RUHR
ECONOMIC PAPERS

James P. LeSage
Colin Vance
Yao-Yu Chih

A Bayesian Heterogeneous Coefficients Spatial Autoregressive Panel Data Model of Retail Fuel Price Rivalry

RWI #617

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Prof. Dr. Thomas K. Bauer
RUB, Department of Economics, Empirical Economics
Phone: +49 (0) 234/3 22 83 41, e-mail: thomas.bauer@rub.de

Prof. Dr. Wolfgang Leininger
Technische Universität Dortmund, Department of Economic and Social Sciences
Economics – Microeconomics
Phone: +49 (0) 231/7 55-3297, e-mail: W.Leininger@wiso.uni-dortmund.de

Prof. Dr. Volker Clausen
University of Duisburg-Essen, Department of Economics
International Economics
Phone: +49 (0) 201/1 83-3655, e-mail: vclausen@vwl.uni-due.de

Prof. Dr. Roland Döhrn, Prof. Dr. Manuel Frondel, Prof. Dr. Jochen Kluve
RWI, Phone: +49 (0) 201/81 49-213, e-mail: presse@rwi-essen.de

Editorial Office

Sabine Weiler
RWI, Phone: +49 (0) 201/81 49-213, e-mail: sabine.weiler@rwi-essen.de

Ruhr Economic Papers #616

Responsible Editor: Manuel Frondel

All rights reserved. Bochum, Dortmund, Duisburg, Essen, Germany, 2016

ISSN 1864-4872 (online) – ISBN 978-3-86788-717-5

The working papers published in the Series constitute work in progress circulated to stimulate discussion and critical comments. Views expressed represent exclusively the authors' own opinions and do not necessarily reflect those of the editors.

Ruhr Economic Papers #617
James P. LeSage, Colin Vance, and Yao-Yu Chih

A Bayesian Heterogeneous Coefficients Spatial Autoregressive Panel Data Model of Retail Fuel Price Rivalry

RWI

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

http://dx.doi.org/10.4419/86788717
ISSN 1864-4872 (online)
ISBN 978-3-86788-717-5

James P. LeSage, Colin Vance, and Yao-Yu Chih [1]

A Bayesian Heterogeneous Coefficients Spatial Autoregressive Panel Data Model of Retail Fuel Price Rivalry

Abstract

We apply a heterogenous coefficient spatial autoregressive panel model from Aquaro, Bailey and Pesaran (2015) to explore competition/cooperation by Berlin fueling stations in setting prices for diesel and E5 fuel. Unlike the maximum likelihood estimation method set forth by Aquaro, Bailey and Pesaran (2015), we rely on a Markov Chain Monte Carlo (MCMC) estimation methodology. MCMC estimates as applied here with non-informative priors will produce estimates equal to those from maximum likelihood, a point we demonstrate with a Monte Carlo experiment. We explore station-level price mark-ups using over 400 fueling stations located in and around Berlin, average daily diesel and e5 fuel prices, and refinery cost information covering more than 487 days. The heterogeneous coefficients spatial autoregressive panel data model uses the large sample of daily time periods to produce spatial autoregressive model estimates for each fueling station. These estimates provide information regarding the price reaction function of each station to neighboring stations. This is in contrast to conventional estimates of price reaction functions that average over the entire cross-sectional sample of stations. We show how these estimates can be used to infer competition versus cooperation in price setting by individual stations. The empirical results reveal a mix of competitive and collusive price setting, with some evidence that stations located near others of the same brand tend toward collusion, while those located near rival brands tend toward competition.

JEL Classification: C11, C23, L11

Keywords: Spatial panel data models; Markov Chain Monte Carlo; spatial autoregressive model; observation-level spatial interaction

May 2016

---

1 James P. LeSage, Texas State University; Colin Vance, RWI; Yao-Yu Chih, Texas State University. *- The authors express their gratitude to Dr. Alexander Kihm, whose contributions made the assembly of the data possible. Colin Vance has been partly supported by the Collaborative Research Center Statistical Modeling of Nonlinear Dynamic Processes (SFB 823) of the German Research Foundation (DFG), within the framework of Project A3, Dynamic Technology Modeling.* *- All correspondence to: James P. LeSage, Fields Endowed Chair, Texas State University, Department of Finance & Economics, San Marcos, TX, USA 78666, e-mail: james.lesage@txstate.edu*

1 Introduction

There is a great deal of literature on regional tax competition between local governments within a country (Allers and Elhorst, 2005, Elhorst and Fréret, 2009), gas station pricing (Pennerstorfer, 2009, Kihm et al. 2016) hospital pricing (Mobley, 2003), research activity competition between economics departments (Elhorst and Zigová, 2014), and so on. Empirical investigations often rely on spatial econometric methods developed to analyze spatially dependent cross-sectional and panel data.

The basic methodology involves comparing behavioral outcomes in one region (or observational unit such as gas station, hospital, etc.) to actions taking place in neighboring regions, or behavioral reactions taken by an individual or institution to actions taken by a more general type of neighbor, a peer group or a set of peer institutions. Spatial autoregressive processes/models represent a parsimonious way to specify a *global* relationship between a sample of (say N) regions/institutions/individuals and the average behavior of neighboring regions/insitutions/peers in the sample. By global, we mean that the sample of size N produces a scalar parameter indicating the average strength, sign, and statistical significance of reaction, where averaging takes place over the sample of size N. This allows an inference regarding the presence or absence of a positive/negative/insignificant reaction by the *typical* observational unit (region, institution, or individual) to actions of neighbors. For example, we might be able to conclude that on average over the sample of N regions we see statistical evidence of a negative and significant reaction function involving tax rates set by the typical region to average tax rates set by neighboring regions. This could be interpreted as evidence in favor of tax competition between regions in our sample.

A more ideal situation would allow inference regarding how each of the *individual* observational units $i = 1, \ldots, N$ react to actions taken by each unit i's neighboring units. Some regions/institutions/individuals might exhibit competitive reactions vis à vis their neighbors, while others react in a cooperative fashion, or do not react at all. This is an ideal situation because competition/cooperation reflect outcomes of institutional/individual decisions, which we might expect to vary across the sample of observational units.

Aquaro, Bailey and Pesaran (2015) make the observation that space-time panel data samples covering longer time spans are becoming increasingly prevalent. If we let N denote the number of spatial units in the sample and T the number of time periods, panel data sets

with sufficiently large T allow us to exploit sample data along the time dimension to produce spatial autoregressive parameter estimates for all N spatial units. In our setting where we wish to examine competition/cooperation between price setting behavior of individual fueling stations, individual estimates of the reaction function for each station to neighboring stations' pricing actions holds a great deal of intuitive appeal. The majority of studies that address competition in the retail gasoline market average over the sample of stations, not allowing for the possibility that some stations interact with their neighbors in a collusive manner while others interact in a competitive manner.

Our heterogeneous coefficient spatial autoregressive model is in contrast to conventional static spatial panel models where a single (scalar) dependence parameter is estimated that relates the NT decision outcomes in the vector y and the $NT-$vector of spatial lags $(I_T \otimes W)y$, representing a linear combination of neighboring unit decisions. The scalar dependence parameter averages the relationship over all N fueling stations and T time periods.[1] Aquaro, Bailey and Pesaran (2015) argue there are a large number of situations where it is plausible to believe that the level of interaction between observational units differs greatly when considering spatial interaction patterns, and our fueling station price setting interaction represents one such situation.

In section 2 we set forth the heterogeneous coefficient SAR model from Aquaro, Bailey and Pesaran (2015), and discuss how it will be applied to our examination of Berlin fueling station price setting behavior. Section 2.1 discusses interpretation of the model estimates, a topic not covered in Aquaro, Bailey and Pesaran (2015).

Section 3 develops a model of (station-level) price conjectural variations that explicitly identifies station-level reaction functions, and shows how this relates to our HSAR specification. Estimates of net spatial spillovers consisting of spill-in plus spill-out effects for each station reflect unobservable conjectural price variations. The relative size of spill-in versus spill-out impacts associated with neighboring station price markup actions determine cooperative versus competitive price markup rivalry at the station-level.

In section 4 we outline a Markov Chain Monte Carlo (MCMC) procedure for estimation of the model parameters. Appendix A validates the algorithm by replicating some Monte Carlo experiments from Aquaro, Bailey and Pesaran (2015).

[1] Of course, the conventional static space-time panel model can allow for station-specific and time-specific fixed effects in an attempt to ameliorate the restrictiveness of the model. This however amounts to allowing for station-specific and time-specific differences in the model intercept.

In section 5 we apply the model to daily price mark-ups of more than 400 fueling stations located in and around Berlin, Germany.[2] Given estimates that we can use to infer competition versus cooperation in price-setting behavior, section 5.3 turns attention to analysis of factors that enhance or deter competition in price setting.

2 The heterogenous spatial autoregressive model

The heterogeneous SAR model of Aquaro, Bailey and Pesaran (2015) (which we label HSAR hereafter) can be written as in (1), where w_{ij} represents the i,jth element of a row-normalized spatial weight matrix with $w_{ii} = 0$.

$$HSAR: \; y_{it} \;=\; \alpha_i + \psi_i \sum_{j=1}^{N} w_{ij} y_{jt} + \varepsilon_{it}, \; i = 1, 2, \ldots, N, \; t = 1, 2, \ldots, T \tag{1}$$

The disturbances ε_{it} are assumed distributed independently, and for our purposes we can assume independent normal distributions, $\varepsilon_{it} \sim N(0, \sigma_i^2)$.[3]

The HSAR model can be written in matrix notation shown in (2) by stacking regional units,

$$y_t \;=\; \alpha + \Psi W y_t + \varepsilon_t \tag{2}$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)'$, $\Psi = \text{diag}(\psi), \psi = (\psi_1, \psi_2, \ldots, \psi_N)'$, $W = w_{ij}, i, j = 1, \ldots, N$, $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{Nt})'$, $\varepsilon_{it} \sim N(0, \sigma^2 I_N)$, $\sigma^2 = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2)'$.

The data generating process for the HSAR model can be written as:

$$y_t \;=\; (I_N - \Psi W)^{-1}(\alpha + \varepsilon_t), \; t = 1, \ldots, T \tag{3}$$

[2] The fueling stations employed are not a sample, but rather the universe of stations in the Berlin region.

[3] Aquaro, Bailey and Pesaran (2015) show that quasi-maximum likelihood (QML) estimates are robust with respect to two error generating processes, a Gaussian $\varepsilon_{it} \sim IIDN(0, \sigma_i^2)$, as well as a non-Gaussian IID chi-square variate where: $\varepsilon_{it}/\sigma_i^2 \sim IID(\chi^2(2) - 2)/2$, with σ_i^2 generated as independent draws from $\chi^2(2)/4 + 0.5$, for $i = 1, \ldots, N$. This requires use of a variance matrix estimate based on the sandwich formula (which they provide) for their QML procedure.

2.1 Interpreting the HSAR model coefficients

The HSAR specification data generating process is shown in (4), where we have replaced the dependent variable y_t with $p_t = \text{price}_t - \text{cost}_t$, where cost_t is the refinery cost of diesel or e5 fuels, making p_t the price markup. This is the dependent variable we employ in our empirical application.

$$p_t = (I_N - \Psi W)^{-1}(\alpha + \varepsilon_t),\ t = 1, \ldots, T \tag{4}$$

Consider the partial derivatives of the reduced form HSAR model shown in (5). These expressions are derived from the reduced form (3), after taking into account the fact that Ψ, α do not change over time. The partial derivatives show how price markup respond to (random) changes in price markups by rival stations.

$$\partial P/\partial \varepsilon = \begin{pmatrix} \partial p_1/\partial \varepsilon_1 & \partial p_1/\partial \varepsilon_2 & \ldots & \partial p_1/\partial \varepsilon_N \\ \partial p_2/\partial \varepsilon_1 & \partial p_2/\partial \varepsilon_2 & \ldots & \partial p_2/\partial \varepsilon_N \\ \vdots & \vdots & \ddots & \vdots \\ \partial p_N/\partial \varepsilon_1 & \partial p_N/\partial \varepsilon_2 & \ldots & \partial p_N/\partial \varepsilon_N \end{pmatrix} = (I - \Psi W)^{-1} \tag{5}$$

Expression (5) is an $N \times N$ matrix, since a change in a single station's price markup could (potentially) impact price markups of all other stations, with the strength of these other-station impacts determined by the levels of dependence between rivals (Ψ) and their relative spatial locations (indicated by non-zero weights assigned to rivals by the matrix W).

The main diagonal of the matrix represent own-partial derivatives ($\partial p_i/\partial \varepsilon_i$), while the off-diagonal elements are cross-partial derivatives ($\partial p_j/\partial \varepsilon_i$) showing external impacts on rival stations. If we sum the off-diagonal elements across row i of the matrix, we have the *cumulative spill-in impact*, that of rival station's ($j \neq i$) price changes on station i behavior. Summing off-diagonal elements down the column i produces a *cumulative spill-out impact* showing how changes in station i price markup impacts rival stations $j \neq i$.

The model provides estimates of $\alpha_i, \psi_i, \sigma_i^2, i = 1, \ldots, N$ for each station i. We can inter-

pret the parameters α_i as station-specific fixed effects that reflect differences that arise from: branding or location advantages, station cost factors, traffic access patterns, etc. We note also that separate variance scalar estimates for each station accommodate heteroscedasticity.

3 A theoretical motivation for the SAR specification

3.1 A price reaction function motivation

A price conjectural variation is firm i's anticipated response from a rival firm when firm i changes its price. Liang (1989) shows that if station i has a price conjectural variation, the equilibrium price and output that results can range from a competitive to monopolistic outcome depending on this conjecture.

We first show that the price conjectural variations of each station can be estimated directly in the station's price reaction function. We specified the demand function for station i as a simple linear relationship:

$$q_i = a_i - b_i p_i + f_i W_i P \tag{6}$$

where q_i is output for station i, p_i is the price of station i and $W_i P$ represents the ith row of the $N \times N$ matrix W that has non-zero elements in row i for each rival station. P is an $N \times 1$ vector of prices, so that $W_i P$ reflects a linear combination of prices by station i's rivals.

Price reaction functions arise in oligopolistic markets as our representative station i realizes that changes in price may provoke subsequent price changes by its rivals. Given this type of situation, station i's first order profit maximization conditions can be expressed in terms of its price reaction function $R_i(W_i P)$ to it rivals prices $(W_i P)$.

Assume station i's cost (c_i) is independent from its pricing strategy.[4] Station i's profit function is:

[4] We use the same refinery cost of fuel for all stations in our empirical implementation.

$$\begin{aligned}\pi_i &= p_i \times q_i - c_i = p_i(a_i - b_i p_i + f_i W_i P(p_i)) - c_i \\ &= a_i p_i - b_i p_i^2 + f_i p_i W_i P(p_i) - c_i \end{aligned} \tag{7}$$

The first-order profit maximization condition is:

$$\frac{\partial \pi_i}{\partial p_i} = a_i - 2b_i p_i + f_i W_i P + f_i p_i (\partial W_i P / \partial p_i) = 0 \tag{8}$$

Let the $N \times N$ adjacency matrix $W = [w_{ij}]$ identify rival stations using $w_{ij} > 0$ and zero values for other stations. If we let the conjectural variation $cv_i = (\partial W_i P / \partial p_i)$, station i's price reaction function $R_i(W_i P)$ from (8) is:

$$R_i(W_i P) : p_i = \frac{a_i}{2b_i - f_i cv_i} + \frac{f_i}{2b_i - f_i cv_i} W_i P, \tag{9}$$

where W_i is ith row vector of W, and P is an $N \times 1$ price vector.

Under the assumption that station i's sales are more responsive to own price than the substitute/rival's price ($b_i > f_i > 0$), the slope of station i's price reaction function is positive (Liang 1989) when $cv_i < 2b_i/f_i$. This condition requires that firm i's price elasticity of sales are more responsive to own price changes than to (an average of) rivals prices, which seems reasonable. We note also that the slope becomes steeper if stations assume ($cv_1 > 0$), and flatter if station i has a conjecture ($cv_1 < 0$).

To illustrate the relationship between station i's conjectural variation about price (cv_i) and its strategic pricing behavior, two cases are considered under general assumptions and under symmetry.

Case 1: station i anticipates rivals' price setting ($cv_i = \partial W_i P / \partial p_i > 0$)

Beginning from the Bertrand equilibrium A with equilibrium price (p_i^1,$W_i P^1$), when station i's rivals decide to increase fuel prices[5], other stations' (which we reference as j) reaction functions shift from $R_j^1(p_i)$ to $R_j^2(p_i)$ (see Figure 1).

[5]We are assuming this is for reasons other than a response to station i's increase in fuel price.

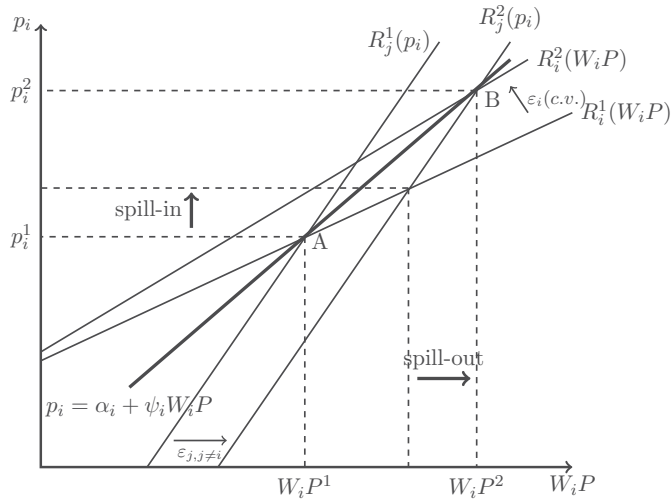

Figure 1: Case 1: Cooperative outcome

In case 1, station i expects that when station i increases price in response to the price change of its rivals, rivals will "cooperate" by not cutting their price in response ($cv_i > 0$). A result of this conjecture is that station i becomes more reliant on rivals' price decisions, and therefore, station i's price reaction function becomes steeper (from $R_i^1(W_iP)$ to $R_i^2(W_iP)$). After this adjustment, the Bertrand equilibrium changes from A to B with equilibrium price (p_i^2,W_iP^2).

By connecting two equilibria, we get station i's equilibrium price strategy as shown by the line $p_i = \alpha_i + \psi_i W_i P$ in Figure 1. Specifically, if station i's price conjectural variation is positive ($cv_i > 0$), we have a positive slope ($\psi_i > 0$) reflecting equilibrium price behavior.

Case 2: station i anticipates rivals price setting ($cv_i = \partial W_i P / \partial p_i < 0$)

In this scenario where station i expects that rivals will reduce their price in response to a station i's price increase ($cv_i < 0$), the slopes of the reaction functions are upward sloping as shown in Figure 2. However, this type of price conjecture ($cv_i < 0$) makes station i less sensitive to price decisions of its rivals. This leads to station i's price reaction function

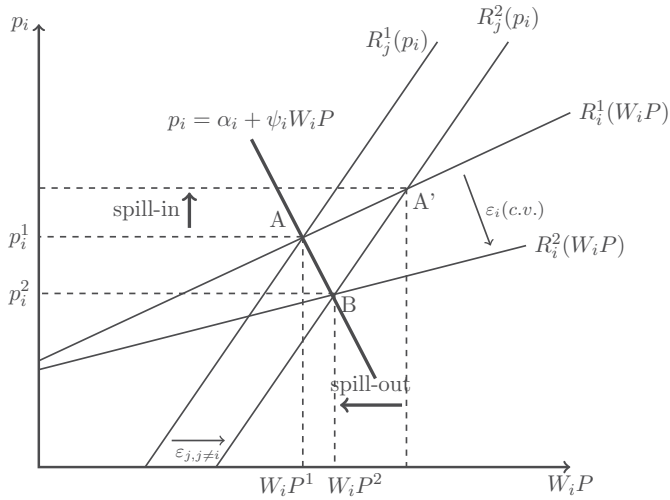

Figure 2: Case 2: Competitive outcome

becoming flatter (e.g., moving from $R_i^1(W_iP)$ to $R_i^2(W_iP)$ in Figure 2), but the reaction function is still positively sloped.

Despite this, there exist conditions where the equilibrium can move from A to B when station i's rivals decide to increase price. In this type of situation, station i's equilibrium price behavior has a negative slope ($\psi_i < 0$), as shown in Figure 2. As can be seen from the figure, the negative slope arises when net spillovers = spillin effects + spillout effects are negative. We note that although the case considered here was that of an increase in price (markup), the model produces a symmetric result for decreases in price.

3.2 SAR as a steady-state equilibrium process

The price reaction function assumes that stations learn over time about the price setting behavior of rivals. An econometric motivation for how the simultaneous autoregression model specification can be viewed as a steady-state process that arises from stations observing/learning from past prices of rivals is given here.[6]

The parameters ψ_i for each station show how the price markup of station i relates to that of a linear combination of price markups by neighboring stations. (We will refer to

[6]This motivation follows along the lines of an argument developed in LeSage and Pace (2009).

price rather than price markups in the sequel for the sake of brevity, but we are referring to the price markup dependent variable p_{it}.)

We begin with a situation where each station sets price on day t after examining neighboring station prices on the previous day, using the partial adjustment relation in (10).

$$p_{i,t} = \alpha_i + \psi_i \sum_{j=1}^{N} w_{i,j} p_{j,t-1} + \varepsilon_{i,t} \tag{10}$$

We can view (10) as a partial adjustment mechanism whereby station i (partially) adjusts ($-1 < \psi_i < 1$) to prices markups of neighboring stations. Station i also adds a station-specific factor α_i into its pricing markup decision.

Rewriting (10) in matrix/vector form shown in (11), we can recursively replace the lagged $N \times 1$ vector p_{t-1} on the right-hand-side. Doing this for q periods leads to:

$$\begin{aligned} p_t &= \left(I_N + \Psi W + \Psi^2 W^2 + , \ldots, + \Psi^{q-1} W^{q-1}\right) \iota_N \alpha + \Psi^q W^q p_{t-q} + u_t \\ u_t &= \varepsilon_t + \Psi W \varepsilon_{t-1} + , \ldots, + \Psi^{q-1} W^{q-1} \varepsilon_{t-(q-1)} \end{aligned}$$

where $p_t, u_t, \alpha, \varepsilon_t$ are $N \times 1$ vectors, Ψ is an $N \times N$ diagonal matrix and W is an $N \times N$ spatial weight matrix. These expressions can be simplified by noting that $E(\varepsilon_{t-r}) = 0, r = 0, \ldots, q-1$, implies that $E(u_t) = 0$, and, the magnitude of elements in the $N \times 1$ vector $\Psi^q W^q p_{t-q}$ becomes small for large q. This magnitude becomes small because spatial autoregressive processes assume that $|\psi_i| < 1, i = 1, \ldots, N$, and that the matrix W is normalized to have row-sums of unity. This results in the matrix W having a maximum eigenvalue of one.

A consequence of this is that we can interpret the empirical spatial autoregressive model relationship as the outcome or expectation of a long-run equilibrium or steady state shown in (11).

$$\lim_{q \to \infty} E(p_t) = (I_N - \Psi W)^{-1} \iota_N \alpha \tag{11}$$

The expectation in (11) is however equal to the expectation of the data generating

process for our spatial autoregressive model of price markup setting behavior. The HSAR specification can be viewed as modeling a *simultaneous dependence process* that arises as a steady state equilibrium when individual stations use a partial adjustment reaction function in response to past period price markup actions taken by neighboring stations.

It should be noted there is no dynamic component to our HSAR specification, which is essential to our exploiting the time dimension of the sample data to produce estimates α_i, ψ_i for each observational unit. Rather, we interpret the coefficient estimates $\hat{\alpha}_i, \hat{\psi}_i$ as reflecting (a sample data-based) estimate of station i's reaction to a change in the steady-state equilibrium markup behavior of neighboring stations. These estimates reflect changes required to ultimately produce a new steady-state equilibrium. The story is one of comparative statics, with no information regarding the dynamics of how we move from one steady-state to another, nor the speed of this movement.[7]

4 MCMC estimation of the model

Bayesian estimation requires that prior distributions be assigned for the model parameters. However, we use normal priors for the parameter $\alpha_i, \psi_i, i = 1, \ldots, N$, with zero prior means and extremely large variances, to produce posterior estimates equivalent to those from maximum likelihood estimation.[8] We also rely on uninformative priors for the parameters $\sigma_i, i = 1, \ldots, N$, as prior information for these is unlikely to be available in applied modeling situations. As is traditional, we assume the priors for the parameters $\alpha_i, \psi_i, \sigma_i^2$ are independent.

An advantage of MCMC estimation may be computational speed. Basically, MCMC estimation decomposes a complicated problem involving $N \times 1$ parameter vectors $\alpha_i, \psi_i, \sigma_i^2$ into a sequence of simpler problems involving conditional distributions which are typically simple. Our MCMC estimation proceeds by sequentially sampling from the complete sequence of conditional distributions for: 1) the N different parameters α_i, the N different scalar parameters ψ_i and N different scalar noise variances σ_i^2. A single pass through the sampler involves evaluating only $3N$ different conditional distributions. Each of these con-

[7] The interested reader should consult LeSage and Pace (2009) for detailed discussion of how simultaneous dependence specifications are interpreted, along with a wide variety of economic motivations for this type of specification.

[8] Use of a normal prior for the parameters ψ_i might be viewed as problematical given a theoretical upper bound of unity for this parameter. However, during MCMC estimation we reject candidate values of ψ_i that exceed unity inside our Metropolis-Hastings sampling scheme.

ditional distributions is relatively simple to sample from, and involves calculations based on matrices or vectors of small dimensions. In contrast, QML estimation (as set forth in Aquaro, Bailey and Pesaran (2015)) requires optimization of the likelihood in (12) over $N \times 3$ parameters.

$$\begin{aligned} \ln\ \mathrm{L}(\psi_i, \alpha_i, \sigma_i^2) &= \frac{-NT}{2}\ln(2\pi) - \frac{T}{2}\sum_{i=1}^{N}\ln\sigma_i^2 + T\ln|I_N - \Psi W| \\ &- \frac{1}{2}\sum_{i=1}^{N}(y_i - \psi_i y_i^*)'(y_i - \psi_i y_i^*)/\sigma_i^2 \\ y_i^* &= \sum_{j=1}^{N} w_{i,j} y_j \end{aligned} \tag{12}$$

In our case, N equals over 400 stations, making this a computationally challenging optimization problem.

Markov Chain Monte Carlo estimation consists of sampling draws from the complete sequence of conditional posterior distributions. These are derived from the log likelihood in (12), considering each parameter sequentially while assuming all others are known. In our case, where extremely large prior variances are used, the prior distributions do not play a material role in the posterior estimates or the conditional distributions. We are simply using MCMC as an alternative to maximizing the likelihood function.

Sampling begins with arbitrary values for the parameters $\alpha^{(0)} = \alpha_i, \psi^{(0)} = \psi_i, \sigma^{(0)} = \sigma_i^2$ for all $i = 1, \ldots, N$. We need to sample updated values for the parameters $\alpha_i, \psi_i, \sigma_i^2$ for each $i = 1, \ldots, N$. One pass through the sampler involves: 1) producing a draw for $\alpha_{i=1}$ used to update the arbitrary starting value $\alpha^{(0)}$, a draw for $\sigma_{i=1}^2$ to replace $\sigma^{(0)}$ and a draw for $\psi_{i=1}$ that updates $\psi^{(0)}$, 2) sampling from the three conditional distributions to update $\alpha_{i=2}, \sigma_{i=2}^2, \psi_{i=2}$, and 3) continue drawing from the three conditional distributions until we have updated $\alpha_{i=N}, \sigma_{i=N}^2, \psi_{i=N}$. One pass through the sampler requires sampling from $3N$ conditional distributions.

A number m such passes are carried out, with draws from some initial number of passes b discarded to allow the sampler to "burn-in". Posterior means, standard deviations, and other summary statistics for these distributions for the parameters are analyzed using the sample of $m-b$ retained draws. The number of passes usually is in the thousands to produce

an adequate sample of size $m - b$ on which to base posterior inference.

Despite the apparent computational intensity of evaluating these $3N$ conditional distributions thousands of times, the conditional posteriors for the parameters α_i, σ_i^2 take distributional forms that are known, and therefore easy to sample from. Further, calculation of values needed to produce samples from these distributions are easy to calculate and involve matrices/vectors of small dimension.

Specifics regarding the $3N$ conditional posterior distributions are presented, where we ignore the subscript i for notational simplicity when presenting the MCMC sampling scheme. It should also be noted that several simplifications arise in the conditional distributions because of our use of zero prior means and very large prior variance settings. This essentially allows us to ignore the prior distributions when constructing conditional distributions.

We calculate the mean and variance-covariance for α using the conditional posterior based on arbitrary starting values, that we label $\psi^{(0)}, \sigma^{(0)} = \sigma^2$ in (13), where we let y^* represent $\sum_{j=1}^{N} w_{i,j} y_j$.[9]

$$\begin{aligned} p(\alpha|\psi^{(0)}, \sigma^{(0)}) &\sim N(\alpha^*, \Sigma^*) \\ \alpha^* &= \Sigma^*(\iota_N'(y - \psi^{(0)} y^*) \\ \Sigma^* &= (1/N)\sigma^{(0)} \end{aligned} \tag{13}$$

An updated value that we label $\alpha^{(1)}$ can be obtained from a univariate normal distribution with mean α^* and variance equal to Σ^*. The updated values $\alpha^{(1)}$ will be used in place of $\alpha^{(0)}$ when calculating the conditional posterior for updating $\sigma^{(0)}$ based on the inverse Gamma distribution shown in (14).

$$\begin{aligned} p(\sigma^2|\alpha^{(1)}, \psi^{(0)}) &\sim IG(a_1, b_1) \\ a_1 &= T/2 \\ b_1 &= (y - \psi^{(0)} y^* - \iota_N \alpha^1)'(y - \psi^{(0)} y^* - \iota_N \alpha^{(1)})/2 \end{aligned} \tag{14}$$

We label the updated value produced by this draw $\sigma^{(1)}$ which replaces the initial value

[9] Recall, we are suppressing the subscript i.

$\sigma^{(0)}$ in the conditional posterior expression for ψ.

While the conditional posteriors for the parameters α_i, σ_i^2 take known distributional forms that are easy to sample from, the conditional posterior for the parameters ψ_i do not have this property. A Metropolis-Hastings (M-H) approach is used to sample these parameters based on the conditional posterior. For (M-H) sampling we require a *proposal distribution* from which we generate a candidate value for the parameter ψ_i, which we label $\psi_i^\star$.

We use a normal distribution as the proposal distribution along with a *tuned random-walk procedure* suggested by Holloway, Shankara, and Rahman (2002) to produce the candidate values for ψ. The procedure involves use of the current value ψ^c, a random deviate drawn from a standard normal distribution, and a tuning parameter c as shown in (15).[10]

$$\psi^\star = \psi^c + c \cdot N(0,1) \tag{15}$$

Expression (15) should make it clear why this type of proposal generating procedure is labeled a random-walk procedure. The goal of tuning the proposals coming from the normal proposal distribution is to ensure that the M-H sampling procedure *moves* over the entire conditional distribution. We would like the proposal to produce draws from the dense part of this distribution and avoid a situation where the sampler is stuck in a very low density part of the conditional distribution where the density or support is low.

To achieve this goal, the tuning parameter c in (15) is adjusted based on monitoring the acceptance rates from the M-H procedure during the MCMC drawing procedure. Specifically, if the acceptance rate falls below 40%, we adjust $c' = c/1.1$, which decreases the variance of the normal random deviates produced by the proposal distribution, so that new proposals are more closely related to the current value ψ^c. This should lead to an increased acceptance rate. If the acceptance rate rises above 60%, we adjust $c' = (1.1)c$, which increases the variance of the normal random deviates so that new proposals range more widely over the domain of the parameter ψ. This should result in a lower acceptance rate. The goal is to achieve a situation where the tuning parameter settles to a fixed value resulting in an acceptance rate between 40 and 60 percent. At this point, no further adjustments to the tuning parameter take place and we continue to sample from the normal proposal

[10] Recall, we are ignoring the subscript i in the expressions that follow.

distribution using the resulting tuned value of c.

We note that Aquaro, Bailey and Pesaran (2015) show that theoretical bounds on the parameters ψ_i are $\sup_i|\psi_i| < \max\{1/\|W_1\|, 1/\|W\|_\infty\}$, where $\|W_1\| = \max_{1\leq j\leq N}(\sum_{i=1}^{N}|w_{ij}|)$, the maximum absolute column sum norm, and $\|W\|_\infty = \max_{1\leq i\leq N}(\sum_{j=1}^{N}|w_{ij}|$, the maximum absolute row sum norm.[11] A sufficient, but not necessary condition for invertability of $(I_N - \Psi W)$ is that $-1 < \psi_i < 1$ for all i. The MCMC algorithm was coded to reject candidate values that fell outside the (-1,1) range, and draw a new candidate value in these cases until a value within the (-1,1) interval arose. However, as a practical matter, estimation did not produce candidate values outside the (-1,1) interval, so there appears to be no issue regarding inference at the boundary of the parameter space.[12]

The candidate value $\psi^\star$ as well as the current value that we label ψ^c are evaluated in the expression for the (logged) conditional posterior for ψ in (16). Note that we use updated sampled values for $\alpha^{(1)}$ and $\sigma^{(1)}$ when evaluating the conditional posterior in (16).

$$\begin{aligned} \ln(p(\psi)|\alpha^{(1)},\sigma^{(1)}) &= -(NT/2)\ln\pi\sigma^{(1)} + T\ln|I_N - \psi W| - \ln(e'e/2\sigma^{(1)}) \qquad (16) \\ e &= (y - \psi y^* - \iota_N\alpha^{(1)}) \end{aligned}$$

If $(\ln p(\psi^\star) - \ln p(\psi^c)) > \exp(1)$, we accept the candidate value $\psi^\star$ as an update for the current parameter ψ^c. If this condition is not true, we compare $\nu(\psi^c, \psi^\star)$ calculated using:

$$\nu(\psi^c,\psi^\star) = \min\left[1, \frac{p(\psi^\star|\alpha^{(1)},\sigma^{(1)})}{p(\psi^c|\alpha^{(1)},\sigma^{(1)})}\right] \qquad (17)$$

with a uniform random deviate (say r), and decide acceptance based on: $r < \nu(\psi^c, \psi^\star)$ (accept), set $\psi^{(1)} = \psi^*$, otherwise (reject). If we reject the candidate value, we simply set $\psi^{(1)} = \psi^c$, that is, we stay with the current value of ψ.

Having completed one pass through of the MCMC sampler updating all parameters: $\alpha^{(1)}$, $\sigma^{(1)}$ and $\psi^{(1)}$, we return to sample a second update of the parameters α, that relies on the most recently updated values $\sigma^{(1)}$, $\psi^{(1)}$, producing an updated vector that we label $\alpha^{(2)}$. This vector is used in the update based on the conditional distribution for σ^2 to produce a

[11] These bounds reduce to the condition in Lemma 2 of Kelejian and Prucha (2010) for the homogeneous case where $\psi_i = \psi$ for all i.

[12] The proportion of MCMC draws outside the (-1,1) interval can be interpreted an estimate of the posterior probability that ψ_i lies outside the interval.

new parameter draw $\sigma^{(2)}$, which is used when updating ψ to $\psi^{(2)}$. This process continues making m passes through the sampler to produce $m - b$ sets of draws for the parameters, where values from the first b (burn-in) passes are discarded to allow the sampler to achieve a steady-state and begin sampling from high density regions of the conditional posterior distributions of the parameters. The set of parameter draws $\alpha^{(m-b)}$, $\sigma^{(m-b)}$, $\psi^{(m-b)}$ can be used to calculate posterior means and standard deviations for the parameters. These draws reflect not conditional distributions of the parameters but rather the joint posterior distribution from which we draw inferences.

Appendix A presents results from a Monte Carlo experiment that follows the design set forth in Aquaro, Bailey and Pesaran (2015). Bias and root mean squared errors from the Monte Carlo experiment using our MCMC estimation approach are similar to those from QML estimation presented in Table A.1 of Aquaro, Bailey and Pesaran (2015).

5 The model applied to Berlin fueling stations

5.1 The sample data

Since September 2013, stations in Germany are legally obligated to post every price change, the precise time stamp, the geographic coordinates of the station, the operating hours and brand on an online portal, the so-called Market Transparency Unit for Fuel (Haucap, Heimeshoff, and Siekmann 2015). To access these data, a script was used to continuously retrieve entries from the site and store these on a server (Frondel, Vance and Kihm 2015). From the raw data, a balanced panel of daily station-level prices was created for over 14,000 filling stations in Germany.

We drew a sub-sample of $N = 414$ diesel and $N = 411$ e5 (unleaded gas) prices for fueling stations located in and around Berlin. Fuel prices were available for these stations over the period from June 1, 2014 to September 30, 2015, or $T = 487$ days. Prices are in nominal terms and include excise and value-added taxes, which were subtracted from the variable we designate as price (price_t). To measure the cost variable (cost_t), we use the daily refined diesel and gas prices reported in Rotterdam, where one of the major pipelines into Germany originates. As already noted, we use $p_t = \text{price}_t - \text{cost}_t$ as the dependent variable in our model.

5.2 Model estimates

In specifying the model, the issue of how to define neighboring stations that are specified in the matrix W arises. This issue was explored by comparing the log-likelihood function for model estimates based on 6, 8 and 10 nearest neighboring stations. Given the large number of stations, and a model that specifies a single matrix W for all stations, use of a weight matrix based on the number of nearest neighboring stations has the virtue of simplicity. A drawback to this approach is that the relevant market/competitor stations might differ for each station, a feature that our heterogeneous SAR model does not allow for. A related point is that LeSage and Pace (2014) show that estimates from spatial autoregressive models are not sensitive to small deviations in the weight structure around one that maximizes the model likelihood.

Table 1 shows the log-likelihood for models based on 6, 8 and 10 nearest neighbors. This involved evaluating the (log) likelihood shown in (12) using posterior means of the parameters $(\alpha_i, \psi_i, \sigma_i^2)$ constructed from 1,000 draws retained from a sample of 2,500 draws.

Table 1: Comparison of results based on 6, 8, and 10 nearest neighboring stations

Model	6 neighbors	8 neighbors	10 neighbors
Log Likelihood diesel	-202161.87	-201896.79*	-202245.01
Log Likelihood e5 (gas)	-202701.91*	-203521.03	-202807.15

* indicates log-likelihood maximum

From the table we see that models based on eight nearest neighbors for diesel and six neighbors for e5 gas exhibited the highest likelihood function values. In the following analysis of estimation results, we rely on estimates for diesel based on eight neighbors and for e5 based on six neighbors.

As already motivated, the model predicts that in cases where the net spillover effects (spill-out plus spill-in) are negative, we should see estimates of the parameters $\hat{\psi}_i < 0$, and where these are positive we have $\hat{\psi}_i > 0$. Figure 3 shows a plot of net spillovers (spillout plus spillin) versus parameter estimates $\hat{\psi}_i$ based on the diesel fuel price model, and Figure 4 a similar plot for e5 fuel. In the figures, stations for which estimates of ψ were not significantly different from zero based on lower 0.05 and upper 0.95 credible intervals are excluded.[13] Although stations where ψ is not significant are eliminated from the figures, we did not

[13]For diesel, 86 of 414 stations were not significant, and for e5, 63 of 411 stations.

suppress stations where net spillovers were not significant. Recall, our theoretical model predicts that negative ψ will result from negative net spillovers, which was one motivation for this decision. Also note that a zero estimate for ψ_i indicates that station i does not react to neighboring station price markups.

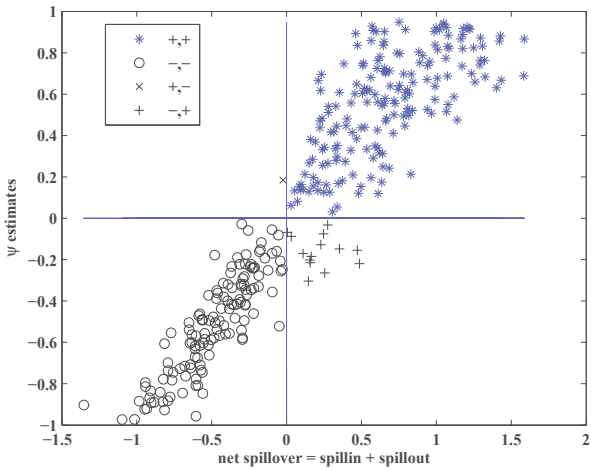

Figure 3: Diesel ψ estimates vs. net spillout effects

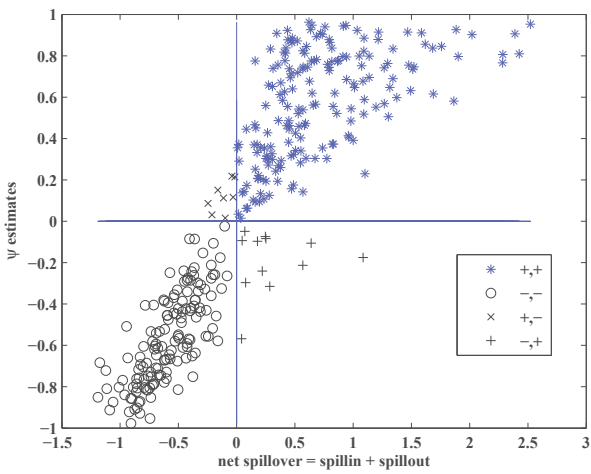

Figure 4: e5 fuel ψ estimates vs. net spillout effects

From figure 3, we see estimates of ψ_i that are highly consistent with our theoretical model. In the case of diesel, there were 328 significant ψ estimates and only a single case/station with a theoretically inconsistent positive estimate for ψ_i and negative net spillovers (the single 'x' in quadrant II). For e5 fuel shown in figure 4 we have 8 of the 348 theoretically inconsistent cases where ψ was positive and significant and net spillovers negative ('x' points in quadrant II). We note that our theoretical model does not rule out results in quadrant IV where net spillovers are positive, but $\psi < 0$. However, our interest centers on outcomes reflecting a cooperation scenario (quadrant I) and competitive scenario (quadrant III). For diesel fuel, of the 328 cases where ψ was statistically significant, there were $180/328 = 0.5488$ cooperation scenario (quadrant I) outcomes, and $133/328 = 0.4055$ competitive (quadrant III) outcomes. In the case of e5 fuel, there were $184/348 = 0.5287$ cooperative cases and $144/348 = 0.4138$ competitive.

To illustrate the locational/spatial aspect of competition (negative and significant net spillovers) versus cooperation (positive and significant net spillovers), these are mapped in Figure 5 and Figure 6 for diesel and e5 fuels. Spatial clustering of stations (along roadway segments) engaged in competitive price rivalry denoted with a '$*$' symbol, and clustering of stations involved in cooperative price setting (those with a 'o' symbols is evident in the maps. This clustering is indicative of interaction between spatially neighboring stations in their price markup decisions.

5.3 Analysis of competition/cooperation

The relationship between net spillover magnitudes and brands of fuel sold by each station was explored. We have six categories of fuel brands, Aral, Jet, Esso, Shell, Total, and a category for independents that we label Other.

A set of 12 regressions were run, six regressions that related net spillovers to the number of own-brand neighboring stations among the 8 neighbors for diesel, and among the 6 neighbors for e5 fuel. For example, net spillovers for the sample of 80 Aral brand stations was regressed on a constant term and a count of Aral stations among the 8 nearest neighbors of these 80 stations, and similar regressions for the samples of stations associated with other brands. We would expect a positive influence of same brand stations on the net spillovers, suggesting more cooperation in cases with higher counts of same brand stations among the neighboring stations. Results from these regressions are shown as the main diagonal

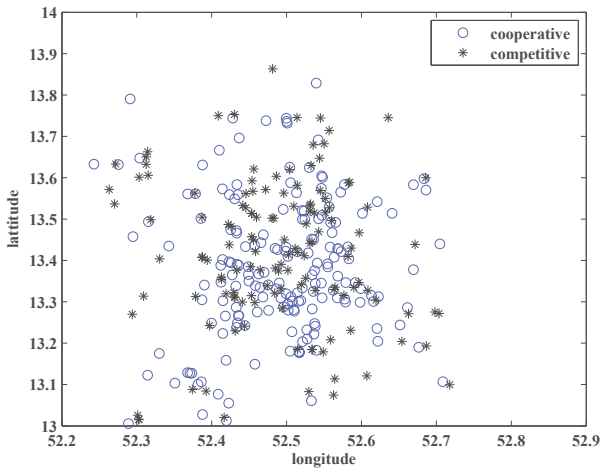

Figure 5: Diesel fuel map

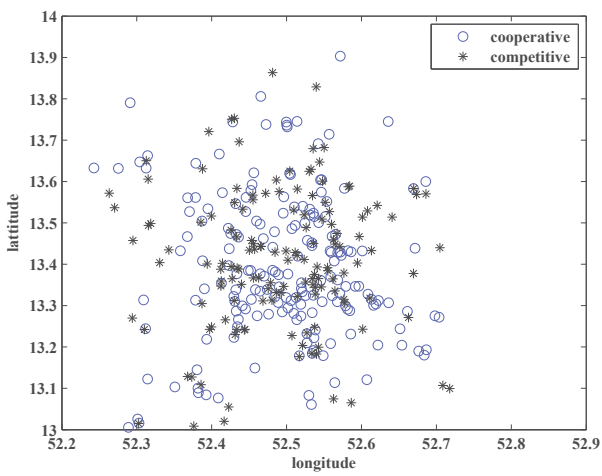

Figure 6: e5 fuel map

elements of the matrix shown in Table 2.

Another set of six regressions were carried out that related each sample of stations of brand i to counts of the number of other brand j stations among the set of 8 neighbors

for diesel (and another six regressions for the set of 6 neighbors for e5 fuel). Estimates from these regressions are presented as off-diagonal elements in the two matrices shown in Table 2.

For off-diagonal elements, we would expect a negative relationship between net spillovers for brand i stations and the counts of other brand j stations. A negative relationship would point to more competitive price markup behavior in situations where neighboring rival other brand stations were present. The regression coefficients on counts of each other brand station should indicate how the presence of one additional station of brand j impacts net spillovers.

Table 2: Brand rivalry regression results

8 neighbors diesel							
Brands	Aral	Jet	Esso	Shell	Total	Other	Sample size
Aral	0.1794***	0.0468	-0.1242	-0.1685*	-0.2720***	-0.1644**	80
Jet	-0.1356	0.1694	-0.1231	-0.0771	-0.0811	-0.1896	33
Esso	-0.1510	-0.4574*	-0.0047	0.0010	-0.0228	-0.0184	30
Shell	0.0525	-0.1131	-0.3957***	0.1039	-0.1946**	-0.1704**	71
Total	-0.1175	-0.1864	0.0959	-0.0188	-0.0087	-0.02258	53
Other	0.0129	0.0610	-0.0494	-0.0509	-0.1241*	0.0081	147
6 neighbors e5 fuel							
Brands	Aral	Jet	Esso	Shell	Total	Other	Sample size
Aral	0.1262	-0.0320	-0.0844	-0.0285	-0.2142	-0.1565	80
Jet	-0.3304	0.5437**	-0.5900*	-0.5218*	-0.3885	-0.4576*	33
Esso	0.2962	0.2403	-0.4015*	-0.6413**	-0.4994*	0.3477	30
Shell	-0.0479	-0.0865	0.0694	-0.0134	-0.0995	0.1467	70
Total	-0.1004	0.1622	-0.0479	0.0189	-0.0215	-0.0182	53
Other	-0.0798	0.1007	-0.0760	-0.1591*	-0.0356	-0.0023	145
* = 90% significance, ** = 95% significance, *** = 99% significance							

From the table main diagonal we see a positive and significant (at the 99% level) coefficient for Aral stations with higher counts of Aral neighbors, pointing to higher net spillovers (cooperation) in these cases for diesel fuel, with no other diagonal coefficients significant. For e5 fuel, we have a positive and significant (at the 95% level) diagonal coefficient for Jet stations with more Jet-brand neighbors, but an unexpected negative and significant (at the 90% level) diagonal coefficient for Esso, with more Esso-brand neighbors. As noted, we would expect more same-brand neighboring stations have a positive impact on net spillovers reflecting more cooperative price markup behavior.

The significant off-diagonal elements in the table are all negative, as we would expect, since higher counts of rival brand stations should lead to a decrease in net spillovers, pointing

to more competition.

Considering the diesel table from a row-wise perspective, no significant off-diagonal coefficients are found for Jet and Total brand stations in the case of diesel fuel, with three negative and significant coefficients in the rows for Aral and Shell rivals, and only a single negative and significant row-coefficient for Esso and the Other brands stations. Taking a column-wise perspective of the diesel table, we see that every column except the first has at least one negative and significant off-diagonal coefficient, pointing to some (competitive) brand rivalry impacting all brands but Aral. Counting negative and significant off-diagonal coefficients down the columns of the matrix indicate that the presence of Total brand neighboring stations produces the largest brand rivalry impact on net spillovers (competitive price markup behavior), in the case of diesel fuel.

The e5 fuel results show no significant off-diagonal row-coefficients for Aral, Shell and Total brand stations, pointing to no brand rivalry here. There are three negative and significant coefficients for Jet brand stations, two for Esso brand stations, and a single negative and significant coefficient for the Other brand stations. Considering the e5 fuel table from a column-wise perspective, we see no significant coefficients in the Aral and Jet brand columns, but at least one negative and significant coefficient in all other brand columns. The greatest amount of (competitive) brand rivalry arises when Shell brand stations are neighbors (three cases).

6 Conclusion

We introduce a heterogeneous coefficient spatial autoregressive (HSAR) panel model that is capable of producing *station-level estimates* of gas station price rivalry. Our approach allows for inferences regarding how each of the *individual* observational units (stations) $i = 1, \ldots, N$ react to actions taken by each unit i's neighboring units (stations). It seems plausible that some stations make decisions to react competitively to neighboring station price markup actions, while others will react in a cooperative fashion, or not react at all. Conventional homogeneous coefficient panel models average over the sample of economic agents to produce estimates that reflect the typical agent's behavior, when in fact we would expect behavior to vary across the sample of observational units (agents).

A Markov Chain Monte Carlo approach to estimation of the HSAR model is set forth,

along with a Monte Carlo study showing that this approach to estimation produces results equivalent to those from quasi maximum likelihood estimation.

Partial derivatives that allow interpretation of estimates from the HSAR model are set forth, a topic not previously considered in the literature. We show that estimates from the HSAR model provide a very different interpretation than those from conventional spatial autoregressive model specifications. Model estimates allow for station-level spill-in and spill-out impacts arising from changes in actions taken by neighbors (nearby stations j). Spill-in impacts show how actions of neighboring stations impact each station i. Spill-out impacts show how actions taken by station i impact neighboring stations j.

We derive a model of (station-level) price conjectural variations that explicitly identifies station-level reaction functions, and show that this takes the form of our HSAR specification. Estimates of net spatial spillovers consisting of spill-in plus spill-out effects for each station reflect unobservable conjectural price variations. The relative size of spill-in versus spill-out impacts associated with neighboring station price markup actions determine cooperative versus competitive price markup rivalry at the station-level. The model is applied to a sample of more than 400 fueling stations in or around Berlin, for which we have daily price information for diesel and e5 fuels covering June 1, 2014 to September 30, 2015, or 487 days.

Estimates from the empirical application are highly consistent with predictions of our theoretical model concerning the station-level price reaction functions. The theoretical model rules out positive spatial dependence estimates in cases where net spillovers estimates (spill-in plus spill-out impacts) are negative, and we find only 1 of 328 such cases for diesel fuel, and only 8 of 348 such cases for e5 fuel. The model also predicts cooperative price markup behavior in cases where both spatial dependence and net spatial spillovers are positive, and competition when both spatial dependence and net spatial spillovers are negative. This result allows us to consider how the presence of own- or rival-brand stations in the neighborhood impact cooperative versus competitive price markup behavior of individual stations.

References

Allers, M.A. and J.P. Elhorst (2005) Tax mimicking and yardstick competition among local governments in the Netherlands. *International Tax and Public Finance*, 12(4), 493-513.

Aquaro, M., N. Bailey and M.H. Pesaran (2015) Quasi Maximum Likelihood Estimation of Spatial Models with Heterogeneous Coefficients. USC-INET Research Paper, (15-17).

Elhorst, J.P. and S. Fréret (2009) Evidence of political yardstick competition in France using a two-regime spatial Durbin model with fixed effects. *Journal of Regional Science*, 49(5), 931-951.

Elhorst, J.P. and K. Zigová (2014) Competition in research activity among economic departments: Evidence by negative spatial autocorrelation. *Geographical Analysis*, 46, 104-125.

Frondel, M., C. Vance and A. Kihm (2015) Time lags in the pass-through of crude oil prices: Big data evidence from the German gasoline market. *Applied Economics Letters*, Nov 11: 1-5.

Haucap, J., U. Heimeshoff and M. Siekmann (2015) Price dispersion and station heterogeneity on German retail gasoline markets (No. 171). DICE Discussion Paper.

Kelejian, H.H. and I.R. Prucha (2010) Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances, *Journal of Econometrics* 157, 53-67.

Kihm, A, N. Ritter and C. Vance (2016) Is the German retail gas market competitive? A spatial-temporal analysis using quantile regression. *Land Economics*, forthcoming

LeSage, J.P. and R.K. Pace. (2009) *Introduction to spatial econometrics.* Boca Raton: CRC Press.

LeSage, J.P. and R.K. Pace (2014) The biggest myth in spatial econometrics, *Econometrics*, 2(4): 217-249. http://www.mdpi.com/2225-1146/2/4/217

Liang, J.N. (1989) Price reaction functions and conjectural variations: An application to the breakfast cereal industry. *Review of Industrial Organization*, 4(2), 31-58.

Mobley, L.R. (2003) Estimating hospital market pricing: an equilibrium approach using spatial econometrics. *Regional Science and Urban Economics*, 33(4), 489-516.

Pennerstorfer, D. (2009) Spatial price competition in retail gasoline markets: Evidence from Austria. *The Annals of Regional Science*, 43(1), 133-158.

Appendix A

To test our MCMC algorithm we replicated a part of the Monte Carlo study carried out by Aquaro, Bailey and Pesaran (2015). This involved generating 1,000 different y vectors based on a fixed set of parameters. We use the HSAR data generating process, relying on an approach set forth in Aquaro, Bailey and Pesaran (2015).

The matrix W was based on a two-ahead, two-behind spatial configuration that is consistent with regions oriented in a line with two left- and right-neighbors.

The 1,000 different sets of disturbances were generated as Gaussian, $\varepsilon_{it}/\sigma_{i0} \sim N(0,1)$, with σ^2_{i0} generated as N independent draws from a $\chi^2(2)/4+0.5$ deviate for $\sigma^2_i, i = 1, \ldots, N$.

A single set of parameters generated from iid uniform (IIDU) and normal (IIDN) distributions: $\psi_i \sim IIDU(0, 0.8)$, $\alpha_i \sim IIDN(1, 1)$, were used in all experiments.

Table 3 shows results comparable to those in Table A1 of Aquaro, Bailey and Pesaran (2015), where W is based on two connections. The table shows (average over the N regions) bias calculated using: $(NR)^{-1} \sum_{i=1}^{N} \sum_{r=1}^{R} (\hat{\gamma}_{i,r} - \gamma(i, 0))$, where $\hat{\gamma}_{i,r}$ represent parameter estimates for $\alpha_i, \psi_i, \beta_i$ and $\gamma(i, 0)$ the true values, and $R = 1000$ Monte Carlo runs. RMSE was calculated (also averaged over the N regions) using: $N^{-1} \sum_{i=1}^{N} [\sqrt{R^{-1} \sum_{r=1}^{R} (\hat{\gamma}_{i,r} - \gamma_{i,0})^2}]$.

As in the results from Aquaro, Bailey and Pesaran (2015), we see that bias and RMSE decline with increases in T from 25 to 200. Changes in N do not seem to reduce bias and RMSE, a result consistent with theoretical results from Aquaro, Bailey and Pesaran (2015). These results are also consistent with the intuition that we are exploiting the time dimension of the space-time panel data to produce estimates, so increases in this dimension of the sample size should improve accuracy and precision.

Aquaro, Bailey and Pesaran (2015) present Monte Carlo results for bias and RMSE for sub-samples of individual observations to demonstrate that the average (over regional units) bias and RMSE statistics are representative of accuracy and dispersion that would be found at the observation-level in applied practice. An examination of our observation-level results showed the same pattern.

Table 3: Monte Carlo results for α, ψ HSAR model

		$N = 25$		
Average Bias, RMSE	α bias	α RMSE	ψ bias	ψ RMSE
$T = 25$	-0.0707	0.3922	0.0520	0.2153
$T = 50$	-0.0657	0.3538	0.0386	0.1623
$T = 100$	-0.0244	0.2550	0.0179	0.1332
$T = 200$	-0.0057	0.1738	0.0066	0.0929
		$N = 50$		
Average Bias, RMSE	α bias	α RMSE	ψ bias	ψ RMSE
$T = 25$	-0.1288	0.3945	0.0688	0.1919
$T = 50$	-0.0591	0.3503	0.0291	0.1668
$T = 100$	-0.0130	0.2460	0.0122	0.1379
$T = 200$	-0.0087	0.1813	0.0046	0.1007
		$N = 75$		
Average Bias, RMSE	α bias	α RMSE	ψ bias	ψ RMSE
$T = 25$	-0.1153	0.4086	0.0717	0.1972
$T = 50$	-0.0526	0.3092	0.0318	0.1563
$T = 100$	-0.0223	0.1992	0.0154	0.1249
$T = 200$	-0.0076	0.1611	0.0066	0.0956
		$N = 100$		
Average Bias, RMSE	α bias	α RMSE	ψ bias	ψ RMSE
$T = 25$	-0.0861	0.3835	0.0625	0.1984
$T = 50$	-0.0551	0.3578	0.0291	0.1671
$T = 100$	-0.0303	0.2309	0.0197	0.1262
$T = 200$	-0.0076	0.1594	0.0055	0.0915