

## Centrality and Pricing in Spatially Differentiated Markets: The Case of Gasoline

Weiss, Christoph; Pennerstorfer, Dieter; Firgo, Matthias

*Published in:*  
International Journal of Industrial Organization

*DOI:*  
[10.1016/j.ijindorg.2015.03.009](https://doi.org/10.1016/j.ijindorg.2015.03.009)

Published: 01/01/2015

*Document Version*  
Peer reviewed version

[Link to publication](#)

*Citation for published version (APA):*  
Weiss, C., Pennerstorfer, D., & Firgo, M. (2015). Centrality and Pricing in Spatially Differentiated Markets: The Case of Gasoline. *International Journal of Industrial Organization*, 40, 81 - 90.  
<https://doi.org/10.1016/j.ijindorg.2015.03.009>

# Centrality and Pricing in Spatially Differentiated Markets: The Case of Gasoline\*

Matthias Firgo<sup>†</sup>, Dieter Pennerstorfer<sup>‡</sup> and Christoph Weiss<sup>§</sup>

Accepted Manuscript

## Abstract

We highlight the importance of 'centrality' for pricing. Firms characterized by a more central position in a spatial network are more powerful in terms of having a stronger impact on their competitors' prices and on equilibrium prices. These propositions are derived from a simple theoretical model and investigated empirically for the retail gasoline market of Vienna, Austria. We compute a measure of network centrality based on the locations of gasoline stations in the road network. Results from a spatial autoregressive model show that prices of gasoline stations are more strongly correlated with prices of central competitors.

Keywords: Network Centrality, Spatial Differentiation, Gasoline Prices

JEL code: C21, D43, L11, L81, R12

---

\*Acknowledgments: The authors gratefully acknowledge helpful comments from Giuseppe Arbia, Harald Badinger, Robert D. Weaver, and the participants of the World Conference of the Spatial Econometrics Association (2010), the EBES 2010 Athens Conference, the International Industrial Organization Conference (2011), the Agricultural Food and Industrial Organization Seminar at the Toulouse School of Economics (2011), and the 38th Annual Conference of the European Association for Research in Industrial Organization (EARIE 2011). The work was generously supported by funds of the Oesterreichische Nationalbank (Anniversary Fund, project number: 12974)

<sup>†</sup>Corresponding author, Austrian Institute of Economic Research (WIFO), Vienna, Austria; E-mail: matthias.firgo@wifo.ac.at.

<sup>‡</sup>Austrian Institute of Economic Research (WIFO).

<sup>§</sup>Vienna University of Economics and Business (WU).

# 1 Introduction

In his seminal book *The Theory of Monopolistic Competition*, Chamberlin (1948) refers to the gasoline market as a prototype for what he calls ‘localized competition’. At the retail level consumers face transportation (time) costs when switching between gasoline stations; this introduces spatial product differentiation into an otherwise homogeneous product market. The importance of spatial product differentiation for market outcomes is typically investigated in economic models in the tradition of Hotelling (1929), and Salop (1979). The present paper examines an important dimension of spatial product differentiation and market power that has been widely ignored by the literature: the centrality of firms.

Centrality, defined as the extent to which agents are connected to other agents, is among the most fundamental concepts in the social network literature. In networks of agents connected via friendship, acquaintanceship, or professional links, researchers found centrality to be associated with an agent’s social status, power, and influence. In their description of a widely studied star-shaped network structure, for instance, Brass and Burkhardt (1992, p. 191) note that “most people would simply look at the diagram and declare [the central agent] the most powerful”.

Whether firms characterized by a more central position in a network unfolded in space are more powerful than other firms, however, has not yet been investigated in detail in economic models. The canonical model of spatial competition formulated by Salop (1979), for example, assumes that firms are distributed equidistantly (symmetrically) in a circular market. Per definition, the number of direct competitors (two adjacent neighbors) and the distances between them – and thus the extent to which firms are interconnected – is the same for all firms. Firms are different, but they are ‘equally different’ and spatially homogeneous so the specific location of an individual firm in space is irrelevant. This simplified assumption reduces the complexity of spatial models considerably but at the same time precludes an analysis of the importance of spatial heterogeneity and centrality for firms’ pricing behavior.

The aim of the present paper is to implement the concept of centrality in a simple theoretical model of spatial product differentiation and empirically investigate its importance for market power and firms’ pricing behavior. We compare price setting for ‘central’ and ‘remote’ firms in a modified version of Chen and Riordan’s (2007) spokes model.<sup>1</sup> By

---

<sup>1</sup>A few other studies also deviate from the traditional approach of a linear or circular market space and consider alternative spatial structures (e.g. Braid, 1989; Fik, 1991). Balasubramanian (1998), among others, studies a market in which consumers buy either from firms located at a Salop circle or from a firm located at the center of the circle. However, none of these studies provide empirical evidence of the importance of centrality for firm pricing and market performance.

analyzing the retail gasoline market in the metropolitan area of Vienna econometrically, we provide first empirical evidence of the importance of centrality for pricing.<sup>2</sup>

A key advantage of the retail gasoline market for this purpose is the fact that the concept of centrality is based on a definite and easy-to-visualize physical foundation: Gasoline stations are connected through a network of roads and intersections and can be characterized by different degrees of centrality (interconnectedness) within this network. More central stations (a) directly compete with more rivals and (b) are more important competitors for each of these rivals. Borrowing a measure of network centrality from the social network literature, we demonstrate that the correlation between gasoline stations' prices is significantly related to this measure, as predicted by our extended spokes model.

## 2 The Model

Following Chen and Riordan (2007), we describe the market as a set of  $N \geq 2$  spokes with a common core (the market center or the central intersection). Consumers are uniformly distributed along each spoke with unit density. When consuming at the location of a specific firm, their net utility equals the utility of the product ( $s$ ) minus the price charged ( $p$ ) and minus transportation costs that depend linearly on the distance consumers have to travel to the firm of their choice. We set per unit transportation costs to  $t = 1$ , without loss of generality.<sup>3</sup> The locations of firms are exogenously given and fixed. In contrast to Chen and Riordan (2007) we assume that each consumer attributes a value ( $s$ ) to all (not only two) firms (varieties) of a local market and we extend their model by allowing for firms to be located at heterogeneous distances ( $d$ ) from the center. There is always exactly one central firm ( $C$ ) and a finite number of  $1 \leq n \leq N - 1$  remote firms ( $R_i$ ), with  $i = 1, \dots, n$ . The central firm is the supplier closest to the market center, thus  $d_C < d_i, \forall i$ .<sup>4</sup> The lengths of the spokes hosting firms are denoted by  $l_i$  and  $l_C$ , whereas  $l$  is the length of all empty spokes (i.e. spokes not hosting a firm). Firms sell a spatially differentiated but otherwise homogeneous product ( $s_C = s_i = s$ ) at constant marginal costs ( $c_C, c_i$ ). Fixed

<sup>2</sup>Empirical research on competition in gasoline markets has mainly focused on the impact of spatial differentiation on prices and price dispersion (e.g. Netz and Taylor, 2002; Barron et al., 2004), on market concentration and the role of independent stations (e.g. Pennerstorfer, 2009; Houde, 2012; Pennerstorfer and Weiss, 2013), as well as on the existence of asymmetries in price adjustment (e.g. Borenstein et al., 1997; Verlinda, 2008) and Edgeworth price cycles (e.g. Eckert, 2003; Noel, 2007; Lewis, 2012). An excellent survey of this literature is available in Eckert (2013).

<sup>3</sup>In the framework of inelastic demand (as assumed below) and differentiated products limit equilibrium prices are infinite as  $t \rightarrow \infty$ . If  $t \rightarrow 0$  spatial differentiation (and space itself) becomes irrelevant. With  $t = 0$  the firm with the lowest marginal cost will serve the entire market at prices slightly below the marginal costs of the firm with the second lowest marginal costs.

<sup>4</sup>The concept of centrality in our analysis focuses on firm locations relative to competitors and not to other factors such as consumer clustering as in Anderson et al. (1997).

costs are normalized to zero for convenience.

To avoid discontinuities in the demand curve, we assume that the net utility of consumption is strictly positive and that each consumer purchases exactly one unit of the product per period, i.e. that the market is covered. We further restrict the parameters of the model so that in equilibrium (a) the market area of  $C$  exceeds its own spoke and (b) all firms sell at least to some consumers. Sufficient conditions for the existence of such an equilibrium can be stated in terms of upper and lower bounds for differences in marginal costs between the central and remote firms relative to their locations as well as to the number of empty spokes. More specifically, (a) and (b) are satisfied if:  $-\frac{9d_i+\bar{d}_i}{2} - d_C - \frac{2}{n} [l_C + l(N - n - 1)] + 3l_i - \bar{l}_i < c_C - \frac{3c_i - \bar{c}_i}{2} < \frac{3d_i - \bar{d}_i}{2} - d_C - \frac{2}{n} [l_C + l(N - n - 1)] + 3l_i - \bar{l}_i$ , where  $\bar{d}_i = \sum_{i=1}^n d_i/n$ ,  $\bar{c}_i = \sum_{i=1}^n c_i/n$ , and  $\bar{l}_i = \sum_{i=1}^n l_i/n$ . A detailed discussion of this restriction is provided in the online appendix. This implies that (a) all consumers located at empty spokes buy at the central firm and (b) the central firm competes (shares common market boundaries) directly with all other (remote) firms in the market, whereas all remote firms directly compete with the central firm only. This stronger degree of connectedness to competitors associated with the central position of firm  $C$  results in a special role for  $C$  in the determination of market prices. Figure 1 illustrates a simple network for the case of three firms (a central supplier  $C$  and  $n = 2$  remote suppliers  $R_i$ , with  $i = 1, 2$ ) in a market of  $N = 4$  spokes.

[Figure 1]

A marginal consumer located at  $x_i$  is indifferent between  $C$  and  $R_i$  if  $s - p_C - (d_C + x_i) = s - p_i - (d_i - x_i)$ , which can be rearranged to

$$(1) \quad x_i = \frac{p_i - p_C + d_i - d_C}{2}.$$

Profits ( $\pi$ ) for the central and the remote firms are given by

$$(2) \quad \pi_C = (p_C - c_C) \left[ \sum_{i=1}^n x_i + l_C + l(N - n - 1) \right],$$

$$(3) \quad \pi_i = (p_i - c_i)(l_i - x_i).$$

Maximizing profits with respect to  $p_C$  and  $p_i$  leads to

$$(4) \quad p_C = \frac{1}{2} \left[ \frac{\sum_{i=1}^n p_i}{n} + \frac{\sum_{i=1}^n d_i}{n} - d_C + c_C \right] + \frac{1}{n} [l_C + l(N - n - 1)],$$

$$(5) \quad p_i = \frac{1}{2} [p_C + d_C - d_i + c_i] + l_i.$$

A comparison of the price reaction functions for central and remote firms reveals three effects of centrality on firms' pricing decisions.

**Proposition 1.** *Centrality implies an asymmetry in the firms' strategic pricing behavior: Firms respond more strongly to price changes by a central firm than to price changes by a remote firm.*

This can easily be verified since  $\partial p_i / \partial p_C = 1/2 > 1/2n = \partial p_C / \partial p_i$ ,  $\forall n > 1$  and  $\partial p_j / \partial p_i = 0, \forall i \neq j$ . Figure 1 illustrates that – given the parameter restrictions discussed above – two remote firms never compete for the same customer. Thus, a price change by the remote firm  $i$  has no direct impact on all other remote firms and will directly influence one competitor only: the central firm. According to Proposition 1, the optimal price response of a central firm to a price change by a single remote firm  $i$  decreases with the number of remote firms ( $n$ ). Consequently, if  $n$  is large, a price change by one remote firm will be of relatively minor importance and will trigger a relatively small price response. In contrast, the central firm has  $n$  direct competitors. A price change by this firm has a direct impact on all remote firms, which does not depend on the size of  $n$ . Thus, in terms of the influence of one agent on other agents' actions, the central firm is indeed the most powerful.

**Proposition 2.** *In equilibrium the price of the central firm exceeds the price charged by a remote firm if and only if  $\frac{2}{n} [l_C + l(N - n - 1)] - 3l_i + \bar{l}_i - 2d_C + \frac{3d_i + \bar{d}_i}{2} + c_C - \frac{3c_i - \bar{c}_i}{2} > 0$ .*

The central position does not necessarily result in a higher equilibrium price compared to remote firms. Centrality is associated with two characteristics which exert countervailing effects on equilibrium prices: A central firm holds a larger market share than remote firms, which has a positive impact on its equilibrium price. However, centrality also implies a larger number of direct competitors, which restricts the ability to raise prices. The 'market share effect' is likely to dominate the 'number of rivals effect' if the number of empty spokes relative to the number of competitors ( $N/n$ ) is large, if the central firm is located close to the center ( $d_C$  is small), if the central firm has a large 'hinterland' (a large  $l$  and  $l_C$  for a given  $d_C$ ), and if the remote firms are located at a larger distance from the center ( $d_i$  and  $\bar{d}_i$  are large). The 'number of rivals effect', on the other hand, is stronger if the number of remote competitors  $n$  is large. A formal discussion of this aspect and a proof of Proposition 2 are provided in the online appendix.

**Proposition 3.** *The indirect impact of a price change of the central firm (following some idiosyncratic exogenous shock) on equilibrium market prices is stronger than the indirect impact of the same price change emanating from a remote firm.*

This proposition highlights the importance of centrality for the transmission of shocks on equilibrium market prices. An exogenous shock to firm  $i$  will not only directly affect firm  $i$ 's price but will also indirectly affect prices of its neighbors, which again triggers

price adjustments by the neighbors' neighbors including feedback effects to station  $i$  itself. These indirect effects are higher if the price change (following some idiosyncratic exogenous shock) is induced by a central rather than by a remote firm. A formal proof of proposition 3 is provided in the online appendix.

## 3 Industry, Data and Empirical Specification

### 3.1 The Retail Gasoline Market of Vienna

The retail gasoline market is particularly appropriate for our empirical analysis for several reasons. First, gasoline is a rather homogeneous product and the main source of product differentiation is a gasoline station's location (Barron et al., 2004; Clemenz and Gugler, 2006). Second, we assume in the theoretical model that firm locations are exogenously given. While this is a simplification, decisions on market entry and location choice are long-run decisions and can be considered as predetermined to the (short-run) pricing decision. This perspective is supported by the fact that establishing or closing a gasoline station is a costly endeavor. Third, the network centrality of gasoline stations is fairly easy to conceptualize and measure based on their locations within the network of roads.

For the present analysis we use detailed data on the geographical locations and other characteristics of all 273 gasoline stations in Vienna, which were collected by the company Experian Catalist in August 2003 (see <http://www.catalist.com> for company details). This data set is merged with retail price data for diesel<sup>5</sup> collected by the Austrian Chamber of Labor in the Vienna metropolitan area within one particular day every three months between October 1999 and March 2005 (a total of 22 points in time). Prices were collected for a randomly selected sub-sample of stations by telephone or – if a station refused to pass on information on prices – by driving past the respective station. The number of price observations available ranges from 144 to 152 per period.<sup>6</sup> Retail prices are nominal and measured in Euro cents per liter. The retail fuel price in Austria is determined by

---

<sup>5</sup>Unlike in North America, diesel-engined vehicles are very common in Europe. The share of cars with diesel engines was more than 50% in Austria in 2005 (Statistik Austria, 2006).

<sup>6</sup>An econometric analysis of potential determinants for missing prices supports the claim of the Chamber of Labor that stations are picked randomly: While some station characteristics have statistically significant effects on the probability to observe the price of a station, the marginal effects of each of the significant variables are extremely small. Further, the degree of centrality does not have a significant impact on the probability that prices are observed. Thus, we conclude that prices are missing at random (MAR) (see e.g. Little and Rubin (2002) for a classification). A detailed description of this econometric exercise is available in the online appendix. The group of randomly preselected stations does not change much over time: Out of 273 stations located in Vienna prices are observed for all periods for 121 stations, while prices are not available in any period for 87 stations. Note that missing prices do not affect our centrality measure which is based on the locations of all 273 gasoline stations.

three components (Benigni and Prinz, 2005): The first makes up about one third of the gross price and includes the crude oil price, the refining margin, transportation costs from the refinery to the station, as well as costs associated with the compulsory emergency reserves and storage expenses. The second accounts for about 10% of the gross price and includes station maintenance, advertising and overhead expenses as well as the station's mark-up. The final component is made up of an excise fuel tax (28.02 cents per liter until the end of 2003 and 30.02 cents afterwards) as well as 20% VAT (based on the net price including the fuel tax). In total, these taxes amount to more than half of the gross diesel price.

Table 1 illustrates the structure of the retail gasoline market of Vienna in August 2003. The market consists of branded and unbranded stations. Three major brands account for more than half (52%) of the 273 gasoline stations. BP (28%) has the largest number of outlets followed by OMV (13%) and Shell (11%). Six minor brands account for 31% and 17% are unbranded (independent) stations.

[Table 1]

Most branded stations are owned and operated by the company behind the brand implying that the refiner sets the retail price directly. In some instances company-owned branded stations are leased to a residual claimant who is obliged to purchase wholesale gasoline directly from the refiner but is independent in setting the retail price. In the case of dealer-owned branded stations, the dealer only signs a contract with a branded refiner to sell its brand of gasoline. Unbranded gasoline stations can shop for the lowest wholesale price from any refiner and separately determine their retail prices. They charge relatively low prices but offer few additional amenities such as car washing or service bays. While the assumption that gasoline stations set prices independently is straightforward for unbranded gasoline stations, it may not be so for branded stations. The same considerations apply to dealer-owned versus company-owned stations. However, having applied variance ratio tests, we do not find a smaller variance among branded stations than among unbranded stations. Having differentiated between major brands and minor brands, we only find a slightly smaller variance for the group of major brands compared to unbranded stations, but no difference between minor brands and unbranded stations. Thus, as all these differences are very small, we stick to the assumption that price dispersion within brands is not systematically different from unbranded stations, for which the assumption of independent pricing is straightforward. A variance ratio test for the prices of company- and dealer-owned stations does not reveal significant differences between these two groups.<sup>7</sup>

---

<sup>7</sup>A detailed description of differences in the within-variance of prices for branded and unbranded, company- and dealer-owned stations is available in the online appendix.



While these results do not preclude that some brands (company-owned stations) charge systematically different prices, the findings indicate that branded (company-owned) stations' price setting decisions take local market characteristics into consideration similarly to independent retailers.

To explore the importance of price fluctuations over time as well as permanent and random price movements within time periods we estimate a simple two-way fixed effects model of prices on station and time fixed effects (without any additional controls). Although we do not observe extreme price hikes over the sample period, fixed time effects account for nearly 90% of the overall price dispersion. Station-level fixed effects capture roughly 50% of the remaining cross-sectional price variation. Detailed results are available in the online appendix. To further investigate the stability of the price distribution over time we compute an index of rank reversal from comparing prices of all pairs of stations in a local market. Irrespective of the specific procedure used to define neighbors we observe that on average the usually cheaper station charges a lower price than its neighboring station in roughly 90%.<sup>8</sup> All of these figures show that the permanent part of the cross-sectional price variation plays an important role and that differences in relative prices are not purely random.<sup>9</sup>

Data on prices are merged with data on the geographical locations and other characteristics of all 273 gasoline stations in Vienna. Using ArcGIS Austria and the tool WIGeoNetwork, we link the geographical location of each gasoline station to information on the Viennese road system, which allows us to generate accurate measures of distance (measured in driving time) as well as the centrality of gasoline stations in the road network (see Section 3.2). The distance to the nearest neighbor (DISTANCE NEXT) is used as a proxy for the effect of the degree of spatial differentiation on price levels. To approximate demand and cost differences at different locations, we include variables on population density (POP DENS) as published by the Austrian Statistical Office, and on prices for factory and business premises (PREMISES) at the level of 23 districts collected by the Austrian Chamber of Labor for 2001, the share of commuters among all potential consumers (COMMUTERS)<sup>10</sup> as well as a dummy variable (TRAFFIC) that indicates heavy traffic intensity at the site of a gasoline station according to the Experian Catalist

---

<sup>8</sup>Similar degrees of rank reversals are found by Chandra and Tappata (2011). Summary statistics on rank reversals for the whole period and sub-periods are available in the online appendix.

<sup>9</sup>These numbers, however, also indicate that there is a substantial part of price variation that cannot be explained by (observed or unobserved) station and location heterogeneity. Theoretically, the unexplained part of the price variation can be rationalized by search-theoretic models with consumers being differently well-informed about prices, so-called 'clearinghouse models' (Varian, 1980; Stahl, 1989), or by dynamic pricing games as proposed by Maskin and Tirole (1988) that can lead to Edgeworth price cycles.

<sup>10</sup>A detailed description of this variable is provided in the online appendix.

data.<sup>11</sup> A number of dummy variables are included to account for additional station characteristics. Prices can be expected to be different if a station is company- rather than dealer-owned (COMPANY). Stations offering attendance service usually charge higher prices, as well as large stations which may increase the customers' willingness to pay by offering more comfort and ensuring less risk of congestion. Thus, the dummy variable SERVICE (LARGE) is equal to one if a station offers attendance service (if a station's ground surface is more than 2,000 square meters) and is zero otherwise. Table 2 in Section 3.2 summarizes all of these variables. Our estimations further include dummy variables for each of the nine brands illustrated in Table 1 as well as time fixed effects.

### 3.2 Measuring Centrality

The fact that the spatial structure of gasoline stations can be more complex than suggested by Figure 1 (as for instance, several stations can be located along a particular road) makes it difficult to measure centrality properly. A clear-cut dichotomy between central and remote competitors thus appears inappropriate for an empirical analysis. Rather, the spatial structure of gasoline stations is characterized by different 'degrees' of centrality within the network of roads. In order to calculate the centrality of gasoline stations, the network of roads, intersections and gasoline stations has to be transformed into a network whose only nodes are gasoline stations, whereby the links between these nodes reflect neighborhood relations relevant for strategic interaction. Let  $G$  of dimension  $m \times m$  reflect a network of  $m$  nodes. In  $G$  a link  $g_{ij} = 1$  if  $j$  is among the  $H$ -nearest neighbors of  $i$  in terms of driving time, while  $g_{ij} = 0$  otherwise. In the following, we borrow the measure 'degree centrality', introduced by Freeman (1979), which is frequently used in the social network literature (see e.g. Jackson, 2008). In our case it measures the number of times a particular gasoline station is among the  $H$  nearest neighbors of other gasoline stations. The degree centrality ( $dc$ ) of gasoline station  $j$  in network  $G$  based on  $H$ -nearest neighborhood is given by<sup>12</sup>

$$(6) \quad dc_j^H = \sum_{i=1}^m g_{ij},$$

For our main specifications we set  $H = 5$  but we also experiment with  $H = 2$  and

---

<sup>11</sup>The data contain four categories of traffic levels (very heavy, heavy, medium, low) that each station is assigned to by the surveyors. TRAFFIC is equal to one if traffic is considered to be (very) heavy and is zero otherwise.

<sup>12</sup>As relations are not necessarily reciprocal –  $i$  is not necessarily among the  $H$ -nearest neighbors of  $j$  if  $j$  is among the  $H$ -nearest neighbors of  $i$  – each node has the same number of  $H$  out-degrees (the number of links emanating from each node) but the network is directed. Therefore, our centrality measure is based on a node's in-degree (the number of links leading to a node).

$H = 10$ . More details on these and other centrality measures (used to check the robustness of our results with respect to different concepts of centrality) and a numerical example for a stylized network are provided in the online appendix. Table 2 shows descriptive statistics for the main centrality measure as well as for all other variables used in the empirical model.<sup>13</sup>

[Table 2]

### 3.3 Model Specification and Estimation Method

Our theoretical analysis suggests that centrality influences both, the level of prices as well as the strategic interaction in pricing between competitors. The empirical specification of the model that accounts for both effects is given by the following spatial autoregressive model

$$(7) \quad \mathbf{p} = \rho_1 \mathbf{W} \mathbf{p} + \rho_2 \mathbf{W} \mathbf{C} \mathbf{p} + \mathbf{X} \boldsymbol{\beta} + \gamma \mathbf{C} \boldsymbol{\iota} + \boldsymbol{\epsilon}.$$

In equation (7)  $\mathbf{p}$  is the  $M \times 1$  vector of prices, where  $M$  is the total number of observations in a repeated cross section of  $T = 22$  periods, so that  $M = \sum_{t=1}^T m_t$ , with  $m_t$  being the number of observations in period  $t$ . The matrices  $\mathbf{W}$  and  $\mathbf{C}$  are block diagonal ( $T$  blocks) and of dimension  $M \times M$ .  $\mathbf{W}$  is the spatial weights (distance decay) matrix with element  $w_{ij}$  being the squared inverse of the driving time from station  $i$  to station  $j$ , if station  $j$  is within a critical driving time (5 minutes) from  $i$  and  $i \neq j$ , and  $w_{ij} = 0$  otherwise.<sup>14</sup>  $\mathbf{C}$  is a diagonal matrix with the main diagonal element  $c_{jj}$  measuring the degree of centrality of station  $j$ .  $\mathbf{X}$  is an  $M \times k$  matrix of  $k$  explanatory variables including a constant,  $\boldsymbol{\iota}$  is an  $M \times 1$  unit vector, and  $\boldsymbol{\epsilon}$  is the  $M \times 1$  vector of error terms.  $\rho_1$  and  $\rho_2$  are the coefficients of spatial autocorrelation,  $\boldsymbol{\beta}$  is the  $k \times 1$  vector of coefficients of the exogenous variables in  $\mathbf{X}$ , and  $\gamma$  measures the partial correlation between centrality and a station's

<sup>13</sup>Centrality is positively correlated with company ownership, stations' size, heavy traffic and major brands, and negatively correlated with the driving time to the nearest neighbor, population density, and unbranded stations. However, the coefficients of correlation between the centrality measure and these other variables are rather small. We also tested for differences in the spatial distribution of several station characteristics within the market. The mean degree of centrality and the share of major branded stations do not deviate significantly from the whole market in any of the sub-regions we analyzed. More details are available in the online appendix.

<sup>14</sup>The spatial weights reflect the elasticity of substitution and we expect this elasticity to depend on driving time rather than driving distance. However, substitution patterns might also be influenced by other determinants. Houde (2012) puts forward the idea that commuting patterns influence substitution elasticities. Pennerstorfer et al. (2014) use commuting patterns to delimit local markets. We refrain from doing so as our data on commuting is not very accurate for commuters residing or working in Vienna (see the online appendix for details).

price level. The matrices  $\mathbf{W}$  and  $\mathbf{WC}$  are row-normalized. Row-normalization implies that the influence of a station ( $i$ ) on a rival station's ( $j$ ) price ( $\partial p_j / \partial p_i$ ) decreases with the number of competitors. This is plausible for a number of reasons: First, in the theoretical model row-normalization is suggested by the fact that  $\partial p_C / \partial p_i = 1 / (2n)$ . Second, in a field experiment Barron et al. (2008) show that in a retail gasoline market stations respond less to a (exogenous) price change of an individual 'treated' gasoline station if the number of stations in the local market is high. Third, row-normalization facilitates the interpretation of the spatial autoregressive parameter  $\rho_1$  ( $\rho_2$ ), as  $\mathbf{W}\mathbf{p}$  ( $\mathbf{WC}\mathbf{p}$ ) is the spatially weighted (spatially and centrality-weighted) average price of rivals. To account for unobserved spatial heterogeneity we allow the error term  $\epsilon$  to be spatially correlated so that  $\epsilon = \lambda \mathbf{W}^{error} \epsilon + \mathbf{v}$  (or, equivalently,  $\epsilon = (\mathbf{I} - \lambda \mathbf{W}^{error})^{-1} \mathbf{v}$ ), where  $\mathbf{I}$  is an  $M \times M$  identity matrix. In  $\mathbf{W}^{error}$  element  $w_{ij}^{error} = 1$  if station  $i$  and  $j$  are within the same local market (within 5 minutes' driving time) and  $i \neq j$ , and zero otherwise. The  $M \times M$  matrix  $\mathbf{W}^{error}$  is again block diagonal with  $T = 22$  blocks and is row-normalized to calculate the average residuals  $\epsilon$  of other nearby stations. In  $\mathbf{W}^{error}$  we opt for binary rather than distance-based weights because unobserved local characteristics or shocks are expected to affect all firms within a local market equally. We follow a referee's suggestion and cluster the remaining error  $\mathbf{v}$  at the station level to account for serial correlation.

By definition, the spatial averages of rivals' prices on the right-hand side of the price reaction function in equation (7) are endogenous. Therefore, OLS will lead to biased and inconsistent parameter estimates, unless  $\rho_1 = \rho_2 = 0$  (Anselin, 1988). Two alternative approaches to solve this problem are frequently used in the applied spatial economics literature: First, an estimation of the reduced form of a spatial system of equations applying maximum likelihood (ML) techniques; second, the use of instrumental variables (IV) or similar concepts such as the generalized method of moments (GMM).

In the ML-approach – as proposed by Anselin (1988) and Case (1991) – the reduced form of equation (7) is equal to

$$(8) \quad \mathbf{p} = (\mathbf{I} - \rho_1 \mathbf{W} - \rho_2 \mathbf{WC})^{-1} (\mathbf{X}\boldsymbol{\beta} + \gamma \mathbf{C}\boldsymbol{\iota} + \boldsymbol{\epsilon}).$$

It is important to note that the structural model in (7) can be recovered by estimating this reduced form model based on equilibrium prices.<sup>15</sup> In section 4 we will present results based on ML estimation first, as this method is most commonly used in empirical research applying spatial econometric techniques (Gibbons and Overman, 2012). Lee (2004) shows

<sup>15</sup>In the online appendix we use simulated data based on our stylized theoretical model to show that recovering  $\rho_1$  and  $\rho_2$  is possible when using data on equilibrium prices, as long as the model is correctly specified.

that (quasi-) ML estimators are consistent and asymptotically efficient, as long as some regularity conditions are met.

However, ML-based methods have recently been criticized because the properties of the estimator(s) are based on prior knowledge of the data-generating process which is unusual in empirical work (Gibbons and Overman, 2012). Using instrumental variables (IV) allows for an estimation of the structural equation in (7) directly. Most researchers using IV estimation in a spatial setting, e.g. Pinkse et al. (2002), follow Kelejian and Prucha (1998, 1999) and use spatial lags of (some of the) exogenous variables as instruments, such that prices of nearby competitors are instrumented by their own characteristics. More specifically, Kelejian and Prucha (1998) propose a three-step procedure: First, a two stage least squares (2SLS) estimator is used to obtain consistent estimates for  $\rho_1$ ,  $\rho_2$ ,  $\beta$  and  $\gamma$ . In this step  $\mathbf{W}\widetilde{\mathbf{X}}$  and  $\mathbf{W}\mathbf{C}^{dc}\widetilde{\mathbf{X}}$  are used as instruments excluded from the second stage regression, with  $\widetilde{\mathbf{X}}$  containing all variables (including  $\mathbf{C}^{dc}\boldsymbol{\iota}$ ) that vary at the station level. The spatial process in the residuals can be ignored in this step, as only consistent (and not efficient) estimators of the coefficients are necessary. Second, the autoregressive parameter  $\lambda$  is estimated with a generalized moments procedure suggested in Kelejian and Prucha (1999), using the residuals obtained in the first step.<sup>16</sup> In the third step the estimator of  $\lambda$  ( $\widetilde{\lambda}$ ) is used to perform a Cochrane-Orcutt type transformation to account for spatial correlation.<sup>17</sup> The transformed model is finally estimated by a GMM estimator using  $\mathbf{W}\widetilde{\mathbf{X}}$ ,  $\mathbf{W}\mathbf{C}^{dc}\widetilde{\mathbf{X}}$ ,  $\mathbf{W}^{error}\widehat{\mathbf{X}}$ ,  $\mathbf{W}^{error}\mathbf{W}\widetilde{\mathbf{X}}$  and  $\mathbf{W}^{error}\mathbf{W}\mathbf{C}^{dc}\widetilde{\mathbf{X}}$  as excluded instruments – as proposed by Kelejian and Prucha (1998) – with  $\widehat{\mathbf{X}}$  including all variables with cross-sectional variation. This procedure – which we denote as a generalized spatial (GS)2SLS/GMM estimator – allows us to estimate exactly the same model (including spatial autoregressive processes both in the endogenous variable and in the disturbance term) as in the ML-specifications. In our application the observed firm characteristics are mainly dummy variables and their spatially lagged averages might thus be relatively weak instruments. However, this method addresses issues of identification directly and does not rely on the assumption of normally distributed residuals. To increase the credibility of our findings we therefore use both ML- and IV-based estimation procedures as recommended by McMillen (2012).

<sup>16</sup>Note that Kelejian and Prucha (1998) derive a point estimate for  $\lambda$ , but no standard deviation for the estimated parameter.

<sup>17</sup>Applying this transformation gives the equation  $\mathbf{p}^* = \rho_1\mathbf{W}\mathbf{p}^* + \rho_2\mathbf{W}\mathbf{C}^{dc}\mathbf{p}^* + \mathbf{X}^*\beta + \gamma\mathbf{C}^{dc}\boldsymbol{\iota}^* + \mathbf{v}$  to be estimated, with  $\mathbf{p}^* = \mathbf{p} - \widetilde{\lambda}\mathbf{W}^{error}\mathbf{p}$ ,  $\mathbf{W}\mathbf{p}^* = \mathbf{W}\mathbf{p} - \widetilde{\lambda}\mathbf{W}^{error}\mathbf{W}\mathbf{p}$ ,  $\mathbf{W}\mathbf{C}^{dc}\mathbf{p}^* = \mathbf{W}\mathbf{C}^{dc}\mathbf{p} - \widetilde{\lambda}\mathbf{W}^{error}\mathbf{W}\mathbf{C}^{dc}\mathbf{p}$ ,  $\mathbf{X}^* = \mathbf{X} - \widetilde{\lambda}\mathbf{W}^{error}\mathbf{X}$ , and  $\mathbf{C}^{dc}\boldsymbol{\iota}^* = \mathbf{C}^{dc}\boldsymbol{\iota} - \widetilde{\lambda}\mathbf{W}^{error}\mathbf{C}^{dc}\boldsymbol{\iota}$ .

### 3.4 Identification Issues and Interpretation

Ideally, one could follow Pinkse et al. (2002) and Pinkse and Slade (2010), among others, and interpret the system of equations in (7) as a set of reaction functions obtained from a simultaneous pricing game. In this case the parameter estimate of  $\rho_1$  measures the (spatially weighted) price interaction between neighboring stations. An asymmetry in price adjustment between central and remote firms is then captured by  $\rho_2$ . A positive estimate of  $\rho_2$  would indicate that prices respond more strongly to price changes by more central stations, as suggested by Proposition 1.

However, the identification of strategic pricing interaction between (neighboring) firms in both equations (7) and (8) is impeded (i) by an incomplete data sample, (ii) by the existence of a common time-varying price component, and (iii) by spatially correlated unobservable determinants of gasoline prices. Additionally, disentangling the influence of rivaling stations' prices from the impact of other stations' characteristics (the so-called 'reflection problem', see Manski, 1993) is a further challenge.

First, as our sample of prices is incomplete (prices are unavailable for nearly half of all observations),  $\mathbf{Wp}$  and  $\mathbf{WCp}$  are based on a sub-sample of neighboring stations if prices are not observed for all other stations in the vicinity at a particular point in time. The average price of a sub-sample of rivals stations, however, will most likely differ from the average of all rivals in a local market. As noted by Pinkse and Slade (2010, p. 113) "[t]here is not much work offering a serious solution to this problem". In this paper we ignore stations in the vicinity for which prices are unavailable (listwise deletion). For the ML-estimator and randomly missing dependent variables Wang and Lee (2013) show that the bias of the spatial autoregressive parameter ( $\rho$ ) is rather small and decreases with sample size if all observations with missing dependent variables are excluded from the analysis. For IV-estimators small sample results of the consequences of listwise deletion are not available, but Kelejian and Prucha (2010) show that a two-stage least squares IV estimator remains asymptotically consistent if observations with missing endogenous variables are dropped from the analysis, as long as the share of missing endogenous variables does not get too large. While missing information on prices is expected to reduce the efficiency of the parameter estimates, we are confident that they remain consistent and that potential biases are negligibly small.<sup>18</sup>

Second, fluctuations in gasoline prices over time typically account for a large share of the total price variation. Even if there was perfect competition, prices of neighbo-

---

<sup>18</sup>In contrast to prices, information on the explanatory variables is nearly complete (see Table 2). As the data are missing (completely) at random and the share of missing data is very small, we follow Greene (2008) and replace missing information with zeros in the estimations while including dummy variables that are equal to one if the information is missing for an observation, and zero otherwise.

ring stations would still be correlated because of this common time-varying component.<sup>19</sup> Thus, we include time fixed effects which completely remove price fluctuations common to all gasoline stations. Any remaining correlation between prices of neighboring stations cannot be driven by common exogenous shocks.

Third, it is generally difficult to identify whether (or to what extent) the correlation between the price of a particular station is caused by other stations' prices or (a) by unobserved local characteristics or common local market shocks or (b) by rival stations' characteristics directly rather than indirectly via rival stations' prices. To address (a), which might also violate the assumption of *i.i.d.* errors, we allow the error term to be spatially correlated.<sup>20</sup> If the model (in particular the spatial weights matrices) is specified correctly, it is possible to disentangle the causal effects of one station's price on other stations' prices from the (potential) influence of unobserved local characteristics. However, while we provide arguments for our preferred model specification, the choice of the spatial weights matrices has to be based on assumptions that cannot be tested.

(b) The so-called 'reflection problem', i.e. whether  $\mathbf{p}$  is affected by  $(\mathbf{W}\mathbf{p}, \mathbf{W}\mathbf{C}\mathbf{p})$  or directly by  $(\mathbf{W}\mathbf{X}, \mathbf{W}\mathbf{C}\mathbf{X})$ , is outlined in Manski (1993) and thoroughly discussed in Pinkse and Slade (2010) and Gibbons and Overman (2012). We think that for the explanation of the price charged by a particular station, rival stations' prices play a dominant role compared to other stations' characteristics, as including  $(\mathbf{W}\widetilde{\mathbf{X}}, \mathbf{W}\widetilde{\mathbf{C}\mathbf{X}})$  in the maximum likelihood estimation hardly affects the parameter estimates of  $\rho_1$  and  $\rho_2$  (see specification [2] and [3] in Table 3). Additionally, when performing IV-regressions we use  $(\mathbf{W}\widetilde{\mathbf{X}}, \mathbf{W}\widetilde{\mathbf{C}\mathbf{X}})$  as excluded instruments (see specification [3] in Table 3). In this case a Hansen  $J$ -test indicates that these spatially lagged characteristics are valid instruments and therefore correctly excluded from the main regression. However, disentangling the effect of rival stations' prices from the direct impact of rival stations' characteristics requires the spatial weights matrix  $\mathbf{W}$ , capturing the structure of the 'reference groups' (i.e. market boundaries and functional form of distance decay), to be known a-priori. Again, the specification of the weights matrix is based on assumptions that cannot be fully tested.

Despite the efforts to address these identification challenges we cannot provide experimental evidence by (exogenously) shocking the network at various points. Alternatively, a fully-specified demand system would be necessary for identifying the causal effects of one station's price on rival prices. However, this would require quantity data, which are

<sup>19</sup>This aspect is also visible in the data: The correlation between prices and spatially weighted (spatially and centrality weighted) average prices of rival firms is approximately 0.97. Once common time effects are controlled for the correlation decreases to about 0.67. See the online appendix for further details.

<sup>20</sup>See Section 3.3. A spatial autoregressive process in the residuals is preferred over clustering the residuals at a local level, as clustering restricts local markets to non-overlapping spatial units.

not available in the present context.<sup>21</sup> We therefore refrain from interpreting the parameter estimates of  $\rho_1$  and  $\rho_2$  as measuring the causal effects between stations' prices. Nevertheless, even in a descriptive fashion our results highlight the importance of taking the complex geography of the market into account when analyzing prices in spatially differentiated markets.

## 4 Results

From the theoretical model we expect – once we include the spatially and centrality weighted average price,  $\mathbf{WCp} - \rho_2$  to be positive and statistically significant, while  $\rho_1$  is not significantly different from zero (Proposition 1). Due to countervailing effects of the degree of centrality on prices ('market share' vs. 'number of rivals' effect) we have no clear expectations as to the relation between centrality and price levels  $\gamma$  (Proposition 2). The question whether a price shock induced by a firm with a higher degree of centrality has a stronger effect on equilibrium prices (Proposition 3), is treated in a simulation experiment in Section 4.2 after discussing the regression results in Section 4.1. A sensitivity analysis of the econometric results is discussed in Section 4.3.

### 4.1 Econometric Results

The main econometric results are summarized in Table 3. The parameter estimates of a benchmark model (which does not explicitly control for differences in centrality by assuming  $\rho_2 = 0$  and  $\gamma = 0$ ) are reported in column [1]. Similar to previous studies estimating spatial lags for retail gasoline prices (Netz and Taylor, 2002; Pennerstorfer, 2009), we find a positive parameter estimate of  $\rho_1$ . The parameter is significantly different from zero at the 1%-level. A (spatially weighted) average price increase by 1 cent by all relevant neighbors is associated with a price increase by 0.45 cents per liter.

[Table 3]

Columns [2] to [4] report parameter estimates of the extended model including the centrality measure defined above. Centrality is found to significantly contribute to explaining the spatial correlation of prices. The parameter estimates of  $\rho_2$  are positive and

---

<sup>21</sup>A different approach is taken by Thomadsen (2005). Following the seminal work of Feenstra and Levinsohn (1995), the author accommodates the lack of quantity data by using information obtained from the assumption of consumer utility maximizing behavior. Thomadsen (2005) substitutes the relationship between price and quantity (obtained from consumer demand) into the firms' first-order conditions from static Bertrand competition to jointly estimate the parameters of the indirect utility functions of consumers and the marginal costs of firms.



significantly different from zero at the 1%-level, while  $\rho_1$  is not significantly different from zero in all specifications that include  $\mathbf{WCP}$ . A particular gasoline station’s price is more strongly related to prices of a central competitor than to prices of a remote rival. This is in line with Proposition 1. Further, the inclusion of centrality significantly improves the explanatory power of the models: a likelihood ratio ( $LR$ )-test clearly rejects the ‘restricted’ model [1] in favor of the ML-models including degree centrality (models [2], [3]) at the 1% significance level. Column [3] extends model [2] by including spatially lagged exogenous variables ( $\mathbf{W}\widetilde{\mathbf{X}}$  and  $\mathbf{WC}\widetilde{\mathbf{X}}$ ), as neighboring stations’ characteristics might affect prices directly – and not only indirectly – by altering neighboring stations’ prices. The parameter estimates of  $\rho_1$  and  $\rho_2$  are hardly affected by this alteration: In model [3]  $\rho_1$  ( $\rho_2$ ) is again insignificant (significant and positive). The difference between the point estimates of  $\rho_2$  in [2] and [3] is small and not statistically different from zero. The results are also hardly affected when applying an IV-based method (GS2SLS/GMM), reported in column [4]. The point estimate of  $\rho_2$  is very similar to model [2] and [3] and statistically not different from them, while  $\rho_1$  remains insignificant. As expected, the standard errors of  $\rho_1$  and  $\rho_2$  are somewhat larger in the IV-model. Table 3 also provides specification tests for the IV-estimates: A Hansen  $J$ -test for the overidentification restrictions does not reject the joint null hypothesis that the instruments are uncorrelated with the residuals. This suggests that the excluded instruments are valid instruments and are therefore correctly excluded from the second-stage regression.  $F$ -tests to assess the explanatory power of the exogenous variables excluded from the second-stage regression in both first-stage regressions (on  $\mathbf{Wp}$  and  $\mathbf{WC}^{dc}\mathbf{p}$ ) reject the null hypothesis of no explanatory power at the 1%-significance level for both endogenous variables. To sum up, the instruments appear to be valid and strong enough to strengthen our confidence in our findings.

The two countervailing effects (‘market share’ vs. ‘number of rivals’ effect) associated with centrality that follow from the theoretical model (Proposition 2) seem to balance each other on average. The parameter estimate of DEGREE ( $dc$ ) centrality is not significantly different from zero. Thus, we do not find evidence that centrality explains differences in the price levels.<sup>22</sup>

Our models include a number of control variables, as discussed in Section 3.1. An increase in the distance to the nearest neighbor (DISTANCE NEXT) by one minute is associated with an increase in the price of a station by about 0.2 cents. An increase in the rate of COMMUTERS by ten percentage points is related to an increase in prices by 0.2 to 0.4 cents. The results for both variables are highly significant in all specifications. The local population density (POPDENS) is significantly larger than zero only in the

<sup>22</sup>The lack of significance might also stem from the fact that the empirical model directly controls for differences in consumer demand through the included location characteristics.

IV-specification [4], while the price for PREMISES is significantly positively correlated with fuel prices in two specifications ([3] and [4]). Company- rather than dealer-owned stations (COMPANY = 1) are found to charge significantly higher prices by about 0.8 to 1.1 cents per liter at least at the 10% significance level in all specifications. In contrast, LARGE stations ( $> 2,000m^2$  of surface area) do not significantly charge more than smaller stations. We find some evidence for higher gasoline prices at stations located at roads with heavy TRAFFIC, but the effect is significantly different from zero only in specification [3]. Stations offering attendance SERVICE charge higher prices by about 0.7 cents compared to stations exclusively offering self-service. The three major brands operating in Austria (BP, OMV and SHELL) charge significantly higher prices than unbranded stations. According to column [2], the price differences range from 1.5 cents (OMV and SHELL) to 1.9 cents (BP). The price differences between minor brands and independent (unbranded) stations are more heterogeneous: Some minor brands (AGIP, ARAL and ESSO) charge prices similar to major brands, while the price differences between some other minor brands (AVANTI, JET and STROH) and unbranded stations are not significantly different from zero. All of these results are in line with previous findings on the determinants of station level gasoline prices.

The parameter estimates for the coefficient of spatial autocorrelation in the residuals ( $\lambda$ ) are significantly different from zero in all three ML-models. Including  $\mathbf{W}\tilde{\mathbf{X}}$  and  $\mathbf{WC}\tilde{\mathbf{X}}$  as additional control variables in column [3], however, reduces  $\lambda$ , from around 0.4 to 0.2. These results clearly reject the assumption of spatially uncorrelated residuals and justify our approach of modeling a spatial autoregressive process in the error term. Surprisingly,  $\lambda$  takes a negative sign in the IV-model [4].

## 4.2 Simulation

If equation (7) could be interpreted as a set of reaction functions, then it is important to note that the parameter estimates of  $\rho_1$  and  $\rho_2$  only account for the direct relation between prices of neighboring stations. To illustrate this point and to address the third proposition of our theoretical model, i.e. the effect of centrality on the transmission of shocks to the general price level, we need to consider that each price change also triggers feedback effects to and from all neighbors in the market. Starting with equilibrium prices, an exogenous idiosyncratic shock for a particular station  $i$  will not only change  $i$ 's own price (direct effect) but also the prices of its (first-order) neighbors, which again triggers price adjustments by the neighbors' neighbors (second-order neighbors of station  $i$ ) including feedback effects to station  $i$  itself (indirect effect). To calculate differences in the total effect of idiosyncratic, station-specific shocks on equilibrium prices, we use the estimates

of  $\rho_1$  and  $\rho_2$  from specification [2] in Table 3 and apply a bootstrap simulation technique in which each parameter is drawn randomly from a normal distribution with the mean and the standard deviation obtained from this regression.<sup>23</sup> Figures 2 and 3 illustrate the relationship between centrality and the transmission of shocks via indirect neighborhood effects. More specifically, in Figure 2 the relative impact of the indirect (feedback) effects of an idiosyncratic exogenous shock emanating from one gasoline station on this particular station is measured by the ratio of the total effect (including the direct impact as well as all indirect effects) to the direct effect. This ratio is shown on the vertical axis on the left-hand side of Figure 2. The centrality of the station inducing the shock is measured on the horizontal axis, the density of the different degrees of centrality of gasoline stations in the market is measured on the vertical axis on the right. According to Figure 2, an exogenous shock triggering a direct price increase by 1 cent for an individual station with a median degree centrality of 5 leads to an additional (indirect) increase in its price (after considering all feedback effects to and from neighboring firms) by 4 %. Thus, the total price increase at this station is 1.04 cents per liter. In contrast, the total price increase is not significantly different from the direct increase in case of a remote supplier with a degree centrality of 2, being 1.11 cents per liter for a central supplier with a degree centrality of 11.

[Figure 2]

[Figure 3]

Similarly, Figure 3 illustrates the aggregated increase in prices of all other gasoline stations in the market due to indirect feedback effects following an idiosyncratic shock induced by an individual gasoline station. The horizontal axis again measures the ‘degree centrality’ of the station inducing the shock and the vertical axis on the right again illustrates the distribution of degrees of centrality of gasoline stations in the market. The vertical axis on the left now measures the aggregated (sum of) indirect effects on all other gasoline stations in the market excluding the station inducing the shock. Figure 3 shows that an idiosyncratic shock emanating from a gasoline station with a degree centrality of 5 (11) triggering a price increase by 1 cent at this station, leads to an aggregated increase in prices of all other stations in the market by 0.52 (2.55) cents. The same shock induced by a remote gasoline station with a ‘degree centrality’ of 2, however, does not lead to an aggregated increase significantly different from zero in the price of other stations. Gasoline stations with a higher degree of centrality tend to be neighbors to more stations, to be

---

<sup>23</sup>In the simulations we normalize  $\rho_1$  and  $\rho_2$  so that they sum up to the value of  $\rho_1$  obtained in the benchmark model of specification [1] for each draw. This assumption is justified as we cannot reject the restriction of  $\rho_1$  in specification [1] being equal to  $\rho_1 + \rho_2$  in specification [2].

relatively closer to other stations, and are thus more influential in affecting the prices of neighboring stations. The results of this simulation show that the indirect impact of a price increase induced by a firm on equilibrium prices is positively correlated with its degree of centrality, and therefore support Proposition 3.

### 4.3 Robustness Checks

In order to confirm that the results are not driven by the specific definitions of key variables used in our empirical analysis, the regressions were run using perturbations of these definitions. In a first set of estimation experiments we use different concepts of centrality (‘weighted degree’ and ‘closeness’ centrality), different neighborhood criteria determining a station’s centrality ( $H = 2$  and  $H = 10$  instead of  $H = 5$ ) as well as different specifications of the spatial weights matrix  $\mathbf{W}$  such as the (cubed) inverse rather than the squared inverse distance. Secondly, to show that the degree of spatial autocorrelation in prices is related to station centrality and not to other factors (correlated with our centrality measures), we interact the spatial weights matrix with several station characteristics and with interaction terms of various explanatory variables. Thirdly, we also include district-fixed effects instead of variables that vary only at the district level to control for additional potentially unobserved regional heterogeneity. Fourthly, we analyze the residuals of a two-way fixed effects model (including station level and time period fixed effects only), denoted as ‘residual’ or ‘cleaned’ prices, instead of ‘raw’ prices, to control for all time invariant station and location characteristics. We then investigate whether these residual prices are correlated with spatially and centrality weighed residual prices. Further, we test alternative specifications of the IV estimator that model different forms of spatially clustered errors instead of spatially autocorrelated residuals. Finally, we exclude  $\mathbf{W}\mathbf{p}$  and  $\mathbf{WC}\mathbf{p}$  from the regression to investigate whether the importance of centrality is picked up by  $\mathbf{WC}\widetilde{\mathbf{X}}$ , once  $\mathbf{W}\mathbf{p}$  and  $\mathbf{WC}\mathbf{p}$  are left out.

A thorough description of these robustness checks as well as a detailed discussion of their results is available in the online appendix. All of the results support the findings presented in Section 4.1: The parameter estimate of  $\rho_2$  is positive and significantly different from zero in all models, which suggests that firms’ prices are more strongly correlated with prices of nearby firms characterized by a higher degree of centrality (as predicted by Proposition 1). We do not find a statistically significant relation between centrality and price levels, which is in line with Proposition 2. Simulating the impact of an exogenous shock by a single firm – based on parameter estimates obtained from alternative models – provides very similar results as those summarized in Figures 2 and 3 and supports the prediction of Proposition 3: price changes of firms with a high degree centrality have a lar-

ger impact on rivals' (equilibrium) prices. Using different neighborhood criteria, different concepts to determine centrality, alternative functional forms in  $\mathbf{W}$ , additional interaction terms as endogenous spatial lags as well as different estimation procedures does not affect the main conclusion of our analysis: pricing is significantly related to gasoline stations' centrality.

## 5 Conclusions

The present paper highlights the importance of centrality for pricing. Firms are characterized by different degrees of centrality within a network unfolded in space. The specific position of a firm in the network (its degree of centrality) relative to its competitors determines the intensity of competition between firms. According to our theoretical model central suppliers are found to be more powerful in the sense of (a) exerting a stronger influence on the pricing decisions of their neighbors, and (b) having a stronger impact on equilibrium market prices.

We provide first empirical evidence on the relation between centrality and pricing in a spatially differentiated market by adopting a measure of centrality from the literature on (social) networks. Econometric results based on spatial autoregressive models confirm that prices of a particular gasoline station are more strongly correlated with prices of a central competitor compared to those of remote firms. Simulation experiments suggest that the impact of a price change by an individual gasoline station on equilibrium prices increases with its degree of centrality. We do not find, however, empirical evidence of a significant relation between centrality and the level of prices. This result can be explained by the existence of two countervailing effects: centrality implies a larger number of consumers (higher prices) but, at the same time, is associated with a larger number of direct competitors (lower prices).

A major caveat of our empirical part is that we cannot interpret our findings as causal effects of one station's centrality on another station's price, due to a lack of quantity or (quasi-)experimental data. Extending our analysis along the lines suggested by Thomadsen (2005) is left for future research. Still, our contribution highlights the importance of taking into account the complex geographical reality in many spatially differentiated markets. Therefore, our results have a number of important implications for both economic policy and academic research: First, gasoline stations are often members of a network of multi-station firms (large chains of gasoline stations) coordinating their pricing behavior within the network. The effects of joint ownership or mergers between firms will depend on the specific geographic positions of the gasoline stations involved. Coordination of

prices among a number of remote gasoline stations will have different effects on social welfare than price coordination among central stations.

Second, papers analyzing Edgeworth price cycles find that larger firms (Noel, 2007) or certain major brands (Lewis, 2012) are more likely to initiate ‘price restorations’ (i.e. sharp price increases), while price reductions are primarily started off by small firms (Noel, 2007). It would be interesting to explore the relationship between centrality and price leadership. Similarly, a number of studies observe that retail gasoline prices adjust significantly faster to an increase than to a decrease in wholesale prices. Verlinda (2008) finds that this asymmetry is higher for branded stations and increases with the geographic distance to rivals as well as with the provision of additional amenities. Further research could thus focus on relating this literature to stations’ centrality.

Finally, the present paper also underlines the need to analyze entry and exit decisions in the context of centrality. In contrast to traditional spatial models, in which firms and consumers are distributed symmetrically and the specific location of a firm in space is irrelevant, locational choice in space constitutes an important strategic decision in the present framework. Entry at a central location will have a strong impact on incumbents since, for many of them, it creates additional direct competition for customers. Entry at a remote position, however, will have a minor effect on few incumbents only. We hope that our contribution spurs further research in these directions.

### **Acknowledgments**

We gratefully acknowledge helpful comments by Giuseppe Arbia, Harald Badinger, Klaus Gugler, Michael Pfaffermayr, Ingmar Prucha, Philipp Schmidt-Dengler, Robert D. Weaver, the participants of seminars and conferences in Athens, Boston, Bratislava, Chicago, Salzburg, Stockholm, St.Gallen, Toulouse and Verona, the editor Stephen Ryan as well as three anonymous referees. The work was generously supported by the Austrian National Bank (OeNB) Anniversary Fund (project number: 12974).

## **References**

- Anderson, S. P., Goeree, J., Ramer, R., 1997. Location, location, location. *Journal of Economic Theory* 77 (1), 102–127.
- Anselin, L., 1988. *Spatial Econometrics: Methods and Models*. Dordrecht, Kluwer Academic Publishers.
- Balasubramanian, S., 1998. Mail versus mall: A strategic analysis of competition between direct marketers and conventional retailers. *Marketing Science* 17 (3), 181–195.

- Barron, J. M., Taylor, B. A., Umbeck, J. R., 2004. Number of sellers, average prices, and price dispersion. *International Journal of Industrial Organization* 22 (8-9), 1041–1066.
- Barron, J. M., Umbeck, J. R., Waddell, G. R., 2008. Consumer and competitor reactions: Evidence from a field experiment. *International Journal of Industrial Organization* 26 (2), 517–531.
- Benigni, J., Prinz, S., 2005. *Der Österreichische Kraftstoffmarkt 2004*. Vienna, Bundesministerium für Wirtschaft und Arbeit.
- Borenstein, S., Cameron, A. C., Gilbert, R., 1997. Do gasoline prices respond asymmetrically to crude oil price changes? *Quarterly Journal of Economics* 112 (1), 305–339.
- Braid, R. M., 1989. Retail competition along intersecting roadways. *Regional Science and Urban Economics* 19 (1), 107–112.
- Brass, D. J., Burkhardt, M. E., 1992. Centrality and power in organizations. In: Nohira, N., Eccles, R. G. (Eds.), *Networks and Organizations: Structure, Form and Action*. Boston, Harvard Business School Press, pp. 191–215.
- Case, A. C., 1991. Spatial patterns in household demand. *Econometrica* 59 (4), 953–965.
- Chamberlin, E., 1948. *The Theory of Monopolistic Competition*, 6th Edition. Cambridge [et al.], Harvard University Press.
- Chandra, A., Tappata, M., 2011. Consumer search and dynamic price dispersion: An application to gasoline markets. *The RAND Journal of Economics* 42 (4), 681–704.
- Chen, Y., Riordan, M. H., 2007. Price and variety in the spokes model. *The Economic Journal* 117 (522), 897–921.
- Clemenz, G., Gugler, K., 2006. Locational choice and price competition: Some empirical results for the Austrian gasoline market. *Empirical Economics* 31 (2), 291–312.
- Eckert, A., 2003. Retail price cycles and the presence of small firms. *International Journal of Industrial Organization* 21 (2), 151–170.
- Eckert, A., 2013. Empirical studies of gasoline retailing: A guide to the literature. *Journal of Economic Surveys* 27 (1), 140–166.
- Feenstra, R. C., Levinsohn, J. A., 1995. Estimating markups and market conduct with multidimensional product attributes. *The Review of Economic Studies* 62 (1), 19–52.
- Fik, T. J., 1991. Price competition and node-linkage association. *Papers in Regional Science* 70 (1), 53–69.
- Freeman, L. C., 1979. Centrality in social networks: Conceptual clarification. *Social Networks* 1 (3), 215–239.
- Gibbons, S., Overman, H. G., 2012. Mostly pointless spatial econometrics? *Journal of Regional Science* 52 (2), 172–191.

- Greene, W. H., 2008. *Econometric Analysis*, 6th Edition. Upper Saddle River [et al.], Pearson Prentice Hall.
- Hotelling, H., 1929. Stability in competition. *The Economic Journal* 39 (153), 41–57.
- Houde, J.-F., 2012. Spatial differentiation and vertical mergers in retail markets for gasoline. *American Economic Review* 102 (5), 2147–2182.
- Jackson, M. O., 2008. *Social and Economic Networks*. Princeton [et al.], Princeton University Press.
- Kelejian, H. H., Prucha, I. R., 1998. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics* 17 (1), 99–121.
- Kelejian, H. H., Prucha, I. R., 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review* 40 (2), 509–533.
- Kelejian, H. H., Prucha, I. R., 2010. Spatial models with spatially lagged dependent variables and incomplete data. *Journal of Geographical Systems* 12 (3), 241–257.
- Lee, L.-F., 2004. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* 72 (6), 1899–1925.
- Lewis, M., 2012. Price leadership and coordination in retail gasoline markets with price cycles. *International Journal of Industrial Organization* 30 (4), 342–351.
- Little, R. J., Rubin, D. B., 2002. *Statistical Analysis with Missing Data*, 2nd Edition. New York, Wiley.
- Manski, C., 1993. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies* 60 (3), 531–542.
- Maskin, E. S., Tirole, J., 1988. A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and Edgeworth cycles. *Econometrica* 56 (3), 571–599.
- McMillen, D. P., 2012. Perspectives on spatial econometrics: Linear smoothing with structured models. *Journal of Regional Science* 52 (2), 192–209.
- Netz, J. S., Taylor, B. A., 2002. Maximum or minimum differentiation? Location patterns of retail outlets. *The Review of Economics and Statistics* 84 (1), 162–175.
- Noel, M., 2007. Edgeworth price cycles: Evidence from the Toronto retail gasoline market. *Journal of Industrial Economics* 55 (1), 69–92.
- Pennerstorfer, D., 2009. Spatial price competition in retail gasoline markets: Evidence from Austria. *The Annals of Regional Science* 43 (1), 133–158.
- Pennerstorfer, D., Schmidt-Dengler, P., Schutz, N., Weiss, C., Yontcheva, B., 2014. Information and price dispersion: Evidence from retail gasoline, Mimeo.



- Pennerstorfer, D., Weiss, C., 2013. Spatial clustering and market power: Evidence from the retail gasoline market. *Regional Science and Urban Economics* 43 (4), 661–675.
- Pinkse, J., Slade, M. E., 2010. The future of spatial econometrics. *Journal of Regional Science* 50 (1), 103–117.
- Pinkse, J., Slade, M. E., Brett, C., 2002. Spatial price competition: A semiparametric approach. *Econometrica* 70 (3), 1111–1153.
- Salop, S. C., 1979. Monopolistic competition with outside goods. *Bell Journal of Economics* 10 (1), 141–156.
- Stahl, D. O., 1989. Oligopolistic pricing with sequential consumer search. *American Economic Review* 79 (4), 700–712.
- Statistik Austria, 2006. Kfz-bestand 2005, available online at [http://www.statistik.at/web\\_de/static/kfz-bestand\\_2005\\_021489.pdf](http://www.statistik.at/web_de/static/kfz-bestand_2005_021489.pdf), last accessed 19/07/2013.
- Thomadsen, R., 2005. The effect of ownership structure on prices in geographically differentiated industries. *RAND Journal of Economics* 36 (4), 908–929.
- Varian, H. R., 1980. A model of sales. *American Economic Review* 70 (4), 651–659.
- Verlinda, J. A., 2008. Do rockets rise faster and feathers fall slower in an atmosphere of local market power? Evidence from the retail gasoline market. *Journal of Industrial Economics* 56 (3), 581–612.
- Wang, W., Lee, L.-F., 2013. Estimation of spatial autoregressive models with randomly missing data in the dependent variable. *Econometrics Journal* 16 (1), 73–102.

Figure 1: A stylized network of firms

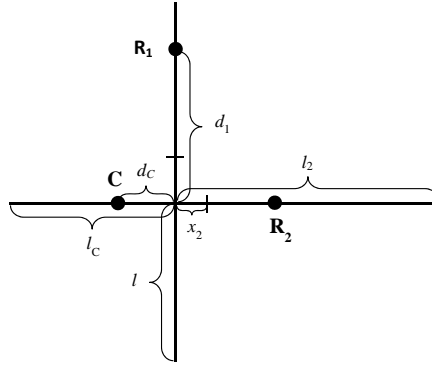


Figure 2: The impact of a price shock on the initiator by centrality

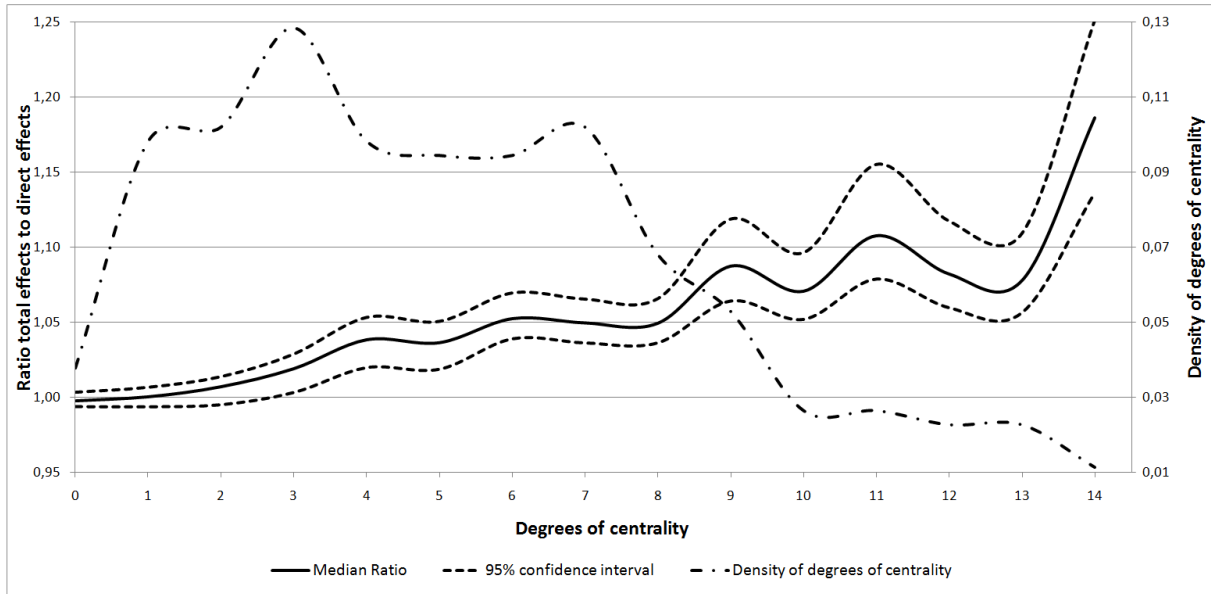


Figure 3: The impact of a price shock on other stations by centrality

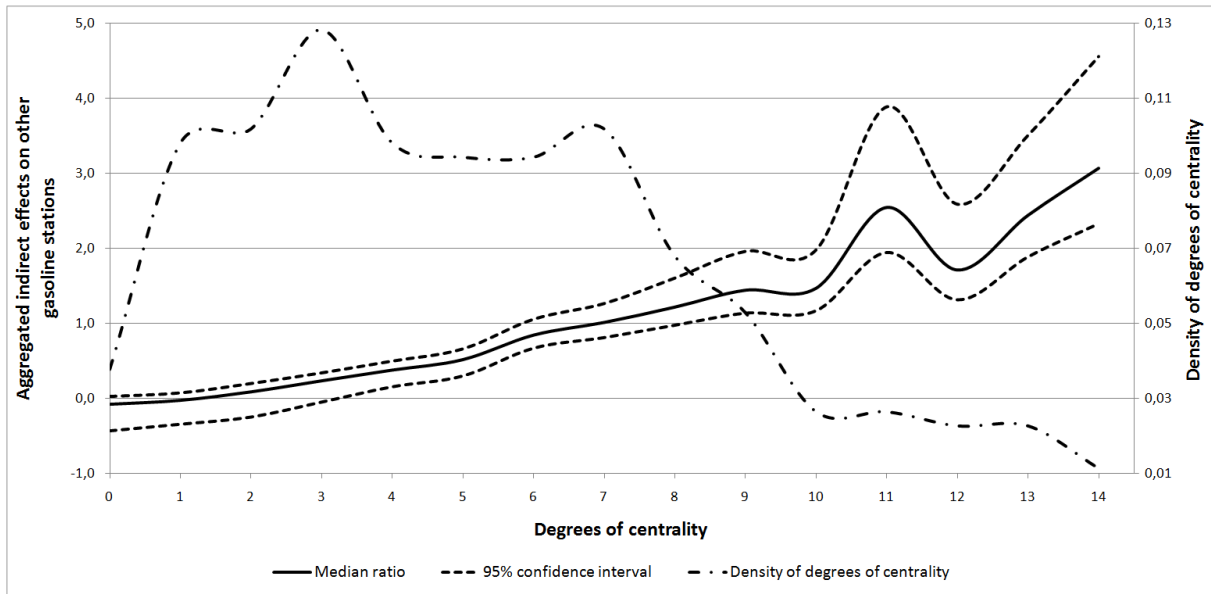


Table 1: Gasoline stations in Vienna by Brand

<b>Brand</b>	<b>Outlets</b>	<b>in %</b>
BP	76	27.8%
OMV	35	12.8%
SHELL	30	11.0%
<i>Major Brands</i>	<i>141</i>	<i>51.7%</i>
ESSO	23	8.4%
AVANTI	21	7.7%
AGIP	18	6.6%
STROH	11	4.0%
JET	10	3.7%
ARAL	2	0.7%
<i>Minor Brands</i>	<i>85</i>	<i>31.1%</i>
<i>Unbranded</i>	<i>47</i>	<i>17.2%</i>
<i>Total</i>	<i>273</i>	<i>100.0%</i>

Table 2: Definition and descriptive statistics of the variables

		Mean	Min	# of
		Std. Dev.	Max	Obs. Share Missing
PRICE	Price of one liter diesel in Euro cents	75.515	62.426	3,051
DEGREE ( <i>dc</i> )	Degree centrality for $H = 5$	6.449	92.900	3,051
DISTANCE NEXT	Driving time to the nearest neighbor in minutes	5.915	0	3,051
COMMUTERS	Ratio of commuters assigned to station over commuters and non-commuters in percent	3.399	17	3,051
log POP DENS	Log of the population density of the district in inhabitants per square km	1.668	0.050	3,051
log PREMISES	Log of the average land price for premises in the district in Euros per square meter	0.980	4.680	3,051
TRAFFIC	Dummy variable set equal to one if road traffic at the location is heavy	45.860	35.538	3,051
COMPANY	Dummy variable set equal to one if the location is owned by the company	5.634	80.155	3,051
SERVICE	Dummy variable set equal to one if the location offers attendance service	8.495	7.196	3,051
LARGE	Dummy variable set equal to one if the ground surface of the site is $> 2000m^2$	0.804	10.127	3,051
<b>Fixed effects</b>				
Brands	9 brands, unbranded stations left out as a reference group	5.316	5.169	2,542
Time Periods	22 periods, first period left out as a reference group	0.110	5.638	0.167
		0.768	0	3,051
		0.782	1	2,951
		0.353	1	0.033
		0.244	1	2,888
			0	0.053
			1	3,017
			1	0.011

Table 3: Regression results on gasoline prices

Specification Method Variables	[1]		[2]		[3]		[4]	
	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.
$Wp$	0.445	0.032	***	0.136	0.028	0.140	0.227	0.187
$WC^{dc}p$								
CONSTANT	29.808	5.414	***	0.138	0.518	0.136	0.652	0.183
DEGREE ( $dc$ )			***	5.122	19.720	5.239	0.678	-2.750
DISTANCE NEXT			***	0.019	-0.001	0.018	-0.008	0.008
COMMUTERS			***	0.068	0.201	0.062	0.169	0.027
log POPDENS			**	0.013	0.037	0.014	0.005	0.005
log PREMISES				0.111	0.011	0.100	0.121	0.043
COMPANY			*	0.847	1.779	0.723	0.695	0.338
LARGE				0.452	1.054	0.410	0.731	0.265
TRAFFIC				0.154	0.147	0.142	-0.029	0.058
SERVICE			***	0.163	0.318	0.144	0.077	0.078
BRAND:			***	0.258	0.604	0.238	0.750	0.208
BP (Major)	1.919	0.494	***	0.481	1.452	0.441	1.747	0.352
OMV (Major)	1.585	0.513	***	0.500	0.924	0.453	1.524	0.345
SHELL (Major)	1.582	0.530	***	0.515	1.079	0.472	1.732	0.361
AGIP	1.161	0.535	**	0.522	0.449	0.490	1.111	0.359
ARAL	2.096	0.746	***	0.724	2.006	0.674	2.087	0.365
AVANTI	0.495	0.503		0.490	0.110	0.448	0.329	0.265
ESSO	1.267	0.544	**	0.530	0.752	0.491	1.144	0.355
JET	0.490	0.589		0.575	0.288	0.545	0.776	0.397
STROH	0.281	0.497		0.484	-0.170	0.440	0.230	0.283
$W_{error}\epsilon$	0.418	0.041	***	0.041	0.186	0.084	-0.293	
Additional interaction effects								
$W\tilde{X}$ and $WC\tilde{X}$		no		no		yes		no
Excluded Instruments								$W\tilde{X}, WC^{dc}\tilde{X}, W_{error}\tilde{X}$ $W_{error}W\tilde{X}, W_{error}WC^{dc}\tilde{X}$
# of Obs.		3,051		3,051		3,051		3,051
$\ell$		-4,387.991		-4,374.812		-4279.177		
$LR(\chi^2)$ -test ( $df$ )				26.359 (2)		217.628 (40)		
$p$ -value				0.000		0.000		
$R^2$								0.975
$J(\chi^2, df = 96)$ -test								95.06 $p$ -value = 0.508
$F(df = 98, 184)$ -test on $Wp$								40.90 $p$ -value = 0.000
$F(df = 98, 184)$ -test on $WCp$								45.73 $p$ -value = 0.000

\*\*\* significant at 1%, \*\* significant at 5%, \* significant at 10%. Inference is based on a variance-covariance matrix of  $v$  that is clustered at the station level.  $LR$ -tests of the respective nested specification without centrality ( $H_0$ ) against the particular HI.  $\tilde{X}$  includes all variables (including  $C_4$ ) that vary at the station level.  $\tilde{X}$  includes all variables with cross-sectional variation. All models include dummies for missing explanatory variables as well as time period fixed effects.