Diskussionspapierreihe
Working Paper Series

A New Price Test in Geographic Market Definition – An Application to German Retail Gasoline Market

Melissa Bantle
Matthias Muijs

Nr./ No. 180
August 2018

Department of Economics
Fächergruppe Volkswirtschaftslehre

Autoren / Authors

Melissa Bantle
Helmut-Schmidt-University Hamburg
Department of Economics
Holstenhofweg 85, 22043 Hamburg
linderm@hsu-hh.de

Matthias Muijs
University of Hohenheim
Department of Economics
70593 Stuttgart
matthias.muijs@uni-hohenheim.de

Redaktion / Editors
Helmut Schmidt Universität Hamburg / Helmut Schmidt University Hamburg
Fächergruppe Volkswirtschaftslehre / Department of Economics

Eine elektronische Version des Diskussionspapiers ist auf folgender Internetseite zu finden / An electronic version of the paper may be downloaded from the homepage:

https://www.hsu-hh.de/fgvwl/forschung

Koordinator / Coordinator
Ralf Dewenter
wp-vwl@hsu-hh.de

Helmut Schmidt Universität Hamburg / Helmut Schmidt University Hamburg
Fächergruppe Volkswirtschaftslehre / Department of Economics
Diskussionspapier Nr. 180
Working Paper No. 180

A New Price Test in Geographic Market Definition - An Application to German Retail Gasoline Market

Melissa Bantle

Matthias Muijs

Zusammenfassung / Abstract

Market delineation is a fundamental tool in modern antitrust analysis. However, the definition of relevant markets can be very difficult in practice. This preliminary draft applies a new methodology combining a simple price correlation test with hierarchical clustering -a method known from machine learning- in order to analyze the competitive situation in the German retail gasoline market. Our analysis reveals two remarkable results: At first, there is a uniform pattern across stations of the same brand regarding their maximum daily prices which confirms the claim that prices are partly set centrally. But more importantly, price reactions are also influenced by regional or local market conditions as the price setting of gasoline stations is strongly affected by commuter routes.

JEL-Klassifikation / JEL-Classification: market definition, gasoline market, price tests, competition, k-means clustering, hierarchical clustering

Schlagworte / Keywords: D22, D40, D43, L10

We would like to thank the participants of the Hohenheimer Oberseminar 50 for their valuable comments and suggestions.

1 Introduction

Studying the competitive situation and defining the relevant market and the closest rivals is fundamental in modern antitrust analysis. A relevant market is needed to assess the anticompetitive effects of mergers, but also for calculating concentration ratios to be able to judge whether a firm has a dominant market position. In addition, for cartel cases it would be helpful to have an easy method to assess the competitive situation in a market.

However, because markets are multidimensional and complex, market delineation is rarely an easy task. Moreover, these analyses must often be undertaken under limited data. In many cases only prices are available which is why methods to analyze the competitive situation based on prices alone would be very helpful in merger analysis.

We are looking at the German retail gasoline market as the competitive situation of this market is object of intensive and recurring discussions. The purpose of this paper is to analyze the competitive situation and try to delineate geographically the German retail market for gasoline by using a simple price test. With this test we hope to gain some insights into the determinants of the price setting of gasoline stations. Gasoline is a relative homogeneous product and therefore, one can analyze the market by using empirical price tests, which are based on the price movement. Two regions belong to the same market when arbitrage is possible. Therefore, it can be checked whether the prices of these areas converge.

Besides that, the opportunity cost of consumers play a crucial role in the definition of local gasoline markets. Since, no one would drive 500 kilometers just to fuel up his or her car, because the opportunity costs would exceed the benefits of this arbitrage operation significantly. From this, the following question arises: How many petrol stations are considered as competitors by a specific station and is it possible to define local markets for gasoline in Germany?

The Federal Cartel Office (FCO) chooses an accessibility model and defines an area according to a driving time of 30 minutes around a gasoline station in a city (60 minutes for gasoline stations in rural areas) to identify the competitors of this specific gasoline station.[1] In doing so, the FCO assumes that a consumer is willing to drive 30 minutes from his or her starting point (preferred station) to be able to refuel at a lower price. We question whether this assumption reflects the actual behavior of consumers. Moreover, we think that it is important to identify the forces that determine the price setting behavior of the gasoline stations. Gasoline stations cannot observe an individual consumer who lives or works around the specific station. However, gasoline stations are able to observe commuter routes and traffic which will possibly affect the price setting of gasoline stations along these routes. Therefore, this paper takes a different approach by using price correlation in conjunction with hierarchical clustering to analyze the competition between gasoline stations in Germany. With this procedure, the determinants

[1] The accessibility model was applied in several merger cases, for example in the case Shell Deutschland Oil GmbH /Honsel Mineralölvertriebs GmbH (B8–31-09), in the case Total Deutschland GmbH/OMV Deutschland GmbH (B8-175-08) or in the case Shell Deutschland Oil GmbH/Hanseatic Petrol Vertriebs GmbH (B8-134-07).

considered and observed by gasoline stations which shape their price setting are taken into account.

These quantitative tests are particularly valuable if they are used in conjunction with a thorough understanding of the industry. Therefore, the paper begins with a detailed description of the German retail gasoline market. This is necessary to assess the outcomes of the statistical tests in chapter 4.3. Chapter 3 introduces already known methods to define a geographic market. At first, general methods like the SSNIP test are presented. Afterwards, several price tests are highlighted in greater detail. In the subsequent chapter we present our new approach combining hierarchical clustering -a method known from machine learning- and correlation tests to identify the closest competitors and define geographic markets for the German retail gasoline market. Results of this analysis reveal that commuter routes indeed play a crucial role in the price setting of gasoline stations. Chapter 5 compares our approach with the market definition of the FCO to verify our findings and highlighting the problems of the accessibility model as this approach cannot capture the complex competitive situation in the German retail gasoline market. The paper concludes with a summary of the findings.

2 The German Retail Gasoline Market

In order to evaluate the statistical results, presented in chapter 4.3, a profound understanding of the market participants, market structure and pricing practices of the German retail gasoline market is necessary.

The German retail gasoline market is characterized by an oligopoly of five vertically integrated oil companies (BP (Aral), ConocoPhilipps (Jet), ExxonMobil (Esso), Shell and Total). These companies are the only ones in the market that have access to their own refining capacities and have a nationwide network of filling stations. The oligopolists have very high market shares, whereas the independent stations (the so-called "Freie Tankstellen") are only regionally active and have rather low market shares. Aral is the largest retailer with a network size of 2,335 gasoline stations in Germany and a market share of 21.5% in terms of the total sales of transport fuel in Germany. Shell has 1,929 gasoline stations and a market share of 20%. Total has 1.136 gasoline stations (9% market share), Esso 992 stations (7.5% market share) and Jet has 821 stations and 10.5% market share. The remaining market is distributed across a large number of independent stations with small and medium-sized networks.[2] This gives these oil companies an outstanding market position compared to their competitors. Moreover, the competing petrol stations on the retail level which belong to smaller networks are dependent on the gasoline deliveries of the vertically integrated oil companies as only these companies have access to refinery capacities (Bundeskartellamt (2011)). Additionally, the German gasoline market is characterized by high barriers to entry. Besides missing places for new stations, newcomers need a high capital to get access to refinery capacities. As a result,

[2]For detailed information on gasoline stations in Germany and their network sizes please refer to Energie Informationsdienst, 2017 `http://www.eid-aktuell.de/inhalt/statistiken/excel-download-strasentankstellen-in-deutschland/`.

the number of gasoline stations is overall very constant over the last years. Thereby, petrol stations know their competitors pretty well.

This market structure essentially allows for collusive behavior since the gasoline market exhibit many factors that facilitate a collusive arrangement between market participants. The existence of vertically integrated oligopolists and their nationwide presence is only one of the reasons why the German gasoline market is exposed to repeated investigations by the Federal Cartel Office.[3]

The pricing behavior of petrol stations further reinforces this suspicion. Many gasoline markets exhibit characteristic price cycles, including the German market where the prices exhibit daily cycles: a high price is placed in the morning, then, prices decrease throughout the day and in the evening prices are increased again. This price-setting pattern is obeyed by (nearly) all gasoline stations in Germany. The cycles are even more pronounced since the implementation of the "Markttransparenzstelle für Kraftstoffe (MTS-K)" of the Federal Cartel Office on August 31, 2013. Petrol stations are committed to report every price change in real-time to the MTS-K. This data is then provided to suppliers of information services where consumers can easily compare prices. It was implemented to increase the transparency for consumers by facilitating comparison of prices. The purpose of such a reporting office was to increase competition between petrol stations through the force of better informed consumers. But gasoline stations likewise have easier access to the prices of their competitors and price changes of competitors can be traced with very low effort. The increased market transparency facilitates collusive behavior: whether firms comply with the agreement or deviate from the collusive (or parallel) behavior can be monitored with little effort. Linder (2018) investigates the price cycles in the German retail gasoline market in great detail and evaluate their competitiveness. Moreover, Dewenter, Bantle, and Schwalbe (2018) discuss the possibility of tacit collusion resulting in this cyclical price setting. As the cyclical behavior plays a crucial role in the German market, the empirical analysis will take this into account which will be explained further in chapter 4.

For a well-founded investigation of this pricing behavior and especially for the evaluation of mergers and possible market power, it would be extremely helpful to define the relevant market and identify the closest competitors. For the German retail gasoline market the relevant product market and the relevant geographical market have to be identified. As the focus of the present paper is the definition of the geographic market, the relevant product market is addressed here only briefly. The assessment of the relevant geographic market and the appropriate methods are analyzed in detail in the remainder of the paper.

For the definition of the relevant product market the Federal Cartel Office makes use of the so-called "Bedarfsmarktkonzept". According to this concept, products or services belong to one market if consumers consider them to be equally suitable to satisfy a certain requirement on account of their properties, purpose of use and price. From

[3]Other reasons are the lack of buying power, the product homogeneity and the repeated interaction and mutual dependencies between the oligopolists. For a further discussion of the market structure and the firms' behavior see Bundeskartellamt (2011), p. 50f.

this perspective, petrol and diesel fuel are not substitutable for consumers. With the purchase of a vehicle the buyer has himself committed to one of the fuel types which makes diesel and petrol not interchangeable from consumers' point of view, at least not in the short run. Consequently, diesel and petrol fuel constitute separate product markets (Bundeskartellamt (2011)).

Another important characteristic of the gasoline market is the product homogeneity which makes price tests suitable for market definition. Fuels are very homogeneous products and innovations are rare. Consumers do not differentiate between petrol or diesel of different fuel stations. The diverse customer loyalty programs also point to a high product homogeneity as gasoline stations try to bind customers with such offers. As a result, the price for gasoline (diesel or petrol) is the most important competition parameter for gasoline stations. Of course, some consumers prefer petrol from brand stations, some might prefer stations with a shop and some have a preference for a station due to their location. But the main parameter for competition remains the price for petrol or diesel. This makes coordination on a collusive outcome even easier as the companies have to agree on only one parameter. But another -for this paper more important- consequence is that price tests are highly suitable for market delineation in gasoline markets. If the product is homogeneous, the possibility of arbitrage leads to uniform prices throughout the market. If consumers switch to another region due to a price increase, this price change will spill over and prices will adapt to each other. When a market is characterized by product homogeneity, the opportunity of arbitrage prevents prices from moving independently. Price tests can be used to test whether the price in one region is exogenous to the price in another region (Slade (1986) and Audy and Erutku (2005)). The subsequent chapter discusses price tests in detail.

3 Methods to Define a Geographic Market

As we are interested in delineating a geographic market for the gasoline industry, methods that are applied to these markets are depicted in the following. First of all, general methods that are applied to define geographical markets in general and especially to gasoline markets will be considered. These include inter alia the accessibility model used by the Federal Cartel Office in various cases. Furthermore, the hypothesis of Weizsäcker about chain of substitutions in the German gasoline market is considered (see Franz, Ramser, and Stadler (2002)).

However, these methods have some weaknesses which will be discussed in the following. For this reason, the second part of this chapter presents price tests which are appropriate to delineate the relevant market. Afterwards, chapter 4 introduces our approach in great detail. This analysis is aimed at giving a reasonable market definition for the German retail gasoline market and shows that the competitive situation is more complex and cannot be captured either by the model of chain of substitutions or the accessibility model applied by the Federal Cartel Office.

However, it should be noted that there is a well recognized distinction between "economic markets" and "antitrust markets". In the literature, it is controversially discussed whether price tests are appropriate to define antitrust markets. The geographic extent of an economic market is defined by arbitrage opportunities and transportation costs, whereas an antitrust market comprises the geographic space within which a hypothetical monopolist can exercise market power (Church and Ware (2000)). Price tests investigate whether price series in different regions are related, but cannot verify whether firms have the facility to raise prices profitably. Price elasticity estimates are an appropriate tool for defining an antitrust market as these estimates provide direct evidence for market power (Massey (2000)). Unfortunately, it is often not possible to estimate these elasticities due to data limitations. However, price tests are not generally inconsistent with the concept of antitrust markets. A tool to define a relevant antitrust market should rank the substitute products and identify the closest substitutes. As an antitrust market comprises the smallest possible set of substitutes that would enable an hypothetical monopolist to exercise market power. The price series of close substitutes are probably linked. Price tests that focus on the size and not only on the existence of price relationships should be able to identify a relevant antitrust market. Nevertheless, price tests should be used in conjunction with a thorough understanding of the industry under consideration. Without institutional knowledge of the relevant market, price tests can be misleading (Boshoff (2012)).

3.1 General Methods

The Federal Cartel Office has repeatedly emphasized regionally separated markets for gasoline stations in Germany and thereby rejects the proposal of the oil companies which plead for a nationwide gasoline market (concept of chain of substitutions according to Weizsäcker). For the delineation of the relevant geographic market the Federal Cartel Office again uses the "Bedarfsmarktkonzept", as it is as well done for the product market. For the geographical dimension, the actual behavior of consumers is a crucial factor. Thus, the geographic market for a gasoline station is determined by the distance consumers are willing to drive to buy gasoline at an alternative station if the target station increases its price. Most of the consumers refuel their cars on the journey between home and their work place. The Federal Cartel Office concludes that a radius of 25 km around the gasoline station in question is sufficient to define the relevant market.

For a precise market definition on a case-by-case basis the Federal Cartel Office makes use of the accessibility model of the "Bundesamt für Bauwesen und Raumordnung".[4] With this model all stations within a specific driving time around one target station can be identified. These gasoline stations are considered to be substitutes in the view of the consumers and thus belong to one market. For rural areas a driving time of 60 minutes is assumed and for urban areas a driving time of 30 minutes. For urban areas the resulting geographic market will probably be too broad. The higher density of gasoline stations in urban areas is an argument for smaller submarkets. The main problem with this approach is the arbitrarily chosen driving time which is based only on consumer surveys.

The alternative petrol stations within one regional market are, however, no equal

[4]The model is based on the digitally recorded existing network of streets.

choices for the demand side. A consumer will always choose the nearer gasoline station if the price difference is smaller than the driving cost and the opportunity cost of time. To account for the different intensities of competition depending on the distances to the target station, a weighted accessibility model is applied to the gasoline market. The stations are weighted according to their distance to the target station (or the center of the regional market). Nearer stations get a higher weight which display their intensity of competition.

The market delineation according to the accessibility model is a clear rejection of the nationwide market suggested by oil companies. Besides, also the pricing behavior of the oil companies points to regional markets. Although oil companies initially specify their prices nationwide, the station dealers submit price requests to align their prices to the local level (Bundeskartellamt (2011)).

Nevertheless, oil companies call for a nationwide market and refer to the concept of chain of substitutions. According to this approach, products (or regions) can belong to the same relevant market without competing directly with each other. This is the case when competition is passed on through a chain of substitutes. This concept has been used by the European Commission in a few cases. In the case Pilkington-Techint/SIV the Commission considers the geographic float glass market as a chain of substitution as the regional markets are determined by the transport cost. Moreover, it was essential to proof that prices for the various regions track each other and also that a substantial amount of sales stems from other regions.

Oil companies claim that the German gasoline market is nationwide as price changes are spread throughout Germany through the chain of substitutions. Price tests would be able to reveal whether gasoline stations are connected via a chain of substitutions as arbitrage would lead to uniform prices across the market as already mentioned in chapter 2. However, our results do not provide evidence for this hypothesis and show a more complex competitive structure in the German retail gasoline market. The greatest danger of the approach of chains of substitutions is that the relevant market is defined too generous since it overstates the influence of distant competitors (gasoline stations with a longer distance from the target station). A hypothetical monopolist (this method is further outlined below) needs not control over the whole substitution chain to raise his price profitably. If a sufficient amount of consumers is left after a price increase, this price increase still is profitable for the hypothetical monopolist (Bishop and Baldauf (2006)).

However, the purpose of delineating a market is to identify the main competitors. This applies in particular for gasoline markets. The geographic market should comprise those gasoline stations whose price setting mutually influence each other. Since this is the question in many antitrust cases. For mergers, for instance, it is necessary to know the main competitors that remain after the merger and which are able to compete effectively with the merging entity. Therefore, the market delineation should examine how many gasoline stations are observed by the target station.

Another method, frequently used to define a geographic market, is the hypothetical monopolist test or SSNIP test (small but significant increase in price). The SSNIP test asks whether a hypothetical monopolist is able to profitably increase its price (by 5%) over a certain period. If a considerable amount of consumers switch to another product (or market) in response to a price increase, these substitutes need to be included to the relevant market.

For assessing the SSNIP test demand elasticities have to be calculated. The data required is sometimes difficult to obtain and implementation of this test can be challenging. The same problems occur with the Critical Loss Analysis (based on the Critical Loss test by Harris and Simons (1989)). The critical loss is the volume that have to shift to other products (geographic areas) to make a hypothetical price increase unprofitable. This analysis requires estimation of the cost function where data is likewise difficult to obtain.

Despite the difficulties in obtaining appropriate data, the SSNIP test is still widely used for market definition and the identification of close competitors (or substitutes). The following chapter introduces price tests as an alternative approach for market delineation. Afterwards, our methodology is presented which combines a commonly known price test with a machine learning method.

3.2 Price Tests

Market definition frequently involves rules of thumb like the driving times assumed for the accessibility model implemented by the Federal Cartel Office. But market delineation should not be dependent on such subjective assumptions. Therefore, econometric methods are in demand.

Methodologies investigating product flows to define the relevant geographic market has been known for a long time (see for example Elzinga and Hogarty (1983)). In contrast, Horowitz (1981) proposes a test based on price differentials and assumes that regions belong to the same geographic market if price differentials are stable. One strong limitation of this approach is the assumption of a specific pattern for the dynamic adjustment process after a shock. Hence, the remainder of this section focuses on less restrictive price tests. Since the price of a product is affected most by changes in the competitive situation, price tests are particularly suitable for investigating them.

There are a large number of tests examining price correlations and interrelationships. Most of these price tests are simple to conduct and the necessary information is often publicly available. Due to the fast and straightforward implementation, these tests are appropriate as a first inquiry for antitrust agencies.

Different econometric tests can be found in the literature that are suitable to define relevant markets and there is as well some literature that applies econometric tests to gasoline markets. Audy and Erutku (2005) and Slade (1986), for example, use price correlation and Granger causality to define geographic markets in the gasoline industry. Moreover, cointegration and stationarity tests as well as econometric models of price responses and co-movements across regions based on natural experiments are proposed for defining the relevant geographic market. However, for antitrust inquiries it is important that these tests are easy to apply with minimal data requirements.

The different price tests ask different questions and might therefore provide different results. But this does not mean that the tests are not mutually consistent. The various price tests focus on different dimensions of a price relationship. Therefore, it is not the purpose of such tests to confirm each other, but to explain different perspectives of the price relationships (Boshoff (2012)).

In general, price tests can be divided into two categories: short-run relationships and long-run relationships. Correlation tests and Granger causality tests belong to the first group, whereas unit root tests and cointegration tests belong to the latter. The remainder of this section will outline in detail the above mentioned price tests. As our method comprises a price correlation test in combination with hierarchical clustering, the other tests are only shortly depicted to demonstrate why a price correlation test is most appropriate for our purpose.

Stationarity

One of the long-run tests is based on stationarity which is applied by Forni (2004), for example, who uses this test to define the relevant product market for the Italien milk industry. His method is linked to cointegration techniques. Two products are substitutes and belong to the same product market if the log of the price ratio is stationary. To check for stationarity two tests are applied: the Augmented Dickey Fuller (ADF) test and the KPSS test (Kwiatkowski-–Phillips—Schmidt—Shin test). The null hypothesis of the ADF test is nonstationarity whereas stationarity is the null hypothesis of the KPSS test. Two markets are considered to be distinct if stationarity (KPSS) is rejected and nonstationarity (ADF) cannot be rejected. In contrast, if stationarity cannot be rejected (KPSS), the two products (or regions) can not be considered as a single market. Stationarity can also be observed if markets are distinct. This is the case when prices themselves are stationary or are affected by common sources of nonstationary variation. To overcome this problem, the sources of variation need to be identified. If prices are nonstationary in levels, a conclusion concerning the relevant market can be reached.

Forni (2004) investigates whether the relative prices of two products revert to a stable long-run value. If the two prices tend to a stable value, the two regions belong to the same relevant market.

A shortcoming of this stationarity test is that it can only compare prices pair-wise. This is specifically troublesome for defining geographical markets as it is often beneficial to compare numerous regions simultaneously (for example when the geographic markets might be rather small). Moreover, it is important to control for common influences that could lead to spurious results.

Another main critic involves that stationarity tests will often lead to very broad markets as the concept of stationary is misleading, especially if a longer time horizon is considered. In markets with homogeneous products that are sold in different geographical areas (like gasoline), nonstationarity will always be rejected with a long enough sample period (e.g. due to cost changes over time). If both, the log of the relative price and the log of the relative cost follow a random walk, the suggested test will result in nonstationarity and we would conclude that the two firms belong to different regional markets (for a more detailed critic on Forni's approach, see Genesove (2004)).

Therefore, although the log difference between two prices must be stationary if both belong to the same market, a stationary difference does not indicate whether the two regions belong to the same market.

Cointegration Tests

Another test for studying long-run relationships are cointegration tests. Cointegration describes a particular kind of long-run equilibrium. For regions belonging to the same geographical market, an equilibrium relationship among prices is expected. Even if the prices of different regions diverge from their equilibrium in the short-run, they are expected to adjust back to their equilibrium in the long-run in case of a common geographic market for these regions. Accordingly, if prices of different regions are cointegrated, it is statistical evidence that these regions belong to the same geographic market, since there exists a systematic equilibrium relationship among them (see for example Engle and Granger (1987), Forni (2004) or Warell (2005)).

The standard cointegration test is the Johansen system approach. This test requires checking the price series for non-stationarity prior to the cointegration test. The ARDL bounds testing approach developed by Pesaran, Shin, and Smith (2001) does not need a prior unit root test and can test for a long-run relationship regardless of the order of integration of input variables. With this test Boshoff (2012) investigates the adjustment speeds for the prices returning to their long-run relationship after a disturbance. These adjustment speeds give guidance whether the long-run relationship is strong enough to be relevant for market definition.

For cointegration tests it is necessary to control for common factors that influence the prices before running the analysis. Moreover, cointegration tests suffer from small-sample power problems. A crucial issue with cointegration tests (as well as with stationarity tests) are their long-run dimension. It is problematic to use these long-run tests for market delineation as they neglect the possibility of consumers to react to price changes and therefore might overstate the size of a market. Gasoline stations in Germany change their prices several times a day and consumers will possibly react to these frequent and repeating changes. The following price tests -including Granger causality and price correlation - are suitable for a short-run analysis which seems to be more appropriate in order to identify the closest competitors and define a relevant geographic market for gasoline stations.

Granger Causality

A Granger causality test is able to infer whether there is a causal relationship between the price series of two products or firms (in our case gasoline stations). Another advantage of Granger causality tests is that they allow for dynamic interaction among price series and that more than only two price series can be investigated simultaneously.

The proposed test was developed by Granger (1969), Sims (1972) and others and is based on standard regression techniques, which assume that there is a cause and effect relationship between the dependent and independent variables. By following the general literature in econometrics, a variable X is said to Granger cause another variable Y, if past values of X can improve the prediction of variable Y in an appropriate regression model. In the context of market delineation, two regions A and B are part of the same relevant geographic market, if the prices of region A affect the prices of region B and vice versa (feedback effect). However, if the causal relationship is only unidirectional[5] or there is no causality at all, then the two areas should not be in the same relevant geographic market. To be precise, Granger causality tests whether past prices of one or more regions significantly explain current prices of another region. If past values of other regions lead to a better prediction of the prices in the region at issue, these regions can be assigned to the same geographic market.

Cartwright, Kamerschen, and Huang (1989) argue that Granger causality tests have some distinct advantages compared to simple price correlation tests. At first, price correlation tests can only provide static information about the linear association between two geographic areas. Granger causality tests, in contrast, take account of the dynamic structure of the price series. Second, the correlation analysis can not validate any causal relationships between the units of observations. The last big advantage is given by the fact that there is no uniquely defined general threshold, which indicates whether the correlation level between two regions is high enough to form a relevant geographic market. Granger causality tests do not need such an arbitrary threshold value.

However, Granger causality tests are very sensible to serial and spurious correlation. Furthermore, the price series have to be stationary in order to be tested for Granger causality. To overcome serial correlation, the time series need to be regressed on the common factor and autocorrelation in error terms has to be eliminated. Moreover, omitted variables can lead to misspecification and biased results (see also Slade (1986)). It therefore is important to control for common causal factors like the prices of input factors (like cost of crude oil).

Moreover, Granger causality tests focus solely on the existence of a relationship rather than its size. But for the definition of a relevant antitrust market it is of great importance whether two price series are meaningfully related. The size of a relationship is crucial to pass the requirements of the definition of an antitrust market. Nevertheless, Granger causality tests can confirm the existence of a dynamic short-run relationship between price series (Boshoff, 2012).

Slade (1986) proposes a geographic market test based on Granger causality to determine whether a disturbance in price in one region have repercussions in another. If the two areas belong to the same market, exogeneity is rejected. The test is applied to crude oil prices of various regions of the United States. The advantage of this approach is that no specific model of price formation has to be presumed.

[5]For instance, prices in region A are influenced by price movements in the other region B, but the reverse is not the case.

Price Correlation

However, Granger causality tests would have to be performed for each gasoline station individually which would be a very time-consuming task. As we are seeking for a method which is easy and fast to implement for competition authorities -as part of an initial market investigation- we will have a closer look at price correlation tests, which are appropriate to identify the closest competitors. Moreover, Cartwright, Kamerschen, and Huang (1989) demonstrate that both, price correlation and Granger causality provide the same results in market delineation and conclude that Granger causality could be used supplementary to a price correlation analysis. Also Audy and Erutku (2005) apply both tests, price correlation and Granger causality, to the wholesale gasoline market in Canada. Both tests indicate that the relevant geographic markets can be larger than cities but can not be bigger than East and West Canada.

Following the reasoning of Stigler and Sherwin (1985), two geographic areas belong to the same market when their relative prices maintain a stable ratio or rather when their prices move together over time. This can be measured statistically by using price correlations. Correlation is a measure of the linear relationship between two variables and indicates the degree of contemporary linear association.

The basic idea of the Stigler and Sherwin test for the relevant geographic market by the similarity of price movement can be summarized as follows: The greater the price correlation between two geographic areas, the greater the likelihood that these two areas are in the same relevant market. A high correlation between the prices of two geographic areas suggests that the cross-price elasticity between those two regions is positive. Following this, there has to be some competitive interaction. For instance, if there are two geographical areas X and Y and the correlation coefficient between the prices of those areas is high (near to 1), than this may indicate that the two areas should be part of the same relevant geographical market. Buyers (or sellers) will shift to area Y if the price in area X increases. The high correlation in prices imply that a relative price increase in region X will lead to an adjustment of the prices in both regions. This adjustment process is caused by arbitrage opportunities in conjunction with homogeneous products. Market participants (buyers as well as sellers) can move from one region to the other without incurring large costs (transportation cost, transaction cost and opportunity cost). Due to this movement, the relative prices in areas X and Y will return to the base level. In contrast to that, a small correlation coefficient implies that a relative price increase in region X should not affect the price in region Y. Thus, a small price correlation is an indication for separate geographic markets. In gasoline markets sellers can not shift to another market, but buyers will drive to another (nearby) gasoline station if the cost (transportation and opportunity cost) are smaller than the price difference. If the distance is too long, buyers do not switch to the other region and regional price differentials can persist, indicating separate geographical markets (see Slade (1986)).

Price correlation tests have played a prominent role in numerous antitrust and merger cases of the last decades and have been used by both the European Competition Authorities and the parties under investigation to define the relevant market.[6]

However, there are several weaknesses associated with the use of price correlation tests to define relevant markets. As pointed out by Audy and Erutku (2005), there may be five shortcomings and statistical complications which need to be considered to be able to apply price correlation tests properly.

The first big issue might be the presence of serial correlation in the price series. A price series is subject to serial correlation if the series is correlated with itself at different points in time. To overcome this problem, first or second differences in prices can be used to measure correlation.

A second shortcoming might be caused by common influencing factors. In such cases, the degree of correlation between two price series is high, but this linear association is caused by parallel movements of common factors and not by competitive constraints. Such a spurious correlation exists when variables are related only through their correlation with omitted variables or with a common trend. In our analysis, such a common influence may be the spot price for crude oil. Since the price of crude oil is an important determinant of the level and movement of retail gasoline prices, neglecting this common influence would result in upward biased correlation results. Spurious correlation can also arise when the prices of products are influenced by seasonal trend. For instance, most Germans would have noticed that the price for gasoline increases at the beginning of school vacations and on holidays, such as Christmas or Easter. To solve this problem, the price data need to be purged of any common factors and the price series should be stationary or should be made so. Controlling for common influences must also be examined for other analysis, like the Granger causality test or cointegration, to avoid spurious results.

Another problem can occur when two series show a common trend but differ in the level of prices. For instance, the movement in prices of gasoline with 100 octane and gasoline with 85 octane may be parallel but they differ in their price level. In this case the result of a correlation analysis would be around one but these two varieties of gasoline are no perfect substitutes. The same problem exists with market power. If a firm has market power in one of the markets, price correlation tests can lead to false conclusions and are therefore not suitable to identify market power.

Fourth, there is no unique threshold, which determines whether the correlation between two price series is large enough to be part of the same relevant market.[7] It is difficult to assess whether a particular correlation is economically meaningful. And as there is no objective threshold, chosen benchmarks remain arbitrary.

The last big issue is given by the nature of the correlation analysis. A correlation coefficient can only measure the contemporary linear relationship between two series. Thus, the market is defined too narrowly if the prices are independent in the short run but not in the long run. When response to price changes is delayed, contemporaneous correlation will be very small. The price correlation test is misleading when the two price

[6]For instance, Case M190 Nestle/Perrier (1997) OJ L356/1 or COMP/M.4439 Ryanair/Aer Lingus (June 27, 2007).

[7]See Bishop and Walker (2010), p. 516f. for a discussion about benchmarking.

series are highly correlated in the long run (see also Slade (1986)). However, gasoline stations in Germany change their prices very frequently and react immediately, so we do not expect delayed responses. Therefore, price correlation tests are appropriate to identify the closest competitors of a gasoline station in the German market.

Despite those weaknesses, price correlations, if used and interpreted correctly, can be a useful tool for the definition of the relevant market, at least in a first step. This is due to the pure simplicity of the method and the modest data requirements, especially when compared to the SSNIP-test. Tirole (1988) argued that a high price correlation between two products or geographic areas is at best a necessary but not a sufficient condition for belonging to the same market.

The following section describes our methodology and explains how we solved the potential issues with price correlations outlined above. As we are interested in identifying the closest competitors and price reactions occur very fast in the German retail gasoline market, an analysis based on short-run price correlations seems to be appropriate.

4 Methodology

The characteristics of the gasoline market, which are outlined in chapter 2, in conjunction with the available price data make price tests an appropriate method to analyze the competitive situation on the German retail gasoline market. As all price tests have some weaknesses, we choose price correlation as it is an efficient method and potential issues can be solved. In addition, correlation itself is an appropriate dissimilarity measure for our hierarchical clustering method. We describe the data used in the following subsection. Afterwards, we present our approach in more detail and analyze our findings.

4.1 Data

We use data from the service provider and consumer information service „Tankerkönig“[8], that makes the price information from the „Markttransparenzstelle für Kraftstoff (MTS-K)“ publicly available. As prescribed by the Federal Cartel Office (FCO), the data set contains every price change of each individual filling station in Germany. Thus we have information about the current price of E5-gasoline, E10-gasoline and diesel for each and every point in time within the observation period. In addition, the data set contains petrol station specific information, such as the name, address, brand or geographical coordinates of all $14,714$ filling stations in the sample. However we do not have to examine all petrol stations, since we assume that the relevant market - and thus the closest competitors - depends heavily on consumer behavior. And most consumers are not prepared to drive hundreds of kilometers to refuel cheaper, as this detour entails additional costs to the customer (opportunity cost of time and transportation cost). For this reason we consider subsamples of the original data set. Since we assume regional markets, we obtain the subsamples by clustering the filling stations according to their

[8]www.tankerkoenig.de

geographical location, i.e. longitude and latitude. We choose the well-known k-means method to partition our population of filling stations into 128 regional subsamples. These 128 subregions are still quite large, so we ensure that the subsamples chosen by the k-means clustering are not too restrictive

The basic idea of k-means clustering is as follows: All N observations are clustered into k clusters, such that within each cluster the average distance of the observations to the respective cluster centroid is minimized. This can be achieved by an algorithm. First the desired number of clusters k and the features p to be clustered over are selected. In a next step, each observation is then randomly assigned a number between 1 and k. This allocation forms the starting point for an iterative descent algorithm. For each of the k clusters, the centroid is calculated as a vector of the mean values of the p features over all observations in the respective cluster. Each observation is then assigned to the cluster whose centroid is closest. The proximity between the observations and the centroids is measured by the squared Euclidean distance of the features. The algorithm proposed by Hartigan and Wong (1979) stops as soon as the within-cluster variance is minimized, i.e. there is no single switch of an observation from one cluster to another that would further minimize the within-cluster variation, which can be measured as

$$\sum_{k=1}^{K} \frac{1}{N_k} \sum_{i,i' \in N_k} \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2, \tag{1}$$

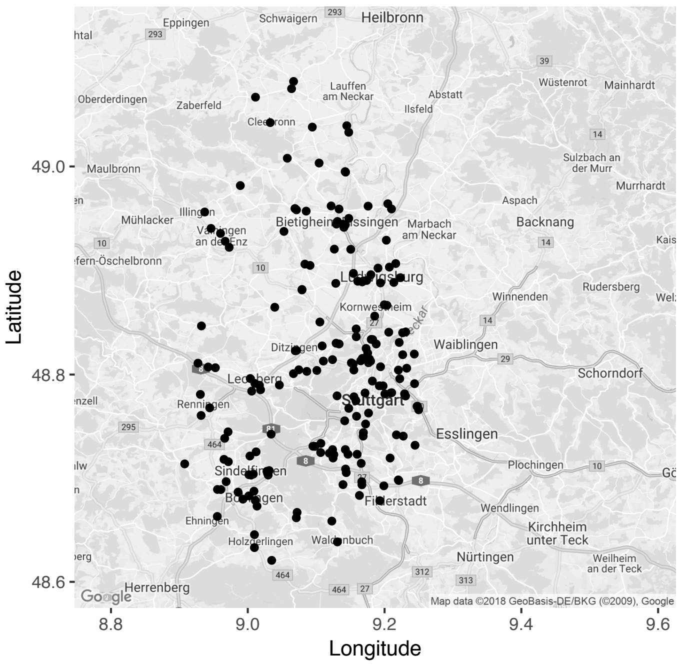
where N_k is the number of observations in cluster k, x_{ij} is the feature vector of observation i, and $x_{i'j}$ denotes the feature centroid of cluster i'[9]. Since we chose k to be 128 and clustered over the two features of latitude and longitude, we end up with 128 non-overlapping regional clusters for our gasoline stations.[10]

For our analysis we selected the cluster that includes all filling stations in the region of Stuttgart. Our subsample covers the period from 8th June 2014 to 6th June 2017 and contains 186 filling stations. We choose the region of Stuttgart as it is well known for its extremely high commuter traffic. Moreover, the cluster contains the metropolitan area which is characterized by a high station density as well as more rural areas with a rather low station density. Therefore, the chosen cluster presents a good mixture of congested urban area, rural regions and commuter routes to test our hypotheses. Chapter 5 shows results for another geographic cluster (the region around Dresden, Leipzig, and Chemnitz) which further confirms our results. The cluster containing the region of Stuttgart covers 65 km from the northernmost to the southernmost point, which corresponds to a journey time of approximately 59 minutes. The exact geographical distribution of the stations in our sample is illustrated in figure 1, whereas the number of stations per brand in the selected sample is depicted in table 1.

[9] See for instance Friedman, Hastie, and Tibshirani (2001) for a detailed discussion of the k-means meathod.

[10] We choose k to be 128 as smaller numbers all resulted in too large clusters as we take consumer behavior into account. The average cluster sizes for the respective number of clusters are: $k = 2$ clusters $\rightarrow 7,357$ stations, $k = 4 \rightarrow 3,678.5$, $k = 8 \rightarrow 1,839.25$, $k = 16 \rightarrow 919.6255$, $k = 32 \rightarrow 459.81255$, $k = 64 \rightarrow 229.91$, and $k = 128 \rightarrow 114.95$.

Figure 1: Sample of stations



Note: All 186 filling stations within the sample.

In the region of Stuttgart, the two major oil companies ARAL and Shell have exactly the same number of filling stations, each with 39. Two other oligopolists follow with ESSO (23) and JET (15), thus TOTAL (4) is the only one of the five major retailers that is underrepresented in the sample we are looking at. Nevertheless, we can argue that the sample is representative in terms of the distribution of stations among the brands.

Table 1: Distribution of brands in the sample

Brand	No. of stations
ARAL	39
Shell	39
ESSO	23
JET	15
Agip	14
OMV	6
HEM	5
TOTAL	5
Others	40
N	186

Note: Number of filling stations per brand.

Descriptive statistics on the prices in the selected sample are shown in table 2. These are the unadjusted prices that were reported by the filling stations to the MTS-K. During the observation period, the 186 stations reported $1,370,994$ prices to the transparency unit with a mean price of E5-gasoline of 137.3 and a standard deviation of 10.27 Euro Cent.

4.2 Our Approach

We apply a method known from unsupervised machine learning to analyze whether the prices of the filling stations in our sample co-move and are therefore in close competition with each other. In this context, it should be briefly mentioned that the present paper is not intended to evaluate the effectiveness of competition in the German retail gasoline market. We refer to competing stations (or competition) when stations react to each other in prices. Whether this can be attributed to effective competition or rather a collusive agreement is not subject of the following analysis. However, the identified relevant markets could be used as a starting point for further investigations of the price setting behavior of gasoline stations whose prices co-move. This method is called hierarchical clustering. In contrast to k-means clustering, the hierarchical clustering method does not require the user to specify a certain number of clusters k, but a dissimilarity measure d_{ij}. Here we use an agglomerative variant of the hierarchical clustering method. This means that we initially assume that each observation forms a separate

Table 2: Descriptive statistics of the prices in the sample

Feature	Statistic
No. of Obs.	$1,370,994$
Mean(price E5)	137.3
Min.(price E5)	110.4
Max.(price E5)	171.9
SD (price E5)	10.27
5%-Q. (price E5)	122.9
25%-Q. (price E5)	129.9
50%-Q. (price E5)	135.2
75%-Q. (price E5)	143.9
95%-Q. (price E5)	157.4

Note: Prices in Euro Cent.

cluster. The basic idea of hierarchical clustering is very simple and intuitive: In a first step, the two observations that have the least dissimilarity are combined into a cluster using a suitable dissimilarity measure. In the next step, the two most similar clusters are merged again[11]. This iterative procedure is carried out until all observations have been combined into one single cluster. While the chosen dissimilarity measure can be used directly with two individual observations, the question arises how the dissimilarity between clusters, containing more than one observation, should be measured. By following the notion of Friedman, Hastie, and Tibshirani (2001) among others a concept of linkage has to be defined, which measures the dissimilarity between two groups of observations, i.e. clusters. To analyze the competition within our subsample of filling stations, we use the method of 'complete linkage'. This linkage method makes use of the maximum intercluster dissimilarity, which can be represented by

$$D_{complete}(A, B) = \max_{\substack{i \in A \\ j \in B}} d_{ij}, \tag{2}$$

where A and B are two distinct clusters and d_{ij} is the chosen dissimilarity measure. As can be seen from equation 2, the method of complete linkage calculates every pairwise dissimilarity between the two clusters and takes the most dissimilar pair from both clusters as the dissimilarity measure between A and B. Thus, this approach to define linkage states that two clusters can only be fused if all members of both groups are very similar. This leads to relative small and compact clusters (Friedman, Hastie, and Tibshirani (2001)).

Boehnke (2017) extensively investigates the pricing strategies of German gasoline stations and using a somehow similar approach. To define local geographic markets for his

[11] Remember that at the beginning each observation forms its own cluster.

demand analysis, he uses hierarchical clustering. However, in contrast to our approach presented below, he based his hierarchical clustering solely on a physical distance measure d_{ij}. It is comparable to our k-means clustering outlined in the preceding chapter. Boehnke (2017) identifies gas stations located close to each other with the hierarchical clustering. However, we go one step further and focus on one of the geographical clusters previously defined by the k-means algorithm to implement hierarchical clustering based on a more refined distance measure - that takes the pricing pattern of gasoline stations into account - to define clusters of gasoline stations that compete with each other. In applying this two-step procedure, we make use of the efficient k-means algorithm to roughly cluster nearby gasoline stations geographically. This makes intuitive sense, since geographical proximity can be seen as a necessary condition for a close competitor in the retail gasoline market. However, this first clustering approach is defined to be not too restrictive, but rather provide a rough framework for the actual analysis. In the second step, hierarchical clustering based on a price-based dissimilarity measure is implemented, as price is one of the main competition parameters in the gasoline market. Therefore, a similar price pattern can be regarded as a sufficient condition for a close competitor in the market.

We focus our analysis on the sample of filling stations in the region of Stuttgart. The analysis focuses mainly on daily average prices, whereby each price is weighted with the time in which this price was active (the longer a price was active during a day the higher his weight). However, based on the idea of Edgeworth cycles we also take a look at the daily highs and lows. Since the vertically integrated companies in particular are very likely to set their prices uniformly, at least on a regional level, we do not expect daily highs to reflect competitive behavior. The maximum prices are more likely to be the result of a centralized pricing by the branded stations, as proposed by Linder (2018). On the other hand, we assume that the analysis on basis of the weighted daily average prices represents the regional competitiveness. That is because the filling stations are in price competition with their closest rivals throughout the day. Gasoline stations decrease their prices sequentially over the day until a sharp and almost simultaneous price increase interrupts this relenting phase. A weighted average price which takes the length of the different active price levels into account is appropriate to depict these characteristic price cycles and the competitive behavior of the filing stations. So our central hypothesis can be formulated as follows: The competitive situation on the market for retail gasoline is determined by the consumers. For that reason, the geographically relevant markets should be relatively small.

The original price data has to be prepared as proposed by Cartwright, Kamerschen, and Huang (1989) among others. This means that the prices should be made stationary and cleared of all common factors (like already pointed out in chapter 3.2). Since the analysis relies heavily on calculating the empirical correlation coefficients of the prices, the price series of each filling station has to be made stationary. This is because the sample moments of stationary series converge to a constant number, whereas the sample moments of random walks converge to random variables. However, the relationship between the price series should be represented by empirical correlation coefficients, which are constant numbers and not random variables. For this reason, the reported prices of the filling stations must be made stationary by forming the log returns. The common factors influencing the retail gasoline price are especially the price for crude oil (Brent Crude) and the exchange rate (USD/EUR), since the crude oil price is usually measured in USD per barrel. In order to pay attention to the problem of serial correlation, we must also take into account the common factors' lagged values. This also makes sense intuitively, as most filling stations have some sort of long-term supply contracts with wholesalers. It is reasonable to include the common factor itself along with three lagged values, as can be seen from equation 3

$$\Delta p_{it} = \beta_1 \Delta C_t + \beta_2 \Delta C_{t-1} + \beta_3 \Delta C_{t-2} + \beta_4 \Delta C_{t-3} + \epsilon_{it}, \tag{3}$$

where Δp_{it} denotes the log difference of the price of station i for one liter E5-gasoline in Euro at time t, ΔC_t is the log difference of one liter Brent crude oil in Euro[12], and ϵ_{it} is the error term of station i at time t. Alternatively, ϵ_{it} can also be interpreted as the price adjusted by the common factor. This is because the error term contains everything that cannot be explained by the regressors. We argue that there are no other common factors affecting the prices[13]. For this reason, these adjusted log returns ϵ_{it} are now used to calculate the empirical correlation coefficients between each and every filling station. As mentioned above, a dissimilarity measure is required to perform hierarchical clustering. We define the dissimilarity between the two filling stations j and k as

$$d_{ij} = 1 - |cor_{ij}|, \tag{4}$$

where $|cor_{jk}|$ is the empirical correlation coefficient of the adjusted price series ϵ_j and ϵ_k in absolute value. Thus the dissimilarity measure of a filling station i to itself is zero and the overall measure is defined between 0 and 1. By choosing correlation as the dissimilarity measure, all advantages and disadvantages associated with correlation as a price test to define markets also apply to this method. As mentioned before, the complete linkage method is used for the following analysis. This means that the dissimilarity measure of two clusters is defined as the maximum distance of all observation pairs from the clusters. Thus, the analysis starts with 186 individual clusters and merges the two most similar clusters in an iterative process until all filling stations are fused into one final group in the end. This step-by-step procedure generates the hierarchical structure, which is usually represented in the form of a dendrogram.

4.3 Results

In the following, results for the daily average price, maximum and minimum price will be presented. As already mentioned in chapter 2, gasoline stations exhibit a cyclical

[12]Both crude oil price and exchange rate data are from the Federal Reserve Bank of St. Louis (`https://research.stlouisfed.org/`).

[13]The distances within Germany are relatively short, which means that there should not be any significant differences in transport costs.

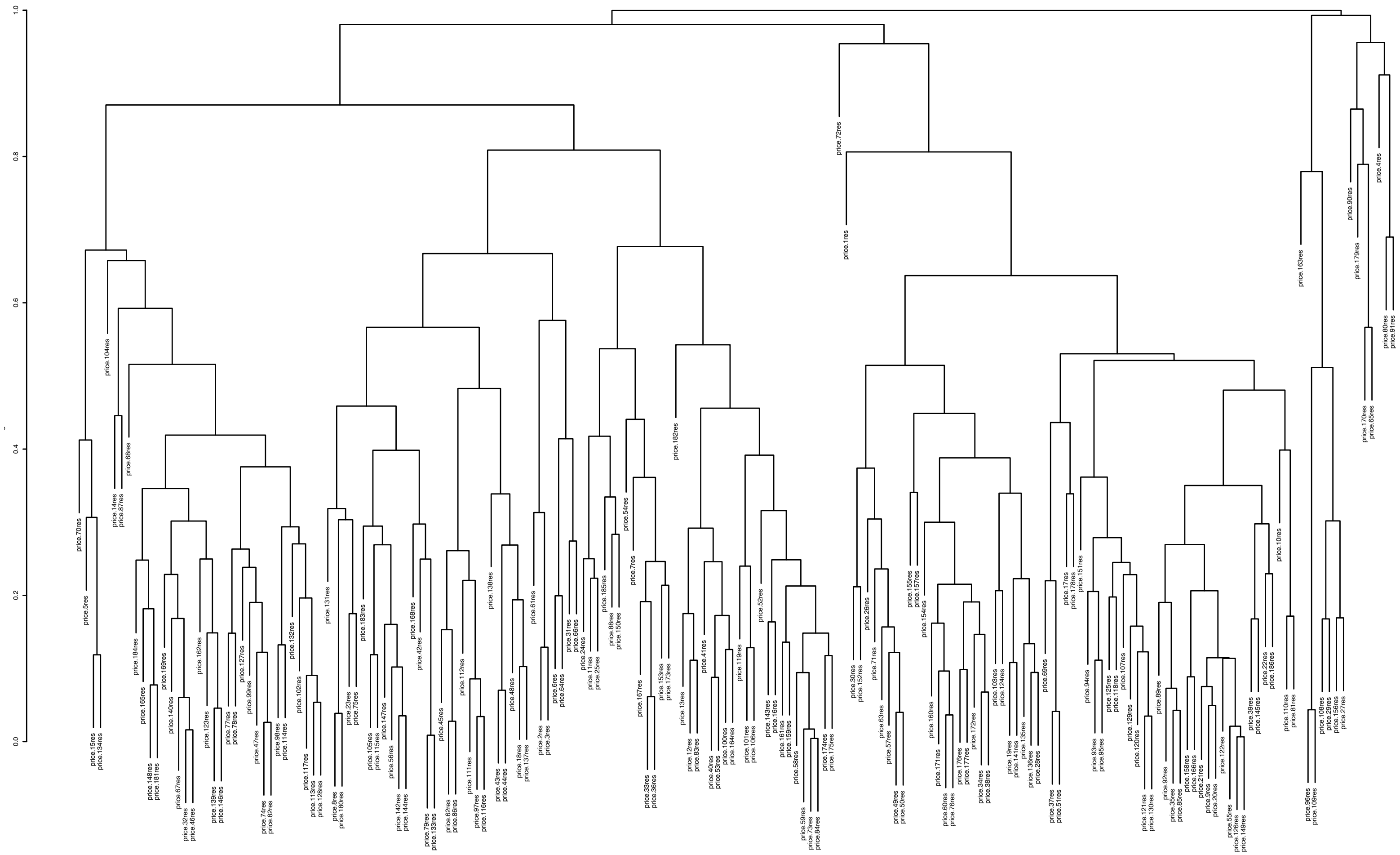
price setting over the day. We use the maximum and minimum prices of a day as limits of these cycles. In addition, to account for the cyclical behavior, we use a weighted daily average price (the longer a price is active on a day, the more weight it gets). Therefore, the weighted average price seems to be most suitable to depict the price reaction of gasoline stations.

Figure 2 shows the hierarchical clustering of the weighted daily average prices for the gasoline stations in our chosen cluster. This tree-like graphic should be read from the bottom to the top due to the agglomerative approach. On the very bottom of the dendrogram, each filling station has its own 'leaf'. Thus, at height equal to zero there are 186 individual clusters. The height is depicted on the far left of figure 2 and is nothing else than the dissimilarity measure defined previously: $d_{ij} = 1 - |cor_{ij}|$. As soon as we move up the tree, some of those leafs begin to fuse into branches. For instance, at a height of 0.0044 the two stations with the id's 73 and 84 form a cluster. This means that the adjusted price series ϵ_{73} and ϵ_{84} show an empirical correlation of 0.9956. These two filling stations are two discount stations in the middle of an industrial area. It is therefore likely that these two stations compete with each other. If we now make a cut at a height of 0.0044, then we no longer obtain 186 individual clusters, but only 185 clusters, since two have merged. In general it can be said that the earlier two clusters merge on the way up, the more similar they are. On the other hand, clusters that fuse near the top of the dendrogram tend to be rather different. One important feature of hierarchical clustering is that once clusters have been formed according to a dissimilarity measure, they can no longer be changed. They can only be merged with other clusters as a whole.

In order to investigate the question of geographically relevant markets, a suitable cutoff value has to be determined. In that context Cartwright, Kamerschen, and Huang (1989) argue that „[...] a correlation coefficient of .5 or higher is consistent with the qualitative statements that are made about market definition". Therefore we choose two cutoff values. One slightly stricter threshold of $\bar{d}_{ij} = 0.4$, which corresponds to a correlation coefficient of 0.6, and a slightly more relaxed threshold of $\tilde{d}_{ij} = 0.6$, which is equivalent to a correlation of 0.4. Since it would be very difficult and incomprehensible to identify the individual clusters in a geographical representation of all 186 filling stations, we have only depicted the largest clusters of the sample. Figure 3 illustrates the seven largest clusters that form at a threshold value of $height = 0.4$, i.e. a correlation of 0.6. In total, the hierarchical clustering with a cutoff value at height 0.4 results in 42 clusters, whereby the seven largest clusters contain 97 of the 186 filling stations of the sample.

Even with this rather strict threshold, there are clearly defined geographical markets. Cluster 14, for example, is the geographically relevant market for retail gasoline in and around Ludwigsburg, whereas cluster 19 is clearly centered on Filderstadt. The two clusters 7 and 26 seem to be less clearly distinguished here. This could be due to commuter routes, since both clusters tend to cover more rural areas. The region around the city of Stuttgart is characterized by heavy commuter traffic from smaller cities. The stations in cluster 26 confirm this as they are all located around the commuter routes which seem to play a role for the competitive situation in the gasoline market.

Figure 2: Dendrogram Daily Average Prices



Note: Result of hierarchical clustering

Figure 3: Map: Daily Average Prices – cut at 0.4

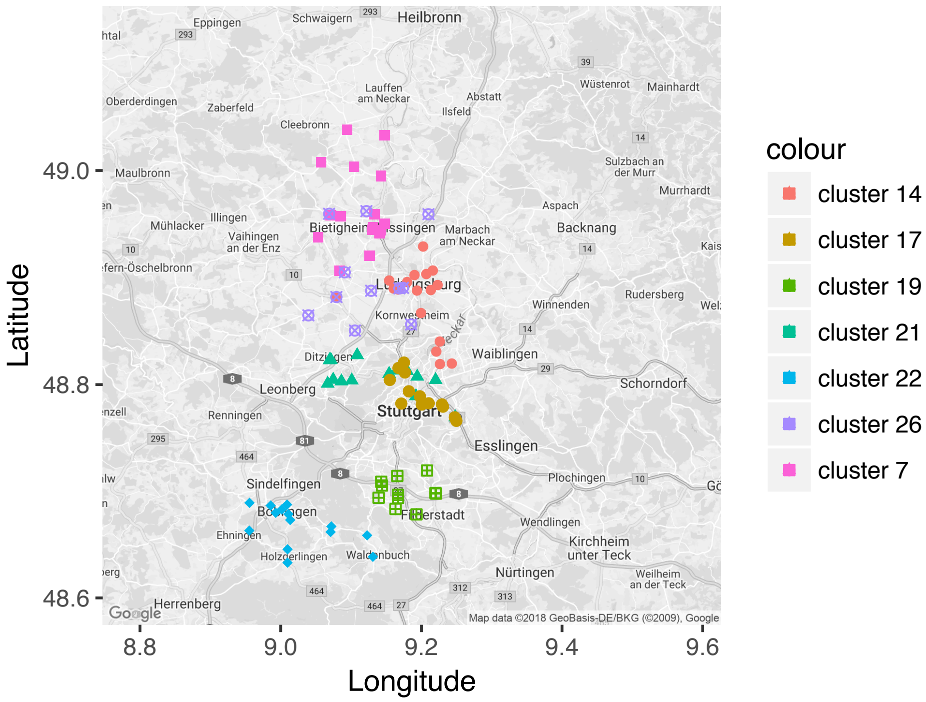


Figure 4: Map: Daily Average Prices – cut at 0.6

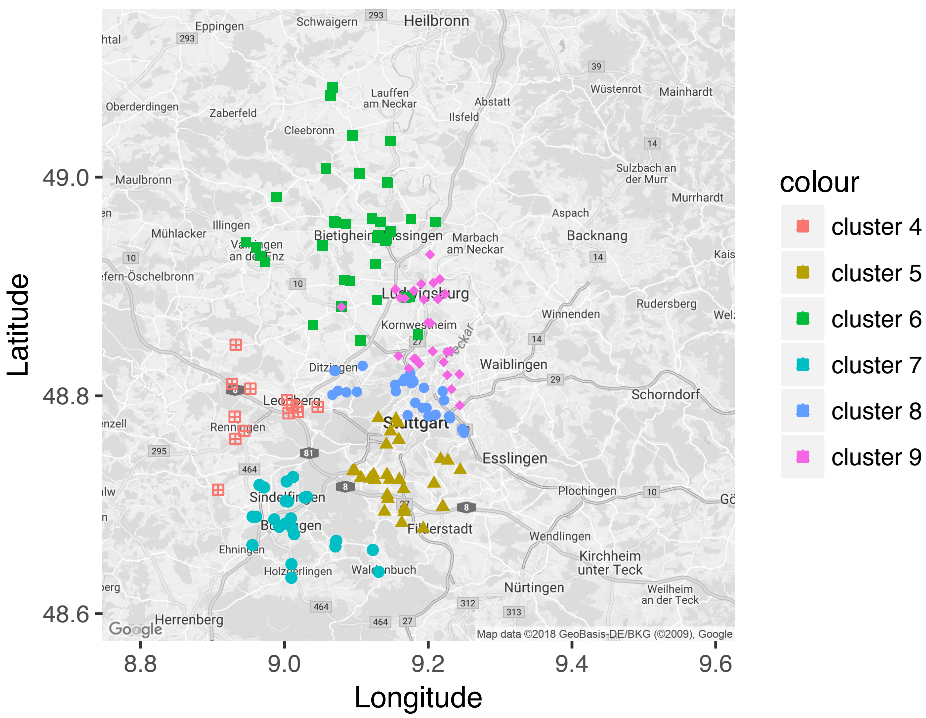


Figure 4 graphically illustrates the situation after loosening the threshold value from 0.4 to 0.6 and thus to a correlation of 0.4. The relaxation of the cutoff value leads to even more clearly defined clusters. For reasons of clarity, only the filling stations that belong to the six largest clusters are depicted. These six geographic groups contain 156 of the 186 stations of the sample. As mentioned above, clusters can no longer be changed once they have been formed. This can be seen at the transition from figure 3 to figure 4. In the former, the two clusters 7 and 26 form two distinct clusters. By loosening up the threshold value, clusters 7 and 26 are fused into cluster 6. However, this cluster still shows the same pattern than the two distinct clusters before and contains the commuter cities around Stuttgart and Ludwigsburg.

The fusion of two clusters into one big cluster when loosening the threshold value reveals an important characteristic of the retail gasoline market. The geographical markets for gasoline stations have certainly no strict boundaries, but are rather overlapping. If we allow for a lower correlation coefficient, larger clusters are formed where some stations will clearly have only limited significance for the gasoline stations in the center of the cluster (where correlation between the prices is high). However, for the purpose of a market definition as part of a merger analysis, smaller clusters seem to be more appropriate as we are interested in identifying the strongest competitors which are able to affect the pricing of the concerned parties (gasoline stations).

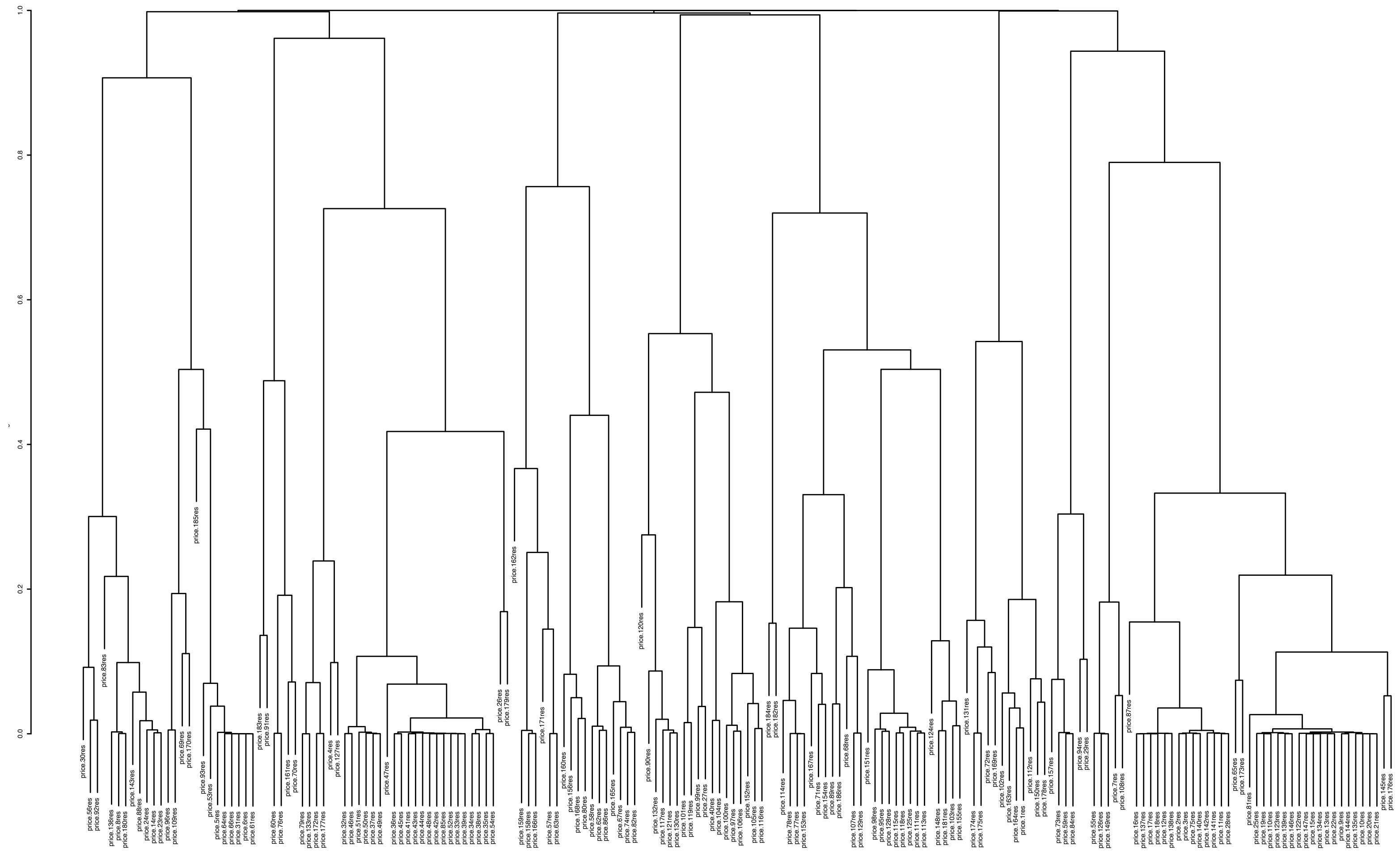
In summary, it can be said that the geographically relevant markets for retail gasoline in the sample tend to be rather small and therefore only the geographical neighbors seem to be close competitors. This is perfectly in line with our hypothesis of small regional markets. Furthermore, we do not have to worry about „special“ filling stations, such as motorway rest stops, because such stations are recognized by the method and are treated accordingly. For instance, the motorway filling station 'Sindelfinger Wald' is isolated in a separate cluster for both variants, a cut at height 0.6 and 0.4, respectively. This is consistent with the general opinion that motorway filing stations form a separate market and do not compete with the other stations.

In the next step, the adjusted daily maximum prices are used as the basis for hierarchical clustering with the complete linkage method. The results of the clustering process is again illustrated as a dendrogram in figure 5. Compared to the results for the weighted daily average prices, the dendrogram looks much more consistent in the sense that most clusters can be fused at the very bottom of the graph (which corresponds to a higher correlation value).

Since the daily highs and lows represent the limits of the Edgeworth-cycles and are not necessarily the long-term results of a competitive situation, it is sufficient to look at the somewhat stricter cutoff height of 0.4. The five largest clusters, resulting from the threshold value of 0.4 are depicted in figure 5. It is very obvious that the individual filling stations were not combined geographically, but rather by brand. The clusters 1 and 5 consists almost exclusively of Shell stations, whereas the clusters 13 and 15 consist mainly of ARAL stations. Therefore, the daily maximum prices for E5-gasoline seem to be set uniformly across the stations of a brand.

Boehnke (2017) reaches a similar conclusion by using k-means clustering in order to

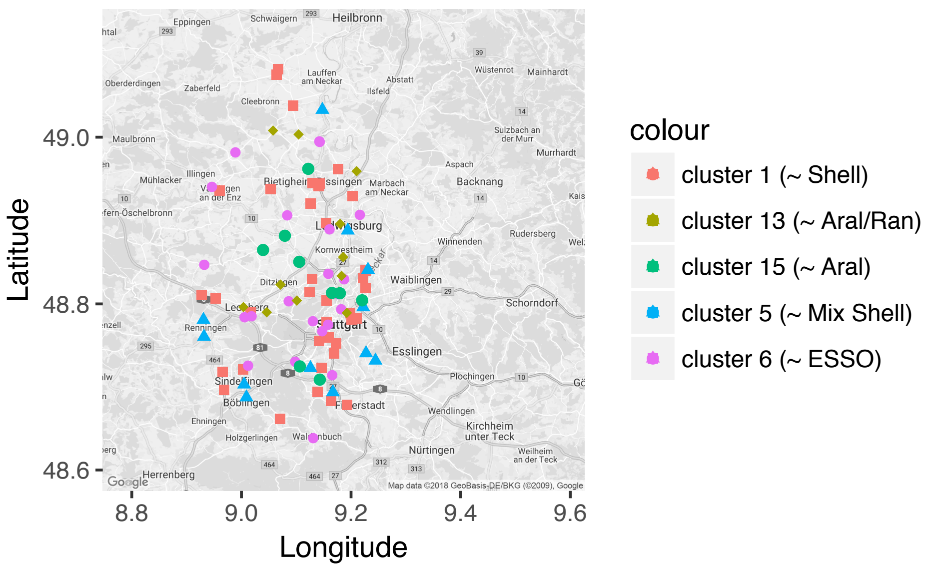
Figure 5: Dendrogram Daily Maximum Prices



Note: Result of hierarchical clustering

detect common pricing behavior between the stations. The hourly prices of gasoline stations in Germany are used to define clusters including stations with a similar pricing pattern. He identifies that three of his five clusters contain a dominant number of gas stations from one network only (Aral, Shell and Esso, respectively) and concludes that these brands exhibit a network-wide pricing behavior (prices appear to be set centrally by a network). This is also in line with the findings of the sector inquiry of the FCO (Bundeskartellamt (2011)) which concludes that many oil companies quote a maximum price for their filing stations. A short descriptive analysis as well reveals that stations of the same brand increase their prices on the same level, which especially holds for distinct regional markets (see also Linder (2018) or Dewenter, Bantle, and Schwalbe (2018)). Therefore, in contrast to the daily average price which results in geographical markets, a hierarchical clustering based on the maximum price reveals the uniform behavior of the branded stations with regard to their daily price increases.

Figure 6: Map: Daily Maximum Prices – cut at 0.4



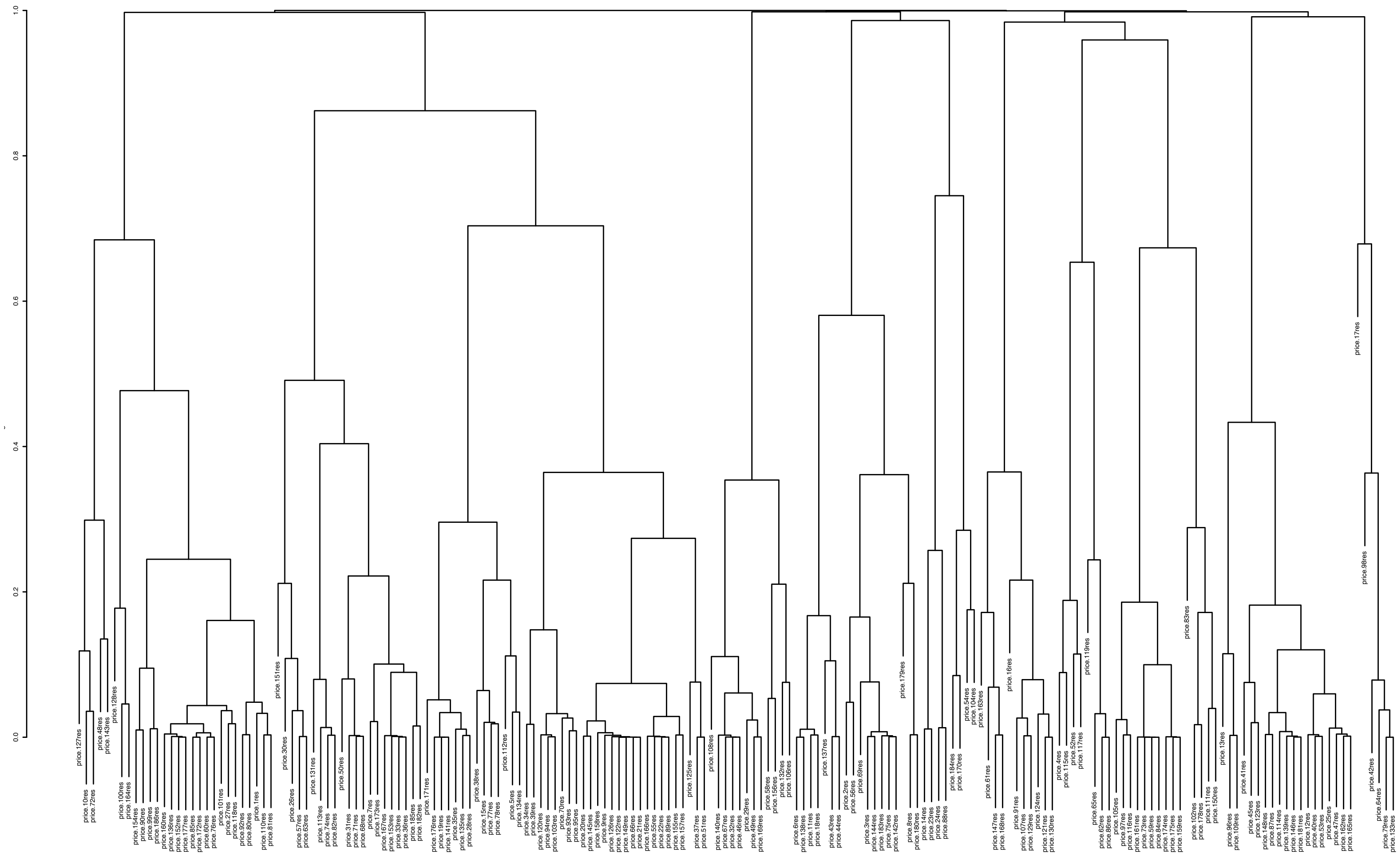
In the last step, the adjusted daily minimum prices are used as the basis for hierarchical clustering. Again the complete linkage method is applied to define the dissimilarity of clusters with more than one member. The results of the clustering process is illustrated as a dendrogram in figure 7. This dendrogram looks not as consistent as the one with daily highs but it is still more balanced than the one with the weighted average prices. As before, we are only looking at the slightly stricter threshold value of 0.4.

The six largest clusters of hierarchical clustering with the daily lowest prices and a cuttoff value of 0.4 are shown in figure 8. Contrary to the analysis of the daily maximum prices, the filling stations are not merged across brands. The geographical clustering, however is not as clear as in the case of the weighted daily average prices. One could argue that the formation of clusters here is largely based on commuter traffic. But this is certainly the most inaccurate clustering result of all three procedures.

Considering the price cycles in the German gasoline market, it becomes apparent, that the daily minimum price is only active for a very short time. First of all, the undercutting phase of the cycles is much more longer than the increasing phase. Moreover, gasoline stations earn a much lower profit during periods of low prices and are interested in increasing prices again. In contrast, the daily maximum price is active during a longer period of time as it is more profitable for the gasoline stations. Gasoline stations increase their prices in the evening to the maximum price and this price stays active during the night until the next morning. On this high level, the decreasing phase starts again on the next day.

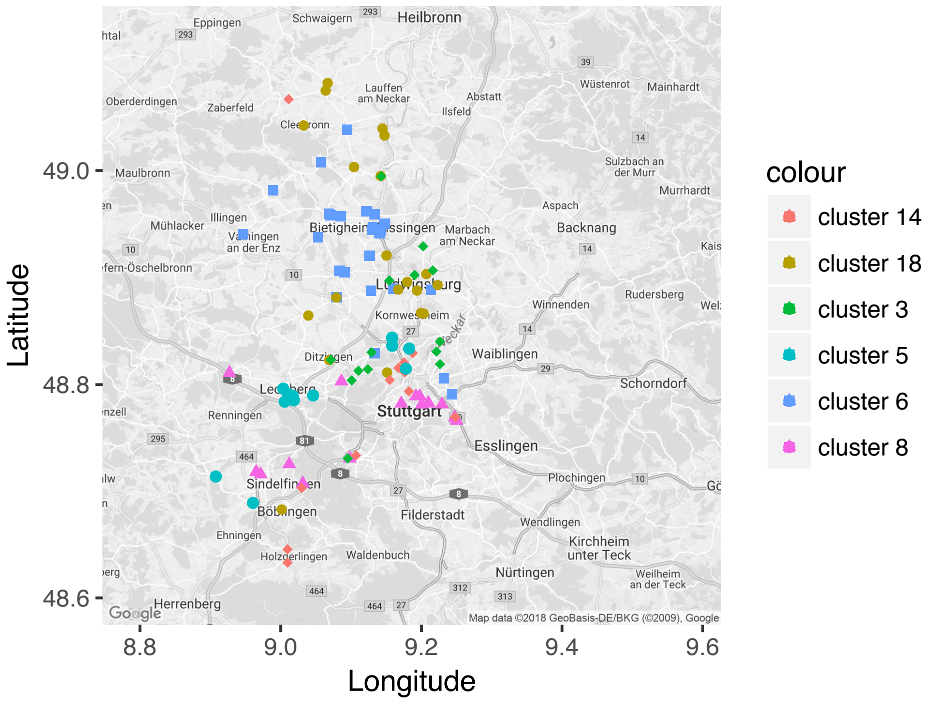
In summary, our analysis of the weighted daily average prices indicates relatively small geographic relevant markets for retail gasoline. Accordingly, the geographical k-means clustering at the beginning of the analysis has no implications for the final results and is not restrictive. The geographical markets are much smaller and the competitive situation more complex and is highly influenced by the demand behavior and the infrastructure. It is obvious that regional factors have an impact on the competitive situation. Especially commuter routes and thereby commuters have a strong influence on the price setting behavior of gasoline stations and determine the price reactions. By analyzing the daily maximum prices, the prices seem to be set uniformly across all stations of a brand. The daily minimum prices on the other hand, seem to depend on a mixture of commuter routes and the brand of a station. However, it should also be noted that the minimum prices are of limited informative value, since it is possible that those prices are only valid for one minute per day, as outlined above. The results based on the maximum prices show the centralized pricing of the gasoline brands for their gasoline stations where the companies seem to dictate the price increases. One could possibly argue that the results of the analysis with the weighted daily average prices are precisely the consequence of the uniform pricing of filling stations across brands and the demand-driven price competition due to commuter routes. Our results show that both, the chain of substitution approach as well as the accessibility model used by the FCO do not properly depict the competitive situation in the German gasoline market which is much more complex.

Figure 7: Dendrogram Daily Minimum Prices



Note: Result of hierarchical clustering

Figure 8: Map: Daily Minimum Prices – cut at 0.4



5 A Comparison with the Market Definition based on the Accessibility Model

In order to affirm our results presented in chapter 4.3 and outline the problems when using the accessibility model applied by the FCO in several cases, we compare our approach with one of the market definitions carried out during a merger analysis by the FCO. One of the main difficulties with the accessibility model is that it neglects commuter routes and thereby ignores an important determinant for the price setting behavior of gasoline stations. However, in the previous section, we found out that commuters determine which filling stations react to each other as price reactions of stations along a commuter route are co-moving.

The accessibility model applied by the FCO has already been briefly described in chapter 3.1. A detailed description of the procedure for the market delineation by the FCO is now explained based on the case Total Deutschland GmbH/OMV Deutschland GmbH (B8-175-08).[14] Information about the final market definition and the exact procedure of the delineation is quite different between the various merger cases of the FCO, whereby the provision of information is in all cases rather sparse. We chose the case Total/OMV as the decision report still is the most informative.

In 2009 Total Deutschland GmbH notify its proposal to acquire 59 filling stations of the OMV Deutschland GmbH located in the federal states Saxony and Thuringia. About half of the filling stations are located in the cities Chemnitz, Dresden, Erfurt and Leipzig whereas the other stations are spread over the rural regions of the federal states concerned.

In general, the accessibility model determines those filling stations that can be reached by car within a certain travel time starting from the target (starting) station. For urban areas, the FCO assumes that consumers are willing to drive 30 minutes to an alternative station to refuel there at a cheaper price. For rural areas the FCO even determines a longer driving time of 60 minutes for relevant stations. In the present case, the FCO only examined the urban areas because it assumes a priori that competitive problems will arise in areas with a higher petrol station density. According to the accessibility model, one would have to determine the relevant market for each of the concerned stations in the case. As the procedure is time-consuming and a multitude of relevant markets should have been determined, the FCO abstains from defining relevant markets for each gasoline station separately, but defines four regional markets for the four cities mentioned above. The four relevant markets are therefore defined around the geographical city centers of Chemnitz, Dresden, Erfurt and Leipzig. With this approach the FCO identifies 104 stations in the relevant regional market Chemnitz, 86 stations for Dresden, 75 stations for Erfurt and 75 stations for Leipzig.[15] The gasoline stations are weighted according to

[14] See Bundeskartellamt (2009).

[15] The report does not include more detailed information on the location of these stations. Moreover, the exact numbers of different brands are not reported. Market shares are only available for Total and OMV. In Chemnitz Total has a market share of 10-15% as well as OMV. In Dresden Total has a market share of 17-22% and OMV has 4-9%. In Erfurt Total has a share of 18-23% and OMV 7-12 % which is comparable to the market shares in Leipzig.

their distance to the city center. On the basis of the market shares resulting from this market definition, the FCO has prohibited the proposed merger.

In the following, we choose a sample from our data set that contains the concerned region mentioned in the case report of the FCO. The FCO performs its market definition for each city separately resulting in four disjuncted geographic regions for the cities Chemnitz, Dresden, Erfurt and Leipzig. However, as we know from our analysis in chapter 4.3 that commuters play a crucial role and as the three cities Chemnitz, Dresden and Leipzig are geographically close together, we choose a sample that contains all three cities. Our sample therefore includes all filling stations up to a distance of 85 km from one of the three cities. Figure 9 shows the location of the 465 stations in our sample.

The composition of brands within the sample is illustrated in table 3. In particular the brands Aral (79 stations), Total (69), Shell (50), Star (41), and Esso (32) must be pointed out. However, Jet as part of the big five oligopolists is somewhat underrepresented in the given sample with only 16 stations.

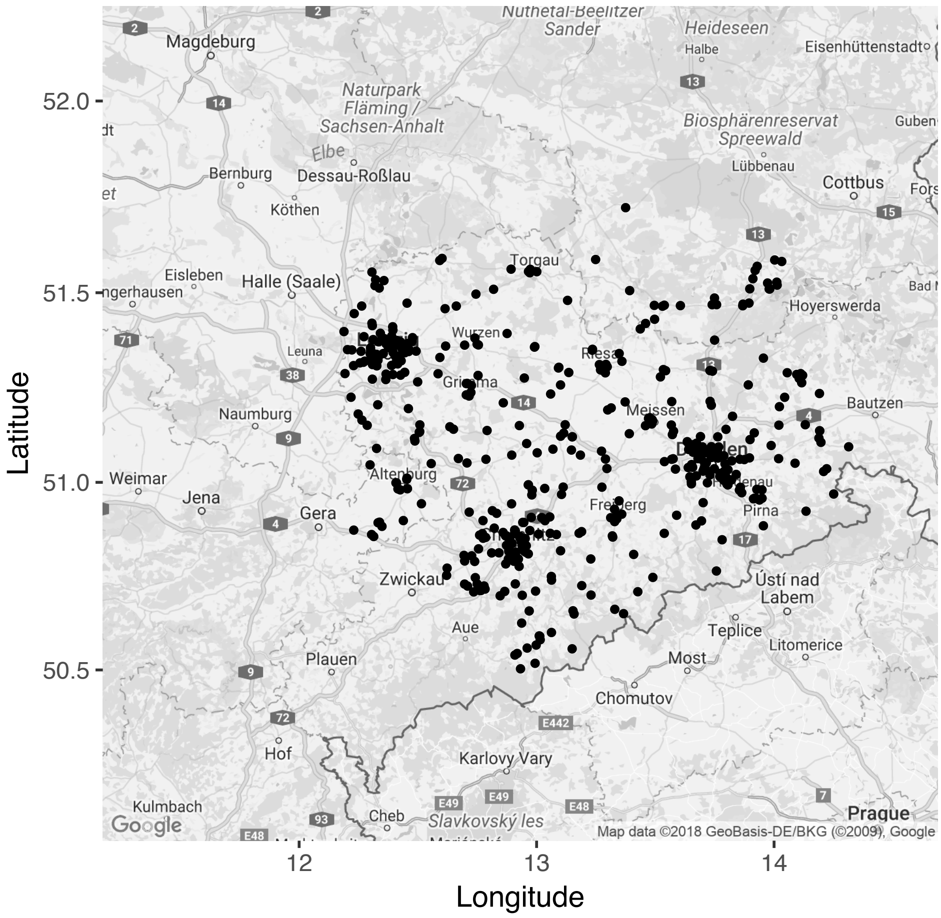
Table 3: Distribution of brands in the BKartA sample

Brand	No. of stations
ARAL	79
TOTAL	69
Shell	50
Star	41
ESSO	32
HEM	19
JET	16
Agip	13
Others	146
N	465

Note: Number of filling stations per brand.

The joint consideration of the three cities allows us to investigate urban as well as rural areas with a single analysis. With the method presented in chapter 4.2 it is possible to run only one model for the whole sample. Therefore, we are able to take commuter routes into account and we can consider areas with low as well as high station density at the same time. Moreover, the approach is not time-consuming and can therefore also be applied by competition authorities during a merger investigation. For a detailed description of the method, please refer to chapter 4. Figure 10 shows the dendogram which presents the hierarchical clustering of the weighted daily average prices for the 465 gasoline stations in our sample. In contrast to chapter 4.3, for this case only results with the weighted average price are presented below. As outlined above, the maximum price leads to a clustering based on the different brands and reflects the centralized price setting of the oil companies. The minimum price is not that meaningful as it is active

Figure 9: Sample of stations (BKartA)



Note: All 465 filling stations within the sample.

only for a very short period. The weighted daily average price, however, is the most appropriate for the analysis and represents the cyclical behavior of the gasoline stations.

The height is depicted on the far left of figure 10 and is nothing else than the dissimilarity measure defined previously: $d_{ij} = 1 - |cor_{ij}|$. Considering the clusters formed in the dendogram, we choose a slightly stricter threshold of $d_{ij} = 0.4$, which corresponds to a correlation coefficient of 0.6. Besides that, a higher correlation coefficient coincides with our intent to identify the closest competitors as the relevant market should contain the closest substitutes. Especially for merger analysis it seems to be appropriate to focus on those rivals which really have an influence on the pricing of the concerned companies. With a threshold value of 0.4 ninety distinct clusters are identified. There is a number of clusters which contain only a few stations (or even only one). These stations are either highway filing stations which obviously have a different price setting behavior and form a separate market or, on the other hand, there are smaller clusters especially in the rural areas where the station density is very low.

Figure 11 looks more closely on the 20 largest clusters that form at a cutoff value of $height = 0.4$, i.e. a correlation of 0.6. As these are regions with a higher station density, we expect that competition problems are more probable in these areas. Moreover, for the present paper an illustration of all clusters in one map would not be very informative. However, in a merger analysis one could easily look at all clusters. The 20 largest clusters include 320 gasoline stations of our 465 stations in the sample. Where the largest cluster contains 43 filing stations and the smallest 9.

At first, it can be noted that a multitude of relevant geographic markets are identified for the geographical area around Chemnitz, Dresden and Leipzig. Many clusters can be identified for each city.[16]

Taking into account the findings of our previous analysis in chapter 4.3, we assume that commuter routes play a major role for the price responses of petrol stations. Unfortunately, we do not have access to specific statistics on commuter flows in this area which could be used directly. However, we use data from the automatic road traffic counting of the „Bundesamt für Straßenwesen“.[17] The automatic counting points record all cars driving on the monitored „Bundesstraßen“ and „Autobahnen“. We are looking at routes with a traffic volume which is above-average. The identified streets are used as a proxy for commuter routes and are highlighted in figure 12. The routes marked in figure 12 all count between 9,000 and 46,000 vehicles per day on average and are therefore appropriate as a proxy for commuter routes. We scrutinised our assumption by investigating the reports of the "Institut für Arbeitsmarkt- und Berufsforschung" which carries out regular studies on commuter flows. The findings of these reports are consistent with the routes identified with the data of the traffic census. Therefore, the traffic-intensity measured by the counting stations offers a good possibility to make ex ante assumptions about commuter flows.

[16] Unfortunately the city names are partly difficult to read due to the high number of gasoline stations located in the cities. However, the map is consistent with the one in figure 9 which can serve as orientation if needed.

[17] `https://www.bast.de/BASt_2017/DE/Verkehrstechnik/Fachthemen/v2-verkehrszaehlung/zaehl_node.html`

Figure 10: Dendrogram FCO–Sample

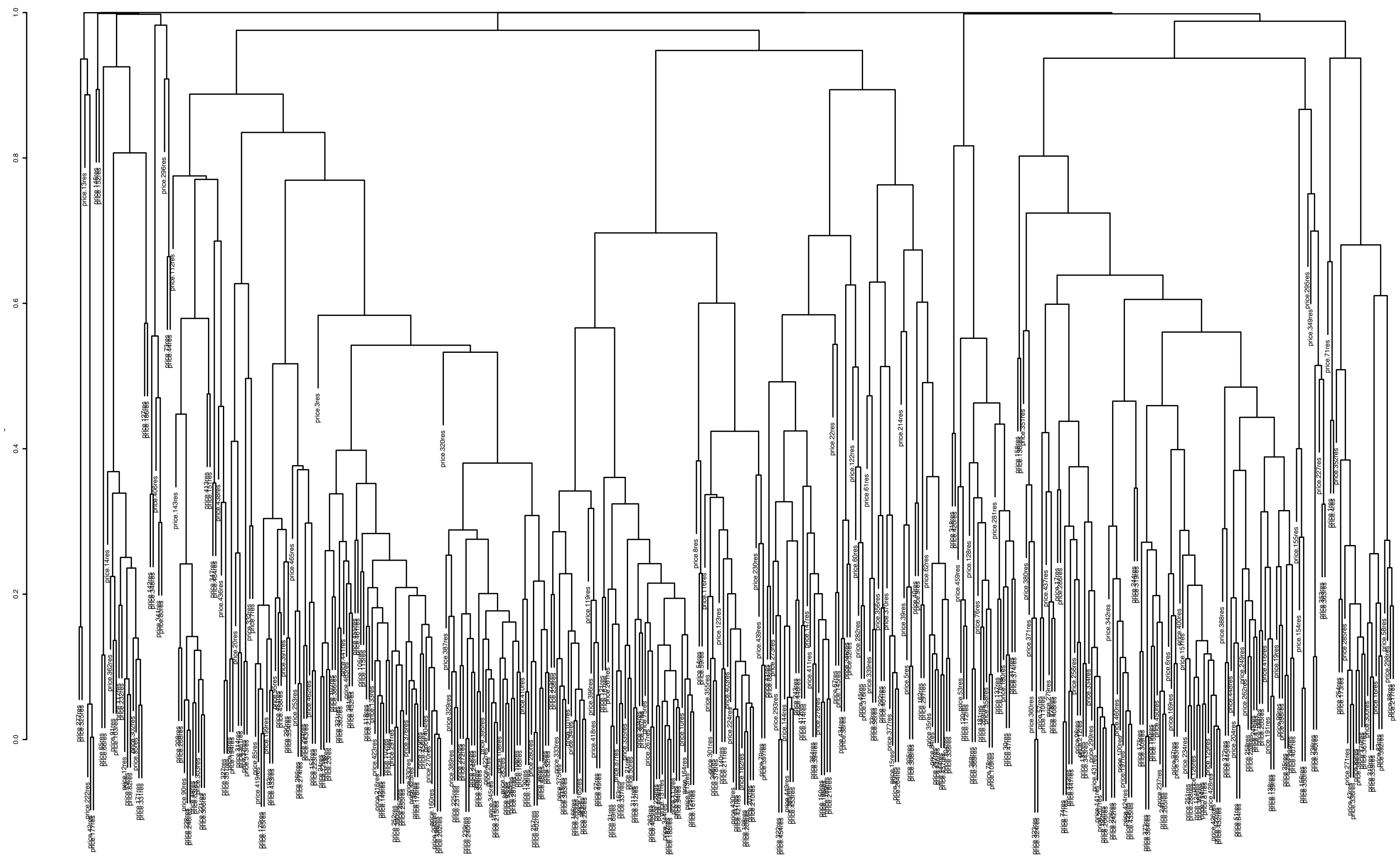


Figure 11: Sample of stations (BKartA)

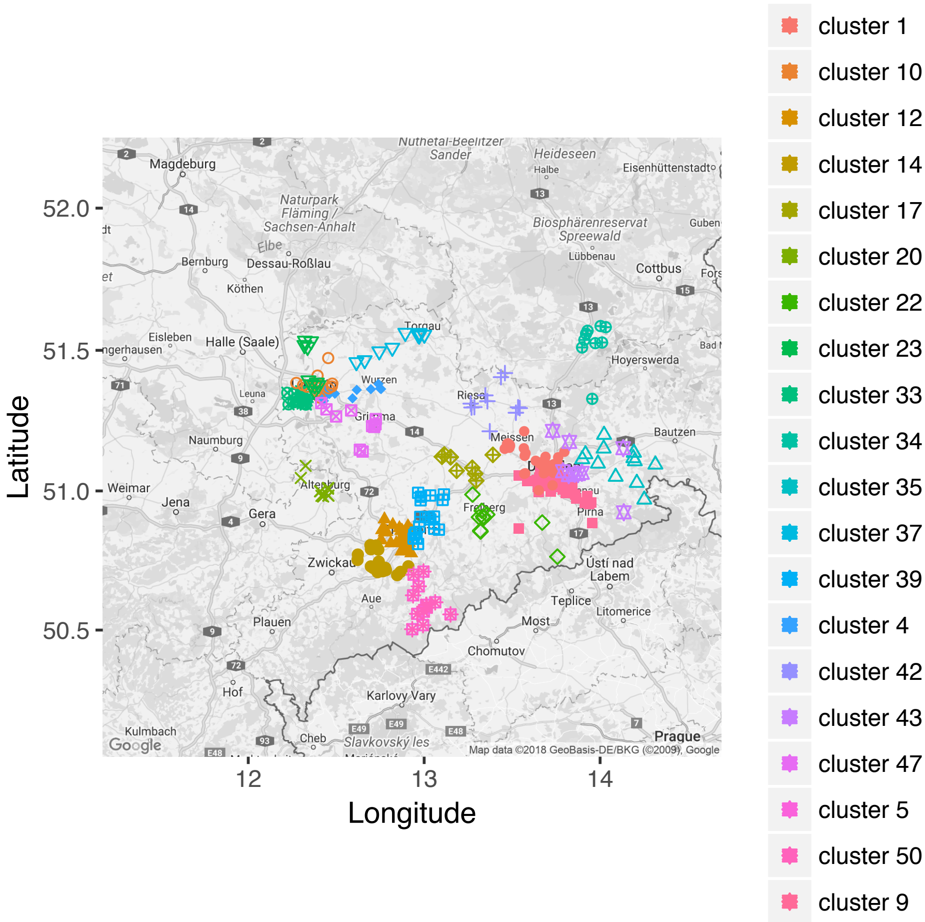
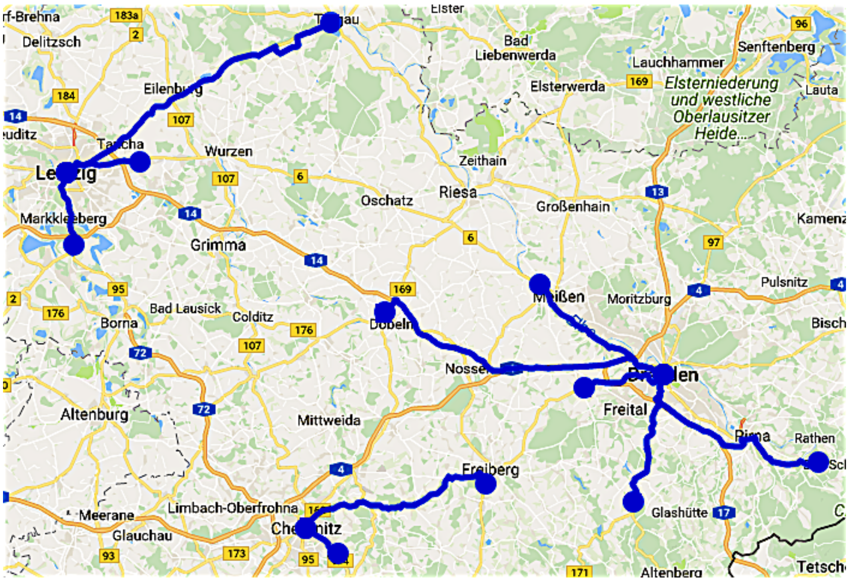


Figure 12: Commuter Routes in Saxony



Comparing the commuter routes highlighted in figure 12 with the results of the hierarchical clustering in figure 11, one can clearly see that several clusters coincide with the commuter routes. If we look at Chemnitz for example (bottom center in figure 11), gasoline stations along the commuter flows from city districts and from nearby smaller cities are clustered together. Another very clear example is the route from Leipzig to Torgau (top-centre of figure 11 and top-left in figure 12). The filling stations along this commuter route are clustered together and demonstrate a very high correlation of their price setting behavior. These results confirm our findings in chapter 4.3 and verify our assumption that commuters are an important determinant for the price setting of gasoline stations as price reactions are triggered by commuter routes. In contrast to the presented approach, the FCO only looks at the competitive situations in the cities. The risk of neglecting commuter routes is therefore high which could possibly lead to wrong conclusions about the relevant rivals for the concerned gasoline stations.

The FCO takes the demand behavior of an individual car driver as a basis for its market definition. However, a driving time around a target (starting) station of 30 minutes does not seem to reflect the actual demand behavior of car drivers. Most car drivers refuel during their travel from work to their home place, which is as well stipulated by the FCO. It does not seem to be a reasonable behavior of a consumer to drive 30 minutes to a cheaper station starting from the target station. Moreover, gasoline stations are not able to observe individual (local) consumers. It seems to be more reasonable that the pricing of gasoline stations is influenced by general fluctuations in demand like commuter flows which can be observed by gasoline stations and can be taken into account in the price setting. The timing of the price cycles further confirms this presumption. Gasoline stations increase prices in the evening when commuter traffic is as well increasing (see for example Dewenter, Bantle, and Schwalbe (2018)) which also shows that commuters are a driving force for the price reactions of gasoline stations. As commuters know the stations and their prices along their route, these gasoline stations will react to each other in their pricing and thereby belong to one market.

6 Conclusion

We combine the simple and intuitive concept of price correlations and hierarchical clustering, a method from the rich toolkit of unsupervised learning, to a new and efficient price test in order to define geographically relevant markets for retail gasoline in Germany and identify the closest competitors.

The method used is very attractive, especially for antitrust authorities, since one can extract any number of clusters from one single dendrogram based on the preferred correlation coefficient. As we are interested in the closest competitors, in most cases a higher correlation coefficient is recommended. Further advantages of hierarchical clustering with price correlations as a dissimilarity measure are that the method is easy to use and very intuitive. Moreover, we do not have to worry about special filling stations, such as motorway rest stops as these are identified as separate markets by our method.

Our analysis reveals that the relevant geographic markets are rather small. Moreover, commuter routes play a crucial role for the competitive behavior in the German retail gasoline market. Gasoline stations cannot observe the refuel behavior of individual consumers as assumed by the FCO. Rather, it is the case that the price reactions are strongly affected by commuter routes along which gasoline stations are competing. The results show that the competitive situation is complex and cannot be captured either by the chain of substitution model or by the accessibility model applied by the Federal Cartel Office. As the accessibility model neglects commuter routes, conclusions based on this approach could be misleading. Gasoline stations cannot observe the behavior of individual consumers which live or work near the station. However, they are able to observe commuting traffic and commuter routes. Our analysis reveals that besides local market conditions, these frequently used routes and thus the commuters are a driving force in the price setting of gasoline stations. Stations along these routes react strongly to price changes of their neighbouring stations.

References

Audy, E. and C. Erutku (2005). "Price Tests to Define Markets: An Application to Wholesale Gasoline in Canda". In: *Journal of Industry, Competition and Trade* 5.2, pp. 137–154.

Bishop, S. and M. Baldauf (2006). *Theoretische Grundlagen und praktische Anwendung wettbewerbsökonomischer Methoden in Bezug auf die Abgrenzung des relevanten Marktes und Fragen zur praktischen Anwendbarkeit des Herfindahl-Hirschman Indexes zur Ermittlung des Konzentrationsgrades*. RBB Economics.

Bishop, S. and M. Walker (2010). *The economics of EC competition law*. Sweet & Maxwell.

Boehnke, J. (2017). "Pricing Strategies and Competition - Evidence from the Austrian and German Retail Gasoline Markets". In: *Harvard University, Center of Mathematical Sciences and Applications*.

Boshoff, W. H. (2012). "Limits and Uses of Price Tests for Market Definition". In: *Stellenbosch Economic Working Papers* 01/11.

Bundeskartellamt (2009). *Total Deutschland GmbH/OMV Deutschland GmbH (B8-175-08)*. Bonn.

— (2011). *Sektoruntersuchung Kraftstoffe, Abschlussbericht*. Bonn.

Cartwright, P., D. Kamerschen, and M-Y. Huang (1989). "Price correlation and granger causality tests for market definition". In: *Review of Industrial Organization* 4.2, pp. 79–98.

Church, J. and R. Ware (2000). *Industrial Organization, A Strategic Approach*. Irwin McGraw-Hill.

Dewenter, R., M. Bantle, and U. Schwalbe (2018). "German Gasoline Retail Markets: Price Discrimination or Tacit Collusion?" mimeo.

Elzinga, K.G. and T. F. Hogarty (1983). "The problem of geographic market delineation in antimerger suits". In: *The Antitrust Bulletin*, pp. 45–81.

Energie Informationsdienst (2017). *Anzahl der Straßentankstellen in Deutschland seit 1964*. URL: `http://www.eid-aktuell.de/inhalt/statistiken/excel-download-strasentankstellen-in-deutschland/`.

Engle, R. and W. Granger (1987). "Co-Integration and Error Correction: Representation, Estimation, and Testing". In: *Econometrics* 55.2, pp. 251–276.

Forni, M. (2004). "Using Stationary Tests in Antitrust Market Definition". In: *American Law and Economics Review* 6.2, pp. 441–464.

Franz, W., H. Ramser, and M. Stadler, eds. (2002). *Fusionen*. 1st ed. Vol. 31. Mohr Siebeck, Tübingen. URL: `https://EconPapers.repec.org/RePEc:mhr:wiseot:urn:isbn:9783161477614`.

Friedman, J., T. Hastie, and R. Tibshirani (2001). *The elements of statistical learning*. Vol. 1. Springer series in statistics New York.

Genesove, D. (2004). "Comment on Forni's "Using Stationarity Tests in Antitrust Market Definition"". In: *American Law and Economics Review* 6.2, pp. 476–478. eprint:

/oup/backfile/content_public/journal/aler/6/2/10.1093/aler/ahh009/2/ahh009.pdf. URL: +%20http://dx.doi.org/10.1093/aler/ahh009.

Granger, C. (1969). “Investigating causal relations by econometric models and cross-spectral methods”. In: *Econometrica: Journal of the Econometric Society*, pp. 424–438.

Harris, B. and J. Simons (1989). “Focusing market definition: How much substitution is necessary?” In: *Research in law and economics*, pp. 207–226.

Hartigan, J. and M. Wong (1979). “Algorithm AS 136: A k-means clustering algorithm”. In: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28.1, pp. 100–108.

Horowitz, I. (1981). “Market definition in antitrust analysis: a regression based approach”. In: *Southern Economic Journal*, pp. 1–16.

Linder, M. (2018). “Price cycles in the German retail gasoline market - Competition or collusion?” In: *Economics Bulletin* 38.1, pp. 593–602.

Massey, P. (2000). “Market definition and market power in competition analysis: some practical issues”. In: *The Economic and Social Review* 31.4, pp. 309–328.

Pesaran, M.H., Y. Shin, and R.J. Smith (2001). “Bounds Testing Approaches to the Analysis of Level Relationships”. In: *Journal of Applied Econometrics* 16, pp. 289–326.

Sims, C. (1972). “Money, income, and causality”. In: *The American economic review* 62.4, pp. 540–552.

Slade, M. (1986). “Exogeneity tests of market boundaries applied to petroleum products”. In: *The Journal of Industrial Economics*, pp. 291–303.

Stigler, G. and R. Sherwin (1985). “The Extent of the Market”. In: *Journal of Law and Economics* 28.3, pp. 555–85. URL: https://EconPapers.repec.org/RePEc:ucp:jlawec:v:28:y:1985:i:3:p:555-85.

Tirole, J. (1988). *The theory of industrial organization.* MIT press.

Warell, L. (2005). “Defining Geographical Coal Markets Using Price Data and Shipments Data”. In: *Energy Policy* 33.17, pp. 2216–2230.

2018

179 Dewenter, Ralf; Linder, Melissa; Thomas, Tobias: Can Media Drive the Electorate? The Impact of Media Coverage on Party Affiliation and Voting Intentions, April 2018

2017

178 Beckmann, Klaus: Bounded rationality in differential games, December 2017

177 Herzer, Dierk; Nagel, Korbinian: The effects of adult and non-adult mortality on long-run economic development: Evidence from a heterogeneous dynamic and cross-sectionally dependent panel of countries between 1800 and 2010, July 2017

176 Dewenter, Ralf; Heimeshoff, Ulrich; Löw, Franziska: Market Definition of Platform Markets, March 2017

2016

175 Dewenter, Ralf; Dulleck, Uwe; Thomas, Tobias: Does the 4th estate deliver? Towards more direct measure of political media bias, November 2016

174 Luik, Marc-André: Child Health, Human Capital and Adult Financial Behavior, November 2016

173 Michael Berlemann; Marc-André Luik: Institutional Reform and Depositors' Portfolio Choice - Evidence from Bank Account Data, November 2016

172 Lauenstein, Philipp; Küster Simic, André: Information Processing in Freight and Freight Forward Markets: An Event Study on OPEC Announcements, September 2016

171 Nagel, Korbinian: A Life Course Perspective on the Income-to-Health Relationship: Macro-Empirical Evidence from Two Centuries, July 2016

170 Dluhosch, Barbara; Horgos, Daniel: International Competition Intensified - Job Satisfaction Sacrified?, June 2016

169 Beckmann, Klaus; Dewenter, Ralf; Thomas, Tobias: Can news draw blood? The impact of media coverage on the number and severity of terror attacks, May 2016

168 Afflatet, Nicolas: Deficit Policy within the Framework of the Stability and Growth Pact - Empirical Results and Lessons for the Fiscal Compact, April 2016

167 Werner, Max: Evaluating Prediction Markets for Internal Control Applications, May 2016

166 Werner, Max; Eißing, Klaus; Langton, Sebastian: Shared Value Potential of Transporting Cargo via Hyperloop, May 2016

165 Werner, Max; Vianelli, A.; Bodek, Mariusz C.: Monitoring Venture Capital Investments through Internal Control Prediction Markets, May 2016

164 Jahn, Vera; Steinhardt, Max Friedrich: Innovation and Immigration - Insights from a Placement Policy, February 2016

163 Beckmann, Klaus; Gattke, Susan; Lechner, Anja; Reimer, Lennart: Lineare dynamische Konfliktmodelle: Ein systematischer Überblick, Februar 2016

162 Beckmann, Klaus; Gattke, Susan; Lechner, Anja; Reimer, Lennart: A critique of the Richardson equations, January 2016

2015

161 Dewenter, Ralf; Schwalbe, Ulrich: Preisgarantien im Kraftstoffmarkt, Oktober 2015

160 Afflatet, Nicolas: Fiscal Policy in a Debt Crisis - A Model, June 2015

159 Beckmann, Klaus; Gattke, Susan; Reimer, Lennart: The Boulding-Richardson Model Revisited, June 2015

158 Jahn, Vera: The Importance of Mittelstand Firms for Regional Apprenticeship Activity - Lessons for Policy, April 2015

157 Im Winkel, Niklas: Rechts? Links? Liberal? Egal? Gründe für die Entstehung verzerrter Medieninhalte und Methoden zur Messung des Bias, February 2015

156 Afflatet, Nicolas: Public Debt and Borrowing. Are Governments Disciplined by Financial Markets?, January 2015

2014

155 Berlemann, Michael; Christmann, Robin: Determintants of In-Court Settlements. Empirical Evidence from a German Trial Court, December 2014