Spatial Patterns in Household Demand for Ethanol

Hayk Khachatryan*
Graduate Research Assistant
Email: **hkhachatryan@wsu.edu**

Eric Jessup
Associate Research Professor
Email: **eric_jessup@wsu.edu**

Kenneth Casavant
Professor
Email: **casavantk@wsu.edu**

Transportation Research Group
School of Economic Sciences
Washington State University
Pullman, WA 99164
Phone: (509) 335-5536
Fax: (509) 335-1173

*Corresponding Author

Abstract

The recent rise in public environmental awareness, concerns of national energy security, and high transportation fuel prices have all served to heighten the interest in alternative fuels. One fundamental issue influencing the economic viability of the ethanol industry is consumers' demand-responsiveness to both gasoline and ethanol price changes. In this paper we present an alternative approach to this problem by estimating the geographic variation of price-elasticity of demand for ethanol across the study area, a departure from previous studies of ethanol demand, in which the price-elasticity of demand is identical across the space. Considering the spatial heterogeneity in household composition and demand preferences, using global estimates to explain the price-demand relationships over a large geographic area may lead to biased results. We demonstrate that the spatially weighted regression technique provides superior estimates over a global regression model. Resulting price-elasticities of demand for ethanol reveal significant geographic variation, suggesting that the use of spatially disaggregated data provides more detailed empirical results and a more thorough understanding for policy determination related to the ethanol industry. In a subsequent paper, these results will be used to simulate the effects of state-level alternative fuel policies on reducing the environmental emissions.

INTRODUCTION AND BACKGROUND

Alternative fuel policies are designed to increase the U.S. national energy independence and to reduce harmful environmental emissions from transportation fuels. According to the Renewable Fuel Standards (RFS),[1] biofuels production and use in the U.S. will reach 36 billion gallons by 2022 (EISA, 2007). To meet the RFS target, the U.S. Department of Energy (DOE) promotes the use of higher blends of ethanol (e.g., E85, 85% ethanol and 15% gasoline) by targeting specific regions and cities to establish high concentration of flexible fuel vehicles (FFV). The DOE also explores the possibility of using low-level blends of ethanol (e.g., E15 – 15% ethanol, 85% gasoline and E20 – 20% ethanol, 80% gasoline) in conventional vehicles. Understanding consumers' demand-responsiveness to ethanol and gasoline price changes at a specific geographical-level is imperative to implementing proposed renewable fuel policies.

In this paper I investigate consumers' demand-responsiveness to fuel price changes across the geographic space. In particular, I estimate the temporal and spatial variations for the own-price and cross-price elasticity of demand for ethanol in Minnesota. In previous studies of ethanol demand, the price-elasticity of demand for fuels was assumed to be constant across the study area (Anderson 2008; Hughes, Knittel, and Sperling 2008; Yatchew and No 2001; Schmalensee and Stoker 1999). I extend the model of household demand for close substitute transportation fuels (ethanol and gasoline) developed in Anderson (2008) to allow spatial variation of price-elasticity.

First, I use monthly price observations and sale volumes by individual E85 service stations in Minnesota to estimate own-price and cross-price elasticities of ethanol demand based on the initial model of household demand for transportation fuels. Then I motivate the problem of spatial non-stationarity in the data structure. The results from an exploratory data analysis show evidence for spatial autocorrelation in the regression residuals from the OLS and 2SLS specifications. The spatial structure in the data indicates that the value of the dependent variable in one spatial unit (a service station in our case) is affected by the independent variables in nearby units. Thus, the assumption of normally and independently distributed error terms when employing ordinary least squares regression is violated with the existence of spatial autocorrelation. This indicates that the non-spatial methods can lead to biased and inefficient parameter estimates. I extend and improve existing models by proposing an alternative model specification that accounts for spatial heterogeneity in data structure and provides superior estimates over global regression models.

I utilize data collected from ethanol service stations in Minnesota, which has been a leader in production and use of ethanol as an additive in gasoline over the last two decades. Prior to the 1990s Minnesota provided a tax credit for blending ethanol into gasoline. However, the tax credit negatively influenced funding for transportation. It was classified as ineffective in increasing ethanol production and was phased out in mid-1990s. Another state financial support program, which started in 1987, provided 20 cents per gallon to in-state ethanol processors for the first 15 million gallons of annual production. Currently, Minnesota provides tax incentives to increase E85 blending by taxing it at a lower rate than E10 or gasoline. Additionally, grants were provided to service station owners for installing E85 dispensing pumps. Many of these service stations that received E85 pump installation grants, participated in a monthly survey (conducted by Minnesota Department of Commerce and American Lung Association of Minnesota). Nearly all of the gasoline sold in Minnesota is required to contain 10% of ethanol (E10). By August 2013, this state law requirement will be increased to 20% (E20), conditional on the increase in the current "10% blending wall" established by the federal government. The combination of these state financial incentives and consumption mandates aim to achieve a broader goal of securing 25% of Minnesota's energy demand from renewable sources by 2025 (Yunker 2009).

The rest of the paper is organized as follows. The next section provides a brief overview of relevant literature. The Theoretical Framework section introduces a basic model of household demand for close substitute fuels (gasoline and ethanol). This section also incorporates spatial patterns in consumer demand-responsiveness to fuel price changes into the model. In the Empirical Framework section I first motivate the problem of spatial dependence and spatial heterogeneity in data. The basic model of household demand for fuels is then extended into a spatial demand model. Data sources are detailed in the Data Sources and Description subsection, including a map that shows the distribution of service stations in relation to five ethanol blending terminals (racks) and major highways in Minnesota. The remaining sections report and compare the basic and spatial model results. The Geographically Weighted Regression (GWR) estimates were used to visualize the variation of price-elasticity estimates across time and space in our study area. I conclude by discussing the implication of our findings for state-level ethanol policies and for continued research in this realm at the national-level.

RELEVANT LITERATURE

Due to the relatively short period of ethanol availability in the marketplace and consequent data limitations, the literature on demand estimation is minimal. Anderson (2008) shows that the household demand for ethanol as a close substitute to gasoline are sensitive to gasoline/ethanol relative prices. The gasoline-price (cross-price) elasticities of ethanol demand were estimated to be in the 2.5 – 3.0 range. The results were applied to study ethanol content standard related policies.

Recently Bromiley et al. (2008), analyzed factors that influence consumer use of E85 in Minnesota. The authors argue that estimating household demand for ethanol for the purposes of understanding their responsiveness to price changes is an important component for the economic viability of the emerging ethanol industry. Schmalensee & Stoker (1999) argue that household composition, demographic characteristics, and demand preferences change considerably over time and geography, and that it is reasonable to expect that not only temporal but also spatial variations will influence the household demand for transportation fuel. Additionally, consumers' environmental perceptions regarding biofuels and their attitudes for prices and performance relative to imported, petroleum-based fuels may vary depending on where they live and purchase fuel (Bromiley et al. 2008).

In contrast, a great deal of attention has been paid to estimating price-elasticities of demand for gasoline. Hughes et al. (2008) analyze U.S. gasoline demand in two time periods – 1975 to 1980 and 2001 to 2006. The short-run elasticities varied from –0.31 to –0.34 for the first period, and from –0.034 to –0.077 for the second, thus providing evidence that short-run price-elasticity of gasoline demand is more inelastic in recent years. These results are consistent with those of recent meta-analytic studies (Espey 1996; Graham and Glaister 2002), which report –0.27 and –0.23 for the short-term price-elasticities, and –0.71 for the long-term. Some recent estimates reported in Brons et al. (2008) showed a slightly higher range, varying from –0.34 for short-run to –0.84 for long-run price-elasticities. Contrary to these findings of inelastic gasoline demand Greene (1989), found own-price elasticity estimates to be over –15.0 (in absolute values).

However, none of these studies explicitly consider spatial attributes and/or provide a geographic comparison for the price-elasticities, which has important policy implications related to local governmental regulations for low-level vs. higher blend of ethanol. Bernstein and Griffin (2006) use a dynamic demand model to investigate the geographic differences in the price-demand relationships at the regional, state and sub-state level. The results showed that there are regional and state differences in the energy demand-responsiveness to price changes. However, their analyses only covered electricity and natural gas in the residential sector, and electricity use in the commercial sector.

Spatial regression techniques are widely used for analyzing data that has spatial characteristics (Case 1991), including hedonic house price spatiotemporal autoregressive models (R. Pace et al. 1998), and transportation spatial demand models (Henrickson and Wilson 2005). Henrickson and Wilson (2005) used a moving-window regression to estimate barge transportation demand elasticities. This approach is conceptually relevant to GWR technique as it produces spatially varying (to some extent) parameter estimates. However, the moving-window regression introduces so-called edge effects, because the data points within each local grid are given a weight equal to 1 (thus, are included in the regression), and those outside of the grid are given a weight equal to 0, which imposes limitations on capturing spatial variation between the two.

THEORETICAL FRAMEWORK

In this section, I first introduce a basic model of household demand for close substitute fuels - gasoline and ethanol. Then I extend the model to account for spatial differences in consumer demand-responsiveness to fuel price changes. I start with a basic model that reflects previous transportation fuel demand estimation models, (Rask, 1998; Anderson, 2008; Hughes et al., 2008). Following the notation in Anderson (2008) the household's utility function in terms of transportation fuels and other goods can be represented as $U = f(E, G, X)$, where E and G are consumption of close substitutes - ethanol and gasoline, and X represents the composite good. Since gasoline and ethanol are close substitutes, the household demand lands at the corner solution, such that the household will purchase ethanol only when $p_e < p_g/r$, where p_e and p_g are per gallon retail prices of ethanol and gasoline respectively, r (alternatively called fuel-switching price ratio) specifies the rate at which the consumer converts gallons of gasoline into ethanol-equivalent gallons, and p_g/r is ethanol-equivalent fuel price. Alternatively, the household will purchase gasoline when $p_e > p_g/r$. In other words, because ethanol has lower energy content (i.e., provides fewer miles per gallon), the fuel type decision is made based on the ethanol-equivalent price (Anderson 2008). The household demand for ethanol can be aggregated by assuming a fraction of households (ϕ) that own flexible fuel vehicles (FFV), and assuming that there are N such households. It is also assumed that each household owns a single vehicle. Further, it is assumed that fuel-switching price ratio r has differentiable cumulative distribution function $H(r)$, which is defined on [0,∞). Because $r < p_g/p_e$, i.e., households choose ethanol only when the switching ratio is less that the relative price, the portion of households that choose ethanol is the function evaluated at $H(p_g/p_e)$. The aggregate demand for ethanol takes the following form

(1)

$$E(p_e, p_g) = N\phi \int_{-\infty}^{p_g/p_e} d(p_e)dH(r) = N\phi H\left(\frac{p_g}{p_e}\right) d(p_e)$$

where the total number of households, N, is multiplied by the fraction that own FFVs, ϕ, multiplied by the fraction of those FFV owners that choose ethanol (which is a function of relative prices), multiplied by the level of ethanol consumption by households that choose ethanol (which is a function of absolute price of ethanol). The logged aggregate demand is

(2)

$$lnE(p_e, p_g) = lnN\phi + lnH\left(\frac{p_g}{p_e}\right) + lnd(p_e)$$

The gasoline-price elasticity of aggregate ethanol demand can be derived by differentiating (2) with respect to p_g and multiplying by p_g

(3)

$$\xi_g = \frac{\partial lnE(p_e, p_g)}{\partial p_g} p_g = \frac{H'\left(\frac{p_g}{p_e}\right)}{H\left(\frac{p_g}{p_e}\right)} \frac{p_g}{p_e}$$

Similarly, the own-price elasticity of aggregate ethanol demand can be derived by differentiating (2) with respect to p_e and multiplying by p_e

(4)

$$\xi_e = \frac{\partial lnE(p_e, p_g)}{\partial p_e} p_e = \frac{p_e d'(p_e)}{d(p_e)} - \frac{H^{'}\left(\frac{p_g}{p_e}\right)}{H\left(\frac{p_g}{p_e}\right)} \frac{p_g}{p_e}$$

Thus, the own-price elasticity (ξ_e) combines the effects of ethanol prices in terms of both reducing/increasing the demand for ethanol (the first term of equation (4)), and in terms of switching to/from gasoline.

The approach described above, however, does not incorporate considerations of spatial patterns in household demand into the model. Schmalensee and Stoker (1999) introduced a model of household demand for gasoline as a function of income, demographics and location. The authors argue that the demographic shift played an important role in increasing overall transportation fuel consumption over the last decades. The same source reports that household structure (number of drivers, household size, and household head age) has strong effects on gasoline demand. In addition to geographically varying household composition, the existence of spatial patterns in demand can be motivated by interdependent preferences. Yang and Allenby (2003) introduce a model of interdependent consumer preferences with data on automobile purchases, in which they found that preferences for Japanese-made cars are attributed to geographically and demographically defined networks. Based on these theoretical priorities, I extend the household demand model introduced above to account for geographic variations in household composition and demand preferences, which in turn influence price-elasticity of demand for fuels.

EMPIRICAL FRAMEWORK

1.1 Basic Model of Consumer Demand for Ethanol

The econometric specification for estimating the ethanol demand basic model described above can be represented by the following equation

(5)

$$y_{it} = \beta_0 + \sum\nolimits_m \beta_m X_{it} + \theta \mathrm{Z}_i + \gamma_t + \psi_t + \varepsilon_{it}$$

where y_{it} is time and location-specific dependent variable (fuel sales volume), X_{it} is a matrix of explanatory variables (county/station-specific characteristics, such as fuel prices, per-capita income, number of vehicles, and number of fueling stations that offer E85), Z_i represents time-invariant station-specific variables, e.g., station distance-to-rack and distance-to-highway, γ_t is the regional dummy (e.g., rural vs. urban), ψ_t represents monthly/seasonal dummy variables, and ε_{it} is a random error term, assumed to be normally distributed. In a classical ordinary least squares specification, these parameters are assumed to be constant across the study area. According to this specification, any geographic variations of the relationships between y_{it} and the parameters are captured in the error term.

1.2 Motivating Spatial Heterogeneity

"There are spatial variations in people's attitudes or preferences or there are different administrative, political or other contextual issues that produce different responses to the same stimuli over space" (Fotheringham, Brunsdon, and Charlton 2002). The utilization of ethanol sales volume and price data across Minnesota for estimating price-elasticity of demand using traditional econometric methods (e.g., OLS regression) involves two types of problems. The first problem is the spatial dependence. In our case, spatial dependence is the extent to which the values of monthly sales volume at one service station depend on the values at another service station in the vicinity. Considering n geographic locations, the spatial dependence can be represented as the following equation

(6)

$$y_i = f(y_j), i = 1, \dots, n \quad j \neq i$$

where y is the value of the variable (e.g., sales volume), and i and j are locations (e.g., service stations). Spatial dependence violates the traditional Gauss-Markov assumption that explanatory variables are fixed in repeated sampling (Lesage and Pace 2009). One reason for the existence of spatial autocorrelation can be the measurement error. Another reason for the spatial dependence can be related to the E85 stations locations (e.g., the proximity to the ethanol blending terminal or to the major highways in the study area). The second problem is the spatial heterogeneity, which violates another Gauss-Markov assumption that a single linear relationship exists across the sample data observations. As shown in equation (7), local relationships can be modeled for each service station in the study area

(7)

$$y_i = X_i\beta_i + \varepsilon_i, \qquad i = 1, \dots, n$$

where y_i is the dependent variable at location i, X_i is a vector of explanatory variables, β_i is the associated set of parameters to be estimated, and ε_i is a stochastic disturbance term.

1.3 Spatial Extension Model of Consumer Demand for Ethanol

In this section I extend the econometric model (5) to a spatially weighted regression model. To address the traditional econometric restrictive assumption of identical or stationary relationships over the space, some of the papers reviewed earlier employed indicator variables. One of the specifications found in Anderson (2008), restricted the data to two relationships by including urban vs. rural dummy variables to observe region effects. However, I do not know if only two dummies for the entire study area is an appropriate disaggregation and are consistent with the data, or if there is evidence to suggest sub-regional dummies to be included. Another approach, market segmentation, is used to reformulate data into small number of mutually exclusive and collectively exhaustive sub-samples (e.g., geographical samples – counties, states; socio-economic samples – income groups, education levels, etc.). Both of these strategies (dummy variables and market segmentation) introduce a problem of discontinuity in data, which eliminates the local spatial variations among different locations (for which data are available) in the study area.

One way to address this issue is to use a relatively recent spatial regression methodology that accounts for spatial non-stationarity in data – GWR (Fotheringham, Brunsdon, and Charlton 2002). The GWR methodology includes a spatial weighting matrix that assigns higher weights to the regressors in the near locations, and gradually decreases the weights as the distance from the regression point increases. In this spatially weighted model, the regression points are service stations. This approach allows estimating a "surface" of location-specific price-elasticity parameters. In our case, the GWR specification will produce local price-elasticities of demand for ethanol throughout the study area. The estimates then can be mapped using Geographic Information Systems (GIS) software. Following the notation in Fotheringham et al. (2002), I represent the demand for ethanol fuel at each of the locations from which the data were drawn as the following

(8)

$$y_{it} = \beta_{ot}(\upsilon_i, \nu_i) + \sum_m \beta_{mt}(\upsilon_i, \nu_i) X_{it} + \sum_k \theta_k(\upsilon_i, \nu_i)\, \mathrm{Z}_i + \varepsilon_{it}$$

where y_{it} is the dependent variable (monthly ethanol sales volume) for each of the ith fueling stations in the study area, X_{it} is a matrix of time and location-specific explanatory variables discussed above, Z_i represents the time-invariant variables, and ε_i is the error term. Coefficients β and θ are to be estimated for each of the fueling station at (υ_i, ν_i) projected coordinates (i.e., converted from geographic coordinates). The expressions for parameters $\beta(\upsilon_i, \nu_i)$ and $\theta(\upsilon_i, \nu_i)$ indicate that the price-elasticity of demand for ethanol and the other estimates are location-specific. The estimator for this model has the following form

(9)

$$\hat{\beta}(\upsilon_i, \nu_i) = (X' W(\upsilon_i, \nu_i) X)^{-1} X' W(\upsilon_i, \nu_i) y$$

where $W(\upsilon_i, \nu_i)$ is a distance-based weighting matrix for expressing potential interaction among spatial units (e.g., fueling stations). The off-diagonal elements of the weighting matrix are zero, and the diagonals denote the geographical weighting of observed data for point i. Denoting the (v_i, v_i) coordinates as (u), the weighting matrix takes the following form

(10)

$$W(u) = \begin{pmatrix} w(u)_1 & 0 & 0 & \ldots & 0 \\ 0 & w(u)_2 & 0 & \ldots & 0 \\ 0 & 0 & w(u)_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & 0 & w(u)_n \end{pmatrix}$$

One way to assign weights to the diagonal elements in the weighting matrix in (10) is to let $w(u) = 1$ if $d_i(u) \leq h$, and $w(u) = 0$ otherwise, where $d_i(u)$ is a measure of Euclidean distance between the ith observation and the location (u) (i.e., a regression point or service station), h is some bandwidth. However, similar to the concept of moving window regression, this strategy introduces some extent of spatial discontinuity. To overcome that problem, we compute the weights as a continuous function of a distance $(d_i(u))$. One possible way of doing it is to calculate the diagonals of (10) according to a kernel that has a Gaussian shape:

(11)

$$w(u) = exp(-0.5 \left({d_i(u)}/{h} \right)^2$$

In this weighting scheme, the $d_i(u)$ is a measure of Euclidean distance as described above, and h is bandwidth. The bandwidth parameter for our distance-based weighting matrix is selected using the following cross-validation procedure

(12)

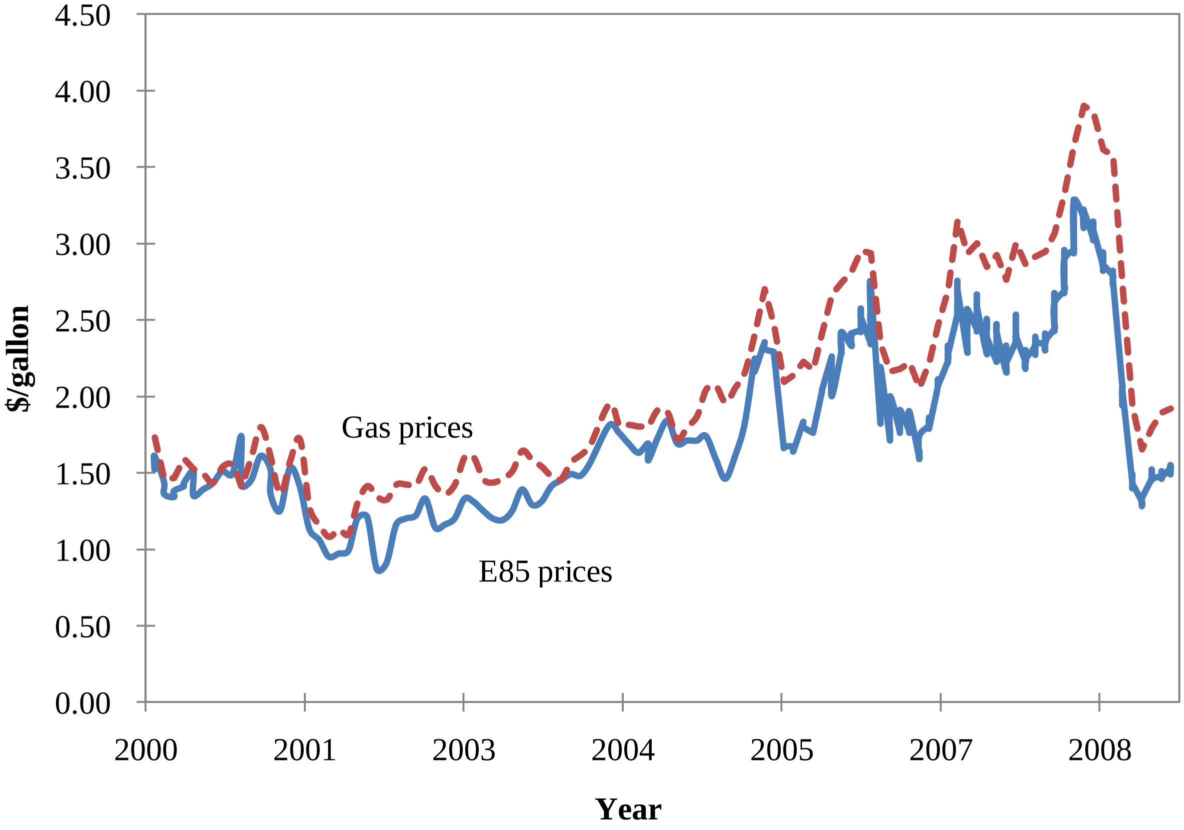
$$CV = \sum_{i=1}^{n} [y_i - \hat{y}_{\neq i}(h)]^2$$

where n is the sample size, $\hat{y}_{\neq i}$ denotes the fitted value of y_i with the observation for point i omitted from the calibration process (Fotheringham, Brunsdon, and Charlton 2002).[2] A value of h that minimizes the *CV* score is then used as the distance-weighting bandwidth. If the ith observation and the location (u) in weighting scheme (11) coincide, i.e., data were observed at location(u), the weight for that point will be unity. Then the weights of other locations around it will decrease according to a Gaussian curve as the distance between the two increases.[3] The the spatial kernel represented in (11) avoids the discontinuity problem by assigning decreasing weights (according to a Gaussian shape curve) as the distance between two locations increases (Fotheringham, Brunsdon, and Charlton 2002).

1.4 Data Sources and Description

Ethanol price information was obtained from a survey conducted by Minnesota Department of Commerce and American Lung Association of Minnesota. The data include monthly price observations and sale volumes by individual E85 service stations in Minnesota from 1997-2009. The number of participating E85 service stations was less than 10 in 1997, then steadily increased up to more than 330 as of mid 2009. As of September 2009, Minnesota has the highest number of E85 stations in the nation (351). This makes up more than 18% of the total number of E85 stations in the U.S. (U.S. DOE Alternative Fuels and Advanced Vehicles Data Center).[4] This information was used to calculate the number of fueling stations (offering E85) in each county for each time period. Monthly observations of retail gasoline prices were averaged from the Minnesota Weekly Gasoline Retail Price Reports provided by the Energy Information Administration (EIA). Wholesale gasoline prices were obtained from the Minnesota Regular Gasoline Wholesale/Resale Price by Refiners database provided by the EIA. Figure 1 shows the relationship between ethanol and gasoline prices in Minnesota from 2000 to mid 2009 period.

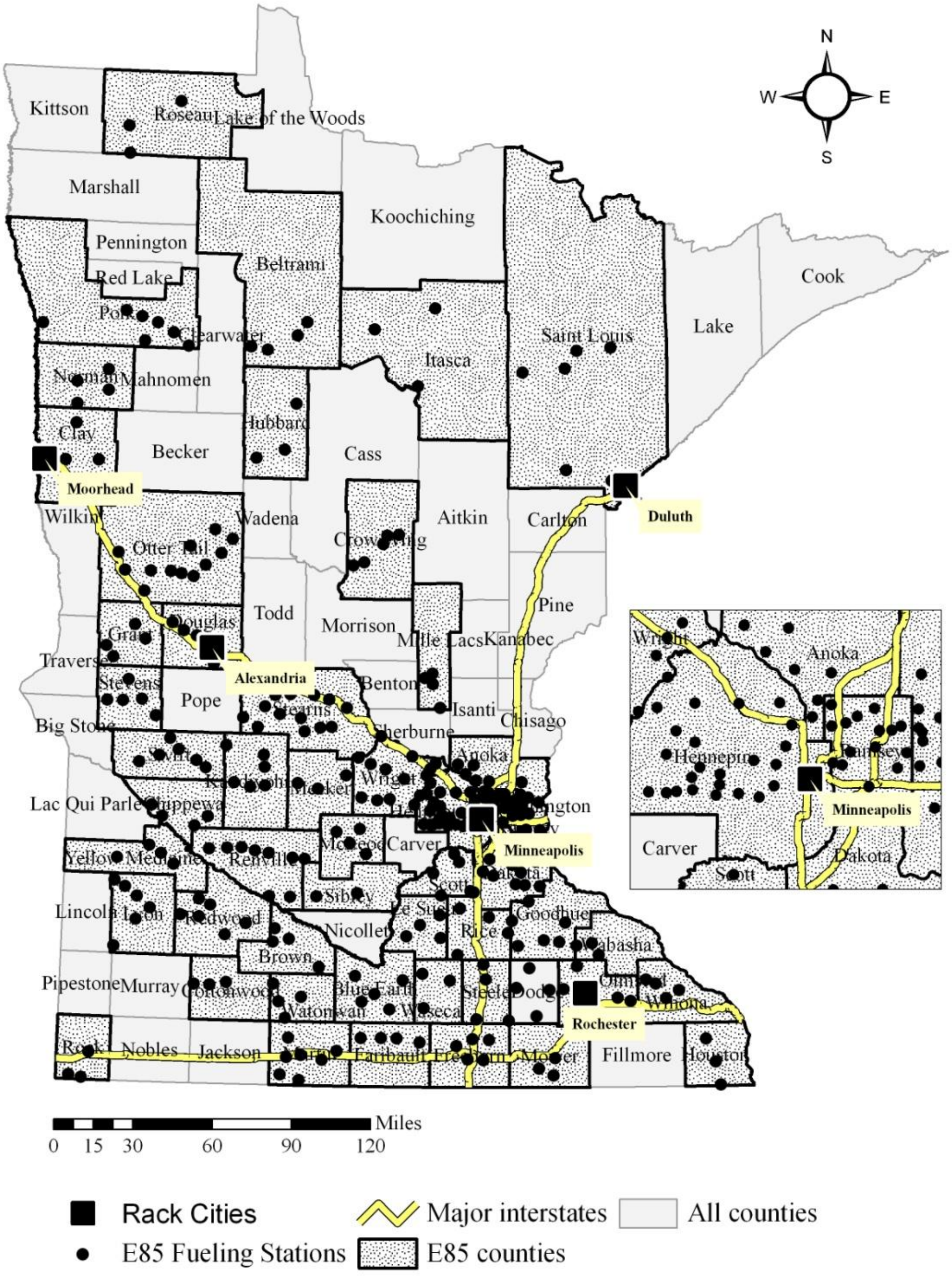
Figure 1: Gasoline and ethanol retail prices



Using historical consumer price index from the Department of Labor, all prices were converted into real 2009 prices. In contrast to service station-level ethanol sales data, the gasoline prices were only available at a county-level, and for only 2000 – 2009. As a result, the number of observations was decreased from 13,339 (1997 – 2009) to 8,542 (2000 – 2008).

Per-capita income information (converted into 2009 dollars) was obtained from the Federal Reserve Economic Data (FRED) state/county-level database. Time series of the number of vehicles per county was obtained from the Driver and Vehicle Services at the Minnesota Department of Public Safety. Unfortunately, in the current research I did not have an access to publicly available FFV ownership data. A small portion of observations were dropped due to missing or not reported price and sales volume observations. The inclusion of the income and the vehicle stock variables restricted the number of usable observations further. As a result, the number of observations was decreased from 8,542 to 6,860. (i.e., the time period was restricted to 2003– 2008.) Figure 2 depicts the spatial distribution of E85 service stations included in our analysis, in which some level of local clustering can be observed around the Twin Cities area.

Figure 2: The geographic distribution of E85 service stations in Minnesota in 2009



Additionally, I used GIS to derive Manhattan distances (in miles) between ethanol fueling stations and five ethanol blending terminals in Minnesota (Minneapolis, Alexandria, Moorhead, Rochester and Duluth). The terminal location information was obtained from the Oil

Price Information Service (OPIS) Rack Cities guide. I also used Minnesota's GIS highway network shapefile[5] and station locations available from the American Lung Association and Clean Air Choice organization.[6] Table 1 provides descriptive statistics for the data used in this paper.

Table 1: Descriptive Statistics

Variables	Mean	Stdev	Min	Max
Ethanol sales volume (gallons/month)	5,186	4,883	11	37,770
Income ($/per-capita)	39,565	6,783	27,274	49,196
Ethanol price (retail; $/gallon)	2.21	0.47	1.02	3.86
Gasoline price (retail; $/gallon)	2.66	0.60	1.64	3.87
Gasoline/ethanol price (retail; $/gallon)	1.20	0.10	0.64	1.99
Gasoline price (wholesale; $/gallon)	1.75	0.61	0.92	3.35
Distance from highway (miles)	22.44	24.51	0.28	144.00
Ethanol pumps in county (number/month)	6	4	1	17
Distance from nearest rack (miles)	34.15	26.32	1.00	100.00
Vehicle stock in county (number/month)	256,533	322,812	10,245	1,115,371

1.5 Basic Model Estimation and Results

First, I estimate the model of aggregate ethanol demand represented in equation (5). I let X_{it} denote ethanol and gasoline prices, per-capita income, number of vehicles, and number of stations offering ethanol, Z_i represents time-invariant distances to racks and to highways, γ_t and ψ_t represent regional and monthly dummy variables. The equation (5) can be represented as the following

(13)

$$\begin{aligned} lnE_{it} &= \beta_0 + \beta_1 \ln(PE_{it}) + \beta_2 \ln(PG_{it}/PE_{it}) + \beta_3 \ln(INC_{it}) + \beta_4 \ln(VEH_{it}) + \\ &+\beta_5 \ln(NSTAT_{it}) + \theta_1 \ln(DISTR_i) + \theta_2 \ln(DISTH_i) + \gamma_1(TC_t) + \\ &+\psi_2(M1) + \cdots + \psi_{11}(M11) + \varepsilon_{it} \end{aligned}$$

where E_{it} is the monthly ethanol sales for all participating E85 stations throughout the time period, PE_{it} is the ethanol price (that was instrumented with wholesale gasoline prices in the 2SLS regression), PG_{it} is the gasoline price, INC_{it} is the per-capita income, VEH_{it} is the number of vehicles in each county, $NSTAT_{it}$ is the number of E85 stations (i.e., service stations having E85 dispensers/pumps) in each county in each time period. $DISTR_i$ represents time-invariant distances from each of the E85 stations to the nearest ethanol blending terminal; $DISTH_i$ is time-invariant distance-to-highway variable representing the distance from each of the E85 stations to the nearest major highway node in the state. TC_t is a regional dummy variable controlling for Twin Cities area. Finally, $M1$ through $M11$ are 11 monthly dummy variables, and ε_{it} is the random error term.

Table 2 provides a summary of the OLS/2SLS estimates from the model described above.[7] I estimated the model for the whole time period, as well as for the prior and post Energy Independence and Security Act of 2007 periods (hereafter, prior to EISA and post EISA). The own-price elasticity of demand was found to be –3.33 for the 2003–2008 period, indicating a one percent increase in the price of ethanol leads to 3.33% decrease in the quantity of ethanol demanded.

Table 2: Basic Model Estimation Results

Dep. Var. = LN(ethanol monthly sales)				
	2003-2008 (OLS)	**2003-2006 (OLS)**	**2007-2008 (OLS)**	**2003-2008 (2SLS)**
Constant	–1.75***	–3.18***	–0.77	–1.84***
	(0.86)	(1.26)	(1.20)	(0.86)
LN(PE)	1.07***	2.11***	0.27***	0.94***
	(0.05)	(0.09)	(0.09)	(0.06)
LN(PG/PE)	4.35***	4.67***	4.36***	4.22***
	(0.12)	(0.17)	(0.18)	(0.12)
LN (INC)	0.41***	0.66***	0.17*	0.44***
	(0.08)	(0.12)	(0.11)	(0.08)
LN (VEH)	0.29***	0.22***	0.43***	0.27***
	(0.01)	(0.02)	(0.02)	(0.01)
LN (NSTAT)	–0.27***	–0.22***	–0.47***	–0.24***
	(0.02)	(0.02)	(0.03)	(0.02)
LN (DISTR)	(0.02)*	–0.01	0.03***	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
LN (DISTH)	0.02***	0.07***	–0.003	0.02***
	(0.01)	(0.01)	(0.01)	(0.01)
Reg. Dummy (Twin Cities area)	2.51***	2.19***	2.88***	2.49***
	(0.05)	(0.07)	(0.09)	(0.02)
Month. Dummies	y	y	y	y
Own-price elasticities	-3.3	-2.6	-4.1	-3.3
N	6860	3163	3697	6860
Adj. R-squared	0.43	0.47	0.45	0.43

***p<0.05, **p<0.1, *p<0.2. Standard errors are in parentheses. Dependent variable is the monthly ethanol sales volume in gallons. Prices are in 2009 dollars; income is the real per capita disposable income in 2009 dollars.

One of the reasons that the change in the quantity of ethanol demanded is proportionately larger than the change in the price (i.e., the demand is elastic) is that consumers have quick access to the close substitute fuel – gasoline, at almost zero search cost (since every service station offers gasoline). Another reasonable explanation for the high elasticity estimate is consumers' concerns related to ethanol's corrosive characteristics. Some service stations in the Midwest advertised gasoline as "ethanol free" fuel, emphasizing that E85 results in a reduced range (miles per tank of fuel) and engine problems because of its moisture content (Galbraith 2008). Considering these conditions, consumers may show high sensitivity to small price increases by decreasing their consumption of ethanol or by switching to gasoline. The estimate for the post EISA period (2007-2008) was estimated to be –4.09, much higher in absolute value compared to the prior to EISA period (2003– 2006) estimate of – 2.56.

Gasoline-price elasticity of ethanol demand was estimated to be 4.35 for the whole period (2003 – 2008); 4.67 and 4.36 for the prior and post EISA periods, suggesting relatively stable, sensitive ethanol demand-responsiveness to gasoline prices changes throughout the study period. (All the logged prices used in the estimate were normalized as $ln(p^*) = ln(p/\bar{p})$, where $p^*$is the normalized price variable, p is the initial price variable, and $\bar{p}$ is the sample mean. Thus, $\hat{\beta}_2$ is interpreted as a gasoline-price elasticity of ethanol demand.)

Income-elasticity of demand for ethanol was found to be 0.41 for the 2003– 2008 period. This estimate is consistent with the results from a recent study that analyzed similar data (Bromiley et al. 2008). The authors found that the influence of income levels on E85 monthly sales is minimal in magnitude and statistically insignificant. These results are also comparable to the estimates found in Hughes et al. (2008), which reports income-elasticity of gasoline demand in the 0.47– 0.54 range.

The estimate for the vehicle stock variable (0.29) for the 2003– 2008 period suggests that every 10% increase in the vehicle stock will lead to only 2.9% increase in ethanol sales. Using the FFV stock variable, I would expect the estimate to be 1 or more, suggesting that doubling the FFV stock will at least double the E85 fuel consumption. However, due to data limitations I am using a conventional vehicle stock variable as a proxy for FFV stock in our analysis. I found the estimates to be 0.22 and 0.43 for the prior and post EISA periods. According to the Minnesota Department of Public Safety registration records, the total number of passenger vehicles in Minnesota reached 3.34 million in 2006, a slight increase from 3.4 million in 2008. Considering 125,000 FFV as of 2006 in Minnesota (as reported in Bromiley et al. (2008)), the ratio of FFV to conventional vehicles is less than 5%. Overall, our estimate is in accordance with our expectation of positive relationship between the stock of vehicles and fuel sales.

The number of ethanol stations per county estimate resulted in –0.27 for the 2003– 2008 period, and –0.22 and –0.47 for the prior and post EISA periods. Consistent with previous findings (Anderson 2008), the negative sign suggests that a 1% increase in the number of ethanol stations in a county will reduce per station E85 sales by 0.22–0.48%.

The distance to a major highway variable showed relatively weak (0.02) influence on the E85 sales volume. Generally, retail gasoline prices are positively correlated with the distances from the source of supply (i.e., refineries, blending terminals, pipelines, ports, etc.) as the distribution costs increase with the distance. However, retail ethanol is primarily shipped to service stations from regional blending terminals, which are usually located near to large consumption areas. Also, major highways are positively correlated with local clusters of regular gasoline stations and relatively dense traffic of conventional vehicles. This suggests that there is more demand for regular gasoline at locations around or in close proximities from major highways. Therefore, the ethanol stations that are near to the major highway may sell less E85 compared to those that are located further away.

The influence of the distances to the blending terminals in Minnesota on E85 monthly sales volume is slightly weaker than that of the distance to highway variable described above. All of the five blending terminals are located within a close distance from major highways in the state. The same reasoning – relatively dense traffic of conventional vehicles on major highways (i.e., higher demand for regular gasoline) may explain the positive influence of the distance to the racks variable on the E85 sales. The estimate for the regional dummy variable TC (Twin

Cities) is positively correlated with the ethanol sales. Lastly, monthly dummy estimates reflect expected seasonal variation in transportation fuel demand.

1.1 Identification issues and spatial autocorrelation

Estimating demand functions that include price among the explanatory variables is often subject to endogeneity issues. In our model, the parameter estimates will be biased if the fuel prices are correlated with the unobserved characteristics embedded in the error term Anderson (2008). argues that many ethanol retail stations in Minnesota price ethanol at a fixed discount to gasoline (specified in a contract with suppliers that lasts several months, and sometimes a year). This indicates that the correlation between ethanol prices and local, short-term ethanol demand shifts is less likely. This pricing behavior implies that the local ethanol demand shifts are not correlated with the individual (i.e., fueling station) price variations. Conditional on the argument above, the OLS estimation results will not be biased.

Another concern is the possible spatial autocorrelation in our data, under the assumption of normally and independently distributed error terms in the model. I calculated the Moran's I statistics (Moran 1950) for residuals from the OLS regression (13). The results showed a moderate spatial correlation in OLS residuals (Moran's I =0.17, Z–score =4.00, p–value =0.00). The GWR model, which allows spatial variation of the underlying processes, should largely eliminate the problem of spatial autocorrelation in the error term. To confirm the validity of the GWR approach, the Moran's I statistics for both OLS/2SLS and GWR model residuals are compared in the next section.

1.2 GWR Estimation and Results

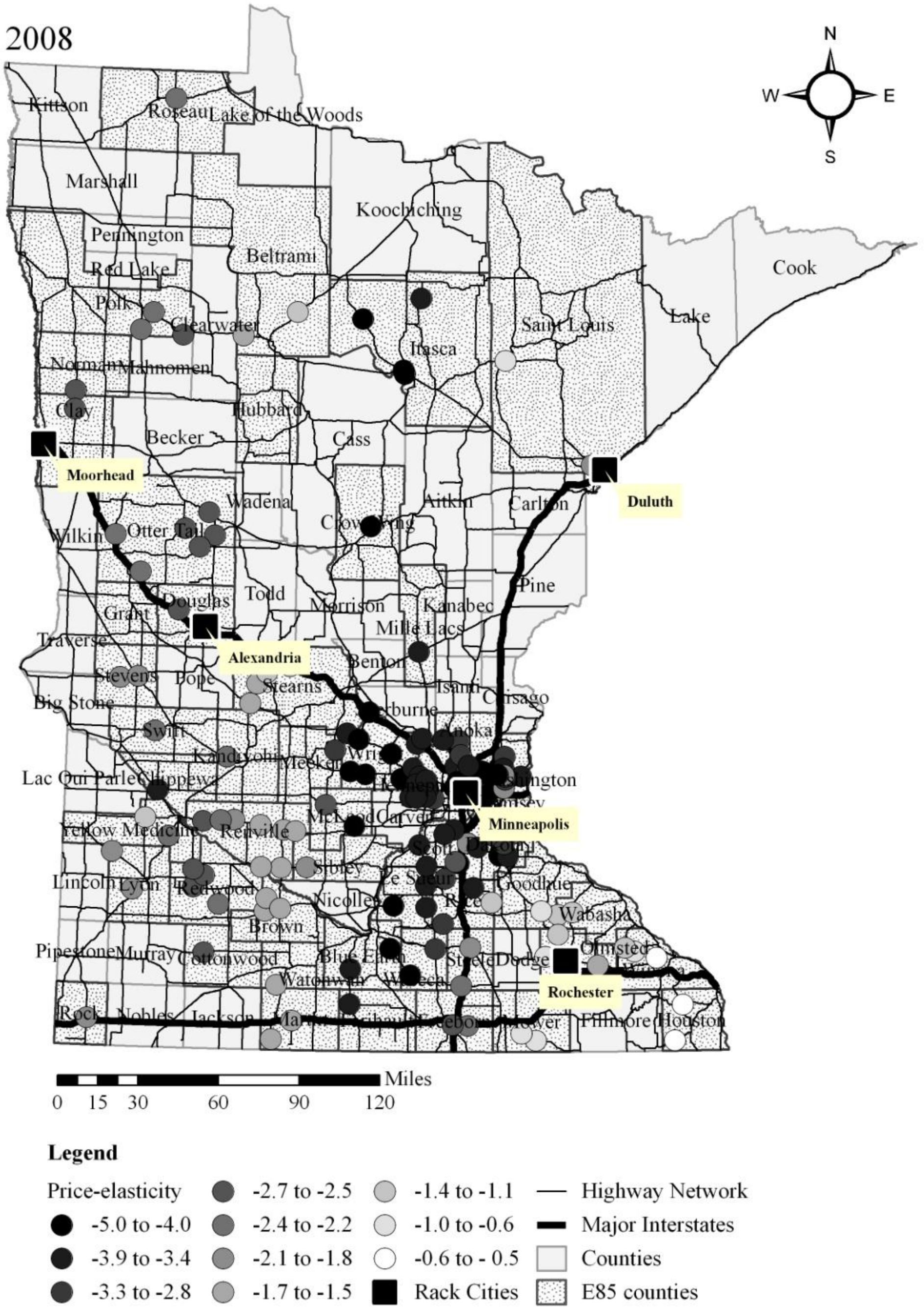
In this section I estimate and visualize the spatial extension of the ethanol demand model described earlier. Considering the variable descriptions provided in the 1.5 section, the GWR model (10) can be represented as[8]

(14)

$$lnE_{it} = \beta_{ot}(\upsilon_i, \nu_i) + \beta_1 ln(\upsilon_i, \nu_i)PE_{it} + \beta_2 ln(\upsilon_i, \nu_i)(PG_{it}/PE_{it}) + \beta_3 ln(\upsilon_i, \nu_i)INC_{it} + \\ + \beta_4 ln(\upsilon_i, \nu_i)VEH_{it} + \beta_5 ln(\upsilon_i, \nu_i)NSTAT_{it} + \theta_1 ln(\upsilon_i, \nu_i)DISTR_{it} + \theta_2 ln(\upsilon_i, \nu_i)DISTH_{it} + \varepsilon_{it}$$

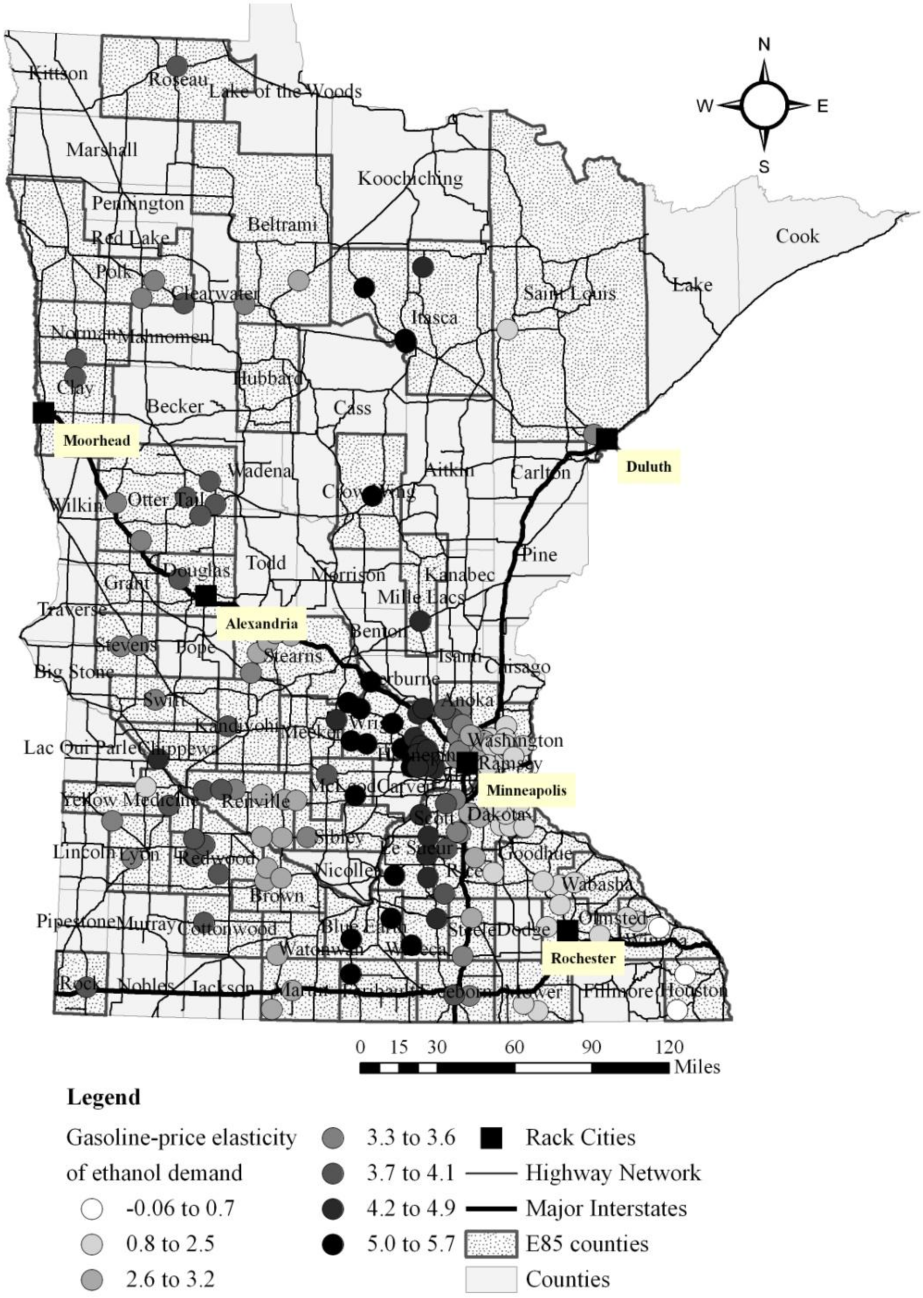
The result of the GWR model is a "surface" of parameter estimates across the ethanol stations in Minnesota that were included in this study. Figure 3 illustrates the spatial changes in the magnitude of the elasticities for the year 2008.[9] With a few outliers in the Itasca county in the Northern part of the state, the figure shows elastic ethanol demand cluster around the Twin Cities area (–5.0 to –2.2). Most of the estimates in the rural areas vary from –0.5 to –2.7. Overall, the estimated high elasticities are consistent with our expectations, explained by the availability of close substitute gasoline at almost zero search cost (since every service station where E85 is available also offers gasoline). The variation in the estimates also supports our motivation of the existence of spatial heterogeneity in the structure of the data.

Figure 3: Spatial distribution of price-elasticity of demand for ethanol in Minnesota



I have also visualized gasoline-price (cross) elasticities of ethanol demand (Figure 4). The estimates widely vary from –0.06 to 5.7 across the space. Because our analysis assumes only gasoline and ethanol fuels, the cross-price elasticity is comparable to the elasticity of Minnesota's ethanol's market share. The estimates in OLS/2SLS estimation showed that consumers are generally more sensitive to relative prices. However, our findings from the GWR model indicate that the consumers' demand-sensitivity to price changes widely varies geographically.

Figure 4: Gasoline-price elasticities of ethanol demand



In addition to visualizing the own- and cross-price elasticities in a map, Table 3 provides a summary of the estimates for comparing the GWR and OLS results side by side. As shown in the table, the OLS cross-price elasticity estimate (4.35) is found between upper quartile and maximum values of the GWR results. The own-price elasticity estimate from the OLS model (-3.3) falls between minimum and lower quartile values of the GWR estimates. Spatial distribution of the own-price and gasoline-price in Figure 3 and Figure 4 reveal that the OLS results represent only a portion of the geographic variation in gasoline-ethanol price-demand relationships.

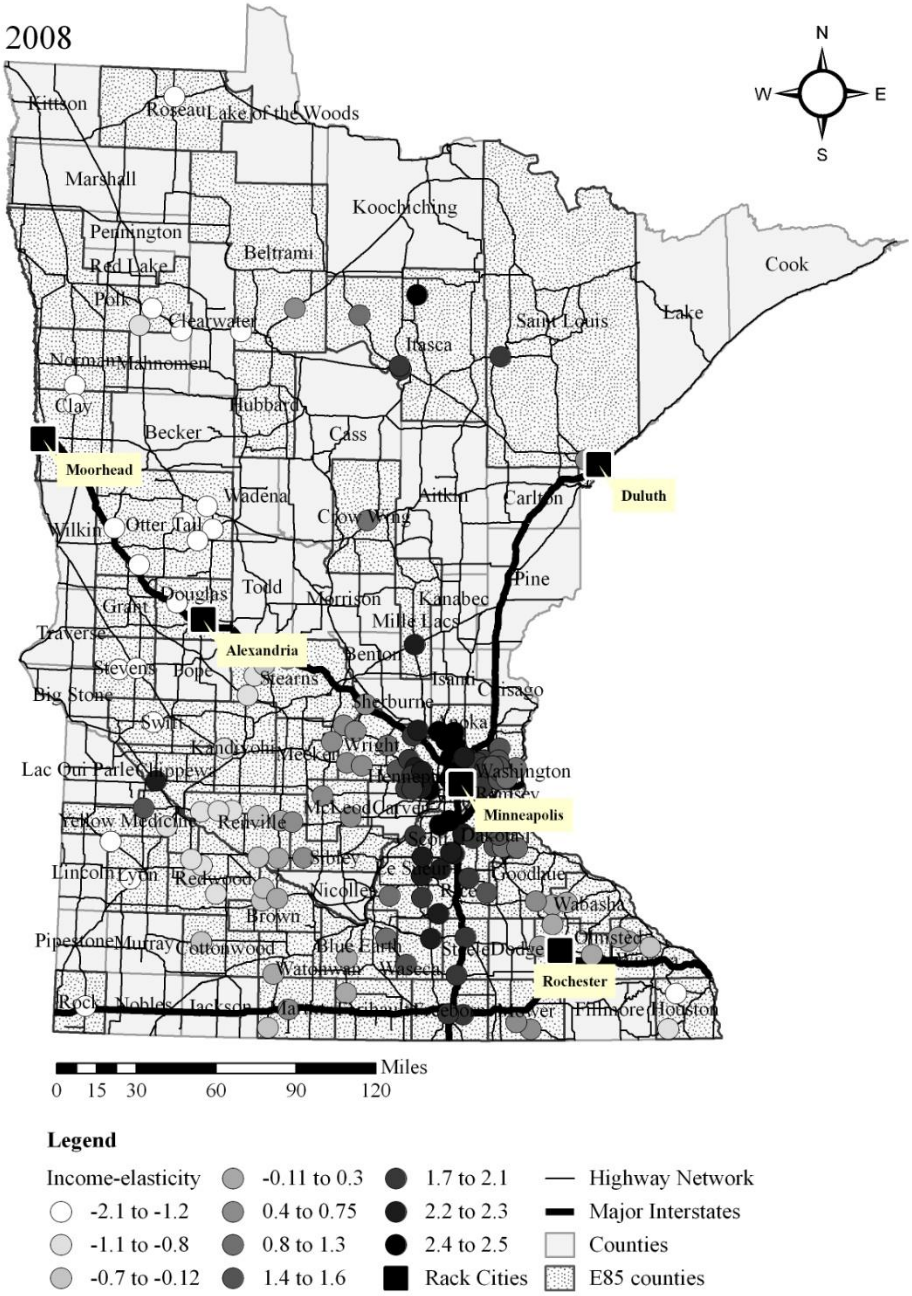
Table 3: GWR parameter summary and comparison with the global (OLS) model

Variables	Min	Lower Quartile (25th percentile)	Median (50th percentile)	Upper quartile (75th percentile)	Max	OLS (2003–2008)	Standard errors–OLS/2SLS (2003–2008)	GWR coefficients variability statistic $(\sqrt{\rho_i})$
ln(PE)	–5.00	–2.70	–2.08	–1.40	–0.50	1.07	0.05	1.06
ln(PG/PE)	–0.06	2.49	3.35	3.93	5.70	4.35	0.12	1.11
ln(INC)	–2.10	–0.48	0.95	2.02	2.50	0.41	0.08	1.36
ln(VEH)	–0.21	–0.02	0.13	0.33	0.59	0.29	0.01	0.21
ln(NSTAT)	–0.51	–0.39	–0.26	-0.14	0.06	–0.27	0.02	0.15
ln(DISTR)	–0.19	–0.08	–0.01	0.07	0.75	0.02	0.01	0.14
ln(DISTH)	–0.22	0.07	0.12	0.20	0.64	0.02	0.01	0.09

Income-elasticities for the Twin-Cities area were found in the 1.4 to 2.5 range (Figure 5), indicating a positive relationship between income levels and ethanol consumption in the urban area. The estimates for the rest of the regions change from negative to positive sign, ranging from –2.1 to 1.3. According to the comparison in Table 3, the OLS estimate (0.41) for income-elasticity falls between lower quartile and median values of the GWR estimates. A close examination of the map provided in Figure 5 indicates that the OLS captured only part of the geographic area outside of the Twin Cities area (to the West and to the Southeast). Overall, as shown in Table 3, GWR estimates show substantial variation in contrast to the OLS estimates. Comparison of the estimates across all of the variables shows that the OLS results are representative of only a segment of the entire range of elasticity estimates.

I test the following hypothesis: $H_0\colon \beta(\upsilon_i, \nu_i) = \beta_{OLS}$ where i indexes the locations, against $H_1\colon \beta(\upsilon_i, \nu_i) \neq \beta_{OLS}$. To test this hypothesis Brundson et al. (1998), suggest to measure the variability of the GWR coefficients (price-elasticities in our case) using the following statistics: $\rho_i = \sum_i (\beta(\upsilon_i, \nu_i) - \beta_{i.})^2/N$, where a dot denotes averaging the GWR coefficients over N locations. The $\sqrt{\rho_i}$ for all of the variables in the model is then compared with the standard errors from the OLS/2SLS model (the last column of Table 3). As shown in Table 3, all of the variability statistics are greater than the standard errors from the OLS/2SLS models suggesting an improvement upon the conventional estimation method. Additionally, I tested the residuals from the GWR for spatial dependence. The test statistics – Moran's I $_{\text{GWR}}$ = 0.07 with a Z-score = 3.33, and p-value = 0.008, compared to Moran's I $_{\text{OLS}}$ = 0.17 with a Z-score = 4.00 and p-value = 0.00 provide additional evidence for the advantage of estimating the price-elasticities with the GWR specification.

Figure 5: Spatial distribution of income-elasticity of ethanol demand in 2008



CONCLUDING REMARKS

The primary objective of this study was to estimate spatially extended version of ethanol demand model. The results of the GWR methodology showed significant spatial variation in the study area. The demand for ethanol was found to be elastic, with the estimates varying from -5.0 to -2.2 in the urban area. Most of the estimates for the rural areas of the state vary from -0.5 to -2.7 (although a few locations with high elasticity levels were found in the northern part of the state). Overall, the temporal variation in the price-elasticity of demand for ethanol was found to be less in magnitude. However, the post EISA (2007, 2008) period estimates showed significant variation, mostly increasing in absolute value around the Twin Cities area. The OLS/2SLS model estimates showed that consumers are more sensitive to relative prices. However, the comparison with the visualized GWR elasticity estimates showed that OLS model results can be attributed to only certain geographic areas.

Our findings of the spatial differences in price-elasticity of demand for ethanol fuel have several useful policy implications. Minnesota has joined several states in the Midwest in adopting the Energy Security and Climate Stewardship Platform Plan, an initiative designed to 1) produce commercially available cellulosic ethanol and other low-carbon fuels in the region by 2012, 2) increase E85 availability at retail fueling stations in the region, 3) reduce the amount of fossil fuel that is used in the production of biofuels by 50%, and 4) replace at least 50% of all transportation fuels consumed by the Midwest by locally-processed biofuels by 2025. As a part of that plan, the Minnesota Environmental Quality Board (EQB) is studying the potential sources of biomass for cellulosic ethanol and other low-carbon fuels production. In contrast to corn-based ethanol, the cellulosic feedstocks are geographically dispersed. So, the cellulosic ethanol costs (and thus retail prices) are sensitive to feedstock transportation and processed fuel (pure ethanol) distribution costs (Khachatryan, Casavant, and Jessup 2009). The ethanol processors face plant location optimization problem. Should the processing plant be located near to feedstock sources or to end-use markets? One component that is necessary for solving this optimization problem is knowing the consumers location-specific demand-responsiveness. Second, knowing spatial patterns in household demand for ethanol is useful for decisions related to increasing the number of E85 dispensing pumps in the state, something that I found to increase the local competition and negatively affects individual, station-specific demand for E85.

On a quantitative side, these findings have useful implications for state-level ethanol policy simulation experiments. Non-spatial econometric models emphasize similarities or regularities of data being analyzed. In contrast, spatially disaggregated estimation approach helps to reveal differences across the study area. Alternative fuel policy simulation requires consideration of a range of price-elasticity estimates to be used in a calibration. The use of disaggregated data in our study allowed obtaining more detailed estimates, which can be used in policy simulations with more certainty.

It is worth mentioning several limitations of this study. Although, our investigation aims to reveal spatial differences in the price-demand relationship, it is geographically bounded. Availability of ethanol fueling stations and price differences outside of Minnesota's borders may influence sales volumes observed in our data. Additionally, a portion of E85 sales can be attributed to the households not residing in Minnesota (since many E85 stations are close to major interstate highways).

In future research, I plan to simulate ethanol policy effects on environmental emissions reductions in Minnesota. From a methodological perspective, it will be useful for future research to develop and use a weighting scheme that accounts for both temporal and spatial effects simultaneously (i.e., spatiotemporal weighting matrix). In a spatiotemporal framework, spatial weights work in a same manner (e.g., decreasing the weights based on the distances between locations, or based on the number of nearest neighbors), however, the temporal weight gives more weight to more recent events, and gradually decreases the weights for previous years.

REFERENCES

Anderson, S. 2008. The Demand for Ethanol as a Gasoline Substitute. *Working paper*.

Bernstein, M.A., and J. Griffin. 2006. *Regional Differences in the Price-Elasticity of Demand for Energy*. Subcontract Report - . RAND Corporation.

Bromiley, P., T. Gerlach, K. Marczak, M. Taylor, and L. Dobrovolny. 2008. *Statistical Analysis of the Factors Influencing Consumer Use of E85*. National Renewable Energy Laboratory.

Brons, M., P. Nijkamp, E. Pels, and P. Rietveld. 2008. A meta-analysis of the price elasticity of gasoline demand. A SUR approach. *Energy Economics* 30, no. 5 (September): 2105-2122.

Brundson, C., S. Fotheringham, and M. Charlton. 1998. Geographically Weighted Regression-Modeling Spatial Non-Stationarity. *The Statistician* 47, no. 3: 431-443.

Case, A. C. 1991. Spatial Patterns in Household Demand. *Econometrica* 59, no. 4 (July): 953-965.

EISA. 2007. *Energy Independence and Security Act of 2007. Energy Security Through Increased Production of Biofuels*. *P.L. 110-140, H.R. 6.*

Espey, M. 1996. Explaining the Variation in Elasticity Estimates of Gasoline Demand in the United States: A Meta-Analysis. *The Energy Journal* 17, no. 3: 49-60.

Fotheringham, S., C. Brunsdon, and M. Charlton. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: Wiley.

Galbraith, K. 2008. In Gas-Powered World, Ethanol Stirs Complaints. *The New York Times*, July 26, sec. Business. http://www.nytimes.com/2008/07/26/business/26ethanol.html?partner=rssnyt&emc=rss.

Graham , D. J., and S. Glaister . 2002. The Demand for Automobile Fuel: A Survey of Elasticities. *Journal of Transport Economics and Policy* 36 (January 1): 1-25.

Greene, D. L. . 1989. Motor fuel choice: An econometric analysis. *Transportation Research Part A: General* 23, no. 3 (May): 243-253.

Henrickson, K., and W. Wilson. 2005. *Patterns in Geographic Elasticity Estimates of Barge Demand on the Upper Mississippi and Illinois Rivers*. http://www.nets.iwr.usace.army.mil/inlandnav.cfm.

Hughes, J. E. , C. R. Knittel, and D. Sperling. 2008. Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand. *Energy Journal* 29, no. 1 (January): 113-134.

Khachatryan, H., K. Casavant, and E. Jessup. 2009. *Biomass Inventory Technology and Economics Assessment: Collection and Distribution Cost Curves*. Final Report. Washington State Department of Ecology.

Lesage, J., and K. Pace. 2009. *Introduction to Spatial Econometrics*. CRC Press Taylor & Francis Group. London, New York: Boca Raton.

Moran, P. 1950. A Test for the Serial Independence of Residuals. *Biometrica* 37: 178-181.

Pace, R., R. Barry, J. Clapp, and M. Rodriquez. 1998. Spatiotemporal Autoregressive Models of Neighborhood Effects. *Journal of Real Estate Finance & Economics* 17, no. 1 (July): 15-33.

Schmalensee, R. , and R. M. Stoker. 1999. Household Gasoline Demand in the United States. *Econometrica* 67, no. 3 (May): 645-662.

Yang, S., and G. Allenby. 2003. Modeling Interdependent Consumer Preferences. *Journal of Marketing Research* 40, no. 3: 282-294.

Yatchew, A. , and J. A. No. 2001. Household Gasoline Demand in Canada. *Econometrica* 69, no. 6 (November): 1697-1709.

Yunker, J. . 2009. *Biofuel Policies and Programs*. Text. State of Minnesota Office of the Legislatieve Auditor. http://www.auditor.leg.state.mn.us/ped/2009/biofuels.htm.

APPENDIX

Table 4: The distribution of E85 service stations in the U.S. (as of September 2009)

State	Number of E85 Stations	State	Number of E85 Stations	State	Number of E85 Stations
Minnesota	351	N. Dakota	31	Idaho	5
Illinois	192	Tennessee	29	Connecticut	4
Iowa	123	Arizona	26	Louisiana	4
Wisconsin	121	Florida	26	Mississippi	4
Indiana	112	Pennsylvania	26	Utah	4
Missouri	95	N. Carolina	17	DC	3
Michigan	91	Washington	15	West Virginia	3
S. Carolina	85	Kentucky	14	Massachusetts	2
S. Dakota	80	Maryland	14	Delaware	1
Colorado	76	Nevada	14	Montana	1
Ohio	63	Alabama	11	Alaska	0
Nebraska	48	New Mexico	11	Hawaii	0

California	40	Oklahoma	11	Maine	0
Texas	40	Arkansas	8	New Hampshire	0
Georgia	37	Oregon	8	New Jersey	0
New York	35	Virginia	8	Rhode Island	0
Kansas	33	Wyoming	6	Vermont	0
Total	1928				

Source: U.S. Department of Energy, Alternative Fuels and Advanced Vehicles Data Center. http://www.afdc.energy.gov/afdc/fuels/stations_counts.html (Updated 09/24/2009)

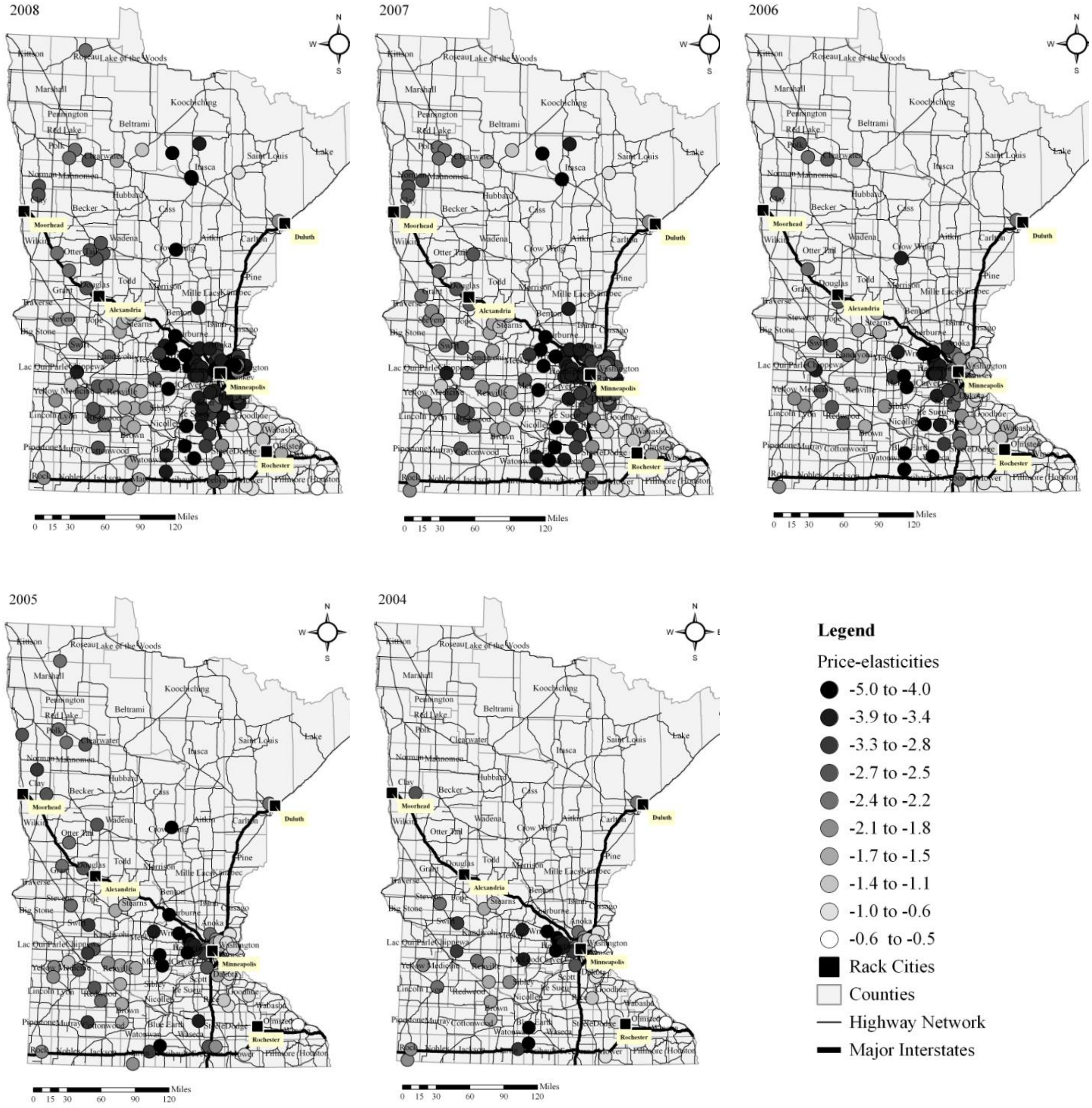
Table 5: Additional estimation results (OLS, linear model)

Dependent variable = LN(Ethanol monthly sales volume

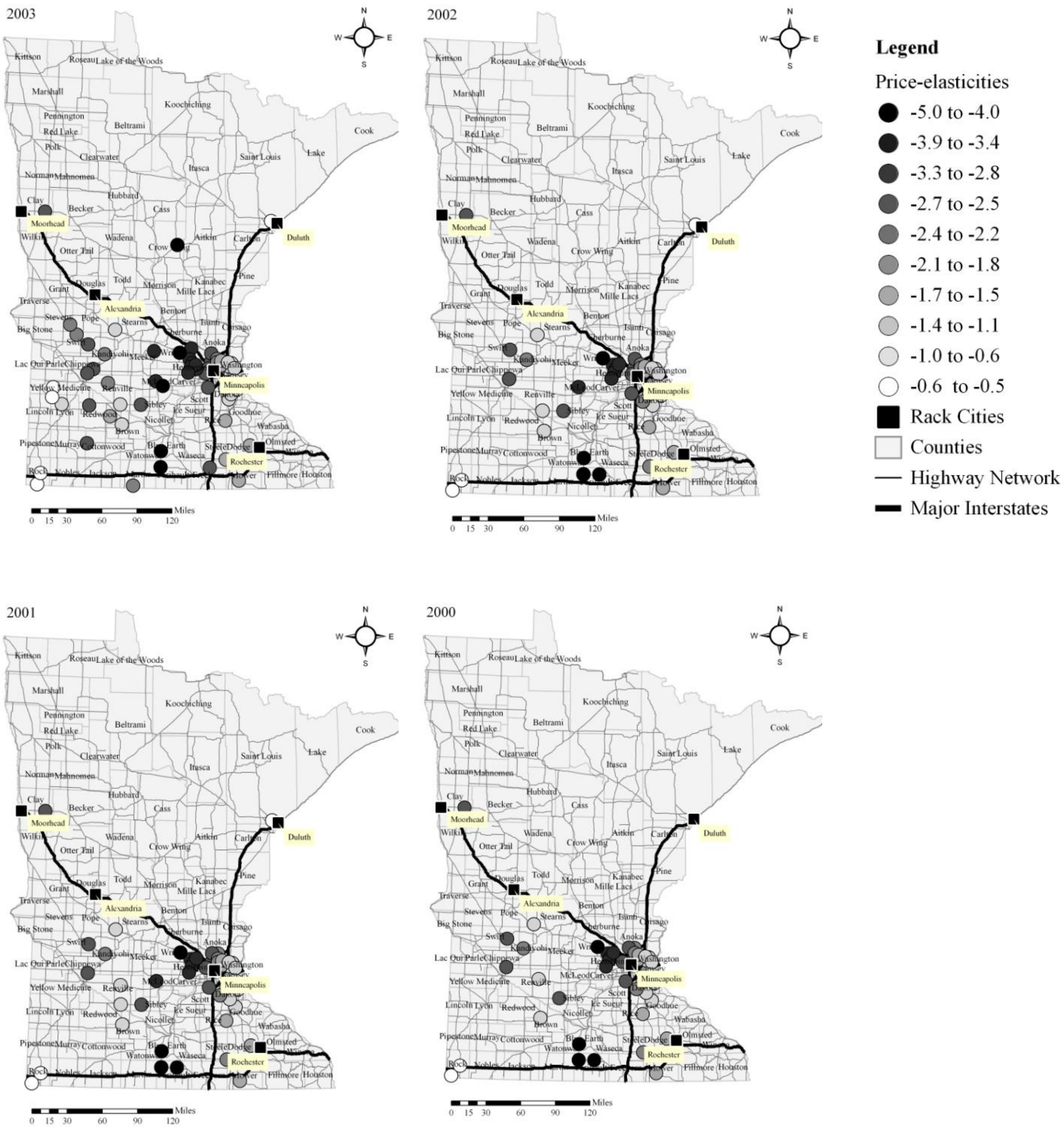
	2003 - 2008	**2003 - 2006**	**2007 - 2008**
constant	-6893.5***	-12065.7***	-3949.4***
	(622.8)	(915.1)	(1036.6)
PE	-6366.0***	-6160.6***	-6945.0***
	(322.2)	(421.9)	(471.7)
PG	6477.9***	9013.8***	5791.3***
	(261.2)	(400.9)	(383.0)
INC	0.07***	0.07***	0.07***
	(0.01)	(0.01)	(0.017)
VEH	0.002***	0.001***	0.005***
	(0.0003)	(0.0004)	(0.0005)
NSTAT	-142.8***	-66.3**	-347.8***
	(23.09)	(36.9)	(33.18)
DISTR	7.97***	-1.09	17.12***
	(2.58)	(3.44)	(3.59)
DISTH	-12.8***	-3.6	-19.8***
	(2.77)	(3.90)	(3.79)
Reg. Dummies			
GM	4878.6***	4071.0***	5513.3***
	(341.3)	(419.6)	(544.7)
TC	7993.7***	6451.9***	9111.9***
	(305.8)	(349.6)	(506.1)
Month. Dummies	y	y	y
N	6860	3163	3697
Adj. R-squared	0.31	0.36	0.32
Durbin-Watson statistic	1.99	2.17	2.03

***p<0.05, **p<0.1, *p<0.2. Standard errors are in parentheses. Dependent variable is the monthly ethanol sales volume in gallons. Prices are in 2009 dollars; income is the real per capita disposable income in 2009 dollars.

Figure 6: Spatial distribution of price-elasticity of demand for ethanol in Minnesota (2000-2008)



Note: The parameter estimates for the 2000 – 2002 period were derived using a specification that does not include the vehicle stock variable (since the vehicle stock restricts our data to 2003-2008). Those maps were included to show the dynamics of the elasticities for the entire period.



Endnotes

---

[1] The Renewable Fuel Standard is one of the key provisions of the Energy Independence and Security Act (EISA) of 2007, a government policy, which is designed to secure roughly one-third of the U.S. transportation fuel consumption.

[2] In the CV equation, omitting the *i*th observation is necessary, otherwise the CV score will be minimized when $\theta = 0$, i.e., as $\theta \to 0$, $\hat{y}_i(\theta) \to y_i$, so the CV score is minimized when $\theta = 0$.

[3] The parameter estimation points are usually coinciding with the points from where data were drawn, but it is not a necessary condition (Fotheringham et al. 2002).

[4] For the distribution of all E85 service stations in the U.S. see Table 4 in the Appendix.

---

[5] Minnesota road networks GIS shapefiles are available from the Minnesota Department of Transportation (http://www.dot.state.mn.us/maps/gisbase/html/datafiles.html)

[6] The map of E85 station locations can be found at: http://www.state.mn.us/mn/externalDocs/Commerce/State-wide_E-85_station_map_121302123133_MinnesotaE85StationsMap.pdf

[7] Wholesale gasoline prices in Minnesota was used as an instrument for E85 prices. Ethanol sales represent a small portion of the gasoline consumption in Minnesota, therefore, wholesale gasoline prices can be considered as exogenous in our model.

[8] Note that I did not include binary variables as the GWR allows explanatory variable coefficients to vary across the study area. Thus, the binary variables are not necessary, and their inclusion will introduce local collinearity.

[9] The estimates covering full time period are included in Figure 6 of the Appendix.