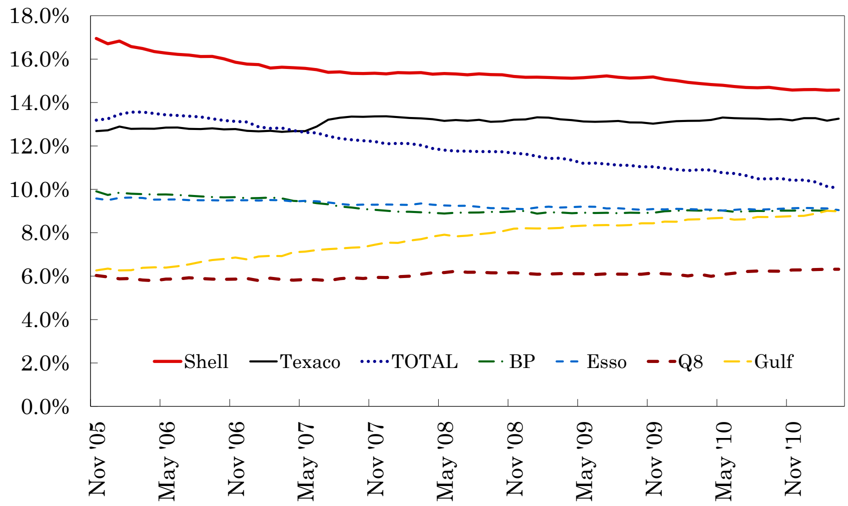


Figure 4: Development of the market shares by firm for the top 7 firms (November 2005 – March 2011).

other players, these six companies still own 88 percent of all highway sites at the end of the sample period. Including off-highway stations, the total market share of the six major firms decreased with 10 percentage points to 56 percent during the sample period; a decrease of 11%. This trend has possibly affected the highway/off-highway price differential because major brands are usually able to sell gasoline at a premium.

Trend 2 *The market concentration of the six major brands has decreased by 11% in the off-highway market in the time period October 2005 to April 2011 while their total market share in the highway-market has been stable around 90 percent.*

4.3 The emergence of unmanned stations

Recent years have seen a significant increase in the number of unmanned fuel stations. Figure 6 shows a very steady growth in the number of unmanned sites off-highways where the market share of unmanned stations has more than doubled from 13.1% to 26.5%. The relative growth has been even larger at highways, but starting from a much lower base: The market share of unmanned stations along highways has increased from 2.1% to 5.9%.

Although the aggregate increase has been very gradual, there is considerable variation in the pace with which individual firms convert stations into unmanned stations. Figure 7 shows that TinQ (since 2004 a sub-brand of Gulf solely operating unmanned stations) has witnessed the largest growth in its market share: from 1.9 to 6.0 percent of the total number of stations. This increase is mostly due to

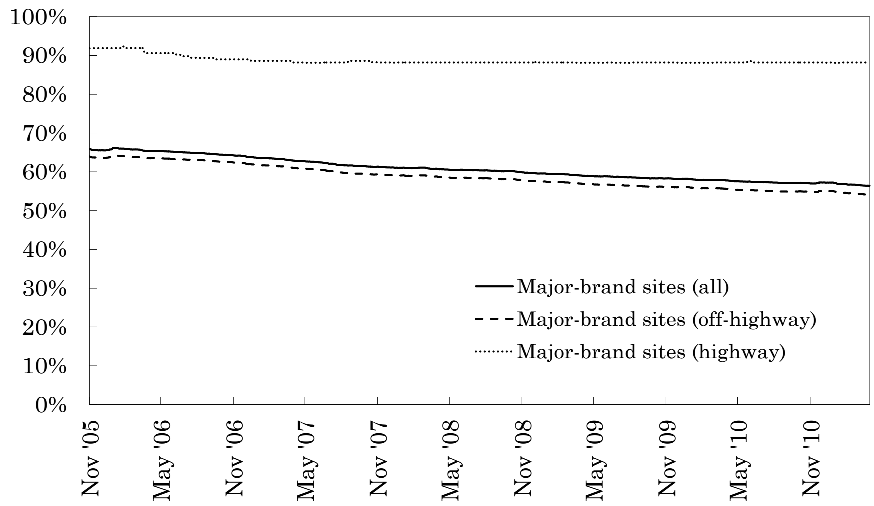


Figure 5: Development of the aggregate market share of the four majors plus TOTAL and Q8 on highway and off-highway markets (November 2005 – March 2011).

Gulf re-branding its stations to operate under the brand name of TinQ. All other brands have also increased the number of unmanned stations in the period considered: Tango's share has been growing steadily and Firezone has become a prominent player in the market. Finally, we observe that Esso and especially BP have only converted a small number of their stations to unmanned stations.

Trend 3 *In the off-highway market, the market share of unmanned stations has more than doubled in the time period October 2005 to April 2011 to 26.5%. The market share of unmanned stations in the highway market has tripled in the same period, but is still modest (about 6%).*

4.4 Research hypotheses

The introduction of unmanned stations is an obvious route to cut cost.[27] On the surface this supports the view propagated by oil companies[28] that competition off-highway has become stronger because of the steeply increasing market share of unmanned sites. However, because of the cost-structure of unmanned stations, one cannot immediately conclude which part of the increased highway/off-highway price differential is due to the increased efficiency of off-highway sites (a direct effect) and which part to more intense competition (a spillover effect). Therefore, we dedicate the next section to examine in more detail the relation between the increased number of unmanned stations and competitive spillovers.

[27] In previous work, Soetevent *et al.*, 2014 have estimated that off-highway, gasoline sold at unmanned stations is approximately 2.6% cheaper.

[28] `http://www.tankpro.nl/brandstof/2014/01/03/prijsverschillen-snelweg-en-onderliggend-wegennet-groeien/`.

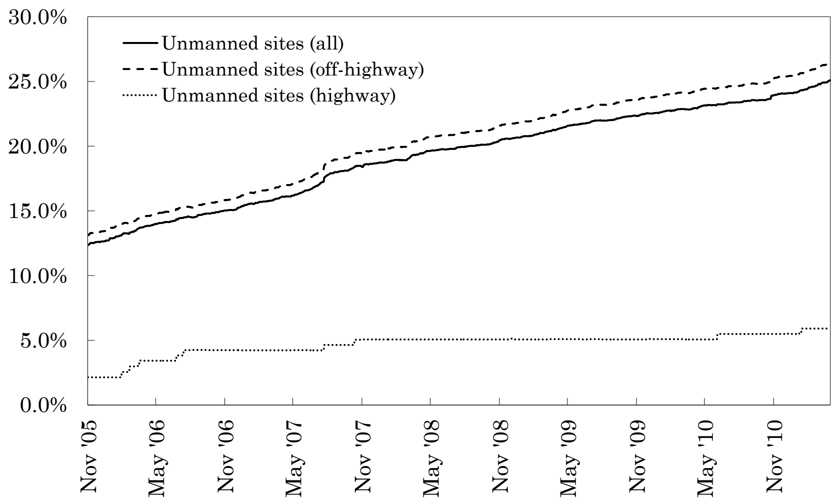


Figure 6: Development of the percentage of unmanned sites (November 2005 – March 2011).

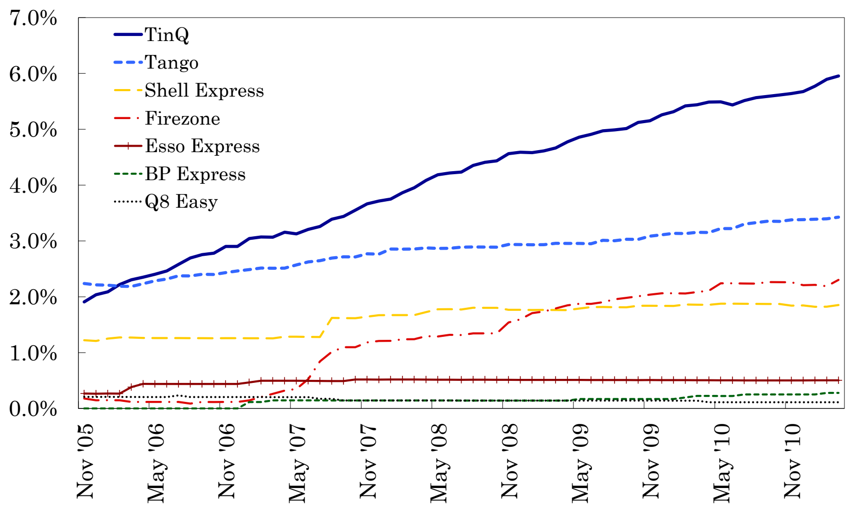


Figure 7: Development of the market shares of unmanned stations by brand (November 2005 – March 2011).

The hypotheses that we will test are based on the trends identified in the previous subsections. In particular, we aim to test the extent to which the observed divergence in the highway/off-highway price differential (Trend 1) is caused by the decreased share of major-brand stations (Trend 2) and/or the increased share of unmanned stations in the highway and the off-highway market (Trend 3).

To formulate our hypotheses, we make the following two assumptions. First, we assume that the retail gasoline market can be characterized as a differentiated Bertrand market, meaning that stations compete in prices and that their products are viewed as imperfect substitutes. Location is probably the most important dimension of differentiation but gasoline produced by different oil companies also slightly differs in the additives used. Moreover, the majority of stations offers additional services such as a shop or a car wash. For these reasons, consumers may not be indifferent between different stations. Second, we assume that unmanned sites will operate at lower costs. Neither of these assumptions is controversial. Given these assumptions, we expect the conversion from manned to unmanned to have two effects on prices.

First, we expect a site that converts to an unmanned site lowers its price, because of lower operating cost. This is the direct effect. Second, due to prices being strategic complements, we expect an indirect/competitive effect as well: The lower price charged by their unmanned neighbors will put a downward pressure on the prices of nearby sites. This competitive effect will typically be much smaller in magnitude than the direct effect. In March 2011, a typical off-highway site faced an average number of 0.89 unmanned off-highway competitors within 2-kilometer distance, up from 0.46 in November 2005. For the typical highway site, the corresponding numbers are 0.32 and 0.16 (within a 5 km radius). The sheer number of sites that has converted in recent years may render the competitive effect statistically and economically significant.

Of course, other developments than the increased number of unmanned stations may also have contributed to the widening price gap in the period considered. These include (unobserved) asymmetric changes in the cost structure of highway and off-highway stations and changes that have led to increased competition off-highway (or, decreased competition among highway sites) such as the entry of new firms/brands and the opening of new locations. In our empirical analysis, we will therefore contrast the direct and competitive effects of manned to unmanned conversions with the effects of entry, exit and re-branding (see Trend 2).

In sum, we arrive at the following research hypotheses:

Hypothesis 1 (Direct effects) *Sites that convert from a manned to an unmanned site will lower their prices.*

Hypothesis 2 (Competitive spillovers) *The price level at a station is decreasing in the number of unmanned stations in its neighborhood.*

Hypothesis 3 (The highway/off-highway price differential) *The rapid growth in the market share of unmanned stations in the off-highway market explains part of the increased gap between average highway and off-highway prices.*

We can test Hypothesis 1 because the size of our data set gives us sufficient observations of gasoline outlets that converted from manned to unmanned (a small number experienced a conversion in the opposite direction, see Table 2); because of the high velocity with which new price quotes arrive, we also have for each outlet sufficient price quotes before and after the conversion. Using geocoded information on the exact location of the sites, we can identify an outlet's local competitors and track on a daily basis changes in the number of unmanned outlets in a site's neigbhorhood. Together, this allows us to test Hypothesis 2. Hypothesis 3 is evaluated by means of a counterfactual analysis. Using the parameter estimates of reduced-form-regression model and fixing the values of the right-hand-side variables at the initial level we can determine the extent to which changes in the level of local competition have affected the gap between highway and off-highway prices.

5 Analysis

The large N large T panel data set we employ ($N = 3,820$, $T = 1,977$) allows for a statistical model that includes both day-specific fixed effects (because of the high volume of observations at a given day), and station-specific fixed effects (because of the high number of observations per station). The time fixed effects capture the time-varying price components common to all highway and off-highway firms. The station-specific fixed effects absorb all unobserved variables at the station level that may be correlated with the other regressors. Including such a rich set of fixed effects implies that in a reduced-form regression of prices on a number of explanatory variables, only the coefficients of the time-varying regressors will be identified. Fortunately, as we saw in Section 3.2, a considerable number of such changes have taken place. Next to site-level characteristics such as brand name and an "unmanned" dummy variable, we include local market characteristics as explanatory variables, such as, for off-highway (highway) stations, the log of the number of highway and off-highway sites within 2 km (5 km), the log of the number of sites of a major brand within 2 km (5 km) and the log of the number of unmanned sites within 2 km (5 km).[29]

We estimate the following two-way fixed effect model:

$$p_{it} = \begin{cases} c_i + c_t^{H_1} + \boldsymbol{\beta}^{\boldsymbol{H_1}}\mathbf{x}_{it} + \varepsilon_{it} & \text{if} \quad H_i = 1 \\ c_i + c_t^{H_0} + \boldsymbol{\beta}^{\boldsymbol{H_0}}\mathbf{x}_{it} + \varepsilon_{it} & \text{if} \quad H_i = 0 \end{cases}$$

where H_i is a dummy variable equal to 1 if station i is located on a highway. The dependent variable is the price p_{it} at station i at day t (measured in euros per liter). On the right-hand side, c_i and $c_t^{H_1}/c_t^{H_0}$ denote station-level and time fixed effects, respectively. The time fixed effects capture the time-varying price components common to all stations, such as variation in the price of crude oil and

[29]We use the natural logarithm of the number of neighbors (plus 1) because the price effects of the 1$^{\text{st}}$ and the 10$^{\text{th}}$ neighbor will in general not be the same. The implicit assumption in this specification is that new competitors will have a larger impact on prices the lower the initial number of local competitors.

taxes (excise duty and VAT) and we expect this effect to differ for highway ($c_t^{H_1}$) and off-highway ($c_t^{H_0}$) stations. The vector $\mathbf{x}_{it} = (x_{it}^1, x_{it}^2, \ldots, x_{it}^K)'$ includes the time varying variables. We allow these to have a different effect on highway ($\boldsymbol{\beta}^{\boldsymbol{H_1}}$) and off-highway ($\boldsymbol{\beta}^{\boldsymbol{H_0}}$) price levels. Throughout, we cluster the errors ε_{it} at the *location* level to account for the fact that, despite the inclusion of station-level and daily time fixed effects, price observations at a given station may be characterized by serial correlation or heteroskedasticity.[30]

Statistical tests clearly favor the above specification with station- and time-fixed effects over a model with random station-specific effects. A Sargan-Hansen test for overidentifying restrictions in the random effects specification indicates that our specification with fixed effects is strongly preferred ($\chi^2(41) = 208.45$, $p < 0.001$).[31] We also test for the time fixed effects and conclude that we should include them ($F(64,\ 3819) = 1.7 \cdot 10^5$, $p < 0.001$).[32]

5.1 Estimates

Table 3 presents the estimates. We find that both on- and off-highway stations are cheaper when unmanned. The size of the effect is 4.5 cpl and 3.1 cpl, respectively, both significant at the 5% and 1% level. This supports Hypothesis 1 that stations will lower their price after they have converted into an unmanned station (with less service and a lower cost-base). We cannot reject the null hypothesis that the direct effect is the same for highway and off-highway sites.[33]

Off-highway sites Second, for off-highway sites we find support for Hypothesis 2 that an increased number of unmanned competitors in station's local environment will lead the station to set lower prices on average. Off-highway stations decrease their prices with on average 0.22 cpl (significant at the 5% level) when the number of off-highway unmanned stations within a 2 km radius doubles because of the conversion of existing stations.[34] When these unmanned stations open at new locations, the average off-highway station decreases price with on average 1.1 cpl.[35] We also see that the density

[30] A Wooldridge test (Wooldridge, 2002, 282-283) on the residuals from the first-differenced regression indeed finds significant ($p < 0.001$) serial correlation in the the disturbances warranting the use of clustering. We emphasize 'location' because there are different firms that may operate the same site at different points in time. Therefore, it is reasonable to assume that price observations are correlated not only at each station (firm), but at the location level as well.

[31] We use the user-written Stata command `xtoverid`.

[32] The null of no time fixed effects is also firmly rejected in a joint test for common and highway-specific (monthly) time fixed effects ($F(128,\ 3819) = 1.4 \cdot 10^5$, $p < 0.001$). The F-statistics are very large because the average daily price level has varied widely over the sample period.

[33] $F(1,\ 3819) = 0.65$, $p = 0.420$

[34] The number of off-highway unmanned stations within a 2 km radius doubles in 42.2% (735 cases) of all increases for this variable.

[35] This total effect of entry by an unmanned station is calculated as $0.88 + 0.22 = 1.1$ cpl, implicitly assuming that the impact of entry is similar for manned and unmanned stations.

Table 3: Fixed-effects regression of p_{it} on explanatory variables.

	Highway		**Off-highway**	
	coeff.	**s.e.**	**coeff.**	**s.e.**
Local market characteristics				
ln(# hw. sites + 1)	0.0225***	(0.0077)	0.0143	(0.0122)
ln(# hw. unmanned sites + 1)	-0.0035	(0.0062)	-0.0047	(0.0119)
ln(# hw. six major sites + 1)	0.0117	(0.0072)	-0.0033	(0.0081)
ln(# off-hw. sites + 1)	-0.0090**	(0.0040)	-0.0088***	(0.0018)
ln(# off-hw. unmanned sites + 1)	0.0019	(0.0012)	-0.0022**	(0.0009)
ln(# off-hw. six major sites + 1)	0.0060**	(0.0028)	0.0021	(0.0013)
Site characteristics				
Unmanned	-0.0454**	(0.0178)	-0.0309***	(0.0019)
Major	0.0138***	(0.0050)	0.0200***	(0.0022)
TOTAL	0.0101	(0.0069)	0.0267***	(0.0039)
Q8	0.0163***	(0.0054)	0.0169***	(0.0029)
Station fixed effects	YES			
Day fixed effects	YES			
Obs.	4,153,898			
R^2 within	0.1055			

Standard errors are clustered at the site level. *** p<0.01, ** p<0.05, * p<0.1

of highway stations or changes in the brand name or operation mode of highway sites does not affect off-highway prices. This fact supports the view that the highway and off-highway market can be considered separate markets. Regarding the brand-name effects we find that all six major brands (Shell, Texaco, BP, and Esso including TOTAL and Q8) off-highway are on average 1.7-2.7 cpl more expensive than their competitors. For off-highway sites, we do not identify significant spillovers of (major) brand effects.

Highway sites At highway stations, fuel prices are 1.4 cpl higher when they serve one of the major brands and 1.6 cpl when their brand is Q8.[36] Other than off-highway sites, prices at highway sites do not seem to experience any impact from changes in the number of local unmanned stations, neither on-highway nor off-highway. So, where we identified significant competitive spillover effects of the conversion of off-highway sites to prices of other off-highway sites, prices of nearby highway outlets seem unaffected. A formal test of equality of the indirect effects for highway and off-highway stations also supports this finding by rejecting the null ($F(1,\, 3819) = 7.89$, $p = 0.005$).[37] The fact that prices at highway stations seem rather immune to the significant increase in the number of unmanned off-

[36]Note however that Table 2 shows that identification in the latter case rests on two stations that were re-branded from Q8 to another brand, with no station experiencing a change in the opposite direction.

[37]Note that if we also consider the indirect effects of unmanned *highway* sites, the joint test is significant only at the 5% level ($F(2,\, 3819) = 3.96$, $p = 0.019$).

highway stations is consistent with Hypothesis 3 stating that this off-highway development may be responsible for part of the increased highway/off-highway price differential.

Entries and exits in the off-highway market however do impact highway prices: Sites that experience a doubling of their number of off-highway rivals within 5 km distance, decrease their price with on average 0.9 cpl.[38] This suggests that highway sites do compete with off-highway sites conditional on being located sufficiently close to the smaller roads.[39] If the entering or exiting station is one of the six major brands, this competitive effect is less ($-0.9 + 0.6 = -0.3$ cpl). Note that prices at highway sites increase with the number of highway competitors within 5 km distance. We believe that this is a result of the fact that entry in the highway market is very regulated and almost exclusively permitted at highways that expand. For this reason, the estimate is likely to pick up underlying increases in local highway traffic, allowing stations to increase their highway premium.

We have tested the robustness of our estimates by using different sub-samples for estimation. For example, we dropped days if we observe less than one half of the maximum number of daily price quotes ($\frac{1}{2}\max(N_t)$), because highly-frequented stations may be over-represented on these days. In another test, we also excluded the bottom 10% of stations in terms of the number of price quotes. Finally, because of our data collection method, we observe less price quotes on Wednesdays and Thursdays in the first two years of our sample.[40] The data set is sufficiently long to allow us to drop the first years of observations without losing precision. None of these modifications changed our results either quantitatively or qualitatively. The most significant change in the estimates is that for highway sites, the number of highway competitors within 5 km distance remains barely significant at the 10% level.

In sum, the evidence presented in Tables 2 and 3 shows that, compared to highway sites, considerably more off-highway sites have converted to unmanned stations thereby not only significantly reducing the prices at the converted sites but also at sites in their direct competitive environment.

5.2 Counterfactual analysis

The previous section estimated at the station-level the price effect of changes in selected site and local market characteristics. Our broader objective however is to find out to which extent these changes can explain the increased highway/off-highway price differential shown in Figure 3. How large is the effect of the manned/unmanned conversion on prices compared to the effects of entry/exit and major/non-

[38] For the highway sites that experience local entry (exit) off-highway, the median number of off-highway sites within 5 km is 9 (13) prior to entry (exit).

[39] We have noticed that some off-highway sites have placed big signs on roofs of farms next to a highway-exit to direct drivers to their premises.

[40] The reason is that until February 2007, we were not collecting the data during weekends. When we started to do so, the average number of price quotes per day jumped from roughly 1,400 to more than 2,000.

major conversions? In this section, we use the estimates in Table 3 to calculate the aggregate effect of these changes on the highway and off-highway price level by means of a counterfactual analysis.

Our approach is as follows. To assess the aggregate price effect of changes in a time-varying variable x_{it}^k, we first construct counterfactual prices by fixing the value of this variable at the initial (January 2006) level to predict for each station the prices $\hat{p}_{it}^k$ that would have been observed under this counterfactual market constellation.[41] That is, we calculate $\hat{p}_{it}^k$ as:

$$\hat{p}_{it}^k = \begin{cases} \hat{c}_i + \hat{c}_t^{H_1} + \hat{\boldsymbol{\beta}}^{\boldsymbol{H_1,-k}} \mathbf{x}_{it}^{-k} + \hat{\beta}^{H_1,k} x_{i0}^k \varepsilon_{it} & \text{if} \quad H_i = 1 \\ \hat{c}_i + \hat{c}_t^{H_0} + \hat{\boldsymbol{\beta}}^{\boldsymbol{H_0,-k}} \mathbf{x}_{it}^{-k} + \hat{\beta}^{H_0,k} x_{i0}^k \varepsilon_{it} & \text{if} \quad H_i = 0 \end{cases}$$

with x_{i0}^k the value of variable k fixed at $t = 0$ and $\mathbf{x}_{it}^{-k} = (x_{it}^1, x_{it}^2, \dots x_{it}^{k-1}, x_{it}^{k+1}, x_{it}^K)'$ the vector of remaining time-varying variables (with corresponding coefficients $\hat{\beta}^{\cdot,k}$ and $\hat{\boldsymbol{\beta}}^{\boldsymbol{\cdot,-k}}$, respectively).

In the second step, we compare, for highway and off-highway sites, the average of the observed prices at the final date T (March 2011) of the sample period ($\bar{p}_T = \sum_{i=1}^N p_{iT}/N$) with the average counterfactual price $\bar{\hat{p}}_T^k = \sum_{i=1}^N \hat{p}_{iT}^k/N$ at this date. The difference between the two shows the net effect fixing the k^{th} variable at its initial values has on prices at time T.[42]

Table 4 presents the results of this exercise. With regard to the number of unmanned stations, the estimates show that off-highway prices in March 2011 are on average 0.39 cpl lower than under the counterfactual. Nearly 90 percent of this difference can be attributed to the direct effect of converting to an unmanned station, the remainder is due to competitive spillovers. The table indicates that manned-to-unmanned conversions did not have any substantial effect on the average highway price level. The net effect of major to non-major re-branding (and *vice versa*) is -0.26 cpl for off-highway stations and again negligible for highway stations. This is no surprise because Table 2 shows that in the period considered, most action has been off-highway and most changes involved the conversion from a major, Total or Q8 to a non-major brand station. Table 2 also shows that the number of market entries and exits has been modest, hence the effects of entry and exit on average prices have been small.

In total, we find that the level of off-highway prices decreased by 0.66 cpl relative to highway prices due to observable fluctuations in the market. We calculated highway/off-highway price differential increased by 1.9 cpl in the sample period. Changes in the share of unmanned stations on- and off-highway are responsible for 22% of this increase. In total, changes in the observed time-varying factors

[41] We choose January 2006 as the initial period because this month is by far the most representative in terms of average number of daily observations until February 2007.

[42] Figure A.2 plots the development of the subtotal effects of unmanned, major (including Total and Q8) brand, and entry/exit variables for off-highway sites.

Table 4: The effects of the changes in the number of unmanned sites, brand name, and station density on the average highway and off-highway prices (in *cents* per liter).

Variable	Effect	**Highway**	**Off-highway**
Unmanned	Direct	-0.04	-0.35
NO ↔ YES	Competitive	0.07	-0.05
	Subtotal	**0.03**	**-0.39**
Major	Direct	-0.01	-0.25
NO ↔ YES	Competitive	-0.01	-0.01
	Subtotal	**-0.02**	**-0.26**
Entry/Exit	Direct	*n/a*	*n/a*
	Competitive	-0.03	-0.01
	Subtotal	**-0.03**	**-0.01**
	Total	**-0.01**	**-0.67**

Notes: *n/a* – not applicable; The variable "major" includes the four traditional majors plus TOTAL and Q8.
Subtotals and totals do not add up due to rounding errors.

explain about 35% of the increased price differential. The remaining 65% cannot be accounted for by changes in any of the time-varying regressors in our data.

6 Event study approach

6.1 Motivation

The regression analysis in the previous section primarily considered the average direct and competitive price effects of a change in one of the station characteristics. Also, because of the small number of observed changes for some of the variables (see Table 2), we implicitly assumed the impact of a change to be the same independent of the direction of the change. For example, the price effect of a nearby exit was assumed to be exactly the negative of the effect of a nearby entry. The aim of this section is to look more in detail what happens to prices in the time period surrounding a change in local market constellation. The high frequency characteristic of our data allows us to apply event study analysis to this end.

Next to shedding light on possible asymmetries in the effects of manned/unmanned conversions, re-branding from and to a major, and entry and exits, event study analysis enables one to see how quickly prices react to the change and whether prices show any change *prior to* the event. The latter is a real possibility. For example, competitors may anticipate the entry by a new competitor by lowering their prices in advance. Also, the decision to convert a manned station into an unmanned one may be related to particular negative price developments in the local market of that station. In these cases, a simple comparison of price levels before and after the change would underestimate the direct effect of converting into an unmanned station and the competitive effect of entry, respectively.

6.2 Analytical setup

In this section we explain the empirical setup of the event study framework we are using. Our setup is almost entirely based on the framework described by MacKinlay (1997).

Retail gasoline prices are almost perfectly correlated with highly volatile crude oil prices. For any event study, we have to isolate the influence of all factors except the event itself. Hence, we consider price deviations from the average highway or off-highway market price, whichever is relevant:

$$\tilde{p}_{it} = \begin{cases} p_{it} - \bar{p}_t^{H_1} & \text{if} \quad H_i = 1 \\ p_{it} - \bar{p}_t^{H_0} & \text{if} \quad H_i = 0 \end{cases}$$

As before, this transformation accounts for all possible time fixed effects such as taxes, seasonal trends, or weekday-specific effects specific to the set of highway or off-highway stations. In event studies on financial data, the interest is in detecting abnormal returns in stock prices. Given our context, we redefine returns as follows. The first difference of $\tilde{p}_{it}$ is taken as the 'return' of station i on day t. That is,

$$R_{it} = \Delta\tilde{p}_{it} = \tilde{p}_{it} - \tilde{p}_{i,t-1} = \mu_i + \zeta_{it} \tag{1}$$

where ζ_{it} is a disturbance term with a mean zero and variance $\sigma^2_{\zeta_i}$. Following the terminology in MacKinlay (1997), this specification has the form of a market model with the imposed constraint that prices at individual stations will move with the general market and therefore show the same volatility.[43] Other than in a financial market context where μ_i measures a stock's excess return (as compensation for the risk borne), we have reason not to expect individual μ's to significantly differ from zero.[44] There are a couple of reasons for using this specification.[45] Firstly, the time series of prices at individual outlets reveal that for most stations, the difference between the station's price and the national average price is very stable over time. Due to inflation, this difference will automatically grow somewhat in time, but this effect is negligible in our case because the periods considered are relatively short (up to a couple of months) and inflation has been benign in the period 2005-2011. Secondly, even though we observe a considerable number of changes in local markets, individual stations are not very often exposed to an event with subsequent events in most cases sufficiently separated in time[46]

[43]The specification is identical to MacKinlay (1997, p. 18) equation (3) with $\beta_i = 1$ for all stations i.

[44]We indeed do not find any stations with a μ_i significantly different from zero at $p = 0.10$.

[45]The common alternative is the market model which assumes a stable linear relation between, in our case, national price changes and station-level price changes. For details, see MacKinlay (1997, pp. 18, 20-21).

[46]Apart from case were they happen simultaneously, e.g. a brand name change going together with a conversion into

for their effects on price returns not to interfere.[47]

Our interest is in estimating abnormal price changes in the periods surrounding the event. To that end, we define an 80-days event window as the time period $[-10, 70)$, containing the price observations from 10 days before to 70 days after the event. In order to identify whether an observed price change in the event window is "abnormal", we have to specify what price changes we would normally expect. To do this, we employ the period $[-90, -10)$ prior to the event as an estimation window. In this configuration, the estimation window is of the same length as the event window and long enough to obtain consistent estimates of normal returns; the event window is sufficiently wide to see the full adjustment of prices after the event day.

Having defined returns, the estimation and event window, abnormal returns at dates τ within the event window are calculated as a difference between the actual and predicted return:

$$AR_{i\tau} = R_{i\tau} - \hat{\mu}_i \tag{2}$$

The variance of $AR_{i\tau}$ is simply equal the variance $\sigma^2_{\zeta_i}$. For a period $[\tau_1, \tau_2]$ within the event window, we can calculate the cumulative abnormal return (CAR) for station i by aggregating abnormal returns from τ_1 to τ_2:

$$CAR_i(\tau_1, \tau_2) = \sum_{\tau=\tau_1}^{\tau_2} AR_{i\tau} \tag{3}$$

Since in our setup the return is a price change, CAR_i can be interpreted as the *abnormal price level* of station i.

To estimate the price effect of a particular event-type (e.g. manned-to-unmanned conversion), we need to aggregate for this event all abnormal return observations for the event window and across observations of the event. In doing this, we assume that the different events are independent. In other words, we rule out overlap between the different event windows at a given station and higher order competitive spillovers. The first assumption is not stringent.[47] The second assumption is somewhat stronger because it imposes that, for example, a station that sees one of its competitors convert into an unmanned site, only responds to that event and not to the possible price response of the converted station's other competitors.

Given N events of a given type, e.g. manned-to-unmanned or non-major-to-major, we compute

an unmanned station.

[47] The maximum number of different events per station is 2 (at 5 sites only). Those events are at least 7 months (217 days) apart hence the estimation and the event windows do not overlap.

the sample aggregated abnormal returns at day τ as:

$$\overline{AR}_\tau = \frac{1}{N_\tau}\sum_{i=1}^{N_\tau} AR_{i\tau} \tag{4}$$

Note that in aggregating abnormal returns, we assume that the events are independent. This assumption is not unreasonable if events of a given type are sufficiently separated in time or space.[48] The variance of average abnormal return depends on the length of the estimation period (MacKinlay, 1997). Since we have a reasonably wide estimation window, we can use the asymptotic variance estimate:

$$\mathrm{Var}(\overline{AR}_\tau) = \frac{1}{N_\tau^2}\sum_{i=1}^{N_\tau} \sigma^2_{\zeta_i} \tag{5}$$

Using equations (4) and (5) the cumulative (average) abnormal return and its variance for any interval in the event window can be calculated as follows:

$$\overline{CAR}(\tau_1,\tau_2) = \sum_{\tau=\tau_1}^{\tau_2} \overline{AR}_\tau \tag{6}$$

$$\mathrm{Var}(\overline{CAR}(\tau_1,\tau_2)) = \sum_{\tau=\tau_1}^{\tau_2} \mathrm{Var}(\overline{AR}_\tau) \tag{7}$$

In aggregating abnormal returns, we assume that the abnormal returns of the included events in a given region do not overlap in calendar time such that the covariances across events can be reasonably assumed to be zero.[49] Finally, to test the significance of $\overline{CAR}(\tau_1,\tau_2)$ against the null hypothesis $\overline{CAR}(\tau_1,\tau_2) = 0$, MacKinlay (1997) derives the following test statistic

$$\theta_1 = \frac{\overline{CAR}(\tau_1,\tau_2)}{\sqrt{\mathrm{Var}(\overline{CAR}(\tau_1,\tau_2))}} \sim N(0,1) \tag{8}$$

with the asymptotic distribution being the limiting distribution with respect to the number of gasoline stations experiencing a certain event and the length of the estimation window.

6.3 Results

We calculate the cumulative abnormal return $\overline{CAR}(\tau_1,\tau_2)$ with $\tau_1 = -10$ and $\tau_2 \in [-10, 70)$. Setting τ_1 equal to the beginning of the event window yields wide confidence intervals and thus a conservative test of the significance of the cumulative abnormal returns.[50]

[48] None of the sites in our data set experience multiple events of a given type.

[49] Taking the 80-day event window $[-10, 70)$ we have 24 instances (12.5% of the total) of overlap within a *5 km* distance for the event manned-to-unmanned; 6 (4.8%) for major-to-non-major, and 0 for non-major-to-major.

[50] Recall from equation (7) that variance is accumulated day-by-day. Therefore, if any results are significant for this choice of τ_1, we can be sure that they will remain significant in different specifications of $\overline{CAR}(\tau_1^*,\tau_2)$ with $\tau_1^* \in (\tau_1,\tau_2]$.

Our main interest is in the events of manned/unmanned conversions and re-brandings from a major to non-major station (or *vice versa*).[51] To disentangle the effects of different event-types, we exclude in our analysis all 99 (out of 457 in total) cases where these two events coincide at a site (e.g. sites that experience a manned-unmanned conversion combined with a re-branding). Another 60 sites are excluded because they have extremely few (≤ 5) observations in the estimation window. Finally we exclude all highway stations from our analysis because of a lack of events. Consequently, the number of events included in the sample is 151 for manned-to-unmanned, 1 for unmanned-to-manned, 37 for non-major-to-major, and 109 for major-to-non-major conversions.

Figure 8 depicts the cumulative abnormal returns for these events in the event window. A comparison of panels (a) and (b) clearly shows that the impact of a conversion from manned into unmanned is not symmetric. Stations becoming unmanned reduce prices by roughly 2.7 cpl on the day of event.[52] The standard error (s.e.) of this one day abnormal return is 0.096 and the $\theta_1 = 27.71$, such that the null hypothesis of no impact is strongly rejected. After the event date, prices consistently stay at this lower level, despite an upward jump at event day 7 (0.6 cpl, s.e.=0.113), possibly because of the ending of first-week discounts. In contrast, panel (b) of Figure 8 shows that site that converts into manned does not experience any significant price change.[53]

The impact of a major to non-major brand change is approximately 1.4 cpl at the event date. However, Figure 8(c) shows that this amounts to less than a half of the cumulative price decrease of about 4.0 cpl realized at the end of the event window. Indeed, in the 40 days following the event, we observe 4 days with significant price decreases in the range 0.2-0.4 cpl and only 1 day with a significant increase of the same order of magnitude.[54] Panel (d) of the figure further shows that this effect is asymmetric as well with re-branding from non-major to major having no significant long-term impact; there is a significant drop of 0.6 cpl at the event day (s.e.=0.106), but this drop is only temporary.

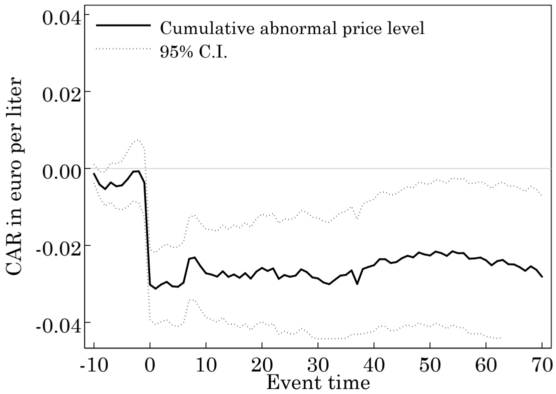
In other words, where both the conversion into an unmanned station and the change to serving a non-major brand lead to a significant decrease in a station's price level, the price adjustment patterns of the two events are very different, with the new price level being reached much more gradually in case of a re-branding. Moreover, whereas the magnitude of the cumulative effects of both events are in line with the fixed effects estimates of Table 3, Figure 8 reveals that the impact is very asymmetric with prices responding to changes from manned to unmanned or major to non-major but no measurable

[51] Again, in this section the variable "major" includes Total and Q8 besides the four traditional majors.

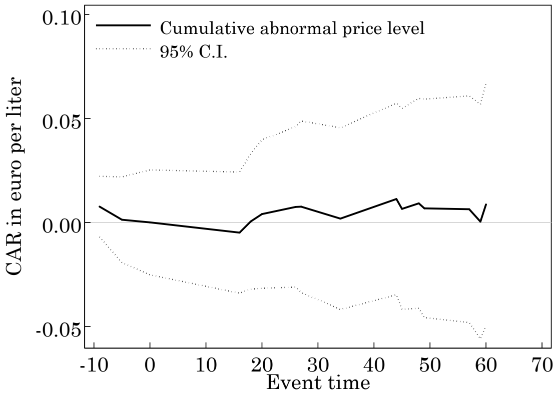
[52] Table A.2 in the Appendix shows for the different events the abnormal returns per event date.

[53] The confidence intervals are wide because only 1 station qualifies for the event-study analysis. Other sites experience other events at the same time or lack observations in the estimation window.

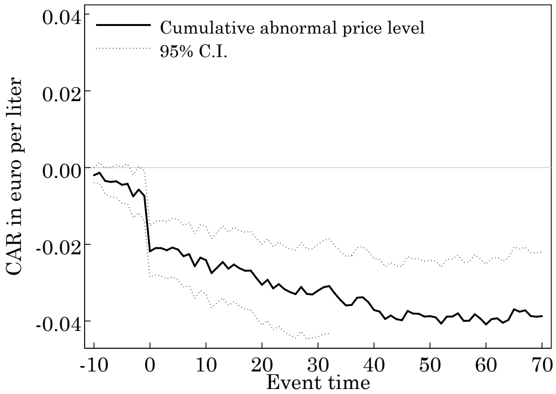
[54] At the 5% significance level. These days are event day 8, 11, 22, and 40, and 9, respectively.



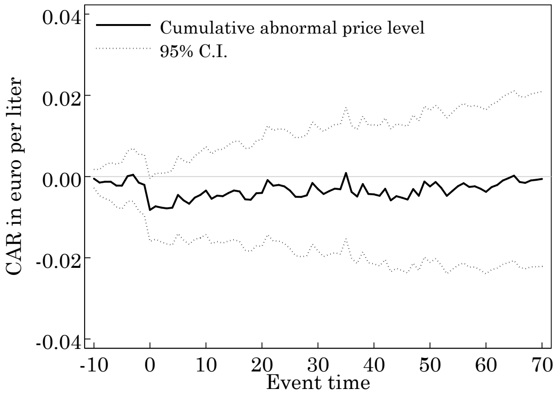
(a) . . . manned into unmanned



(b) . . . unmanned into manned



(c) . . . major into non-major



(d) . . . non-major into major

Figure 8: The cumulative abnormal returns for off-highway sites converting from . . .
Note: All figures for event day -10 to event day 70.

long-term impact of changes in the opposite direction.

Finally, we analyze local spillover effects of the manned/unmanned conversions and major/non-major re-brandings. We do so by considering the events that the number of unmanned (or major) sites within a 2 km radius of a site changes from 0 to 1, from 1 to 2, or *vice versa.* None of the spillover effects are significant at the 5-percent level (see Figures A.3(a)-(d) in Appendix A).

7 Summary and conclusions

Data availability plays a significant role in empirical research and the "big data" revolution has made micro-level data available at an unprecedented scale. For the retail gasoline market, we have pictured the tremendous increase in the volume and frequency of price data since the 2000s and the shift in research orientation triggered by this. The increase in frequency has contributed most to redirecting research in this field. The availability of daily and even hourly price data has allowed researchers to answer new questions as well as to provide the new insights on old problems.

The second part of the paper illustrated some of the opportunities offered by detailed data. An extensive panel data set of almost daily price quotes for more than 85 percent of all gasoline stations in the Netherlands revealed that the highway/off-highway price differential has increased by roughly 50 percent to 6 cpl in the period 2005-2011. We formulated a number hypotheses relating this price trend to underlying trends in market structure, especially the sharp increase in unmanned retailing.

Using a rich set of day and station fixed effects, we found that stations that convert from manned to unmanned reduce prices with 4.5 and 3.1 cpl on- and off-highways, respectively. For off-highway sites we find that these conversions lead to significant competitive spillovers: For off-highway sites, a doubling of the number of off-highway unmanned stations lead to a price decrease of 0.22 cpl. The direct and competitive effects of the increase in the number of unmanned stations together explain 22% of the increase in the highway/off-highway price gap. Another 13% is explained by the re-branding of stations from a major to a non-major brand, a development that has been more prominent in the off-highway market.

We subsequently applied event study analysis to consider more in detail how prices responded in the time period immediately before and after the events of a manned/unmanned conversion or a major/non-major re-branding. Consistent with the regression estimates, we found price drops at the event day when a station converts to an unmanned station or is re-branded as a non-major. However, in both cases, these effects are asymmetric: Changes in the opposite direction do not have a long-term price impact. Whereas the adaptation to the lower price level is almost immediate in case of manned-to-unmanned conversions, it is much more gradual in case of major-to-non-major re-brandings, taking one to two months.

While the event study approach is widely used in finance to measure for example the effect of financial statement announcements on a firm's stock market price, this is the first application of event study analysis to non-financial retail price data. Our results show the added value of this approach in increasing our understanding of pricing. We envision that this, combined with the increased availability of detailed price data, will lead to an increased use of this tool. We hope that this study will stimulate the use of similar research approaches in empirical work on other retail markets that are characterized by frequent price changes and consumer search, such as the markets for groceries, financial products and online markets.

References

Abrantes-Metz, Rosa M., Luke M. Froeb, John F. Geweke, and Christopher T. Taylor, "A Variance Screen for Collusion," *International Journal of Industrial Organization*, May 2006, *24* (3), 467–486.

Alm, James, Edward Sennoga, and Mark Skidmore, "Perfect Competition, Urbanization, And Tax Incidence In The Retail Gasoline Market," *Economic Inquiry*, January 2009, *47* (1), 118–134.

Asplund, Marcus, Rickard Eriksson, and Richard Friberg, "Price Adjustments by a Gasoline Retail Chain," *Scandinavian Journal of Economics*, March 2000, *102* (1), 101–121.

Atkinson, Benjamin, "On Retail Gasoline Pricing Websites: Potential Sample Selection Biases and Their Implications for Empirical Research," *Review of Industrial Organization*, September 2008, *33* (2), 161–175.

____ , "Retail Gasoline Price Cycles: Evidence from Guelph, Ontario Using Bi-Hourly, Station-Specific Retail Price Data," *Energy Journal*, 2009, *30* (1), 85–110.

____ **, Andrew Eckert, and Douglas S. West**, "Price Matching And The Domino Effect In A Retail Gasoline Market," *Economic Inquiry*, July 2009, *47* (3), 568–588.

Bachmeier, Lance J. and James M. Griffin, "New Evidence on Asymmetric Gasoline Price Responses," *Review of Economics and Statistics*, August 2003, *85* (3), 772–776.

Bacon, Robert W., "Rockets and Feathers: The Asymmetric Speed of Adjustment of UK Retail Gasoline Prices to Cost Changes," *Energy Economics*, July 1991, *13* (3), 211–218.

Balmaceda, Felipe and Paula Soruco, "Asymmetric Dynamic Pricing in a Local Gasoline Retail Market," *Journal of Industrial Economics*, September 2008, *56* (3), 629–653.

Barron, John M. and John R. Umbeck, "The Effects of Different Contractual Arrangements: The Case of Retail Gasoline Markets," *Journal of Law and Economics*, October 1984, *27* (2), 313–328.

____ **, Beck A. Taylor, and John R. Umbeck**, "A Theory of Quality-Related Differences in Retail Margins: Why There Is a "Premium" on Premium Gasoline," *Economic Inquiry*, October 2000, *38* (4), 550–569.

____ **,** ____ **, and** ____ , "Number of Sellers, Average Prices, and Price Dispersion," *International Journal of Industrial Organization*, November 2004, *22* (8-9), 1041–1066.

____ **, John R. Umbeck, and Glen R. Waddell**, "Consumer and competitor reactions: Evidence from a field experiment," *International Journal of Industrial Organization*, March 2008, *26* (2), 517–531.

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

____ **, A. Colin Cameron, and Richard J. Gilbert**, "Do Gasoline Prices Respond Asymmetrically to Crude Oil Price Changes," *Quarterly Journal of Economics*, February 1997, *112* (1), 305–339.

____ **and Andrea Shepard**, "Dynamic Pricing in Retail Gasoline Markets," *RAND Journal of Economics*, 1996, *27* (3), 429–451.

Brown, Stephen J. and Jerold B. Warner, "Using Daily Stock Returns: The Case of Event Studies," *Journal of Financial Economics*, March 1985, *14* (1), 3–31.

Castanias, Rick and Herb Johnson, "Gas Wars: Retail Gasoline Price Fluctuations," *Review of Economics and Statistics*, February 1993, *75* (1), 171–174.

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

Chen, Li-Hsueh, Miles Finney, and Kon S. Lai, "A threshold cointegration analysis of asymmetric price transmission from crude oil to gasoline prices," *Economics Letters*, November 2005, *89* (2), 233–239.

Clemenz, Gerhard and Klaus Gugler, "Locational Choice and Price Competition: Some Empirical Results for the Austrian Retail Gasoline Market," *Empirical Economics*, 2006, *31* (2), 291–312.

Cooper, Thomas E. and John T. Jones, "Asymmetric Competition on Commuter Routes: The Case of Gasoline Pricing," *Southern Economic Journal*, October 2007, *74* (2), 483–504.

Deltas, George, "Retail gasoline price dynamics and local market power," *Journal of Industrial Economics*, September 2008, *56* (3), 613–628.

Doyle, Joseph J. Jr. and Krislert Samphantharak, "$2.00 Gas! Studying the Effects of a Gas Tax Moratorium," *Journal of Public Economics*, 2008, *92* (3-4), 869–884.

Duso, Damien J. Neven Tomaso and Lars-Hednrik Röller, "The Political Economy of European Merger Control: Evidence using Stock Market Data," *Journal of Law and Economics*, August 2007, *50* (3), 455–489.

Eckert, Andrew, "Retail Price Cycles and Response Asymmetry," *Canadian Journal of Economics*, February 2002, *35* (1), 52–77.

——, "Retail Price Cycles and the Presence of Small Firms," *International Journal of Industrial Organization*, February 2003, *21* (2), 151–170.

——, "Empirical Studies Of Gasoline Retailing: A Guide To The Literature," *Journal of Economic Surveys*, February 2013, *27* (1), 140–166.

—— **and Douglas S. West**, "Retail Gasoline Price Cycles across Spatially Dispersed Gasoline Stations," *Journal of Law and Economics*, April 2004, *47* (1), 245–273.

—— **and** ——, "Price Uniformity and Competition in a Retail Gasoline Market," *Journal of Economic Behavior and Organization*, February 2005, *56* (2), 219–237.

Einav, Liran and Jonathan Levin, "The Data Revolution and Economic Analysis," in Josh Lerner and Scott Stern, eds., *Innovation Policy and the Economy*, Vol. 14, National Bureau of Economic Research Inc, forthcoming.

Erutku, Can, "Testing post-cartel pricing during litigation," *Economics Letters*, 2012, *116* (3), 339–342.

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

Heijnen, Pim, Marco A. Haan, and Adriaan R. Soetevent, "Screening for Collusion: A Spatial Statistics Approach," *Journal of Economic Geography*, 2014.

Hilbert, Martin and Priscilla López, "The Worlds Technological Capacity to Store, Communicate, and Compute Information," *Science*, April 2011, *332* (60), 60–65.

Hosken, Daniel S., Louis Silvia, and Christopher T. Taylor, "Does Concentration Matter? Measurement of Petroleum Merger Price Effects," *American Economic Review*, May 2011, *101* (3), 45–50.

——**, Robert S. McMillan, and Christopher T. Taylor**, "Retail Gasoline Pricing: What Do We Know?," *International Journal of Industrial Organization*, November 2008, *26* (6), 1425–1436.

Houde, Jean-François, "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline," *American Economic Review*, August 2012, *102* (5), 2147–2182.

Karrenbrock, Jeffrey D., "The behavior of retail gasoline prices: symmetric or not?," *Review, Federal Reserve Bank of St. Louis*, July 1991, pp. 19–29.

Lewis, Matthew S., "Temporary Wholesale Gasoline Price Spikes Have Long-Lasting Retail Effects: The Aftermath of Hurricane Rita," *Journal of Law and Economics*, August 2009, *52* (3), 581–605.

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

——, "Price Leadership and Coordination in Retail Gasoline Markets with Price Cycles," *International Journal of Industrial Organization*, 2012, *30* (4), 342–351.

—— **and Howard P. Marvel**, "When Do Consumers Search?," *Journal of Industrial Economics*, September 2011, *59* (3), 457–483.

—— **and Michael D. Noel**, "The Speed of Gasoline Price Response in Markets With and Without Edgeworth Cycles," *Review of Economics and Statistics*, 2011, *93* (2), 672–682.

Livingston, S. Morris and Theodore Levitt, "Competition and Retail Gasoline Prices," *Review of Economics and Statistics*, May 1959, *41* (2), 119–132.

MacKinlay, A. Craig, "Event Studies in Economics and Finance," *Journal of Economic Literature*, 1997, *35* (1), 13–39.

Manuszak, Mark D., "Predicting the Impact of Upstream Mergers on Downstream Markets with an Application to the Retail Gasoline Industry," *International Journal of Industrial Organization*, January 2010, *28* (1), 99–111.

—— **and Charles C. Moul**, "How Far for a Buck? Tax Differences and the Location of Retail Gasoline Activity in Southeast Chicagoland," *Review of Economics and Statistics*, November 2009, *91* (4), 744–765.

Marion, Justin and Erich Muehlegger, "Fuel tax incidence and supply conditions," *Journal of Public Economics*, October 2011, *95* (9-10), 1202–1212.

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

——, "Competition and Price Levels in the Retail Gasoline Market," *Review of Economics and Statistics*, May 1978, *60* (2), 252–258.

Maskin, Eric and Jean Tirole, "A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles," *Econometrica*, May 1988, *56* (3), 571–599.

Maurizi, Alex R., "The Effect of Laws Against Price Advertising: The Case of Retail Gasoline," *Economic Inquiry*, September 1972, *10* (3), 321–329.

McKenzie, Andrew M. and Michael R. Thomsen, "The Effect Of E. Coli O157:H7 On Beef Prices," *Journal of Agricultural and Resource Economics*, December 2001, *26* (2), 431–444.

Myers, Caitlin Knowles, Grace Close, Laurice Fox, John William Meyer, and Madeline Niemi, "Retail Redlining: Are Gasoline Prices Higher In Poor And Minority Neighborhoods?," *Economic Inquiry*, July 2011, *49* (3), 795–809.

NMa, "Zonder Kritische Klant Geen Concurrentie," Press release, June 17, 2004.

Noel, Michael D., "Edgeworth Price Cycles: Evidence from the Toronto Retail Gasoline Market," *Journal of Industrial Economics*, March 2007a, *55* (1), 69–92.

___ , "Edgeworth Price Cycles, Cost-Based Pricing and Sticky Pricing in Retail Gasoline Markets," *Review of Economics and Statistics*, May 2007b, *89* (2), 224–234.

___ , "Do Retail Gasoline Prices Respond Asymmetrically to Cost Shocks? The Influence of Edgeworth Cycles," *RAND Journal of Economics*, 2009, *40* (3), 582–595.

Peltzman, Sam, "Prices Rise Faster than They Fall," *Journal of Political Economy*, June 2000, *108* (3), 466–502.

Png, Ivan P.L. and David Reitman, "Service Time Competition," *RAND Journal of Economics*, 1994, *25* (4), 619–634.

Radchenko, Stanislav and Hiroki Tsurumi, "Limited Information Bayesian Analysis of a Simultaneous Equation with an Autocorrelated Error Term and its Application to the U.S. Gasoline Market," *Journal of Econometrics*, July 2006, *133* (1), 31–49.

Sen, Anindya and Peter G.C. Townley, "Estimating the Impacts of Outlet Rationalization on Retail Prices, Industry Concentration, and Sales: Empirical Evidence from Canadian Gasoline Markets," *Journal of Economics and Management Strategy*, September 2010, *19* (3), 605–633.

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

___ , "Contractual Form, Retail Price, and Asset Characteristics in Gasoline Retailing," *RAND Journal of Economics*, 1993, *24* (1), 58–77.

Simpson, John and Christopher T. Taylor, "Do Gasoline Mergers Affect Consumer Prices? The Marathon Ashland Petroleum and Ultramar Diamond Shamrock Transaction," *Journal of Law and Economics*, February 2008, *51* (1), 135–152.

Skidmore, Mark, James Peltier, and James Alm, "Do State Motor Fuel Sales-Below-Cost Laws Lower Prices?," *Journal of Urban Economics*, January 2005, *57* (1), 189–211.

Slade, Margaret E., "Vancouver's Gasoline-Price Wars: An Empirical Exercise in Uncovering Supergame Strategies," *Review of Economic Studies*, April 1992, *59* (2), 257–276.

___ , "Strategic Motives for Vertical Separation: Evidence from Retail Gasoline Markets," *Journal of Law, Economics and Organization*, April 1998, *14* (1), 84–113.

Soetevent, Adriaan R., Marco A. Haan, and Pim Heijnen, "Do Auctions and Forced Divestitures Increase Competition? Evidence for Retail Gasoline Markets," *Journal of Industrial Economics*, September 2014, *62* (3), 467–502.

Statistics Netherlands, http://statline.cbs.nl/, website consulted May 15, 2014.

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

Taylor, Christopher T. and Daniel S. Hosken, "The Economic Effects of the Marathon-Ashland Joint Venture: The Importance of Industry Supply Shocks and Vertical Market Structure," *Journal of Industrial Economics*, September 2007, *55* (3), 419–451.

___ , Nicholas M. Kreisle, and Paul R. Zimmerman, "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California: Comment," *American Economic Review*, June 2010, *100* (3), 1269–1276.

The Economist, "Infographics – Winds of Change," http://www.economist.com/news/books-and-arts/21580446-revolution-taking-place-how-visualise-information-winds-change/, July 6, 2013.

Varian, Hal R., "Big Data: New Tricks for Econometrics," *Journal of Economic Perspectives*, Spring 2014, *28* (2), 3–28.

Verlinda, Jeremy A., "Do Rockets Rise Faster and Feathers Fall Slower in an Atmosphere of Local Market Power? Evidence from the Retail Gasoline Market," *Journal of Industrial Economics*, September 2008, *56* (3), 581–612.

Wang, Zhongmin, "(Mixed) Strategy in Oligopoly Pricing: Evidence from Gasoline Price Cycles Before and under a Timing Regulation," *Journal of Political Economy*, December 2009, *117* (6), 987–1030.

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, Cambridge, Massachusetts: MIT Press, 2002.

A Appendix

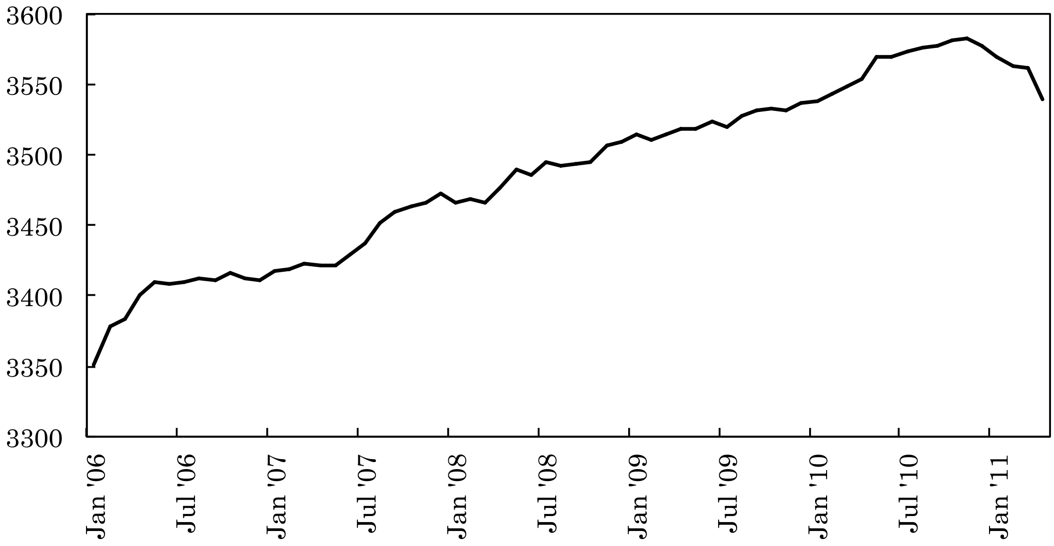


Figure A.1: Number of sample sites (January 2006 – April 2011).

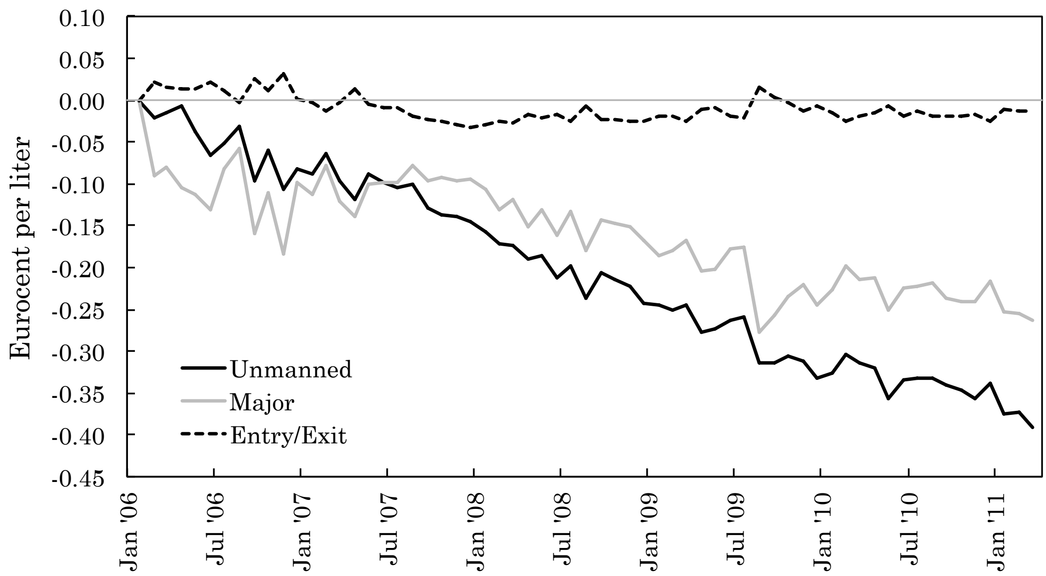
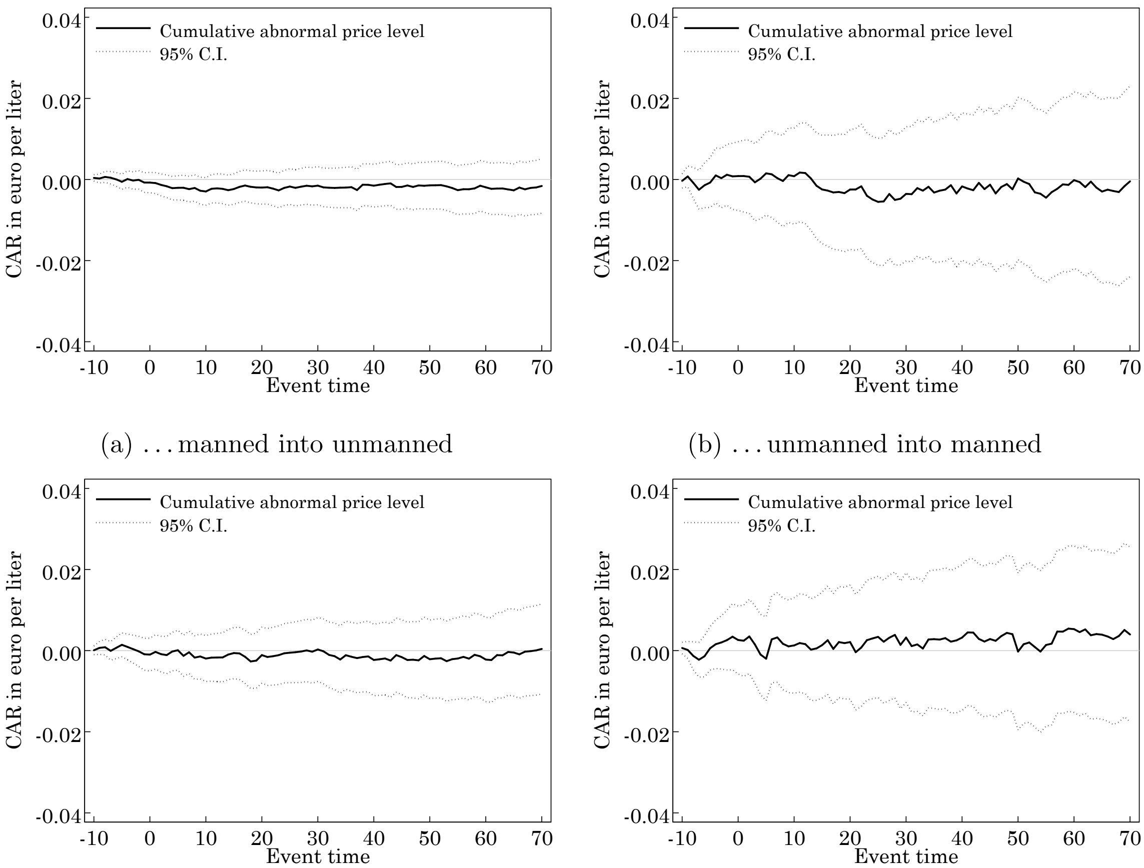


Figure A.2: Development of subtotal effects of different factors on off-highway prices (January 2006 – March 2011).

Note: The variable "major" includes the four traditional majors plus TOTAL and Q8.



(a) . . . manned into unmanned

(b) . . . unmanned into manned

(c) . . . major into non-major

(d) . . . non-major into major

Figure A.3: Spillover price effects of neighboring off-highway station converting from . . .

Note: All figures for event day -10 to event day 70.

Table A.1: Market shares in number of stations, Catalist and Athlon data.

	Catalist		Athlon		Data comparison per 30/6/2010 Market share		% stations
	2005	06/2010	31/12/2005	30/06/2010	Catalist	Athlon	included
ANWB	**23**	**0**	**28**	**31**	**0.0%**	**0.9%**	**n.a.**
Argos	**28**	**51**	**26**	**44**	**1.2%**	**1.2%**	**86%**
Autofood	**n.a.**	**46**	**37**	**37**	**1.1%**	**1.0%**	**80%**
Avia	249	240	144	178	5.7%	5.0%	74%
Avia Express	n.a.	n.a.	0	12	n.a.	0.3%	n.a.
Avia	**249**	**240**	**144**	**190**	**5.7%**	**5.3%**	**79%**
BOL	**n.a.**	**n.a.**	**11**	**13**	**n.a.**	**0.4%**	**n.a.**
BP	393	376	330	312	8.9%	8.7%	83%
BP Express	n.a.	n.a.	0	9	n.a.	0.3%	n.a.
BP	**393**	**376**	**330**	**321**	**8.9%**	**9.0%**	**85%**
Brand Oil	38	30	34	31	0.7%	0.9%	103%
Amigo	17	32	11	23	0.8%	0.6%	72%
Brand Oil	**55**	**62**	**45**	**54**	**1.5%**	**1.5%**	**87%**
De Fakkel	**n.a.**	**n.a.**	**9**	**7**	**n.a.**	**0.2%**	**n.a.**
Easy Fill	**n.a.**	**n.a.**	**4**	**3**	**n.a.**	**0.1%**	**n.a.**
Esso	360	335	313	306	8.0%	8.6%	91%
Esso Express	n.a.	17	9	18	0.4%	0.5%	106%
Esso	**360**	**352**	**322**	**324**	**8.4%**	**9.1%**	**92%**
Gulf	168	130	140	112	3.1%	3.1%	86%
TinQ	64	232	70	196	5.5%	5.5%	84%
BIM	43	5	37	12	0.1%	0.3%	240%
Gulf	**275**	**367**	**247**	**320**	**8.7%**	**9.0%**	**87%**
De Haan	**28**	**33**	**27**	**32**	**0.8%**	**0.9%**	**97%**
PicoBello	**n.a.**	**n.a.**	**12**	**9**	**n.a.**	**0.3%**	**n.a.**
Q8	165	144	117	100	3.4%	2.8%	69%
Tango	79	118	74	118	2.8%	3.3%	100%
Q8 Easy	n.a.	n.a.	7	4	n.a.	0.1%	n.a.
Q8	**244**	**262**	**198**	**222**	**6.2%**	**6.2%**	**85%**
Shell	590	470	522	459	11.2%	12.8%	98%
Shell Express	34	67	42	67	1.6%	1.9%	100%
Shell	**624**	**537**	**564**	**526**	**12.8%**	**14.7%**	**98%**
Tamoil	**79**	**149**	**89**	**144**	**3.5%**	**4.0%**	**97%**
TOTAL	584	439	451	381	10.4%	10.7%	87%
Elan	n.a.	n.a.	25	43	n.a.	1.2%	n.a.
FINA	n.a.	n.a.	10	4	n.a.	0.1%	n.a.
TOTAL	**584**	**439**	**486**	**428**	**10.4%**	**12.0%**	**97%**
Texaco	543	476	427	395	11.3%	11.0%	83%
Firezone	24	77	5	80	1.8%	2.2%	104%
Texaco	**567**	**553**	**432**	**475**	**13.1%**	**13.3%**	**86%**
TankS	**n.a.**	**n.a.**	**16**	**19**	**n.a.**	**0.5%**	**n.a.**
other	**761**	**740**	**326**	**376**	**17.6%**	**10.5%**	**51%**
total	**4270**	**4207**	**3353**	**3575**			**85%**

Notes: 1. Our market share of ANWB stations in 2010 exceeds the market share as reported by Catalist. The reason is that Catalist considers ownership, we instead consider the brand name shown on the site's premises: Gulf bought the ANWB stations in 2006 but re-branding only started in 2009; 2. Marees is counted as AVIA since all Marees stations were taken over by AVIA at the end of 2005. Rebranding of these sites however took place at a later - unknown - date.

Table A.2: Abnormal returns (in eurocents) for different event-types.

Event	manned to unmanned			non-major to major			major to non-major		
Event day	AR	s.e.	θ_1	AR	s.e.	θ_1	AR	s.e.	θ_1
-10	-0.138	0.123	1.119	-0.052	0.116	0.446	-0.199**	0.099	2.015
-9	-0.278**	0.117	2.382	-0.096	0.117	0.819	0.063	0.093	0.672
-8	-0.121	0.147	0.827	0.021	0.134	0.155	-0.212**	0.106	1.998
-7	0.172	0.131	1.313	0.000	0.109	0.002	-0.027	0.087	0.312
-6	-0.099	0.148	0.668	-0.095	0.122	0.774	0.017	0.101	0.168
-5	0.025	0.117	0.213	-0.002	0.119	0.013	-0.091	0.108	0.847
-4	0.158	0.179	0.883	0.233**	0.115	2.028	0.025	0.103	0.246
-3	0.197	0.132	1.496	0.034	0.108	0.314	-0.326***	0.112	2.899
-2	0.010	0.148	0.070	-0.193	0.123	1.573	0.175	0.114	1.533
-1	-0.293	0.182	1.611	-0.055	0.139	0.395	-0.162	0.126	1.283
0	-2.652***	0.096	27.711	-0.618***	0.106	5.809	-1.449***	0.075	19.388
1	-0.106	0.109	0.972	0.089	0.120	0.744	0.086	0.091	0.950
2	0.108	0.120	0.898	-0.031	0.123	0.251	-0.001	0.101	0.006
3	0.070	0.097	0.714	-0.020	0.116	0.175	-0.056	0.097	0.580
4	-0.117	0.115	1.015	0.017	0.144	0.120	0.070	0.087	0.809
5	-0.011	0.109	0.102	0.314**	0.123	2.561	-0.052	0.094	0.549
6	0.109	0.108	1.010	-0.133	0.118	1.127	-0.174*	0.090	1.941
7	0.617***	0.113	5.466	-0.084	0.108	0.771	0.055	0.088	0.620
8	0.034	0.116	0.290	0.151	0.117	1.292	-0.315***	0.086	3.677
9	-0.220**	0.111	1.991	0.066	0.130	0.504	0.224**	0.091	2.451
10	-0.188	0.120	1.567	0.108	0.119	0.907	-0.067	0.097	0.693
11	-0.034	0.116	0.294	-0.203*	0.108	1.875	-0.339***	0.093	3.655
12	-0.053	0.104	0.510	0.077	0.128	0.601	0.146	0.091	1.600
13	0.139	0.096	1.443	-0.014	0.116	0.118	0.145	0.103	1.406
14	-0.145	0.156	0.928	0.077	0.130	0.596	-0.174**	0.084	2.068
15	0.062	0.169	0.370	0.063	0.134	0.474	0.108	0.090	1.208
16	-0.086	0.099	0.870	-0.024	0.117	0.202	-0.099	0.090	1.104
17	0.120	0.103	1.167	-0.195*	0.117	1.664	-0.063	0.099	0.636
18	-0.142	0.135	1.055	-0.009	0.111	0.084	0.004	0.098	0.043
19	0.182	0.150	1.215	0.160	0.121	1.325	-0.193**	0.092	2.110
20	0.097	0.144	0.674	0.008	0.121	0.065	-0.177*	0.093	1.911
21	-0.079	0.115	0.686	0.316**	0.132	2.396	0.131	0.086	1.523
22	0.065	0.096	0.681	-0.139	0.117	1.187	-0.225**	0.093	2.429
23	-0.270***	0.105	2.581	0.019	0.120	0.163	0.111	0.093	1.198
24	0.099	0.110	0.895	-0.036	0.132	0.276	-0.130	0.092	1.423
25	-0.045	0.126	0.360	-0.102	0.132	0.778	-0.076	0.094	0.808
26	0.030	0.162	0.183	-0.154	0.131	1.172	-0.054	0.098	0.551
27	0.163	0.113	1.442	-0.004	0.124	0.030	0.187*	0.099	1.884
28	-0.071	0.116	0.609	0.039	0.121	0.320	-0.181*	0.100	1.813
29	-0.139	0.104	1.339	0.305***	0.106	2.889	-0.016	0.101	0.163
30	-0.027	0.098	0.273	-0.150	0.121	1.241	0.101	0.091	1.108
31	-0.107	0.119	0.903	-0.121	0.119	1.015	0.088	0.093	0.951
32	-0.042	0.109	0.387	0.071	0.126	0.566	0.031	0.095	0.324
33	0.111	0.127	0.876	0.061	0.120	0.506	-0.192**	0.085	2.252
34	0.111	0.096	1.148	-0.014	0.132	0.110	-0.169*	0.102	1.657
35	0.021	0.111	0.186	0.400***	0.125	3.191	-0.143	0.092	1.560
36	0.115	0.116	0.989	-0.467***	0.116	4.014	0.012	0.096	0.122
37	-0.353***	0.107	3.303	-0.112	0.118	0.948	0.190*	0.113	1.684
38	0.390***	0.106	3.677	0.307**	0.135	2.284	0.011	0.091	0.119
39	0.053	0.100	0.533	-0.244*	0.142	1.717	-0.128	0.093	1.386
40	0.045	0.115	0.392	-0.010	0.123	0.081	-0.205**	0.098	2.100

Note: AR is the sample average abnormal return of given type of event at the indicated event day.
*** p<0.01, ** p<0.05, * p<0.1