VILNIAUS UNIVERSITETAS

MATEMATIKOS IR INFORMATIKOS FAKULTETAS

Magistro darbas

ERDVINĖ KAINŲ KONKURENCIJA MAŽMENINĖSE VOKIETIJOS DEGALŲ RINKOSE

SPATIAL PRICE COMPETITION IN GERMAN RETAIL GASOLINE MARKETS

Mykolas Šuminas

VILNIUS 2018

Darbo vadovas   prof. Remigijus Leipus

Darbo recenzentas   dr. Dmitrij Celov

Darbas apgintas __2018 m. sausio 15 d.___
Darbas įvertintas ______________________

Registravimo Nr. 111000-9.1-5/________________

Summary

This paper investigates the globality of German retail gasoline market by studying price reaction functions of neighbouring gasoline stations. The data under analysis originates from a very detailed Germany gasoline station price database for the year 2015. Models of reaction functions were estimated using Bertrand-Hotelling framework; the spatial component was introduced via a weighing matrix. Three main methods were employed to construct the weighing matrix and thus examine the nature of spatial competition: nearest neighbours, border neighbours and Euclidean distance neighbours. The results strongly point to local competition as local models describe the reaction functions most accurately. Furthermore, temporal reaction functions (containing lagged independent variables) were investigated and persistence in price forming behaviour was found. Lastly, hypothesis of vertical (brand level) price competition was rejected.

Keywords: price competition, spatial model, fuel prices, monopolistic competition

Santrauka

Šis straipsnis tiria Vokietijos benzino mažmeninės rinką pasitelkdamas gretimai esančių degalinių kainų reakcijos funkcijų analizę. Nagrinėjami 2015-ųjų metų duomenys paimti iš labai detalios Vokietijos degalinių kainų bazės. Reakcijos funkcijų modeliai buvo įvertinti remiantis Benrand-Hotelling teoriniu modeliu, o erdvinė komponentė į modelius inkorporuota pasitelkiant svorių matricą. Svorių matricai apibrėžti buvo naudoti trys pagrindiniai metodai: artimiausių kaimynų, kraštinių kaimynų ir Euklidinio atstumo kaimynų. Modelių rezultatai pritaria lokaliai konkurencijai, nes lokalūs modeliai geriau aprašo kainų formavimo lygtis nei globalūs. Be to buvo įvertintos reakcijos funkcijos su nepriklausomų kintamųjų vėliniais. Šis metodas atskleidė, kad egzistuoja kainų formavimo nelankstumas. Galiausiai buvo atmesta vertikalios (prekinių ženklų lygio) konkurencijos hipotezė.

Raktiniai žodžiai: kainų konkurencija, erdvinis modelis, degalų kainos, monopolistinė konkurencija

CONTENTS

Introduction

It is not unusual to observe two gasoline stations on opposite sides of a road, literally a few meters apart, charging different prices. While gasoline station cohesion is a direct consequence of Hotelling (1929) model[1] the price difference is more of a perplexing occurrence. Gasoline and other types of fuel are a regulated type of good meaning they are homogeneous[2] not by firm choice. As fuel is a not a differentiated good per se its nonhomogeneity arises from location and transportation costs. The gasoline stations are specialized outlets that sell only a limited variety of goods in besides fuel (if any at all). Consequently visiting a gasoline station further away than the closest to the individual is costly. It requires not only the consumer's time but also knowledge on the pricing patterns of gasoline stations[3]. It is natural to assume that gasoline stations are selling a differentiated product and can adjust their profits by changing prices thus capturing clientele from neighbouring outlets.

It is easy to realize the topicality of gasoline price: numerous articles in the media debate on the topics of the main gasoline price determinants: oil and excise taxes, the rocket and feather hypothesis and whether collusive agreements exists among major market players. The attention given to the retail gasoline industry is not uncalled for. For example in the United States in 2012 average yearly household spend on gasoline stood at 2,912 dollars which amounted to 4% of average income before taxes according to Energy Information Administration's estimates[4]. Furthermore, the nature that gasoline prices are presented to the consumer is rather different that most of other goods. The gasoline prices are shown on billboards for commuters to see, but what is more important they can be altered seemingly any time and in a split second which is not usually not the case for most goods with a few notable exceptions being airline tickets, other modes of transportation (like Uber) and goods traded in exchanges. This type 'dynamic' pricing based on supply and demand is spreading mainly with regard to technological advances while gasoline stations have been implementing such methods for quite a while so analysis of gasoline prices can be an insight of how companies could compete in the future.

Another interesting aspect of the of the retail gasoline market, namely in Germany, is the enforcement by Bundeskartellamt of full price disclosure. The collected prices are later on used by software firms to build tools for easier location of better prices. The fact that prices can be tracked with such accuracy entails that the menu costs (competitor price tracing costs) for the firm are greatly reduced as well as search costs for the consumer. Naturally, the assumption of rationality and full utilization of all available data might be too far-fetched as not all consumers will use the available tools to get the information. However, some research already exists comparing the price levels before and after mandatory price reporting: Dewenter, Heimeshoff, and Lüth (2016) find that gasoline prices have actually increased.

In addition to spatial differentiation gasoline stations try to distinguish themselves by using

[1]see Netz and Taylor for a treatment on spatial differentiation

[2]Of course companies try to differentiate themselves by selling branded gasoline with special additives

[3]It is known from Diamond (1971) model that this lack of information and additional search costs can drive the prices away from perfect competition equilibrium

[4]https://www.eia.gov/todayinenergy/detail.php?id=9831

other methods be in order to be perceived of either higher quality or to provide a bundled good option for the consumer. Although spatial competition is at the core of this analysis of this paper other features of gas stations can significantly affect both their price setting mechanism and the level of prices (Pennerstorfer (2009)).

The rest of the paper organized as follows: Section 1 discusses the relevant literature of price and spatial competition and empirical literature of gasoline price competition. Next Section 2 describes the models and estimation methods used in the paper. Three types of spatial weighing matrices are presented in this section. In section 3 some important facts about the data used in the modeling process are highlighted. Section 4 presents and discusses modeling results and finally Section 5 concludes.

1. Literature review

Gasoline retailing sector frequently draws attention of the academia. Eckert (2013) finds that more than 70 papers have been published in the 2000-2011. The attention given to the topic is understandable for the reasons mentioned in the introduction, but the popularity of such scientific articles also stems from the relative ease of obtaining relevant data for such analysis. The most common feature of gasoline retail market is the analysis of fuel as homogeneous good that is spatially differentiated. A simplifying assumption is often made that oil price or gasoline wholesale prices are taken as unvarying cost determinant across gasoline stations. Although in reality the delivery costs of gasoline do vary as unbranded and branded varieties of gasoline can be purchased from the rack. Moreover, the way that upstream gasoline market operates might be very different: wholesale gasoline transportation could be dominated by refiners themselves or exhibit a more competitive type of market structure and be delivered by jobbers (independent contractors).

Price competition analysis extends as far as Bertrand's (1883) original model of price competition which is the basis for all the models discussed in this thesis. Bertrand's original model is based on the assumptions of uniform, frictionless market (no search, transaction, or fixed costs) of homogeneous goods. It is well known if any of the somewhat unrealistic assumptions are relaxed a richer variety of markets can be described. A theoretical spatial framework that is still employed to this day comes from Hotelling (1929) in which spatial components are introduced: consumers are uniformly distributed on a unit line. Then the consumer maximizes his indirect utility $v(x,p) = r - \tau|l_i - x| - p_i$. Here r is the reservation price, l_i the product's location, x the consumer's location and p_i the price. The consumer buys one unit of the product from the seller which maximizes his utility. $\tau|l_i - x|$ are the transportation or costs where τ can be interpreted as the aversion to make the needed transportation - it shows how costly it is to search for a better price so it is closely related with search costs. A lower τ will characterizes a consumer who is willing to compare more seller prices and wander off further from his original location for a better deal. This model is richer then first meets the eye as the spatial component can be interpreted as general 'similarity' of goods.

The types of models described above are often employed for retail gasoline price modeling because gasoline can largely be considered a homogeneous good at least when it comes to its chemical properties. The spatial component can be interpreted directly - gasoline stations are selling a homogeneous good but they are differentiated spatially. Another important aspect of the models is type of competition, namely, the number of competitors. In the models of Hotelling's linear city, Salop's circular city (1979) or Gabszewicz and Thisse (1979) vertical competition a single firm faces only two competitors that are in the closest vicinity of the firm. This means that firms that are further neighbours have no effect on the firms quantities or prices. These models are a type of (very) local competition that can be contrasted with models of Chamberlain (1933) and Dixit and Stiglitz (1977) that state that all of the firms are direct and global competitors. Finally, models of differentiated products could be seen as a middle ground of these models.

Differentiation comes into play both as spatial (horizontal) and non-spatial (vertical). Spatial differentiation is a property of the gasoline station location while other types of differentiation is described are determined by the gasoline station brand and extra amenities that might be available at a gas station. Often researchers tackle these properties separately, but often their interplay can lead to interesting findings Clemenz and Gugler (2006).

The question of competition globality is often addressed in the literature. Hotelling's model might be viewed as simple and restrictive, but empirical evidence (Benson, Faminow, and Fik (1992)) suggests, that often only the nearest neighbours affect each other prices the prices. Seminal work by Pinkse, Slade, and Brett (2002) who analyze the competition of the wholesale gasoline market by employing various measures of spatial competition: from competing with only with the closest neighbour to completely global competition which allows for asymmetric effects. Authors conclude that competition is indeed local, in fact more local than Hotelling's original model as neighbours which the firm shares their market boundary are the principal determinants of their competitor prices. Furthermore, authors recognize that the competition between neighbours is described that is a restrictive way (by linear reaction functions) so they employ a semiparametric approach to estimate competition.

Generally, there is a lack of information how consumers purchase gasoline i.e. how far is consumer travels or willing to travel for cheaper gasoline. Because of this researchers often employ some ad-hoc cut-off ranging from 0.5 miles (Atkinson, Eckert, and West (2009)) to 15.5 kilometers (Pennerstorfer (2009)). The chosen distance depends on the type of location analyzed, the former being in Los Angeles area and the latter is set in Lower Austria which is a predominantly rural area. Alderighi and Baudino (2015) analyze Cuneo (an Italian town of more than 50,000 inhabitants) gasoline market find that spatial dependence is best described with a cut-off of 1.1 km. Authors describe the Italian retail gasoline market as having an abundance of gas stations in fact the concentration for 10,000 inhabitants of stations is double that of the German market.

Furthermore, the retail gasoline market is often characterized by a few large firms (which are often are large vertically integrated companies) and a lot of smaller companies which could also be vertically integrated, but their station network is small Haucap, Heimeshoff, and Siekmann (2015). Siekmann groups the German stations into three groups: Aral (BP), Shell, Total, Esso (ExxonMobil), and Jet (ConocoPhilipps) are called 'Oligopolists' - they have nationwide gasoline networks and are vertically integrated. The second group consists of Star (Orlen), Agip (ENI), HEM (Tamoil), and OMV which have smaller networks and also direct access to refineries. Lastly, there are small brands and unbranded gasoline stations or so-called 'Independents'. This type of vertical differentiation is also important when analyzing spatial dependence: it is found that unbranded stations typically charge a lower price Hastings (2004)). Furthermore, unbranded (or independent) stations are the ones who are not only charging lower prices but also enforce lower prices in their proximity (Hastings (2004)) . This raises an interesting question: are independents and branded stations really competitors if consumers perceive them as differenciated goods in spite of gasoline's chemical composition being the same? Pennerstorfer claims that there are actually two effects of differentiation: he acknowledges that unbranded gasoline station prices are cheaper (Competition effect), however the presence of independent gasoline stations reduces the tension in the higher level of gasoline stations due to the lower number of direct competitors (Composition effect). The sum of the mentioned effects is what drives the prices. Pennerstorfer (2009) finds that independents have only a small effect on the majors (branded stations) supporting the claims that independents are viewed as inferior by consumers, thus not competing directly with majors. Although a lot of contrary evidence exists Hastings (2004), Meerbeeck (2003), Clemenz and Gugler (2006).

2. Models, methods and estimation

This section discusses for the model that will be used for spatial analysis. As mentioned in the literature review the main family of models that will be employed in this paper come from Bertrand's price competition and Hotelling's linear city. The nature of address type models is interpreted quite directly in this paper. In the last part of this section some extensions to vertical competition between major and lesser brands is introduced.

For the basic model let's say there are n firms and each firm maximizes its profits by choosing a price $p = (p_1, p_2, ..., p_n)^T$ for its differentiated good and selling the equilibrium quantity $q = (q_1, q_2, ..., q_n)^T$ in the market:

$$max\pi_i = (p_i - c_i)q_i(p_i, p_{-i}), \tag{1}$$

The quantity (demand) is a function of own price (p_i) as well as the competitor prices $p_{-i} = (p_1,...,p_{i-1},p_{i+1},...,p_n)$ and other factors that affect supply and demand (income, population density) company strategy (such as margin requirements) are all housed in X matrix.

$$q_i(p_i, p_{-i}) = \alpha_i + \gamma_{ii}p_{ii} + \sum_{i \neq j}^{n} \gamma_{ij}p_j + X\beta \tag{2}$$

Solving the problem with respect to the i-th firm's first order conditions:

$$p_i = R(p_{-i}) = -\frac{1}{2\gamma_{ii}}(\alpha_i - \gamma_{ii}c_i + \sum_{i \neq j} \gamma_{ii}p_{ij} - X\beta) \tag{3}$$

The γ parameters are price elasticicities and $-\frac{\gamma_{ij}}{\gamma_{ii}}$ is called the diversion ratio. This ratio quantifies the share of customers i-th firm lost to j-th firm due to changes in in i-th firms price raise. It is natural to assume that diversion ratios to be related with the proximity of competing gasoline stations. In order to generalize the model rewrite it in matrix notation as system of equations where G is a weighing matrix which has zero entries in the off diagonal (g_{ii}) and non-negative entries everywhere else:

$$p = R(p) = A + X\beta + Gp + u \tag{4}$$

Matrix G elements g_{ij} are assumed to be linearly proportionate to Euclidian distance as in Taylor, Kreisle, and Zimmerman (2010), Kalnins (2003), Kim and Lee (2014) whereas Pinkse, Slade, and Brett (2002) estimate $g(.)$ nonparametrically.

First, the conventional methods of model set-up and estimation will be discussed. G matrix can be parametrized in advance. If errors are assumed to be iid $u_i \sim N(0, \sigma^2)$ and G is consists of weighing matrix $\mathcal{W}$ and scaling by a parameter ρ ($G = \rho\mathcal{W}$) then the loglikelihood function is:

$$l = -\frac{N}{2} ln(2\pi\sigma^2) + |I_n - \rho\mathcal{W}| - \frac{1}{2\sigma^2}(p - \rho\mathcal{W}p - X\beta)^T(p - \rho\mathcal{W}p - X\beta) \tag{5}$$

The function when maximized with respect yields parameters β, ρ, σ, futhermore the maximization can be simplified by noting that:

$$|I_n - \rho\mathcal{W}| = \prod_j (1 - \rho\lambda_j) \tag{6}$$

Where λ are the eigenvalues of the $\mathcal{W}$ matrix.

The weighing matrix can be constructed in many ways, however there is a need to restrict it due to the full matrix having $n^2 - n$ parameters and a symmetrical one having $(n^2 - n)/2$ parameters. Although the latter approach might be feasible, it is not very informative and fails to provide any generalizable insights to the model. Three main approaches of $\mathcal{W}$ restriction are implemented in this paper: by common boundary, nearest neigbhours and Euclidian distance.

The first approach is used by Case (1991) who composes the $\mathcal{W}$ matrix by attributing to 1 to neighbours who share a common boundary and 0 otherwise. To establish a common boundary the space must be partitioned and attributed to each individual gasoline station. There are numerous ways to accomplish the space partitioning, but perhaps the most natural for this problem is Voronoi's method. Voronoi partitioning assigns points in space to prespecified elements in such a way that that a point is assigned to the closest element by a given metric. Defining station territories in this manner assumes that individuals would buy gasoline from the station closest to their location. The assumption might seem strong as individuals would tend to buy from the station that offers the best combination of price and logistics that maximizes his utility, however price differences are not detrimental to common boundary neighbours definition. The common boundary neighbours are assumed to have a uniform effect on a gasoline station price. Furthermore, the common boundary method can help to investigate the competition remoteness - by analyzing effects of second order neighbours i.e. stations that are neighbours of neighbours of the original gasoline station, but are not first order neighbours themselves (see Figure 1 for an illustration). In short, using first and second order neighbours lets account for indirect competition which is not as easily discerned by nearest neighbour or Euclidean distance. Furthermore, the contiguity based metric takes into account the density of population and gasoline station network implicitly: stations will be more dispersed in a less urbanized area, but will still remain neighbours regardless of distance.

The second approach is to assign 1 to $\mathcal{W}_{ij}$ if stations are nearest neighbours and zero otherwise. This approach can extended to N-nearest neighbours if 1 are attributed so second, third neighbours and so on. The method is less restrictive as different neighbours can have varied effects on the price yet it introduces more variables (which at the very least are somewhat correlated) and parameters in the model.

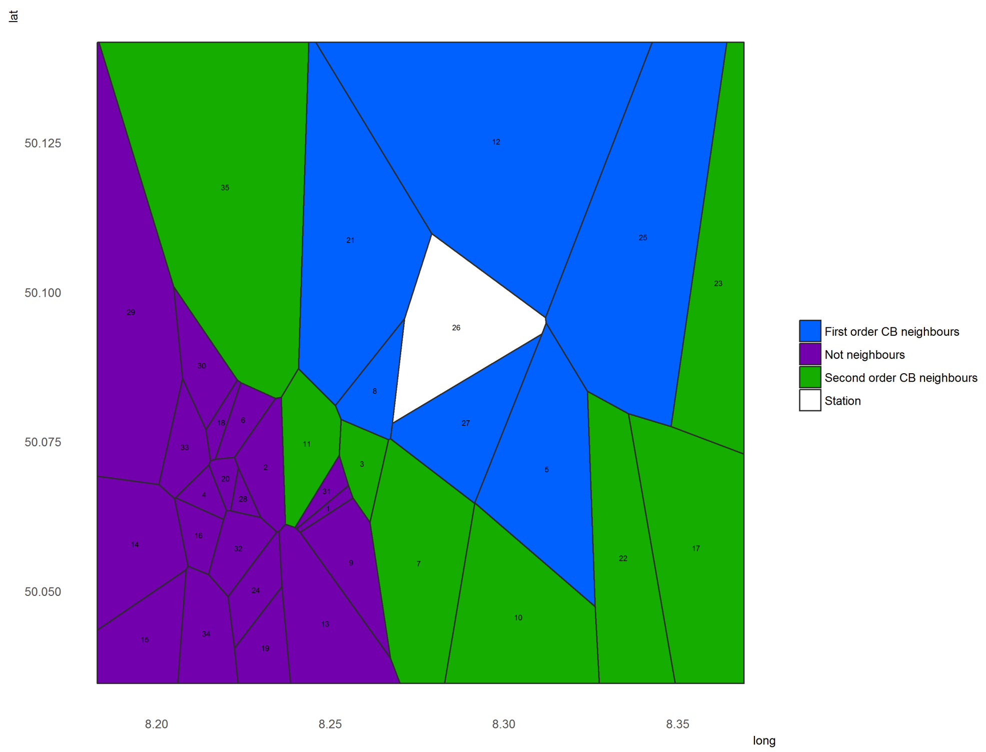

Figure 1. Voronoi partitioning, first and second order neighbours

The third method involves designing $\mathcal{W}_{ij}$ by utilizing Euclidean distances directly. First of all distances (d_{ij}) between gasoline stations in a city are calculated and D matrix is obtained. Then these distances are inverted to obtain weights for the matrix $\mathcal{W}_{ij}$:

$$\bar{w}_{ij} = \frac{1}{d_{ij}}, w_{ij} = \left(\frac{\bar{w}_{ij}}{\sum_{j=1}^{n} \bar{w}_{ij}}\right) \tag{7}$$

The above approach is very global, however, meaning that it is somewhat unrealistic to expect that stations very far away from i can be of even minute importance to i station prices. The issue can be alleviated by selecting a cut-off threshold d^{TH} which, if surpassed, results in a zero weight for the $\mathcal{W}_{ij}$ matrix:

$$w_{ij} = \begin{cases} \left(\frac{\bar{w}_{ij}}{\sum_{j=1}^{n} \bar{w}_{ij}}\right), & \text{if } d_{ij} \leq d^{TH} \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

This method seems to have the best of both worlds: the weighing matrix is restricted to nearest neighbours, but it also leaves some flexibility in the weights so they represent relative distance between stations. Row normalization also helps with the interpretation with ρ. If the $\mathcal{W}$ is not row normalized the weighed price is affected by not only competitor prices and the distances to competitor locations but also to the actual number of competitors. Consequently if stations have a lot of competitors in their threshold vicinity, the final price effect will be inflated due to more summands in the $\mathcal{W}p$ product.

Furthermore, there need not be a single matrix that describes price competition. As mentioned in the literature review competition may occur terms vertical as well as horizontal competition. Using station differentiation by Haucap, Heimeshoff, and Siekmann (2016) gasoline stations are grouped into 'Oligopolists' and 'Independents'. Following Pennerstorfer (2009) the weighing matrix $\mathcal{W}$ is also split to $\mathcal{W}^O$ and $\mathcal{W}^I$ which are composed to distances to gasoline stations owned by 'Oligopolists' ($\mathcal{W}^O$) and 'Independents' ($\mathcal{W}^I$) respectively. Then equation 4 can be rewritten as:

$$p = R(p) = A + X\beta + \rho_1 \mathcal{W}^O p + \rho_2 \mathcal{W}^I p + u \tag{9}$$

where $\mathcal{W}^O + \mathcal{W}^I = \mathcal{W}$

As gasoline station prices are jointly determined in an imperfect competition setting they are endogenous so appropriate instruments must be used to account for endogeneity. These variables have to vary by location, however. There are two ways to instrument prices: by using own stations characteristics (x_i Pinkse, Slade, and Brett (2002)) or use competitor station characteristics x_j (Berry, Levinsohn, and Pakes (1995)). I follow Slade et al. approach using own station location specific instruments for p_i. Furthermore all gasoline stations share a common price determinant - oil price or more accurately wholesale gasoline prices. To eliminate the common cause for prices the gasoline prices are also regressed on wholesale gasoline prices (p^c) to obtain filtered gasoline prices p^f (or otherwise the residuals):

$$\begin{aligned} p_t &= \alpha + \eta p_t^c + \epsilon_t \\ p_t^f &\equiv \epsilon_t \end{aligned} \tag{10}$$

The filtered prices are a convenient way of carrying out further modeling - they are orthogonal to the common price determinants and by regressing each station separately can accommodate for profit margin strategies on an individual station level. Additionally, any regression using filtered prices, as a result of Frisch–Waugh–Lovell theorem, provide the same estimates as on unorthogonalized prices.

To summarise, the estimation starts with price filtering of individual stations, then the prices are instrumented on location specific variables, lastly final regressions are run with station specific characteristics as controls.

3. Data

This section briefly describes the data that was available in order to carry out the analysis with some highlights on the stylized facts. The primary data of interest are gasoline station prices made available by the German Bundeskartellamt (State Cartel Office) legislation. Price data is collected for nearly 15,000 gasoline stations. The main dataset on gasoline price data and gasoline station locations that is used in this paper comes from Bottled Software company which shared the data with the author. There are three types of fuel that are tracked in the stations: Super 5, Super 10 and Diesel. The two former fuel are gasoline types with the number indicating of how much ethanol as percentage is added to them. I restrict the analysis to Super 5 prices only. Very comprehensive gasoline price data is collected - each price change is registered, and on average 5 price adjustments made per day are made. The variable *p* that will is referenced in this paper is the mean price in the respective day. The whole year of 2015 is in used meaning there are 364 observations for each individual stations. Lastly, to properly reveal the complexity of the spatial price competition the analysis focuses on the stations in larger towns or cities, so any settlements that have less than 20 gasoline stations are omitted. Most of the station networks are small, however (Table 1).

Table 1. Station network size distribution

	Bins	Number of tows	Share, %
1	[20, 30)	24	39.34
2	[30, 40)	17	27.87
3	[40, 50)	3	4.92
4	[50, 60)	5	8.20
5	[60, 70)	5	8.20
6	[70,287]	7	11.48

Aside for the price data, station location is used in the analysis. Gasoline stations on average are spaced 9 kilometers apart (with 7.8 being the median distance). However, this statistic is skewed due to the large cities included in the sample. According to most of the literature discussed the gasoline markets tend to be very localized so it is more rational to assess how many gasoline stations are in a circle of X radius of a typical station 2 or the median distance to the nearest station.

Table 2. Average number of stations within selected radii

1 km	2 km	3 km	5 km
0.69	2.11	3.79	6.62

Another important factor is how far away are the N nearest neighbours of a typical gasoline station (Table 3). As seen from 4 almost all (around 94%) gasoline stations have a neighbouring station in 3km radius. To illustrate how the station networks looks like in a major German city Dresden's station network is visualized in Figure 2 in which stations are connected if they are within 3km radius of each other. A fuller Google map view is presented in the Appendix 1. This type of

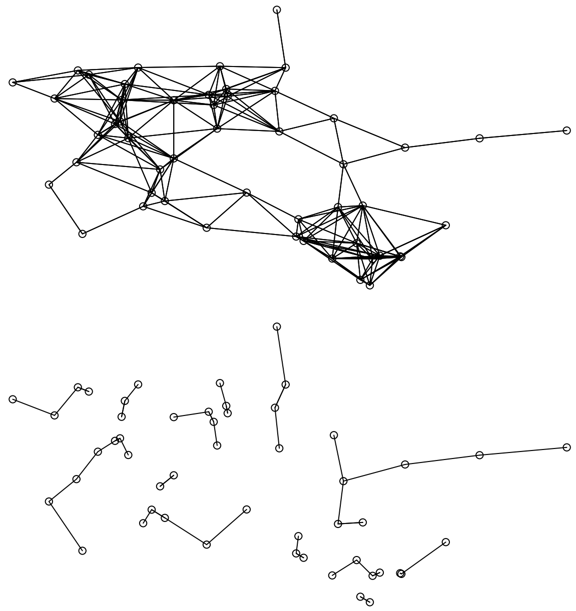

Figure 2. Dresden gasoline 'markets': upper graph 3 km threshold market,
lower graph nearest neighbour market

station network connection can be use to answer one of the most important questions in economics: 'What constitutes a market?'. However even with a relatively small radius of 3km gasoline markets in a densely populated city such as Dresden where a lot of gasoline stations are clustered in a small area are very complex and overlap considerably. Compare it with a most restrictive case of the market definition - one nearest neighbour (1-NN) Figure and the markets are noncummunicating and fail to demonstrate city-wide contingency and have minimal overlap. It is very practical to think of market in terms of how many nearest neighbours or how many kilometers the market covers. The price manager can easily make informed decisions by looking at just a narrow portion of gasoline stations in the city.

Table 3. Median distance to n-th nearest neighbour

1st neighbour	2nd neighbour	3rd neighbour	4th neighbour	5th neighbour
0.66	1.13	1.48	1.77	1.99

In the common boundary context a little more than 5 direct competitors are found for a typical gasoline station and almost 10 gasoline stations can be considered second-order competitors. What inclusion of second order neighbours provides is insight on how consumers are willing go out of their way and avoid several gasoline stations and buy from the one they consider to provide a good deal. In addition to this it shows how far the influence of gasoline price setting can travel within one day. To reflect on these numbers one must look again to the station size distribution (Table 1). Having in mind that a typical town in the sample has at least 20 stations (with 33 being the median value) the common boundary measures are indeed local. Many of the smaller towns do not have a full set of third order neighbours so this metric will not be included in the models.

Table 4. Percentage of nearest neighbours within selected distances

	NN 1	NN 2	NN 3	NN 4	NN 5
1 km	72.64	40.66	18.87	7.88	2.47
2 km	94.15	85.47	74.44	61.79	50.66
3 km	97.97	95.18	91.25	86.11	81.32
5 km	99.47	98.95	98.33	97.28	96.57

The last part of station specific characteristics are station working hours and brand affiliation. Station affiliation (brand name) differentiates the stations into categories of large chains with country-wide presence and vertical integration (level 1 'Oligopolists') smaller chains usually confined within a few administrative regions (level 2 'Oligopolists') and all other gasoline stations which either do not have any brand affiliation or their network is of minor importance. The five firms called by Siekmann level 1 'Oligopolists' make up 57% of the market and the top 10 brands constitute 73% of the market Tables 5 and 6. The firm concentration of course varies city to city, but German retail gasoline market indeed demonstrates features of an oligopolistic market.

Some consumer demand characteristics were obtained: GDP per capita (a proxy for income) of German NUTS3 regions from Eurostat and population density of each square kilometer from NASA's Gridded Population of the World database. Furthermore, driving distances and times are obtained for some of the cities analyzed from www.google.com/maps.

Table 5. Market shares by Siekmann (2016) definitions

Group	share
Independent	30.42
Oligopolist-1	57.01
Oligopolist-2	12.56

It is also beneficial to state what important price determinants are absent from the analysis due to high costs or other circumstances they were unavailable to the author. There is no information on the extra amenities in the gasoline stations and according to Haucap, Heimeshoff, and Siekmann (2015) they are quite important price determinants: 'convenience store is a range of about 0.4 to 1.2 Eurocents/liter, while having a car wash facility, ceteris paribus, is associated with a price increase of close to another 0.2 to 0.4 Eurocents/liter'. What is probably more important that the actual amenity price effects is their ability to interfere with horizontal competition due to additional layer of differentiation introduced. Although population density in income variable will be used from instrumenting gasoline station prices Haucap, Heimeshoff, and Siekmann (2015) finds than traffic intensity and possession of cars are stronger indicators of gasoline demand.

Lastly, due to the fact that some time series analysis will be employed in this paper i.e. series of gasoline prices will be used to estimate temporal reaction functions. It is widely known that refined wholesale gasoline price (and thus oil price) is one of the largest components of variable stations costs. In order to eliminate cost based pricing variation the gasoline station prices are regressed

Table 6. Top brand shares

	Brand name	Station number	Share, %	Group
1	Aral	858	19.60	Oligopolist-1
2	Shell	637	14.55	Oligopolist-1
3	Esso	382	8.73	Oligopolist-1
4	JET	323	7.38	Oligopolist-1
5	Total	296	6.76	Oligopolist-1
6	star	245	5.60	Oligopolist-2
7	bft	140	3.20	Independent
8	HEM	128	2.92	Oligopolist-2
9	Agip	114	2.60	Oligopolist-2
10	SB-Tank	93	2.12	Independent

on wholesale gasoline prices to obtained filtered prices which will later be used in a setting that involves time series modelling. The proxy cost determinants used are Brent oil price and wholesale gasoline traded in New York's stock exchange price. The filtering regression concerns only cost inputs of concurrent periods meaning there are no lagged cost determinants included. This may be looked with suspicion because certainly it takes quite some time for the actual oil or wholesale gasoline to reach the consumer. However the pricing of the local wholesalers (the prices of actual inputs of retail gasoline stations) are dependent on the daily prices of oil. It is common practice that retailers do not know the actual price of input until the end of the month[5] and it is up to the retailer to make price corrections accordingly. In this paper it is assumed that the retailer does not gamble on wholesale prices and adjusts its prices on current input price. Moreover, a constant margin requirement is assumed. From Figure 3 it is apparent that average gasoline price and its cost determinants follow similar patterns. What is more, after applying Granger's causality for the one day lagged [6] gasoline price futures and aggregated gasoline price the p-value is found to be <0.01 which can justify the proposed transformation.

[5]As an example buyers may pay spot prices at the time of order, but at the end of billing period rolling average prices are applied through adjusting accounts

[6]NYSE operates on later time zone that Germany gasoline stations, otherwise it can be assumed that the temporal affect can take one day to manifest

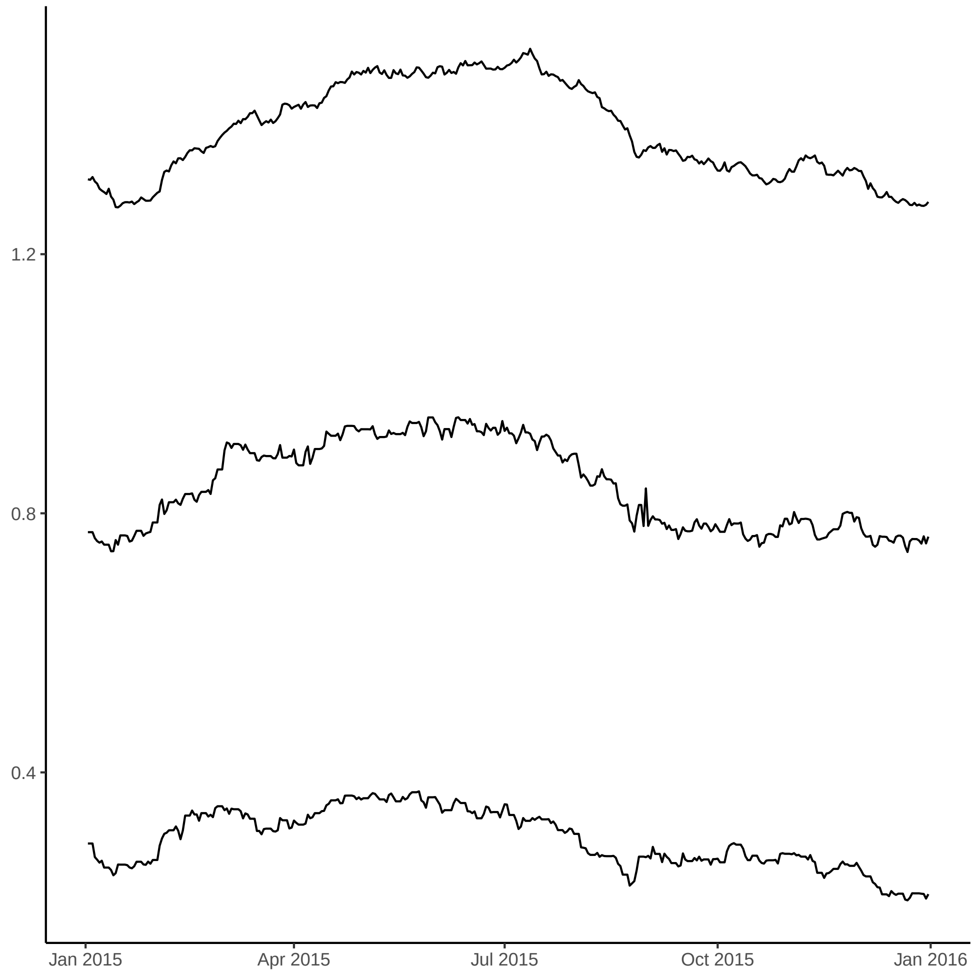

Figure 3. Average gasoline price in German stations (upper graph) and their cost determinants: wholesale gasoline price (middle graph), Brent oil price (bottom graph), euro/l

4. Practical modeling and results

This section outlines the results obtained from models described in Section 2. There are three temporal specifications tested: contemporary prices, one day lagged prices and priced lagged by one week. Due to 4 containing spatially lagged (therefore endogenous) variables they are instrumented on location specific variables (population density and income) and temporally lagged own prices. Moreover city, brand and working hour dummy variables were used as controls. The models were estimated for crossectional data for 364 days of 2015, the coefficients and standard errors presented in the tables are averages of the models. The significance of the models was assessed by making sure that at least 90% of the models have variables that are significant at 10% level. All models presented in the tables had variables with greater than 10% of significance at in at least 90% of the cases (models).

The nearest neighbour approach is employed in two ways. First of all, only the first neighbour is included in the model which is the most common approach used in the literature. Afterwards, a different modeling strategy was used: nearest neighbours were added to until an additional neighbour's price was an insignificant variable in the model. The results presented in Table 7 show that the cumulative effect of 4 or 5 nearest neighbours (this number of neighbour proved to be the most feasible and significant) while separate equations for the models are represented in Equations 11 and 12. The cumulative effect of the neighbour price is as twice as large as one nearest neighbour, however the coefficient of determination hardly increasing with the inclusion of additional neighbours. It begs to question whether inclusion of more neighbours is of any merit and are the results of individual neighbour's prices explainable by their location. In model of lagged (spatially or otherwise) endogenous variable it is usual to assume that the most intimate lags would have the most profound effect on the dependent variable. This assumption is affirmed by looking at Figure 5 where histograms of individual price variables are plotted: the closest neighbours has the largest effect and the remaining nirghbours have diminishing effects on gasoline price. The coefficients fail Kolmogorov-Smirnov and Jarque-Bera normality tests indicating that the process of price setting might be more complex than described with a simple nearest neighbour model. However the empirical distributions closely resemble normal-like histograms so the assumption that the described process is a close approximation to real one is not far-fetched. In summary, the nearest neighbours approach produces quite reasonable result with parsimonious amount of variables in the model.

$$p^F = \underset{(0.0174)}{0.177} p^F_{nn1} + \underset{(0.0177)}{0.098} p^F_{nn2} + \underset{(0.0177)}{0.085} p^F_{nn3} + \underset{(0.0172)}{0.070} p^F_{nn4} \tag{11}$$

$$p^F = \underset{(0.0174)}{0.172} p^F_{nn1} + \underset{(0.0177)}{0.094} p^F_{nn2} + \underset{(0.0177)}{0.080} p^F_{nn3} + \underset{(0.0172)}{0.065} p^F_{nn4} + \underset{(0.0177)}{0.063} p^F_{nn5} \tag{12}$$

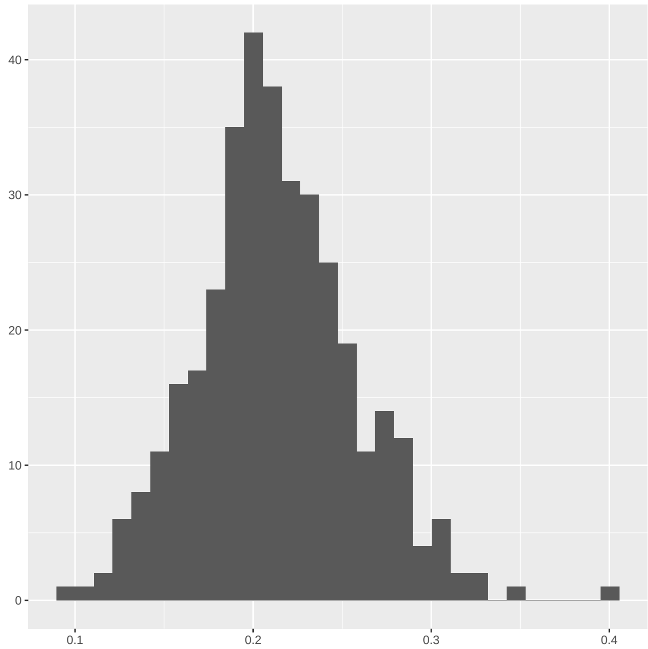

Figure 4. Equation 11 β_1 empirical distribution

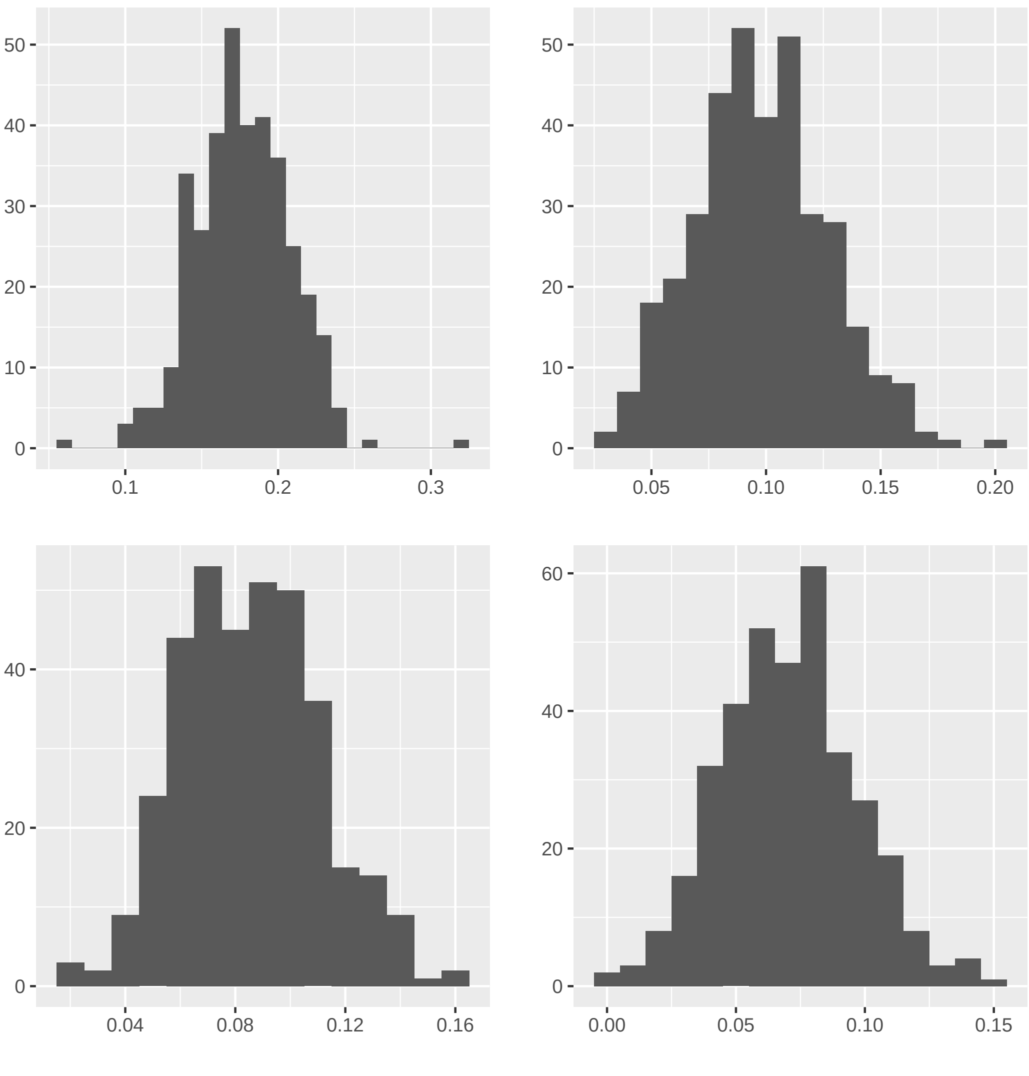

Figure 5. Equation 12 β_i empirical distributions

The second method deals with contiguity measures. In Table 7 second and third rows show the direct and second order boundary neighbour effects. There are on average five first order competitors in the gasoline retail markets so the cumulative effect of direct neighbours is as large as one nearest neighbour. What perhaps drives down the coefficient of determination from 5-NN or even 1-NN is that in this case the restriction that all of the neigbours should possess the same competitive effect magnitude. The second order neighbours (there are 10 of them on average) have an even lesser effect which could be expected. However a larger R^2 measure vis-à-vis the first order neighbours is more puzzling. There could be two plausible explanations of this result. It was already shown that the first nearest neighbour has the strongest effect on rival prices, all other neighbours have slowly diminishing effect. Not only that but the effects of the fourth and fifth neighbour is statistically insignificant, hence the restriction imposed on the second order neighbour effect sign is less severe than in the first order neighbour case. The second reason may lay in the fact that the price effect is already transferred to further neighbours within a day and the increased determination is simply a consequence of more varied information in the second order stations that in the first order ones.

Table 7. Spatial model results

Method	Estimate	S.d.	R.sq
NN-1	0.213	0.0170	0.497
NN-4	0.430*	-	0.514
NN-5	0.411*	-	0.516
Common boundary	0.042	0.0088	0.388
Second order common boundry	0.016	0.0054	0.413
ED threshold 3km	0.500	0.0730	0.267
ED	0.925	0.1673	0.242

* Cumulative effect of all neighbours

Lastly, the Euclidean distance weighing method is assessed. To specify competition in the most global way is to fully specify the weighing matrix without any restrictions save for setting off-diagonal to zero. The measure is rarely used when vast geographical regions are considered it is more appropriate in the analysis of a single city. However it is suggestive in the analysis. Table 7 includes results for both unrestricted Euclidean weighing matrix and the restricted one. There is no clear method for determining the optimal restriction, however authors like Alderighi and Baudino (2015) and Kim and Lee (2014) use information criteria to approximate the threshold distance. Figure 6 displays AIC behaviour under different thresholds. It is quite clear that after the 5th kilometer almost no new information is included in the models hence the stations further away are nuisance to the model. Furthermore, AIC even starts to slowly increase after 7th kilometer. Despite the 'optimal' (minimal) being somewhere between 5km and 7km, the 3km was chosen as the final threshold due to it being at the 'elbow' of the plot i.e. having the last significant decrease in AIC. The 3 km threshold is quite adequate considering other reaserch: for example larger cities

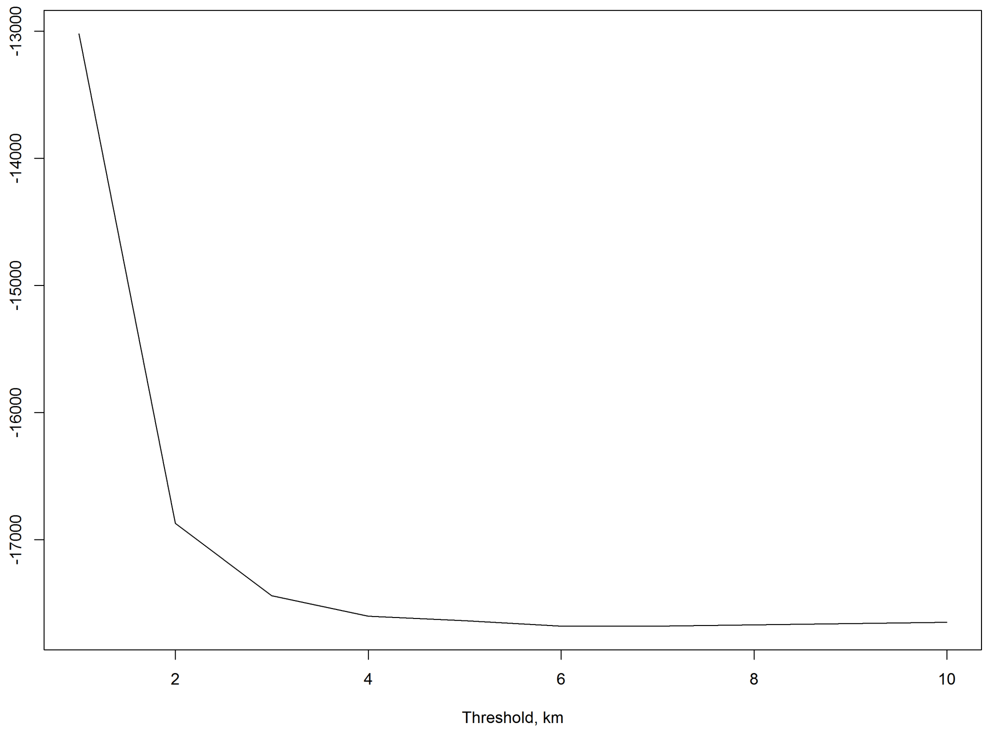

Figure 6. Model Akaike information criteron by restricting $\mathcal{W}$ by different Euclidean distance thresholds

such as Seoul Kim and Lee (2014) and Los Angeles (Atkinson, Eckert, and West (2009)) are found to have radii of close to 1 km while more rustic places like Lower Austria extend as far as 15 km.

It is interesting to compare the different measures in Table 7. When moving from the more local measures nearest neighbour (NN-1) to more global (NN-4, NN-5) the effect of the price tends to increase as well as the R^2. The added benefit of an additional 5th neighbour is negligible, however. Common boundary measures behave the similarly, however the coefficient of determination is higher for second order common boundary measure. As the weighing matrix is row normalized the most global measure ED has a magnitude close to one as all of the relevant stations can have an effect on each other prices. It is quite clear that the nearest neighbour approach is the superior in describing the spatial competition: even at one neighbour the R^2 is the highest of all measures.

One claim that I made on the incoherent behaviour of common boundary measures is that there may be some latency in the adjustment of prices and by this notion the price should be temporally lagged. The estimates of such temporally and spatially lagged models are presented in Table 8. As expected the effects on price drop in the one day lagged model. Furthermore the price magnitude effects and descriptive power drops even further when one week lag is considered. Nonetheless, the common boundary models remain a puzzle even in these models. What stands out in these models is the inversion of determination for the local and global measures: after a week the contiguity measures outperform nearest neighbours. This finding implies that stations reacts to its closest rival readily (within a day) whereas it might take a considerable amount of time to react to competitors that are more remote.

Table 8. One day lag

Method	Estimate	S.d.	R.sq
NN-1	0.347	0.016	0.429
NN-4	0.446*	-	0.275
NN-5	0.383*	-	0.280
CB	0.032	0.0089	0.417
SOCB	0.012	0.0054	0.424
ED 3km	0.382	0.076	0.343
ED	0.695	0.172	0.286

Table 9. One week lag

Method	Estimate	S.d.	R.sq
NN-1	0.192	0.017	0.196
NN-4	0.302*	-	0.214
NN-5	0.261*	-	0.216
CB	0.022	0.0089	0.437
SOCB	0.007	0.0054	0.437
ED 3km	0.267	0.076	0.358
ED	0.468	0.173	0.346

*Cumulative effect of all neighbours

The last part of the analysis concerns a type of vertical competition. Model described by equation 9 was estimated in order to investigate whether independent (small brands or unbranded gasoline stations) can impact overall prices. And if the latter is true, does the impact coefficient differ between Oligopolists and Independents. The specification was tested for all weighing matrices, the results are presented in Tables 10 and 11. Two main findings are represented in these tables. First of all, on average the price coefficient does not differ in the models. Furthermore, the hypothesis of Oligopolist and Independent price coefficient equality in the models is seldom rejected. However the standard deviation for the Independent price coefficients is significantly higher than Oligopolists.

Table 10. Vertical competiton model price estimates

	Independent		Oligopolists		
Method	Estimate	S.d.	Estimate	S.d.	Hypothesis rejection rate*
NN-1	0.213	0.018	0.216	0.023	8.3%
CB	0.048	0.005	0.043	0.008	21.1%
SOCB	0.017	0.003	0.015	0.005	15.8%
ED 3km	0.412	0.037	0.386	0.050	17.2%

*Share of models in which the null hypothesis of coefficient equality was rejected under the 5% significance level

Table 11. Vertical competiton model price estimates: four nearest neighbours case

Measure	NN-1-O	NN-1-I	NN-2-O	NN-2-I	NN-3-O	NN-3-I	NN-4-O	NN-4-I
Estimate	0.171	0.193	0.103	0.088	0.086	0.085	0.074	0.069
S.d.	0.020	0.032	0.020	0.032	0.020	0.033	0.020	0.033

These results heavily imply there is hardly any difference between brand affiliation and price impact on other gasoline stations. There is, however, more variation in the reaction function to Independent gasoline stations which is quite reasonable as there more than 60 brands clumped in the ’Independent’ definition. There may be possible to group the gasoline stations differently or to take a more local and experiment like approach as in Hastings (2004) to arrive at different result.

5. Conclusions

Gasoline prices remains a subject of intense study in the academia. There are numerous papers and little consensus in the topic, however, the spatial models are one of the most usual ways to assess the retail gasoline market competition. This paper analyzed the construction of proximal measures of spatial price competition of gasoline stations in Germany. The spatial competition was introduced by exogenously defining a weighing matrix in three different ways: nearest neighbours, border neighbours and by Euclidean distance. By having limited information on the cost, pricing strategy and auxiliary services for the gasoline stations it was found that competition can be described quite locally and parsimoniously. To be more exact there are no more than five neighbours that a gasoline station adjusts its prices to, or in distance terms price adjustment is made in 3 kilometer radius around a gasoline station. These results are reasonable considering literature relevant literature and the features of the markets analyzed. However, it is important to consider that even though the cities in the sample had at least 20 gasoline stations within their borders, the markets can still be quite heterogeneous.

This paper was made possible only by the fact that all gasoline price changes in Germany have to be recorded and made accessible to the public. This legislative move of the Bundeskartellam was made in order to inform consumers and protect them against cartel agreement yet it had a controversial impact the overall price level of gasoline. The cost of tracking their competitors has dropped significantly for the owners/managers of gasoline stations thus prices are readily adjusted (5 times per day on average). This seems to be enough time to react to competitor price changes so average daily prices were used. To assess temporal effects models with lagged prices were estimated and indeed persistence in the pricing behaviour can observed meaning that adjustment to the competitor prices takes longer than a day.

As mentioned the features of gasoline stations, cost determinants and superior demand variables were unavailable to the author, thus the models rarely reach 50% variance explained. This is a sign that there may be other types of competition involved. One of these types (vertical competition by brand) was introduced in the models yet brand name affiliation was found to be of no statistically significant importance in the models. However this does not mean brand affiliation is not a determinant of price levels themselves as brand affiliation enters all the models in this paper as controlling variables and they are certainly significant price determinants.

There are many ways to extend this analysis. Without the obvious inclusion of more market features there are other variables that impact the demand for gasoline such as traffic intensity. Even more, intraday pricing behaviour could be explored - it could answer questions on how quick the gasoline stations are to react to their rivals, how price changes travels spatially and whether the magnitude of the reactions differs with respect to rival or market features.

References

Alderighi, M. and M. Baudino (2015). „The pricing behavior of Italian gas stations: Some evidence from the Cuneo retail fuel market“. In: *Energy Economics 50*, 33–46.

Atkinson, B., A. Eckert, and D. West (2009). „Price Matching And The Domino Effect In A Retail Gasoline Market“. In: *Economic Inquiry Vol. 47, No. 3* July, 568–588.

Benson, B., M. Faminow, and T. Fik (1992). „Conduct in spatial markets: An empirical analysis of spatial pricing behavior“. In: *Papers in Regional Science* Volume 71, Issue 1, 15–30.

Berry, S., J. Levinsohn, and A. Pakes (1995). „Selling gasoline as a by-product: The impact of market structure on local prices“. In: *Econometrica* July, Vol. 63, No. 4, pp. 841–890.

Case, A. (1991). „Spatial Patterns in Household Demand“. In: *Econometrica* July, Vol. 59, No. 4, pp. 953–965.

Clemenz, G. and K. Gugler (2006). „Locational choice and price competition: some empirical results for the austrian retail gasoline market“. In: *Empirical Economics*, 31:291–312.

Dewenter, R., U. Heimeshoff, and H. Lüth (2016). „The impact of the market transparency unit for fuels on gasoline prices in Germany“. In: *DICE Discussion Papers* 220, pp. 156–168.

Diamond, P. (1971). „A Model of Price Adjustment“. In: *Journal of Economic Theory* 3, pp. 156–168.

Eckert, A. (2013). „Empirical Studies of gasoline retailing: a guide to the literature“. In: *Journal of Economic surveys* Volume 27, Issue 1, February, 140–166.

Hastings, J. (2004). „Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California“. In: *The American Economic Review* March, pp. 317–328.

Haucap, J., U. Heimeshoff, and M. Siekmann (2015). „Price Dispersion and Station Heterogeneity on German Retail Gasoline Markets“. In: *University of Düsseldorf, Düsseldorf Institute for Competition Economics (DICE)*, number 171.

Haucap, J., U. Heimeshoff, and M. Siekmann (2016). „Selling gasoline as a by-product: The impact of market structure on local prices“. In: *University of Düsseldorf, Düsseldorf Institute for Competition Economics (DICE)*, number 240.

Kalnins, A. (2003). „Hamburger Prices and Spatial Econometrics“. In: *Journal of Economics and Management Strategy* Winter, Volume 12, Number 4, 591–616.

Kim, D. and J. Lee (2014). „Spatial Price Competition in the Korean Retail Gasoline Market“. In: *Environmental and Resource Economics Review* December, Volume 23, Number 4, pp. 553–581.

Meerbeeck, W. Van (2003). „Competition and Local Market Conditions on the Belgian Retail Gasoline Market“. In: *De Economist Volume 151, Issue 4* December, 369–388.

Pennerstorfer, D. (2009). „Spatial price competition in retail gasoline markets evidence from Austria“. In: *The Annals of Regional Science* Volume 43, Issue 1, March, 133–158.

Pinkse, J., M. Slade, and C. Brett (2002). „Spatial Price Competition A Semiparametric Approach“. In: *Papers in Regional Science* Volume 70, Issue 3 May, 1111–1153.

Taylor, T., Nicholas M. Kreisle, and Paul R. Zimmerman (2010). „Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California: Comment“. In: *The American Economic Review* June, Vol. 100, No. 3, pp. 1269–1276.

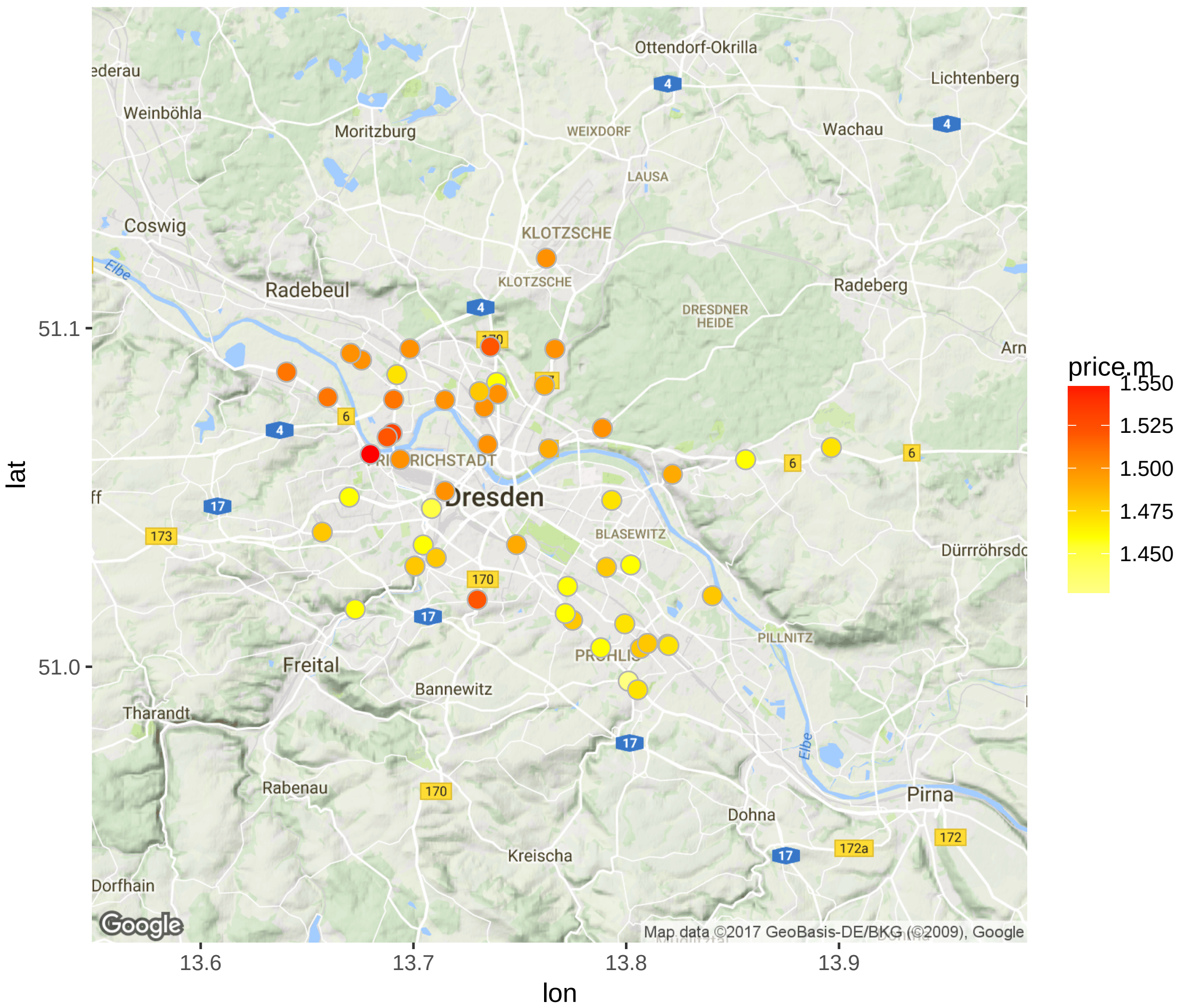

Figure 7. Dresden gasoline station network

Appendix nr. 1

Appendix