The Impact of Religion on Environmental Performance: Applied Spatial Econometrics

Omer Kara[†]

Latest Version

March 30, 2023

1^{st} Draft

Abstract

The objective of this study is to reveal how the cultural background of societies such as religion affects natural environment while accounting for the possible spatial dependencies between observation characteristics. More specifically, using spatial econometric models and an extensive county–level U.S. data in 2010, this study investigates the impact of religion on the environmental performance of a county after controlling for the other important determinants. The environmental performance is measured using seven criteria air pollutants: *CO*, NH_3, NO_x, SO_2, PM_{10}, $PM_{2.5}$, and *VOC*. For each air pollutant, the analyses are conducted separately using various spatial models to account for the global and local spillover effects as well as the global diffusion effect inherited in the data. The extensive spatial analyses suggest that there is a strong and positive spatial autocorrelation between observations and religion appears to be associated with environmental performance. Specifically, religions of Judeo–Christian beliefs such as *Evangelical Protestants*, *Black Protestants*, *Mainline Protestants*, *Catholics*, *Orthodox Christians*, and *Jews* exhibit a negative association with environmental performance, whereas *Hindus* and *Buddhists* are the only ones that show a positive association. Therefore, I find strong support for White's Thesis by providing spatial econometric evidence.

Keywords: Spatial Econometrics, Religion, Environmental Performance, Pollution
JEL Classification Codes: C31, C51, Z12, Q50

I am grateful for the tireless advising and support given by my advisors Barry Goodwin, Jeffrey Prestemon, and Walter Thurman. All errors are my own. The latest version is available here. Please address correspondence to okara@ogu.edu.tr.

[†]Faculty Member, PhD in Department of Economics, Eskisehir Osmangazi University. E-mail: okara@ogu.edu.tr

Contents

List of Tables

List of Figures

> "Everything is related to everything else, but near things are more related than distant things. (Tobler, 1970)"

1 Introduction

A growing community of theologians, scholars, and practitioners have recently started to examine the relationship between religion and environmental behavior[1]. The increasing attention to this relationship is at the intersection of three current trends. The first of these is the growing importance of worship regarding the number of adherents of major religions. Recent statistics show that over 80% of the world population adheres to some religion. Christianity and Islam are the two largest ones (Britannica, 2010). According to Barrett and Johnson (2001), despite the decreasing trend of religious belief in Europe, the number of believers of major denominations have increased in almost all regions of the world since the 1990s.[2]

The second trend of interest is the gradual increase in environmental awareness, especially in developed countries. Figure 1 shows the donations to environmental organizations in Australia. Statistics demonstrate that the public donated more than $140 million in 2009, compared to less than $5 million in 1993, to environmental organizations to help to protect and enhance the natural environment.

The last trend relates to the first two trends and also to the increasing awareness of cultural and spiritual elements in the environmental issues. Lynn White gives the first reference to this trend. In a groundbreaking yet controversial essay, "The Historical Roots of Our Ecological Crisis," White (1967) notes:

> "What people do about their ecology depends on what they think about themselves in relation to things around them. Human ecology is deeply conditioned by beliefs about our nature and destiny—that is, by religion."

According to Iranian–American philosopher Seyyed Hossein Nasr, an environmental crisis is fundamentally a crisis of values (Nasr and Chittick, 2007). Al Gore, an environmental activist, states that the human–enhanced climate change is ultimately a moral issue (Gore, 2006):

> ". . . global warming is not just about science and . . . it is not just a political issue. It is really a moral issue."

[1]The distinction and relationship between attitudes and behavior have been a matter of debate in psychological and sociological studies for decades. Similarly, the economics literature distinguishes environmental attitudes and behavior from one another even though they are closely related. Thus, throughout this paper, to prevent confusion, environmental attitudes and behavior are considered as the same notion and called *environmental behavior*. Eilam and Trop (2012) discuss the definitions and differences between these notions in detail.

[2]Between 1990 and 2000, average annual growth rates for Christianity, Islam, Judaism and world population are 1.36%, 2.13%, 0.91% and 1.3% respectively. See URLs: `https://goo.gl/HKSLhz` and `https://goo.gl/9bBnLe` (visited on Jan. 30, 2017).

[3]All values are in nominal prices. Source: Australian Government, Department of the Environment, Water, Heritage and the Arts. "Donations to Environmental Organizations." Annual Report 2009–2010 (2010), pp. 288. URL: `https://goo.gl/NR3VyU` (visited on Jan. 30, 2017).

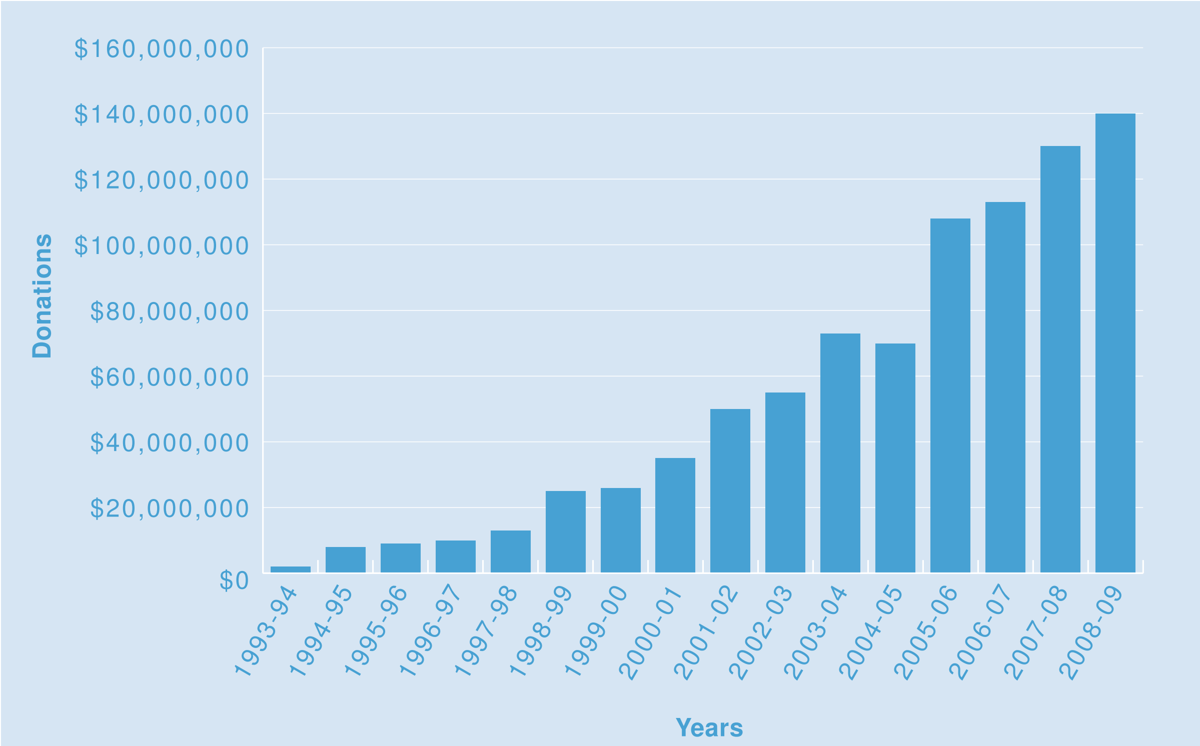


Figure 1: Donations to Environmental Organizations in Australia[3]

"... this crisis is not really about politics at all. It is a moral and spiritual challenge."

In 1986, President of the World Wildlife Fund International invited leaders of the five major religions (i.e., Abrahamic[4], Buddhism, and Hinduism) to meet and discuss how their faiths can help to save the natural environment. The meeting occurred in Assisi, Italy since it is the birthplace of St. Francis, the Catholic saint of ecology. In the Assisi Declarations, leaders of the major religions outlined their distinctive traditions and approaches in caring for nature and declared that how religions as a whole and separately can help to save the environment (Rinpoche, 1986). Moreover, in the subsequent declarations of Assisi, when the Alliance of Religions and Conservation formed in 1995, leaders of the five initial faiths issued more detailed statements, and leaders of the other six significant faiths also made their initial statements about how their holy books or doctrines include environmentally friendly components.

Nevertheless, one might ask that "If every religion claims itself to be environmentally friendly, then why do we see such a difference in environmental performance[5] between countries with similar religion intensity?" For instance, Hsu et al. (2016) use Environmental Performance Index (EPI)[6] and show that the Islamic countries have small EPI compared to western Christian countries.

[4]Abrahamic religions are one of the major divisions in comparative religion. They emphasize and trace their common origin to Abraham. In chronological order, the largest Abrahamic religions are Judaism, Christianity, and Islam.

[5]Throughout this paper, the term *environmental performance* is used to refer to quantitatively measurable and broader concept of environmental behavior.

[6]The EPI is a method of quantifying and numerically benchmarking countries' performance on high–priority environmental issues in the protection of human health and ecosystems.

The earlier studies approach the above arguments from an individual level and try to test the relationship between the degree of religiosity and environmental behavior of a person who adheres only to Judeo–Christian[7] tradition. This study deviates from the previous literature in several ways. First, environmental behavior is a loose concept and admittedly difficult to measure since individuals can make environmentally friendly choices at one point while polluting the environment at another point. Therefore, human–sourced air pollutants from all sources[8] are employed as a quantitatively measurable and broader concept of environmental behavior in the interest of area and called environmental performance. Second, the study departs from the extent of believing (i.e., the degree of religiosity) to belonging in measuring religion. Thus, percentage adherents[9] is used to measure religion. Third, unlike the earlier studies, this research utilizes an extensive county–level spatial dataset which includes not only Judaism and Christianity but also other major religions. Fourth, departing from the previous studies which have mostly used simple bivariate analysis, this study aims to apply spatial econometrics approaches suggested by the literature by Anselin (1988), Cressie (1993), and Lesage and Pace (2008). Fifth, probably one of the most intriguing parts of the present research is the fact that it looks at finer spatial scales (i.e., county–level) than a study of country–level data would. Thus, this study uses the variation in adherents of religions across counties to obtain more information about the relationship between religion and environment.

The objective of this study is to test the assertions explained above by providing spatial econometric evidence and using a large county–level U.S. data. The central question that the present research tries to answer is: Does religion have any impact on the environmental performance of a county after controlling for the other important determinants of environmental performance?

The paper proceeds as follows: Section 2 presents an extensive review of previous studies relevant to the development of this research; Section 3 presents the determinants of environmental performance; Section 4 presents the relevant data sources; Section 5 gives a brief information about the selected contiguity types and weight styles; Section 6 presents the spatial autocorrelations tests; Section 7 provides the spatial autocorrelation test results for dependent variables; Section 8 presents the empirical methods; Section 9 presents and discusses the estimation results; Section 10 concludes.

2 Literature Review

Sociologists identify five key underlying components of culture: language, moral, law, ethnicity, and religion (Macionis, 2011). Since many religions have some predetermined doctrines

[7]Judeo–Christian is a term employed in a historical sense, especially in the U.S., to refer to the connections between precursors of Christianity and Rabbinic Judaism in the Second Temple period.

[8]Referred sources are point, non–point, on–road and non–road. See Section 4 for more information.

[9]It is the adherents of religions as a percentage of total population.

concerning morals and norms, it would be reasonable to say that religion is a vital component in shaping our culture. In an often–cited article, Sapienza et al. (2006) conclude that cultural background plays a crucial role in important economic choices; and thus, cultural differences should be taken into account in any examination of economic phenomena. In line with these results, Barro and Mccleary (2003) provide substantial evidence that religious beliefs influence individual traits which enhance economic performance. Consequently, it appears that religion affects economic choices of a person via culture. Moreover, the theory of environmental economics suggests that there is a strong relationship between economic choices and environmental behavior of an agent (Baumol and Oates, 1988). From the above evidence, it can be inferred that religion indirectly, even if not directly, affects environmental behavior as demonstrated in Figure 2.

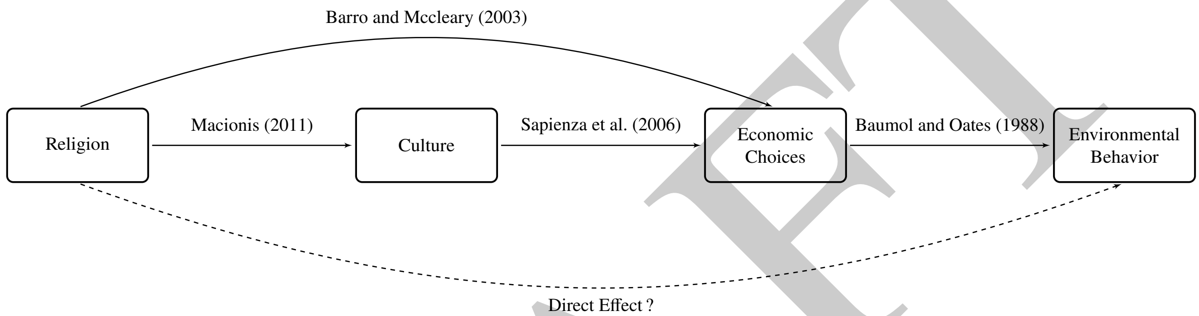


Figure 2: Connection Between Religion and Environment

Historically, literature about the possible relationship between religion and environment can be separated into two approaches. In the first approach, historians, theologians, and sociologists have tried to unravel the causal chain of events, which have resulted in the environmental crisis, back to a religious root. Lynn White Jr. is among the first to follow this approach. With his controversial article "The Historical Roots of Our Ecologic Crisis," White plays a significant role in starting the debate about the causal relationship between religion and environmental crisis (White, 1967). White claims that the Judeo–Christian tradition is responsible for the present ecological crisis; and hence, the roots of our trouble are so largely religious. This statement has become known as the "White's Thesis."

White's Thesis has not only drawn attention from historians and theologians but also given rise to a large number of empirical studies by social scientists who addressed the debate from a micro–perspective[10] (Vonk, 2012). Consequently, the second approach taken in the literature comes from social scientists who carry out extensive surveys to quantitatively measure the correlation between religion variables (e.g., church membership, biblical literalism[11], dominion

[10]The micro–perspective term is used here since the earlier studies have provided evidence for the White's Thesis by investigating the relationship between religion and environmental behavior in individual level. This study presents macro–perspective approach by examining the effect of percentage adherents of religions on the environmental performance of a county.

[11]Biblical literalism is the interpretation or translation of the explicit and primary sense of words in the Bible.

belief[12], and religious salience[13]) and environmental variables (e.g., environmental behavior and willingness to pay for environmental conservation).

Over the last three decades, almost twenty surveys on religion and environmental behavior have been published. The vast majority of previous studies have tried to provide some statistical evidence either in favor of or against the White's Thesis.[14] Thus, our knowledge of the previous literature is based only on the relationship between Judeo–Christian beliefs and environmental behavior as none of the previous studies have considered the other major religions in their analyses. Therefore, to see the effect of other faiths on environmental performance, the following major religions are also considered in our analysis: Orthodox Christianity, Protestantism, Islam, Buddhism, Hinduism and Mormonism.

2.1 Evidence in Favor of the White's Thesis

In an investigation of the White's Thesis, Hand and Liere (1984) find that non–Judeo–Christians are slightly more likely to show greater concern for environmental issues than Judeo–Christians; thus, provide support for the White's Thesis. Eckberg and Blocker (1989) utilize a dataset from a general telephone survey of adult residents in the Tulsa, Oklahoma metropolitan area in the Spring of 1985. They find a negative relationship between biblical literalism and environmental concerns. Using an extensive version of the Eckberg and Blocker (1989) data, Greeley (1993) confirms the White's Thesis by indicating that low levels of environmental concern correlate with biblical literalism.

Five years later, Eckberg and Blocker (1996) not only confirm their previous result that Christian theology has anti–environmental effects but also assert that religious participation has environmental friendly effects. By using a survey data carried out among university students in fourteen countries, Schultz et al. (2000) conclude that respondents with higher biblical literalism scored higher on anthropocentric[15] environmental concerns but lower on ecocentric[16] environmental concerns and the New Ecological Paradigm (NEP)[17] scale.

2.2 Evidence Against White's Thesis

Kanagy and Nelsen (1995) provide some statistical evidence against the White's Thesis. Their results conclude that religious individuals are no less likely than others to identify themselves as environmentalists. Wolkomir et al. (1997) claim that neither biblical literalism nor religious

[12]Dominion belief is the belief that Christians have a holy right to fill and seize positions of power in the society and government to rebuild them in an exclusively Christian way.

[13]Religious salience is a subjective indicator of the importance of religion to a person. It is a way that social scientists measure the degree of religiosity or religious commitment.

[14]Interested readers should read Vonk (2012) for a comprehensive literature review.

[15]Having a position of view which regards human beings as the most significant entity of the universe.

[16]Having a serious concern for environmental issues.

[17]NEP is a measure of endorsement of a "pro–ecological" worldviews. For more information, see Catton and Dunlap (1978) and Dunlap et al. (2000).

salience shows any negative impact on environmental attitudes. Woodrum and Wolkomir (1997) provide a firm rejection of the White's Thesis, which is a rarely seen result in the literature. They assert that religious affiliation has positive effects on environmental concerns and religious attendance has positive effects on environmental behavior. Hayes and Marangudakis (2001) conclude that there is no significant difference between Christians and non–Christians concerning environmental behavior. In the most recent research on this topic, Vonk (2012) compiles a useful overview of these studies—a total of twenty-one distinct surveys. By and large, she finds that adhering to Christian faith does not significantly or systematically lead to less or more environmentally friendly behavior.

2.3 Shortcomings of the Previous Research

The methodology of the previous literature has several weaknesses. This study tries to remedy some of these either by using different variables and data or by drawing from other literature. The shortcomings of the previous research are fourfold.

The first of these is the way of measuring religion. Vonk (2012) claims that the more explicitly religious beliefs are questioned, the less clear the difference between denominations can be determined. Therefore, her suggestion is not to focus on isolated items such as church attendance, biblical literalism, and religious salience but religion as a broader concept (i.e., a content of comprehensive worldview). Hence, this research utilizes the percentage adherents of religions of a county as explanatory variables.

The second shortcoming is the way environmental behavior is measured. Vonk (2012) asserts that environmental behavior is a quite loose concept. Her logic is that environmental behavior is determined by a complex number of factors and rarely unambiguous. For instance, individuals can make environmentally friendly choices at one point while polluting the environment at another point. Therefore, she suggests focusing on a broader concept of environmental behavior. Hence, following Vonk (2012), this study assumes that human–sourced air pollutants from all activities represent the broader concept of environmental behavior as a whole by accounting for all environmental aspects of a county.

A third shortcoming is that the main aim of all earlier studies is to test White's Thesis. Since White's Thesis is Judeo–Christian based, none of the previous studies have incorporated other major religions. Moreover, most of the survey data used in the previous studies are gathered from predominant Judeo–Christian communities which cover only one or at most couple of cities. International Social Survey Project (ISSP)[18], used by Barro and Mccleary (2003) to investigate the relationship between religion and economic growth, is the only country–level project as of my knowledge. However, the number of the participating predominant non–Judeo–Christian

[18]ISSP is a continuing annual program of cross–national collaboration on surveys covering topics which are essential for social science research. The outcomes of the surveys bring a cross–national and cross–cultural perspective to individual national studies.

countries is at most five out of forty-eight countries as of November 2016.[19] Therefore, the previous studies have neither provided evidence on the relationship between non–Judeo–Christian religions and environmental behavior nor used a comprehensive data with a finer spatial scale. The paper overcomes these problems by incorporating other major religions in a large county–level spatial analysis.

The last shortcoming relates to model specification. Especially the earlier works of the previous literature have not accounted for any determinants of environmental behavior but tried to provide evidence based on bivariate correlation only. The suggested remedy for this problem is to incorporate other important determinants of environmental performance by drawing from other literature.

3 Determinants of Environmental Performance

Although this research is interested primarily in the impact of religion on environmental performance, we need to control for some other factors that are identified mainly by economists as important determinants of environmental performance. For this purpose, the research uses four sets of variables which are religion, economic, education, and others.

Since the primary aim of this paper is to test the effect of religion on environmental performance, the way of measuring religion and how we interpret it is critical. Historically, religion has been measured based on isolated items such as church attendance, biblical literalism, and religious salience. This type of religion measurement suffers from the fact that the more explicitly religious beliefs are questioned, the less clear the difference between denominations can be determined (Vonk, 2012). This study departs from the previous literature in the way religion is measured. In testing the impact of religion on environmental performance, percentage adherents of religions are employed instead of focusing on isolated items about religion. By using percentage adherents, we can identify the impacts of different religions on environmental performance more clearly. One important caveat that the reader should be aware of is that this study does not investigate the effect of religiosity level of individuals on environmental performance but the effect of being connected to a specific religion on environmental performance.

There is a large body of theoretical and empirical literature that focuses on the economic determinants of environmental performance. In their comprehensive analysis of determinants of environmental performance, Esty and Porter (2001) show that there is a strong positive association between income and environmental performance. Moreover, Grossman and Krueger (1994) provide evidence that some forms of environmental degradation follow an Environmental Kuznets Curve (EKC) pattern. They find an inverted U–shape relationship between income per capita and pollution.[20] Esty and Porter (2001) also report that artificially low energy prices and

[19]URL: `https://goo.gl/IFKdIr` (visited on Jan. 30, 2017).

[20]The standard interpretation of EKC is that environmental quality is a luxury good in the initial stages of income. As income grows, people satisfy their most immediate necessities and tend to care gradually more about

subsidies are associated with low environmental performance. According to their conclusion, low energy prices and subsidies can give rise to massive inefficiency in energy usage which in turn decrease the environmental performance. On the other hand, Ohlan (2015) points out that developing alternative energy sources such as renewable energy, and using green and clean technologies assist in reducing CO_2 emissions. Following the literature, this study includes income per capita, the quadratic functional form of income per capita, gas price excluding taxes and fees, gas taxes and fees, and renewable energy consumption per capita[21] as the economic determinants of environmental performance.

The increasing concerns about environmental sustainability have led policymakers in developing countries to modify their policies in an environmentally sensitive way. Education is regarded to be the ultimate tool that increases awareness, knowledge, and understanding of the environmental issues. Hence, providing environmentally sensitive education to rising generations is considered as protecting the natural environment in the long run. On this point, Woodrum and Wolkomir (1997) and Hayes and Marangudakis (2001) provide some evidence that the better–educated and scientifically more knowledgeable individuals are more likely to express a pro–environment stance. Hence, education level (i.e., a proxy for the rate of well–educated society) is employed as an explanatory variable for environmental performance.

High population density is likely to increase pressure on finite natural resources and results in poor environmental management; and hence, a possible low environmental performance. For instance, high population density may result in more energy–intensive industrial agriculture and higher rates of deforestation (Papyrakis, 2012). Lamla (2009) and Dulal et al. (2011) point out a close relationship between population pressure and adverse environmental outcomes. Moreover, Ohlan (2015) and Gudipudi et al. (2016) find that population density has a statistically significant positive effect on emissions. Jones and Kammen (2014) report that household carbon footprint (HCF) increases until a population density of about three-thousand people per square mile is reached, after that HCF declines. However, the net effect of this relationship is not statistically significant for all U.S. zip codes. On the other hand, Makido et al. (2012) find a U–shape relationship such that higher population density may be better for lower emissions in general sense but too dense and mono–centric urban settlement can increase residential emissions. In line with this result, Gately et al. (2015) find that high population density correlates with low CO_2 emissions. Although a great deal of previous research has focused on the relationship between population density and emissions, results vary significantly. Thus, population density is employed as an explanatory variable for environmental performance to test the previous results in a spatial analysis context.

Ambient temperature and relative humidity can have a substantial impact on emissions and emission processes. For instance, low temperatures are associated with high start emissions for

environmental quality. Thus, the EKC theory suggests a U–shape relationship between income and environmental quality.

[21] Since data for renewable energy prices are not available for U.S. counties or states, renewable energy consumption per capita is used as a proxy for it.

many pollutants whereas high temperatures are associated with higher evaporative emissions. Moreover, high temperatures and high relative humidity are associated with greater running emissions due to the increase in the heat index and resulting higher engine load for air conditioning (EPA, 2011b). Therefore, to control for weather events, precipitation and temperature variables are employed as explanatory variables of environmental performance.

4 Data and Shapefile Descriptions

This study utilizes a dataset created by combining five county–level and three state–level data sources along with two types of shapefiles for U.S. counties in 2010. However, all analyses are conducted at the county–level.[22] Table 1 presents the descriptive summary of data and shapefiles along with sources and number of observations.

In measuring environmental performance, National Emissions Inventory (NEI) 2011 from U.S. Environmental Protection Agency (EPA) is used as a proxy.[23] The NEI is a complete and detailed estimate of criteria air pollutants (CAPs), criteria precursors, and hazardous air pollutants (HAPs) from air emissions sources. The NEI data is released every three years based primarily upon data provided by many different sources, including industry, state, local, and tribal agencies, and supplemented with data developed by EPA. Some emissions data are based on actual measurements whereas others are estimates. This study uses NEI 2011 data. EPA compiles the NEI 2011 data from eight major source sectors: agriculture, dust, fires, fuel combustion, industrial process, miscellaneous, mobile and solvent; then, categorizes it into five main data groups: point, non–point, on–road, non–road and event. The NEI 2011 data covers all fifty states, District of Columbia, Puerto Rico and the Virgin Islands.

In order to protect the environment and public health, the Clean Air Act requires EPA to set and regulate National Ambient Air Quality Standards (NAAQS) for six principal air pollutants, also known as CAPs. These pollutants include lead (Pb)[24], carbon monoxide (CO), nitrogen oxide (NO_2), sulfur dioxide (SO_2), particulate matter 10 microns or less (PM_{10}), particulate matter 2.5 microns or less ($PM_{2.5}$) and ground–level ozone (O_3). Four of these pollutants (i.e., CO, Pb, NO_2, and SO_2) are emitted directly from various sources. O_3 is not directly emitted but formed when nitrogen oxides (NO_x) and volatile organic compounds (VOC)[25] react in the presence of sunlight. PM_{10} is directly emitted while $PM_{2.5}$ is mostly formed when emissions of NO_x, SO_2, VOC and ammonia (NH_3)[26] react in the atmosphere. Thus, EPA provides the NEI 2011 data for only *CO*, NH_3, NO_x, SO_2, PM_{10}, $PM_{2.5}$, and *VOC* as CAPs.

[22] See Appendix A for the reasoning behind using county–level data.

[23] An increase in emissions can be considered as a decrease in environmental performance.

[24] While Pb is a CAP for the NAAQS, due to toxic attributes and inclusion in the previous National Air Toxics Assessment, it is classified as a HAP in the NEI 2011.

[25] VOCs are organic chemicals that have a high vapor pressure at room temperature and emitted as gases from certain solids or liquids. VOCs include various chemicals, some of which may have short and long–term adverse health effects.

[26] NH_3 is technically not a CAP but an important precursor.

In the NEI 2011, HAPs are reported from state, local and tribal agencies on a voluntary basis whereas CAPs has to be reported from every agency. Therefore, only the CAPs as mentioned above are used as a dependent variable in separate regressions. In line with this approach, data from the *event* category[27] are omitted since reporting emissions from these activities are encouraged but not required. Moreover, while compiling the NEI 2011 data for our analysis, emissions from natural resources are omitted to have anthropogenic[28] emissions only. As it is suggested by EPA (2011b), tribal emissions are dropped since they may duplicate emissions already accounted for at the county–level. Also, to prevent double counting problem, county–level data is generated by aggregating emissions not from main data groups but major sectors. Moreover, since the emissions data for condensable and filterable components of PM_{10} and $PM_{2.5}$ are not complete, only the primary PMs are used (EPA, 2011b). In some states[29] there are portable emission sources which do not belong to any county. Thus, emissions from these sources are distributed among the counties by population weights. Finally, this research follows the literature and clears off the scale effect in emissions data by using emissions tons per capita.[30]

Religion data is collected from Association of Religion Data Archives (ARDA) website. The data collection is conducted by U.S. Religion Census: Religious Congregations on Membership Study 2010 (RCMS 2010) which is designed and carried out by Association of Statisticians of American Religious Bodies (ASARB).[31] In the RCMS 2010, each participating religious body supplied the number of congregations[32] and adherents[33] for each group in all counties of fifty states and District of Columbia. The RCMS 2010 reports that 150 million Americans (48.8% of the population in 2010) is associated with the 236 reporting religious bodies. The RCMS 2010 categorizes some religious bodies under adherents of five major religions which are *Evangelical Protestants*, *Black Protestants*, *Mainline Protestants*, *Catholics* and *Orthodox Christians*. Along with the adherents of these major religions, this research uses adherents of other prominent religions such as *Muslims*, *Jews*, *Hindus*, *Buddhists*, and *Mormons*.[34] In exploratory spatial data analysis, two other categories are employed, which are *Protestants* for all Protestants and *Adherents* for adherents of all religions. One problem with the RCMS 2010 data is that, for some counties, adherents total exceed the population count of the Census 2010. Possible explanations include the Census 2010 undercount, congregation membership overcount, and individuals' county of residence differs from the county of congregation membership (Grammich et al., 2012).

[27]Covers wildfires and prescribed burns.

[28]Caused or produced by human activity.

[29]Alaska, Colorado, Florida, Kentucky, Michigan, Minnesota, Nevada, Ohio, Texas, and Wisconsin.

[30]The Census 2010 population counts are used in per capita calculation.

[31]Interested readers should see the ARDA web site for further information about the RCMS 2010. See URLs: `https://goo.gl/IPYX5d` and `https://goo.gl/dnQmwS` (visited on Jan. 30, 2017).

[32]A congregation may be defined as a group of people who meet regularly at a pre–announced time and location. Congregations may be mosques, temples, churches, or other meeting places.

[33]The adherent figure is a complete count of people affiliated with a congregation. Adherents may include children, members, and attendees who are not members.

[34]Since these religions, other than Islam, have several sub–religious bodies but not categorized into a main category in the raw data, one overall group is created for each religion by combining several subgroups.

Data for these counties are left as is since neither ARDA nor ASARB suggests any approach to mitigate this problem. Finally, variables for adherents of religions are measured in percentages of the total population[35] to clear off the scale effect.

From American Community Survey (ACS) 2006–2010 (5–Year Estimates), utilized datasets are income per capita data as *income* measured in 2010 prices and education data as a percentage of highest educational attainment for population twenty-five years and over. Selected categories for education are *some college or more*, *bachelor's degree or more* and *doctorate degree*.[36] From the Census 2010, population data measured as *population density*[37] is employed. Datasets taken from Census and ACS cover all fifty states, Puerto Rico and District of Columbia.

Data for gas prices excluding taxes and fees in 2010 is gathered from U.S. Energy Information Administration (EIA). It is measured as *gas price* per gallon in 2010 prices. It represents the average motor gasoline[38] prices in all grade levels such as regular, midgrade and premium of the sales[39] made directly to the consumer of the product. From American Petroleum Institute (API), data for total gas taxes and fees in 2010 are collected. It is measured as *gas tax/fee* per gallon in 2010 prices. It includes the total of state excise tax, other taxes and fees, and federal excise tax for gasoline.[40] State Energy Data System from EIA is used to collect renewable energy consumption in 2010. It is measured as *renewable energy consumption* in billion British thermal units per capita[41]. Renewable sources of energy include fuel ethanol, wood, waste, hydroelectric, geothermal, solar, and wind energy. These three datasets cover all fifty states and District of Columbia. Since the finest spatial scale available for these datasets is state–level, the respective state–level observation value is used for each county in that state.

Precipitation and temperature data are taken from North American Land Data Assimilation System 2010. Precipitation is measured as *mean daily precipitation* in millimeters (mm). For temperature, instead of employing the commonly used measures such as mean daily maximum or minimum temperature, *mean daily maximum heat index* in °F is used. Heat Index is an index that combines relative humidity and air temperature, also known as felt air temperature. Precipitation and temperature datasets cover only the forty-eight contiguous states and District of Columbia. Missing observations, one in precipitation and eight in heat index data, are imputed with the respective means of each county's state.

This research uses two types of shapefiles[42] that match with the Census 2010. Both of

[35]The Census 2010 population counts are used in the calculation of percentage adherents of religions.

[36]Note that the education variables are directly taken as percentages from ACS 2006–2010 (5–Year Estimates).

[37]That is population per square mile. In the calculation, only the land area of each county is used.

[38]Motor gasoline includes conventional gasoline, all types of oxygenated gasoline but excludes aviation gasoline.

[39]Includes bulk consumers such as agriculture, industry, and utilities, as well as residential and commercial consumers.

[40]In the calculation of total *gas tax/fee*, API uses the weighted average of local taxes and fees by the population of each municipality to come up with an average tax and fee for the entire state. Then this value is summed up with the federal excise tax. Total *gas tax/fee* may include any of the followings: excise taxes, environmental fees, storage tank taxes, other fees or taxes, and general sales tax.

[41]The Census 2010 state population counts are used in per capita calculation.

[42]By default, both shapefiles use the North American Datum of 1983, which is the horizontal control datum for the U.S. based on a geocentric origin, and the Geodetic Reference System 1980, which uses longitude and latitude

them are downloaded from Census Bureau website.[43] The TIGER/Line Shapefile (i.e., with full detailed borders) contains the legal boundaries of counties which extend three miles out to the ocean or water area whereas in Cartographic Boundary Shapefile (i.e., less detailed) borders end at the shoreline. TIGER/Line Shapefile with full detailed county borders layer is used for spatial analyses such as constructing neighbor objects, weight matrices, and performing spatial regressions. However, Cartographic Boundary Shapefile is employed for plotting purposes only. The reason for this approach is that for a better thematic mapping and complete spatial analysis we need to have all counties with at least one neighbor. This condition is only satisfied with TIGER/Line Shapefile since the borders do not end at the shoreline[44]. However, embedding the plots generated with it to any output document increases the final file size tremendously. Therefore, Cartographic Boundary Shapefile with low–resolution level (20m = 1:20,000,000) is used for thematic mapping only. Figure 3 and Figure 4 show the differences between two shapefiles.

Moreover, to use in spatial analysis, mean center of population of each county in longitude and latitude decimal degrees (i.e., *population centroid*[45]) is taken from the Census 2010. Small red dots in Figure 5 presents the population centroids.

All variables mentioned above, if not in percentages in the raw data, are processed via using the natural logarithm transformation, and only the transformed forms are used throughout the study.[46] One reason for this approach is to ease the interpretability of coefficient estimates in regression analyses. Another reason for the logarithmic transformation is that the Maximum Likelihood estimation of spatial models requires normally distributed dependent variables. Moreover, Tabachnick et al. (2001) suggest using logarithmic transformation in case of substantially positively skewed variables which applies to our data.

Finally, all datasets mentioned above are combined by county FIPS and embedded to both shapefiles. The merged dataset and both shapefiles are subsetted to cover only the counties of forty-eight contiguous states and District of Columbia, which yields contiguous spatial data (i.e., all counties have at least one neighbor). Table 2 presents the summary statistics of raw variables including the transformed versions.[47] As it can be seen from Table 2, the transformed dependent variables are in the range of rule of thumb for normality[48]. Discussion of the thematic maps for

in decimal degrees. European Petroleum Survey Group code is `EPSG:4269`.

[43] URLs: `https://goo.gl/9IHK3B` and `https://goo.gl/P6GGxE` (visited on Jan. 30, 2017).

[44] For instance, Richmond County, NY is an island and connected to the mainland by four bridges. Due to the *borders end at the shoreline* feature of Cartographic Boundary Shapefile, it is a county without any neighbors. However, TIGER/Line Shapefile yields one or more neighbors by using the *legal borders* of the county.

[45] Population centroids are generated by excluding the water areas from the boundary of counties. So, all of the centroids are on land.

[46] Note that all religion and education variables are in percentages. Thus, they are not transformed and used in percentages throughout the study.

[47] To prevent repetition, in the text, tables, and figures of the subsequent sections, all dependent and explanatory variables are labeled without units.

[48] The conservative guidelines of Tabachnick et al. (2001) suggest using ±2 for both skewness and kurtosis. West et al. (1995) is more liberal and recommend using ±2 for skewness and ±7 for kurtosis. Note that throughout this paper, *kurtosis* refers to *excess kurtosis* which is the kurtosis minus three and provides a comparison to the normal

all dependent and explanatory variables are saved for Section 7 where spatial autocorrelation tests for transformed dependent variables are covered.

All analyses performed in this study, including data cleaning and merging, are done by using open source software R. Appendix J presents the R version information along with the used packages.

5 Contiguity and Weight Matrices

Spatial analyses for areal data[49] require a specification of contiguity type (i.e., what a neighbor is) for areal units as well as weight style (i.e., how neighbors influence each other). Therefore, specifying contiguity type and weight style is a necessary step for using areal data. The first step is to define which type of relationship between areal units are to be given a non–zero neighbor link (i.e., choosing the contiguity type). The second step is to assign weights to the identified non–zero neighbor links (i.e., selecting the weight style).

Contiguity and weighting frameworks are *ad hoc* assumptions that incorporate the prior structure of spatial dependence or its weaker form spatial autocorrelation between areal units.[50] Since these frameworks can dramatically affect our inferences, they need to be theoretically defended.[51] Nevertheless, in the spatial econometrics literature, scholars rarely discuss theoretical reasons for their choice of contiguity type and weight style. Instead, researchers test several contiguity type–weight style combinations (i.e., weight matrix) through some information criterion or hypothesis testing procedures. These are necessary procedures due to the insufficient information in specifying an $n \times n$ weight matrix from observations of a single cross–section data, where n is the observation number. For a detailed literature review of contiguity types and weight styles, see Anselin (1988), Lesage and Pace (2008), Bavaud (2010), Zhukov and Stewart (2012), and Anselin and Florax (2012).

5.1 Creating Contiguity Matrix

Although the spatial econometrics literature suggests a wide range of contiguity types, three categories of contiguity types are commonly used: adjacency[52], interpoint distance[53], and graph–based[54] contiguities. Since these categories are quite broad and encompass a wide

distribution.

[49]See Appendix A for the details of areal data.

[50]In the spatial econometrics literature, spatial dependence and spatial autocorrelation are used interchangeably. Apparently, the two terms are not identical, but the weaker form is used more often.

[51]See Assunção and Krainski (2009) for a thorough discussion about how contiguity and weighting frameworks affect variance–covariance matrix of explanatory variables, inference, and spatial coefficient parameters in spatial regressions.

[52]Linear (common eastern or western border), Rook (common border), Bishop (common vertex), and Queen (common border and vertex) contiguities.

[53]Minimum Distance and K Nearest Neighbors contiguities (KNN).

[54]Delaunay Triangulation, Gabriel Graph, Relative Neighbor Graph, and Sphere of Influence contiguities.

range of possibilities, the present study initially considers only six contiguity types which are representatives of the categories they belong. Used contiguity types are 1st Order Queen, 2nd Order Queen, Minimum Distance, 6 Nearest Neighbors (6NN), 10 Nearest Neighbors (10NN), and Sphere of Influence contiguities. For detailed information about the selected contiguity types, see Appendix B.

Each contiguity type specifies links between individual areal units with an $n \times n$ binary contiguity matrix $\mathbf{C}$ in which entry $c_{ij} = 1$ if two areal units are considered connected assuming $i \neq j$ and $c_{ij} = 0$ if they are not, where $i = (1, 2, \ldots, n)$ and $j = (1, 2, \ldots, n)$ are the areal units. By convention, diagonal elements of contiguity matrix $\mathbf{C}$ are set to zero since no areal unit can be viewed as its own neighbor.

Figure 6 plots the neighbor links of the selected contiguity types for North Carolina counties, where counties of one state are used for better visualization. Figure 7 shows the frequency distribution of the number of neighbors by contiguity type, where 6NN and 10NN contiguities are dropped since they have fixed number of neighbors. Figure 8 presents the differences between 1st and 2nd Order Queen contiguities, and between 6NN and 10NN contiguities. Figure 9 and Figure 10 plot the counties with the maximum and minimum number of neighbors respectively, and their neighbors for 1st Order Queen, Minimum Distance, and Sphere of Influence contiguities.

5.2 Creating Weight Matrix

Once the contiguity type is defined, the researcher must transform binary contiguity matrix $\mathbf{C}$ into weights using a weight style. This transformation forms the elements of an $n \times n$ weight matrix $\mathbf{W}$. In spatial autocorrelation tests and spatial regressions, the matrix $\mathbf{W}$ is used to calculate spatially lagged observed values. For instance, vector $\mathbf{WY}$ (i.e., cross–product of matrix $\mathbf{W}$ and the vector of observed dependent variable $\mathbf{Y}$) is called spatially lagged $\mathbf{Y}$ and interpreted as the weighted sum of neighboring dependent values.

Several weight styles appear in the spatial econometrics literature, but this study incorporates five of the most commonly used ones: binary (B), row–standardized (W), global–standardized (C), variance–stabilizing (S), and minmax–normalized (MINMAX) weight styles. Distance decay weight styles (e.g., inverse distance, inverse distance squared, and negative exponentials of distance) are not considered since we have no prior information about whether weight depends on exact distance.[55] For detailed information about the selected weight styles, see Appendix C.

Figure 11 presents row sums of weights for each number of neighbors by contiguity type and weight style. As it can be seen from Figure 6 and Figure 11, Minimum Distance contiguity creates far more than an excessive number of neighbor links (i.e., more than 100 links for some counties). Thus, to prevent the noise generated by possible irrelevant links, Minimum Distance contiguity is dropped at this point.

[55] If we believe that the strength of neighbor links decreases with distance, distance decay weight styles can be applied to contiguity matrix $\mathbf{C}$. However, if we have no reason to assume any more knowledge about neighbor relations than their existence or absence, this step is potentially misleading.

The most intuitive and commonly used weight matrix in the literature is generated with a combination of 1st Order Queen contiguity and row–standardized weight style. However, in this study, all possible combinations are employed since we have no prior knowledge about the contiguity type and weight style. As a result, while testing spatial autocorrelation, twenty-five different contiguity type–weight style combinations (i.e., weight matrix **W**) are employed. These combinations are produced by separately applying the five weight styles to each of the five contiguity types.

6 Global and Local Tests for Spatial Autocorrelation

Use of spatial econometric methods can help to correct for inefficient and possibly biased estimators when spatial autocorrelation is present in the data. On the other hand, excessive modeling of spatial processes when there is no evidence for spatial autocorrelation may lead to inefficient estimators and adds unnecessary complexity to a model. Therefore, testing for spatial autocorrelation is a crucial step to determine whether more complicated spatial models are necessary.

Spatial autocorrelation is the correlation among observations of a variable of interest strictly attributable to their relatively close locations. The most common test for the existence of spatial autocorrelation in irregular areal units with continuous data is due to Moran (1950), which is usually referred to as Moran's I. Moran's I statistic measures the intensity of spatial autocorrelation in a spatial process but not the spatial dependence coefficient directly. It is used to determine whether neighboring areas are more similar than would be expected under the null hypothesis. Therefore, Moran's I test is applied to justify the need for spatial methods employed in this study.

For a particular data or residuals of a model, the Moran's I statistics is defined as

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{1}$$

where n is the number of areal units; x_i and x_j are the values of variable of interest for the i^{th} and j^{th} areal units respectively; $\bar{x}$ is the sample mean of variable of interest; and w_{ij} is an element of the weight matrix **W** which specifies the link between areal units i and j.[56] The test statistic can take values between the range of -1 and 1. A test statistic approaching 1 results in a positive autocorrelation (i.e., high values are clustered by high values as well as low values clustered by low values). On the other hand, a test statistic close to -1 indicates negative autocorrelation. A random spatial pattern exists when the value equals to 0. Although not directly, Moran's I

[56]Theoretical moments and some other properties of Moran's I statistic can be found in Appendix D.1.

statistic for spatial autocorrelation is analogous to simple autocorrelation coefficient. However, spatial autocorrelation is more complex than one–dimensional simple autocorrelation since spatial autocorrelation is two–dimensional and multi–directional.

As suggested by Cliff and Ord (1972), under the null hypothesis of no spatial autocorrelation, inference performed on Moran's I can be constructed with normality assumption, randomization, and Monte Carlo simulation (i.e., equivalent to permutation–based bootstrap tests).[57] The computation of Moran's I statistic depends on a given choice of the weight matrix **W**. If in fact the pattern of spatial dependence is generated by a different weight matrix, then the test can give spurious results. Therefore, to ensure robustness, Moran's I test is conducted with the twenty-five different weight matrices explained in Section 5.2.

The Moran's I just introduced above is based on simultaneous measurements from all areal units and assumes that the spatial process is the same over the entire geographic area. Hence, it is a global statistics and often called Global Moran's I. However, as proposed by Anselin (1995), it is possible to decompose a global statistic into its contributions from each areal unit to test for local spatial autocorrelation and detect areal units with high influence on the global statistic.[58] The procedure of decomposing global statistic into its local components is called Local Indicators of Spatial Association (LISA). The global statistic only provides the significance of spatial process. However, local statistics provide a unique number for each areal unit and identify areal units of particular clusters and hot spots (i.e., areal units with different neighbors). Therefore, global and local statistics are often used together for a thorough understanding of the spatial process in the interest of data.

The corresponding LISA of Global Moran's I is known as Local Moran's I, and for the i^{th} areal unit it is defined as (Anselin, 1995)

$$I_i = \frac{(x_i - \bar{x}) \sum_{j=1}^{n} w_{ij}(x_j - \bar{x})}{n^{-1} \sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{2}$$

then the relationship between Global Moran's I and Local Moran's I is defined as

$$I = \sum_{i=1}^{n} \frac{I_i}{n} \tag{3}$$

For detailed information about the computation methods of Local Moran's I, see Appendix E.

The LISA is a useful addition to the toolbox of Exploratory Spatial Data Analysis (ESDA)[59]

[57]For detailed information about these methods, see Appendix D.

[58]Anselin (1995) states two requirements that have to be fulfilled by a sensible local statistic. First, the local statistic for each areal unit indicates significant spatial clustering of similar values around that areal unit. Second, the sum of local statistics for all areal units is proportional (or equal) to a corresponding global statistic.

[59]See Appendix F for details.

techniques along with the Global Moran's I in which two important interpretations are combined: (1) the detection of outliers and influential areal units on the global statistic depicted by LISA Moran scatterplot; and (2) the assessment of significant local spatial clustering around an individual areal unit illustrated by LISA cluster map.

The LISA Moran scatterplot combines the information from Moran scatterplot[60] and the significance of the LISA statistics with a prespecified significance level. It labels areal units by the significance of LISA with a color code, which categorizes spatial association into four different groups corresponding to the quadrants in Moran scatterplot (Anselin and Bao, 1997). The groups are high surrounded by high, low surrounded by low, high surrounded by low, and low surrounded by high. The LISA cluster map visualizes locations of significant clusters on a map by using the same significance level and color code used in LISA Moran scatterplot.

7 Spatial Autocorrelation of The Dependent Variables

In the spatial econometrics literature, it is a common practice to test dependent variables of regression models for spatial autocorrelation.[61] As a preliminary visual check for spatial patterns in dependent variables, ESDA thematic mapping is used.[62] Figure 12 displays thematic maps of all dependent variables.[63] These thematic maps suggest a strong positive spatial pattern since high levels of emissions are clustered in the central part of U.S. and on the direction of south–east to north–west whereas low levels are clustered around south–west, east, and north–east of U.S. Specifically, high levels for all CAPs, except *log* SO_2, are observed in counties of North Dakota, South Dakota, Montana, Wyoming, Nebraska, Kansas, Oklahoma, New Mexico, Texas, Arkansas, and Louisiana. Whereas, counties in the Northeast region of U.S. show low levels for all CAPs, except *log* SO_2.

Table 3 presents Global Moran's I test results for all dependent variables with twenty-five different weight matrices.[64] Results indicate statistically significant positive spatial autocorrelation for all dependent variables with any weight matrix.[65] It seems that strongest spatial autocorrelation occurs for *log* $PM_{2.5}$ and *log* PM_{10} whereas *log* SO_2 has the weakest.

Figure 13 shows LISA Moran scatterplot and LISA cluster map for each dependent variable.[66]

[60]Moran scatterplot of Anselin et al. (1996b) is a visualization of the relationship between observations in the variable of interest and their spatially lagged values. The slope of the regression line between these corresponds to the value of Global Moran's I.

[61]Spatial autocorrelation tests for explanatory variables are often omitted in the literature since spatial patterns are usually incorporated into models through dependent variables or error terms. See Section 8.3 for more information.

[62]See Appendix G for the classification methods used in thematic maps.

[63]See Figure 15 for the thematic maps of other variables.

[64]The presented results are constructed under normality assumption. Results for randomization and Monte Carlo simulation are dropped since they yield the same results. It should not be surprising since the sample size is large enough that the normality assumption should hold.

[65]Throughout the spatial autocorrelation tests, one–sided (i.e., greater than zero) alternative hypothesis is used since negative spatial autocorrelation is rarely observed. Two–sided alternative yields the same results due to the large sample size. However, in small samples, the inference might be affected when the wrong alternative is used.

[66]A weight matrix generated with a combination of 1st Order Queen contiguity and row–standardized weight

These plots confirm our conclusion from ESDA thematic mapping: (1) there is a statistically significant strong positive spatial autocorrelation in all dependent variables; (2) high levels of emissions are clustered in the north, north–west and central part of U.S.; and (3) low levels are clustered around south of California, Florida, and east and north–east of U.S.

8 Empirical Methods

Before performing spatial regression analysis, a base model[67] for each CAP is needed to be determined through sensitivity analysis using Ordinary Least Squares (OLS). This procedure should ease our job when the spatial dimension is added to a base model. Therefore, rest of the paper proceeds as follows: sensitivity analysis for determining all base models; estimating the selected base models through OLS; thematic maps of base model OLS residuals; calculating Global Moran's I, and plotting LISA Moran scatterplot and cluster maps for OLS residuals of base models; literature review of spatial models; sensitivity analysis for spatial model selection; estimating and discussing the base model spatial regressions; and finally visualizing the spatial autocorrelation and some of the equilibrium effects with thematic maps.

8.1 Preliminary OLS Regressions and Sensitivity Analysis

By combining the religion, economic, education, and other determinants of environmental performance, each base model is estimated with OLS as a preliminary analysis. The set of base model OLS regressions in compact form are defined as

$$Y_m = \iota_n \alpha_m + X_m \beta_m + \varepsilon_m \tag{4}$$

where n and m are the numbers of observations and base model OLS regressions respectively; $\mathbf{Y_m}$ is an $n \times 1$ vector of observed dependent variable for the m^{th} regression; $\alpha_{\mathbf{m}}$ is a scalar represents the constant term parameter; $\iota_{\mathbf{n}}$ is an $n \times 1$ vector of ones; $\mathbf{X_m}$ is an $n \times k$ matrix contains n observations on k explanatory variables; $\beta_{\mathbf{m}}$ is a $k \times 1$ vector representing parameters associated with the k explanatory variables in the predictor matrix $\mathbf{X_m}$; and $\varepsilon_{\mathbf{m}}$ is an $n \times 1$ vector of random errors with $\varepsilon_{\mathbf{m}} \overset{iid}{\sim} N(0, \sigma^2 I_n)$.

The primary purpose of the sensitivity analysis is to determine whether to include *log income squared* as an additional covariate and to decide which education variable should be incorporated. Therefore, sensitivity analysis is performed in two parts for each base model.

In the first part, the decision of including *log income squared* as an additional covariate is given by performing Likelihood Ratio (LR) test for nested models while either one of the three

style is used. Other weight matrices are dropped to save space since they are similar to what is presented.

[67] *Base model* refers to a model in which a CAP in the natural logarithmic form is used as a dependent variable in a separate regression. Throughout this study, each base model is labeled with its label of the dependent variable, and the label of each base model changes according to the results of sensitivity analyses in OLS and spatial regressions.

education variables is already added to the model. Table 4 presents the LR test statistics[68]. The LR tests show that H_0 is rejected in favor of H_1 for each education variable in almost all base models, except where *log CO* and *log VOC* are used as dependent variable.[69] In other words, a model with *log income squared* fits significantly better than a model without it. Thus, *log income squared* is added to all base models, except where log *CO* and *log VOC* are employed as a dependent variable.[70] After this point, the base models are labeled as CO, $NH_3^{\dagger}$, $NO_x^{\dagger}$, $PM_{10}^{\dagger}$, $PM_{2.5}^{\dagger}$, $SO_2^{\dagger}$, and VOC, where † indicates a base model including *log income squared*.

In the second part of the sensitivity analysis, one of the three education variables (i.e., *log some college or more*, *log bachelor's degree or more* and *log doctorate degree*) is selected for each base model, while maintaining the results of part one. The choice of these variables is based on two information criteria and two tests for non–nested model selection. Table 5 and Table 6 show the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for each base model by education variable. Both the AIC and BIC indicate that *log some college or more* returns a better fit for base models CO, $PM_{10}^{\dagger}$, and $PM_{2.5}^{\dagger}$. Whereas, *log bachelor's degree or more* yields a better fit for $NH_3^{\dagger}$, $NO_x^{\dagger}$, $SO_2^{\dagger}$, and VOC.

In order to further analyze these results with hypothesis testing, Cox–Pesaran–Deaton (CPD) and Davidson–MacKinnon J (DMJ) tests are employed. The CPD test is suggested by Pesaran and Deaton (1978) and based on the earlier work of Cox (1961, 1962) and Pesaran (1974). The CPD test allows one to test the truth of a regression when there exists a non–nested alternative hypothesis. The idea of the CPD test is as follows: if the first model includes the correct set of regressors, then a fit of the regressors from the second model to the fitted values from the first model should have no further explanatory value. However, if it has, it can be concluded that the first model does not contain the correct set of regressors. Thus, to compare both models, the fitted values of the first model are regressed on the regressors from the second model and vice versa. The DMJ test, proposed by Davidson and MacKinnon (1981), has a similar procedure. If the first model includes the correct set of regressors, then adding the fitted values of the second model into the set of regressors should yield no significant improvement. However, if it does, it can be concluded that the first model does not include the correct set of regressors.[71] Hence, to compare both models, the fitted values of the second model are included in the first model and vice versa.

Table 7 and Table 8 presents the CPD and DMJ test results respectively. In almost all base

[68]The LR test statistic is $\lambda = 2\,[\ln \mathcal{L}(H_1) - \ln \mathcal{L}(H_0)]$. The probability distribution of the LR test statistic is approximately a χ^2 distribution with degrees of freedom equal to the number of parameters that are constrained (Greene, 2003). Degrees of freedom is equal to one for all tests presented in Table 4; and thus, omitted.

[69]H_1 is a model with *log income squared*, and H_0 is a model without it.

[70]Under the Gauss–Markov assumptions, the OLS estimators are unbiased. It implies that including an irrelevant variable in a model does not affect the unbiasedness of the intercept and other slope estimators. However, it generally inflates the variances of the remaining OLS estimators because of multicollinearity when the true parameter of interest is zero. Since adding *log income squared* to base models with *log CO* and *log VOC* as dependent variables yields lower adjusted R_2 and non–significant estimates for both income variables, it is excluded just for these base models. See Greene (2003) for consequences of including an irrelevant variable.

[71]For further details, see Greene (2003).

models, *log doctorate degree* returns a worse fit than *log some college or more* and *log bachelor's degree or more*. Therefore, only the latter two are considered for each base model. Both the CPD and DMJ tests conclude that *log bachelor's degree or more* yields better fit for base models $NH_3^{\dagger}$, $NO_x^{\dagger}$, $SO_2^{\dagger}$, and VOC, which confirms the results of AIC and BIC. However, tests give conflicting and mixed results for base models CO, $PM_{10}^{\dagger}$, and $PM_{2.5}^{\dagger}$.[72] Finally, combining all results from AIC, BIC, and tests for non–nested models, *log some college or more* is selected for base models CO, $PM_{10}^{\dagger}$, and $PM_{2.5}^{\dagger}$. Whereas, *log bachelor's degree or more* is chosen for models $NH_3^{\dagger}$, $NO_x^{\dagger}$, $SO_2^{\dagger}$, and VOC. After this point, the base models are labeled as CO^c, $NH_3^{\dagger,b}$, $NO_x^{\dagger,b}$, $PM_{10}^{\dagger,c}$, $PM_{2.5}^{\dagger,c}$, $SO_2^{\dagger,b}$, and VOC^b, where ^{c} and ^{b} indicate a base model including *log some college or more* and *log bachelor's degree or more* respectively.

8.2 Base Model OLS Results

Table 9 presents the base model OLS results along with Koenker's studentized version of Breusch–Pagan (BP) test for heteroskedasticity (Breusch and Pagan, 1979; Koenker, 1981). The BP test suggests statistically significant heteroskedastic errors in all base models which might arise from the remaining spatial pattern in the residuals. Therefore base model OLS residuals are tested for spatial autocorrelation to determine whether the OLS model is justified.

Visual check for remaining spatial autocorrelation in base model OLS residuals is performed with ESDA thematic mapping which is displayed in Figure 16. Although the intensity of spatial pattern is decreased compared to the spatial pattern in the dependent variables, it is clear that there is still a spatial pattern remaining in all base model OLS residuals. Table 18 presents the Global Moran's I test results for all base model OLS residuals with twenty-five different weight matrices. Results are similar to the Global Moran's I tests for dependent variables. However, spatial autocorrelation is weaker as concluded in our visual check. The same conclusion is made for the LISA Moran scatterplot and LISA cluster maps which are presented in Figure 17.

It is important to note that none of the base model OLS results account for the statistically significant positive spatial autocorrelation observed in the dependent variables. Disregarding the spatial autocorrelation acts as an omitted variable in regression analysis; and thus, causes errors to be heteroskedastic and spatially correlated (Greene, 2003). The violation of "spherical errors"[73] assumption in OLS causes the interest of coefficients to be no longer efficient even though they might or might not be unbiased (Greene, 2003). Hence, failure to take the suggested spatial autocorrelation into account leads to incorrect inferences that are generally biased toward from rejecting the stated hypotheses.

As a result, all base model OLS estimates presented in Table 9 are invalid, and it would be appropriate not to make the necessary inference now but to save it for the results of spatial

[72] See Godfrey and Pesaran (1983), Greene (2003), and Ghali et al. (2011) for the critiques and restrictions of CPD and DMJ tests.

[73] Errors that meet the assumptions of homoskedasticity and non–autocorrelation are sometimes called spherical errors (Greene, 2003).

models. Estimating the spatial dependence through spatial models can help to correct for a potentially unobservable spatial process that is manifesting itself through spatial dependence in data.

8.3 Spatial Models

The essence of spatial analysis depends on one of the most important premises of geography, also known as "Tobler's First Law of Geography" (Tobler, 1970). It states that "Everything is related to everything else, but near things are more related than distant things." In other words "space matters."

Demographic phenomena are intrinsically spatial since human populations are not randomly placed in space and settlement patterns are dependent on structural attributes of geography. Spatial analysis is necessary for demographic phenomena due to the reasons just explained in the previous section. Therefore, explicit modeling of spatial processes is crucial in any effort to assess "possible" spatial pattern in the demographic phenomena. Although the earlier research tends to ignore the importance of spatial dependence, some recent studies have found spatial dimension of demographic phenomena to be critical in understanding the community characteristics and their impact on the environment (Land et al., 1991; Baller et al., 2001; Işık and Pınarcıoğlu, 2007; Seldadyo et al., 2010; Jones and Kammen, 2014).

ESDA combined with LISA can be a powerful tool to understand the spatial pattern in data, but important issues such as statistical significance and the type of spatial process may substantially bias our results without proper spatial econometric tools. Therefore, the focus of this study is to test the questions as mentioned earlier through a combination of several ESDA and spatial econometric techniques.

Various model specifications for spatial processes have been suggested and empirically implemented in the literature. Manski (1993) shows that three types of spatial interaction effects may explain why an observation associated with a particular location may be dependent on observations at other locations: (1) endogenous interaction effects (e.g., spillovers in dependent variables), where the outcome of an areal unit depends on the outcome of other areal units; (2) exogenous interaction effects (e.g., spillovers in explanatory variables), where the observation of an areal unit depends on the values of other explanatory variables observed by other areal units[74]; and (3) interaction effects among the error terms (e.g., diffusion in unobservables), where similar unobservables result in similar behavior (Elhorst, 2010).

The endogenous interaction effects are appropriate when we believe that the value of dependent variable **Y** in an areal unit i is directly affected by the values of **Y** found in i's neighbors. If we believe that **Y** is not directly influenced, but instead there is a spatially clustered feature that affects the value of **Y** for areal unit i and its neighbors; then, we may consider

[74] If the number of explanatory variables in a linear regression model is K, then the number of exogenous interaction effects is also K, considering the intercept as a separate variable (Elhorst, 2010).

interaction effects among the error terms. Similar to the endogenous interaction effect, the exogenous interaction effects are appropriate when we believe an analogous relationship exists between explanatory variables $\mathbf{X}$.

Spatial models can be constructed by adding these spatial interaction effects either separately or in some combinations into linear regressions. Figure 14 summarizes a taxonomy of linear spatial econometric models with a general–to–specific approach. In Figure 14, Manski Model is on the left–hand side and the non–spatial model (i.e., OLS Model) is on the right–hand side. Each nested spatial model on the right of Manski Model can be obtained by imposing the restrictions specified with the respective arrows.[75]

Some of the models presented in Figure 14 are well known and frequently used in applied research. The most general spatial model is called Manski Model after Manski (1993). It incorporates all types of spatial interaction effects (i.e., spatially lagged dependent variable $\mathbf{WY}$, spatially lagged explanatory variables $\mathbf{WX}$, and spatially lagged error terms $\mathbf{W}\varepsilon$). Lesage and Pace (2008) develop a model with a spatially lagged dependent variable and spatially lagged error terms. It is named as Kelejian–Prucha Model since Kelejian and Prucha (1998) have been the leading advocates of this model. The model with a spatially lagged dependent variable and spatially lagged explanatory variables is introduced by Anselin (1988) and called Spatial Durbin Model (SDM). A model with spatially lagged explanatory variables and spatially lagged error terms is called Spatial Durbin Error Model (SDEM) by Lesage and Pace (2008). The often used and the simplest spatial models are Spatial Autoregressive Model (SAR) and Spatial Error Model (SEM). SAR is a model with a spatially lagged dependent variable whereas SEM is with spatially lagged error terms.

The literature often motivates these spatial models through omitted variable argument either in unobservables or predictors. Spatial heterogeneity for SEM and SDM, and model uncertainty argument for SAR and SDM can also be used to motivate spatial models. See Lesage and Pace (2008) for more information.

8.4 Selected Spatial Models

Figure 14 seems to indicate that the best strategy to test for spatial interactions effects is to start with the most general model (i.e., Manski Model). However, by running a Monte Carlo simulation, Elhorst (2010) shows that endogenous and exogenous effects cannot be distinguished from each other due to identification problem. Thus, spatial dependence parameters cannot be interpreted in a meaningful way in Manski Model.[76]

According to Lesage and Pace (2008), the best option in such circumstances is to omit

[75] Note that Spatially Lagged X Model (SLX) is not considered since it is rarely used in the literature.

[76] Lee (2007) shows that Manski Model is not beyond the bounds of possibility, provided that one employs a different weight matrix for interaction effects among error terms than the endogenous and exogenous interaction effects. However, there is no prior information or theoretical ground in selecting different weight matrices for the same geographical area in this study.

the spatially lagged error terms and employ SDM since the cost of ignoring endogenous and exogenous interaction effects are relatively high. The literature has pointed out that if one or more of the relevant explanatory variables are omitted from a regression, the coefficient estimates for the remaining variables will be biased and inconsistent (Greene, 2003). Hence, if the Kelejian–Prucha Model is taken as the starting point, it will suffer from the omitted variables bias if the true data generating process (DGP) is related to SDM or SDEM. Similarly, if SDEM is taken as the point of departure, it will suffer from the omitted variables bias if the true DGP is related to SAR, Kelejian–Prucha Model, and SDM (Elhorst, 2010). On the other hand, ignoring the interaction effects among error terms will only cause a loss of efficiency, if it exists. Therefore, SDM produces unbiased coefficient estimates even if the true DGP is related to any of the other spatial model specifications presented in Figure 14, except Manski Model. Another advantage of SDM is that it produces correct standard errors or t–values for coefficient estimates even if the true DGP is SEM[77]. It is because SEM is a special case of SDM. There are some other advantages of SDM model which is discussed in Section 8.7.

Since Manski Model has an identification problem and Kelejian–Prucha Model often gives computational errors, this study considers only five of these models: SAR, SEM, SDEM, SDM, and OLS. Starting with OLS, following sections present the selected spatial models briefly. Some of the important details of the spatial models (e.g., impact measures) are discussed after a spatial model is chosen for each base model through sensitivity analysis.

8.4.1 OLS

As it is presented by Eq. 4 in a compact form for all base models, **OLS** for a single base model can be written similarly as in **Eq. 5**.

$$Y = \alpha \iota_n + X\beta + \varepsilon \tag{5}$$

where **Y** is an $n \times 1$ vector of observed dependent variable; α is a scalar represents the constant term parameter; $\iota_\mathbf{n}$ is an $n \times 1$ vector of ones[78]; **X** is an $n \times k$ matrix contains n observations on k explanatory variables; β is a $k \times 1$ vector representing the parameters associated with the k explanatory variables in the predictor matrix **X**; and ε is an $n \times 1$ vector of random errors with $\varepsilon \overset{iid}{\sim} N(0, \sigma^2 I_n)$.

[77]SEM is a special case of classic regression with a non–spherical error variance–covariance matrix in which the off–diagonal elements express the structure of spatial dependence (Elhorst, 2010).

[78]In order to avoid collinearity problems for often used row–standardized weight style, the spatial econometrics literature assumes that the matrix **X** does not contain a constant term, and specifies it separately. This representation is necessary to avoid creating a column vector $W\iota_n = \iota_n$ in **WX** which would duplicate the intercept term (Lesage and Pace, 2008). Therefore, in order to be consistent across models, this representation is followed in all models even if a model does not include matrix **WX**.

8.4.2 SAR

SAR assumes that the spatial pattern is reflected in a regression through dependent variable as an autoregressive process. Adding spatially lagged dependent variable **WY** with autoregressive process into OLS gives

$$Y = \rho WY + \alpha\iota_n + X\beta + \varepsilon \tag{6}$$

and solving Eq. 6 for **Y** gives reduced form **SAR** as shown in **Eq. 7**.

$$Y = (I_n - \rho W)^{-1}\alpha\iota_n + (I_n - \rho W)^{-1}X\beta + (I_n - \rho W)^{-1}\varepsilon \tag{7}$$

where $\mathbf{I_n}$ is an $n \times n$ identity matrix; **W** is an $n \times n$ weight matrix prespecified with a contiguity type–weight style combination; ρ is the spatial dependence parameter, also called spatial autoregressive parameter; and the other notations are as before.

8.4.3 SEM

SEM assumes that the spatial pattern is reflected in a regression through error terms as an autoregressive process. Adding spatially lagged error terms $\mathbf{W}\varepsilon$ with autoregressive process into OLS gives

$$Y = \alpha\iota_n + X\beta + u \tag{8a}$$

$$u = \lambda Wu + \varepsilon \tag{8b}$$

and solving the error specification in Eq. 8b for u gives

$$u = (I_n - \lambda W)^{-1}\varepsilon \tag{9}$$

and substituting Eq. 9 into Eq. 8a gives reduced form **SEM** as shown in **Eq. 10**.

$$Y = \alpha\iota_n + X\beta + (I_n - \lambda W)^{-1}\varepsilon \tag{10}$$

where λ is the spatial dependence parameter, also called spatial autocorrelation; and the other notations are as before.

8.4.4 SDM

SDM assumes that the spatial pattern is reflected in a regression through dependent variable as an autoregressive process and explanatory variables. Adding spatially lagged dependent variable **WY** with autoregressive process and spatially lagged explanatory variables **WX** into OLS gives

$$Y = \rho WY + \alpha\iota_n + X\beta + WX\theta + \varepsilon \tag{11}$$

and solving Eq. 11 for Y gives reduced form **SDM** as shown in **Eq. 12**.

$$Y = (I_n - \rho W)^{-1}\alpha\iota_n + (I_n - \rho W)^{-1}X\beta + (I_n - \rho W)^{-1}WX\theta + (I_n - \rho W)^{-1}\varepsilon \tag{12}$$

where θ is a $k \times 1$ vector representing the spatial spillover parameters associated with the k explanatory variables in the spatially lagged matrix **WX** ; and the other notations are as before.

8.4.5 SDEM

SDEM assumes that the spatial pattern is reflected in a regression through error terms as an autoregressive process and explanatory variables. Adding spatially lagged error terms $\mathbf{W}\varepsilon$ with autoregressive process and spatially lagged explanatory variables **WX** into OLS gives

$$Y = \alpha\iota_n + X\beta + WX\theta + u \tag{13a}$$

$$u = \lambda Wu + \varepsilon \tag{13b}$$

and solving the error specification in Eq. 13b for u gives

$$u = (I_n - \lambda W)^{-1}\varepsilon \tag{14}$$

and substituting Eq. 14 into Eq. 13a gives reduced form **SDEM** as shown in **Eq. 15**.

$$Y = \alpha\iota_n + X\beta + WX\theta + (I_n - \lambda W)^{-1}\varepsilon \tag{15}$$

where the definitions of parameters, matrices, and other notations are as before.

8.5 Estimation Methods for Spatial Models

Least squares methods cannot be used for any of the spatial models presented above since coefficient estimates will be inconsistent in all cases and biased in some cases. For instance, estimating SAR as in Eq. 6 with OLS creates simultaneity between dependent variable **Y** and its spatially lagged values **WY**. Thus, coefficient estimates will be biased and inconsistent. On the other hand, estimating SEM as in Eq. 8a with OLS yields unbiased but inefficient coefficient estimates. As a result, a different estimation method is necessary.

Although various estimation methods for spatial models are proposed by the literature, three methods have been developed extensively and used commonly to estimate spatial models. These techniques are based on the following estimation methods: (1) Maximum Likelihood (ML); (2) instrumental variables or Generalized Method of Moments (IV/GMM); and (3) Bayesian Markov Chain Monte Carlo (MCMC) approach.

The estimation procedure with the best overall performance is ML under the assumption of $\varepsilon \overset{iid}{\sim} N(0, \sigma^2 I_n)$, first outlined in Ord (1975). This procedure has been the most popular and preferred estimation method for over two decades. Thus, ML estimation procedure is used for all spatial models employed in this study. Since the details of the ML estimation for spatial models are complex and beyond the scope of this study, only the most essential parts to mention is discussed. However, an extensive treatment of ML estimation techniques for spatial models can be found in Lesage and Pace (2008).

For the details of ML estimation, consider SDM, for instance. The log–likelihood function for SDM[79] is

$$\ln L = -(n/2)\ln(2\pi\sigma^2) + \ln|I_n - \rho W| - \frac{\varepsilon'\varepsilon}{2\sigma^2} \tag{16a}$$

$$\varepsilon = Y - \rho WY - \alpha\iota_n - X\beta - WX\theta \tag{16b}$$

$$\rho \in (\omega_{\min}^{-1}, \omega_{\max}^{-1})$$

where $\omega_{\min}$ and $\omega_{\max}$ are the smallest and largest eigenvalues of the weight matrix **W** (Lesage and Pace, 2008).

The ML estimation procedure fits spatial regression models by first estimating the spatial dependence parameter (i.e., ρ in SDM) from an explicit maximization of a concentrated likelihood function. Then, the other coefficients are estimated by generalized least squares at that point (Anselin, 1988). Therefore, an estimate of ρ can be found by line search, rather than by optimizing over all of the model parameters at the same time.

Unlike what holds for the classic regression models, the log–likelihood of a spatial model is not equal to the sum of log likelihoods associated with the individual observations (Anselin, 2001). It is due to the two–dimensional nature of the spatial dependence which results in a Jacobian term (i.e., the determinant of a full $n \times n$ matrix $I_n - \rho W$ in SDM). Also, to avoid singularity or explosive processes, the parameter space of ρ is restricted to an interval other than the familiar $[-1, +1]$.

An apparent barrier to calculating Eq. 16a for large n is the $n \times n$ weight matrix **W**. If **W** mainly consists of non–zero elements (i.e., dense matrix), it would require an enormous amount of memory to store this matrix for calculations. According to Lesage and Pace (2008), for instance, a dense weight matrix would require 31.90, 324.80, and 501,659.33 gigabytes of storage for 65,443 tracts, 208,790 block groups, and 8,205,582 blocks respectively in the Census 2000. In contrast, a sparse weight matrix (i.e., a matrix with a significant proportion of zeros) would require less than 0.01, 0.03, and 1.10 gigabytes of storage respectively.

Fortunately, **W** is usually sparse. So, if we could just keep track of non–zero elements, we could save a considerable amount of space. That is what sparse matrix representations do. Instead of keeping track of the entire matrix, they record only the row, column and value of each non–zero element. In the present study, for instance, statistical software needs to keep track

[79]The log–likelihood function for other models can be defined similarly by adjusting Eq. 16b such that $\varepsilon = Y - \rho WY - \alpha\iota_n - X\beta$ for SAR, $\varepsilon = (I_n - \lambda W)(Y - \alpha\iota_n - X\beta)$ for SEM, and $\varepsilon = (I_n - \lambda W)(Y - \alpha\iota_n - X\beta - WX\theta)$ for SDEM.

of only 55, 422 (= 18, 474 × 3) numbers, rather than 9, 665, 881 (= 3109 × 3109) for a weight matrix constructed with 1st Order Queen contiguity, where the number of areal units is 3109. Moreover, calculating matrix–vector products such as **WY** and **WX** take much less time for sparse matrices. For instance, a sparse matrix **W** requires only $\Theta(n)$ operations while a dense matrix **W** would require $\Theta(n^2)$.

Even if the matrix **W** can be represented as a sparse matrix, the $n \times n$ matrix $I_n - \rho W$ is not sparse. However, as it can be seen from Eq. 16a that all we need is the computation of $\ln |I_n - \rho W|$ not the matrix $I_n - \rho W$ itself. Therefore, given a weight matrix **W**, the primary challenge in evaluating the log–likelihood is to compute $\ln |I_n - \rho W|$ which requires $\Theta(n^3)$ operations using dense matrix routines (Lesage and Pace, 2008). For small n, one can compute the log–determinant directly. However, for large n, empirical studies utilize from the fact that the product of the diagonal elements of a triangular or a diagonal matrix yields the determinant. Most of the determinant calculating algorithms exploit this fact by reducing the matrix under consideration to a diagonal or a triangular matrix.

Various approaches can be used to create diagonal or triangular matrices. For instance, by using eigenvalues of matrix **W**, one can compute the determinant of matrix $I_n - \rho W$ which reduces it to a diagonal matrix (Ord, 1975). The advantage of this eigenvalue approach is that the eigenvalues are unique to **W**, so they only need to be computed once. However, it is computationally time–consuming since it requires $\Theta(n^3)$ operations. Also, it becomes numerically unstable and unfeasible for large n (Anselin, 1988; Bivand et al., 2008; Lesage and Pace, 2008).

Alternatively, sparse matrix routines such as LU or Cholesky decompositions reduce the matrix under consideration to a triangular matrix.[80] Although Cholesky decomposition is almost twice as fast as the decomposition, it requires symmetric weight matrix **W** which does not suit to most of our selected weight matrices.[81] As a result, LU decomposition is employed since it does not require symmetric weight matrix **W** and it significantly accelerates the computation of the log–determinant and other quantities of interest.

Although the specification of weight matrix **W** is of vital importance for any spatial model, it cannot be estimated. Thus, it needs to be specified in advance. Moreover, the economic theory underlying the spatial econometric applications often has little to say about the specification of matrix **W** (Leenders, 2002). For these reasons, it has become standard practice to investigate whether the results are robust to the specification of matrix **W**. Fortunately, some weight matrices have known properties which can be used to simplify required calculations in ML estimation (Lesage and Pace, 2008). For instance, due to the nice properties of row–standardized weight style, ρ parameter space is defined as $\omega_{\min}^{-1} < \rho < 1$, where $\omega_{\min}$ is the smallest eigenvalue of

[80] These methods require $n^3/3$ and $2n^3/3$ operations respectively. See Lesage and Pace (2008) for the details of these methods.

[81] The 6NN and 10NN contiguities are not symmetric by construction. Although rest of the contiguities are symmetric, the weight matrix resulting from row standardization is likely to be asymmetric.

matrix **W**. Combining this property with the positive spatial dependence assumption[82] (i.e., $\rho \geq 0$), the interval becomes $\rho \in [0, 1)$ which is a sufficient condition for non–singular $I_n - \rho W$.

In all spatial model estimations, only the row–standardized weight style is considered for several reasons. First, it has some nice computational properties that ease ML estimation. Second, using the row–standardized weight style may yield better results from a statistical perspective (Lesage and Pace, 2008). Third, it increases calculation and interpretability of the impact measures[83] in SAR, SDM, and SDEM since it implies that spatially lagged observed values are averages over the set of neighbors for all areal units. Fourth, it is often preferred over other weight styles and used as the default weight style. Fifth, as it can be seen from Table 3 and Table 18, the Global Moran's I test statistic does not change much across different weight styles given a contiguity type.

After this point, 6NN contiguity is not considered further since Global Moran's I test results slightly differ between 6NN and 10NN contiguities. As a result, the final number of weight styles considered in all spatial models is reduced to four. These are 1st Order Queen, 2nd Order Queen, 10NN and Sphere of Influence contiguities with row–standardized weight style.

8.6 Sensitivity Analysis for Spatial Model Selection

Although the global and local tests for spatial autocorrelation presented in the previous sections suggest a spatial pattern in dependent variables and base model OLS residuals, there is no prior information about the source of the spatial process. In such a situation, many scholars are in two minds about whether to apply the general–to–specific (i.e., from Manski Model to OLS) or the specific–to–general (i.e., from OLS to Manski Model) approach.

Lesage and Pace (2008) argue that SDM is the best starting point of departure (see Section 8.4). However, Florax et al. (2003) show that specific–to–general approach outperforms others in terms of finding the true DGP as well as the accuracy of estimators. Therefore, the standard approach in most empirical studies is to start with a non–spatial linear regression model (i.e., OLS) and then test whether the model needs to be extended with spatial interaction effects as mentioned in Section 8.3.

In specific–to–general approach, sensitivity analysis can be helpful to identify a single best model or to make inferences from a set of multiple competing hypotheses. Up to now, literature has provided several techniques for selecting the best spatial model from a range of model specifications. Anselin (1988) and Anselin et al. (1996a) propose using Lagrange Multiplier (LM) test for OLS versus SAR and SEM. Hoeting et al. (2006) suggest using AIC. Kissling and Carl (2007) develop a model selection procedure and point out the best three spatial model selection criteria: (1) minimum residual spatial autocorrelation (minRSA); (2) maximum model fit (R^2); and (3) AIC. By using a simulation analysis and comparing model parameter estimates

[82] It is essentially Tobler (1970)'s idea of "...near things are more related than distant things".
[83] Direct, indirect, and total impacts. See Section 8.7 for details.

with true values, the authors conclude that using minRSA and AIC jointly gives the best results. On the other hand, Lesage and Pace (2008) suggest using LR tests for nested spatial models as illustrated in Figure 14.

As a result, following the previous literature, the model fitting performance of four spatial model types (i.e., SAR, SEM, SDM, and SDEM) with four weight matrices are evaluated in several steps using LM test, AIC, log likelihood, and LR test for all base models. This procedure allows us to systematically investigate the potential of spatial models with particular types of interaction effects.

The classic LM tests (i.e., LM_ρ and LM_λ) investigate whether SAR or SEM is more suitable to describe the data than OLS (Anselin, 1988). However, robust LM tests by Anselin et al. (1996a) analyze whether SAR is more appropriate than OLS in the possible presence of missing spatially lagged error terms (i.e., LM^*_ρ), and whether SEM is more appropriate than OLS in the possible existence of a missing spatially lagged dependent variable (i.e., LM^*_λ). The latter tests are called robust because the existence of one type of spatial dependence does not bias the test for the other type of spatial dependence (Elhorst, 2010). Both the classic and robust LM tests are based on OLS residuals and follow $\chi^2(1)$.

Table 10 presents the classic and robust LM test results with four different weight matrices for all base models. From the classic LM test results, it is clear that the null hypothesis of no spatial autocorrelation (i.e., OLS) is rejected in favor of SAR and SEM for all base models with any weight matrix. The classic LM test results confirm the statistically significant spatial autocorrelation found in dependent variables and base model OLS residuals by Global Moran's I tests.

In the second step, the robust LM tests are considered since the classic LM tests do not suggest which spatial model should be estimated. Elhorst (2010) states that if the OLS model is rejected in favor of SAR, SEM or both, then SDM model should be estimated since SAR and SEM are the special cases of SDM. As it can be seen from Table 10, OLS is rejected in favor of SAR, SEM or both for all base models with any weight matrix. Therefore, following the literature, it can be concluded that SDM should be estimated for all base models. However, it is important to note that LM tests do not provide any suggestion in the selection of weight matrix.

Since these spatial models are estimated by ML, and some of them are nested in more general ones, LR tests for nested models can be used. The LR test initially used to reconfirm the classic LM test and Global Moran's I results since OLS is nested in all four spatial models. Table 11 shows that using any of the four spatial models increases the model fit since all LR test statistics are statistically significant at the 1% significance level. Therefore, OLS is rejected once again.

The LR test can be used once again to test nested spatial models only. Specifically speaking, using the restrictions presented in Figure 14, it can be shown that SDM nests SAR and SEM, and SDEM nests only SEM. Table 12 shows the LR test results for these nested models for all base models with any weight matrix. From the table, it is clear that in almost all cases SDM is preferred to SAR and SEM whereas SDEM is preferable to SEM. It means that for all base models there is an exogenous interaction effect; and thus, spatially lagged explanatory variables **WX** should be added to every base model. However, the remaining puzzles are whether to add spatially lagged dependent variable **WY** (i.e., using SDM) or to add spatially lagged error terms $\mathbf{W}\varepsilon$ (i.e., using SDEM), and which weight matrix to choose.

Finally, to give an ultimate decision for spatial model and weight matrix selection for each base model, AIC and log likelihood are employed to compare model fitting. Table 13 and Table 14 present AIC and log likelihood results respectively which coincide precisely. These results are used to give final decision of spatial model selection for all base models.

By combining the results of LM tests, LR tests, AIC, and log likelihood the following choices are made: (1) SDM with 1st Order Queen contiguity and row–standardized weight style for the base models $\mathrm{NH_3}^{\dagger,b}$, $\mathrm{PM_{10}}^{\dagger,c}$, $\mathrm{PM_{2.5}}^{\dagger,c}$, and VOC^{b}; (2) SDEM with 1st Order Queen contiguity and row–standardized weight style for the base models CO^{c} and $\mathrm{NO}_x{}^{\dagger,b}$; and (3) SDM with 2nd Order Queen contiguity and row–standardized weight style for only $\mathrm{SO_2}^{\dagger,b}$. After this point, the base models are labeled as $\mathrm{CO}^{c,\circ,q1}$, $\mathrm{NH_3}^{\dagger,b,\bullet,q1}$, $\mathrm{NO}_x{}^{\dagger,b,\circ,q1}$, $\mathrm{PM_{10}}^{\dagger,c,\bullet,q1}$, $\mathrm{PM_{2.5}}^{\dagger,b,\bullet,q1}$, $\mathrm{SO_2}^{\dagger,b,\bullet,q2}$, and $\mathrm{VOC}^{b,\bullet,q1}$, where $^{\circ}$, $^{\bullet}$, ^{q1}, and ^{q2} indicate a base model estimated with SDEM, SDM, 1st Order Queen contiguity, and 2nd Order Queen contiguity respectively.

8.7 Impact Measures

One of the significant advantages of spatial models is their ability to quantify spatial spillovers. A loose definition of spillovers in a spatial context is that changes in an areal unit exert impacts on other areal units. For a researcher, it would be of interest to investigate the presence, magnitude, and extent of spatial spillovers. Therefore, this section briefly covers spatial spillovers only in the selected models (i.e., SDM and SDEM). For comprehensive coverage of all spatial models refer to Lesage and Pace (2008).

Most of the previous empirical studies have ignored the spillover feature of spatial models and misinterpreted parameter estimates by using conventional interpretation methods. Linear regression parameters have a straightforward interpretation which arises from linearity and assumed independence of observations in a model. For instance, coefficients can be interpreted using partial derivatives such as

$$\frac{\partial y_i}{\partial x_{ir}} = \beta_r \quad \text{for all } i \text{ and } r$$

and due to the independence assumption, the cross–partial derivatives are

$$\frac{\partial y_i}{\partial x_{jr}} = 0 \quad \text{for } i \neq j \text{ and all } r$$

where i and j represent indices of the n observations and r represents an index to one of the k explanatory variables ($i, j = 1, \ldots, n$ and $r = 1, \ldots, k$). However, incorporating spatially lagged dependent variable **WY** and explanatory variables **WX** into a linear regression creates spatial spillovers between observations. Thus, it violates the independence assumption and makes the cross–partial derivatives non–zero depending on the spatial model structure. It means that changes to explanatory variables in region i impact the dependent variable values in region j, given that $i \neq j$. Thus, the conventional least–squares coefficient interpretation is not valid, and special methods are necessary.

To illustrate spillovers in spatial models, consider the reduced form SDM in Eq. 12 expressed as

$$Y = (I_n - \rho W)^{-1}\alpha\iota_n + (I_n - \rho W)^{-1}X\beta + (I_n - \rho W)^{-1}WX\theta + (I_n - \rho W)^{-1}\varepsilon$$

To explore the matrix of partial derivatives in details, the inverse matrices in reduced form SDM can be rewritten as the following infinite series

$$(I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots \tag{17}$$

where we assume that $-1 < \rho < 1$ which is satisfied by row–standardized weight style.[84] Hence, $n \times n$ partial derivatives matrix of Y with respect to the r^{th} explanatory variable of X for areal unit one up to n (x_{ir} for $i = 1, \dots, n$, respectively) can be defined as

$$\begin{aligned} \begin{bmatrix} \frac{\partial Y}{\partial x_{1r}} & \dots & \frac{\partial Y}{\partial x_{nr}} \end{bmatrix} &= \begin{bmatrix} \frac{\partial y_1}{\partial x_{1r}} & \dots & \frac{\partial y_1}{\partial x_{nr}} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_{1r}} & \dots & \frac{\partial y_n}{\partial x_{nr}} \end{bmatrix} \\ &= (I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots) \begin{bmatrix} \beta_r & w_{12}\theta_r & \dots & w_{1n}\theta_r \\ w_{21}\theta_r & \beta_r & \dots & w_{2n}\theta_r \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1}\theta_r & w_{n2}\theta_r & \dots & \beta_r \end{bmatrix} \end{aligned} \tag{18}$$

where w_{ij} is the $(i, j)^{\text{th}}$ element of weight matrix **W**.

Due to the contiguity type imposed on the weight matrix **W**, it is clear that at least some of the off–diagonal elements in the last term of Eq. 18 are non–zero. Therefore, even without considering the term $(I_n - \rho W)^{-1}$, it can be said that some cross–partial derivatives are non–zero and local spillovers exist. That means, for instance, if x_{ir} changes, not only y_i changes (i.e., due to respective non–zero diagonal element) but also y_j changes (i.e., due to the respective non–zero off–diagonal elements) given that $w_{ij} \neq 0$. Local spillovers represent a situation where the impacts fall only on nearby or immediate neighbors which is dying out before they impact areal units that are neighbors to the neighbors.

The most striking result is revealed if we multiply the last term of Eq. 18 with the infinite series. As mentioned before, the matrix **W** is sparse, and its diagonal elements are zero by

[84] Our selected weight matrices are all row–standardized. See Section 8.5 for details.

construction. However, higher orders such as matrix $\mathbf{W^2}$ (i.e., *second–order* neighbor of the neighbor) become dense, and the diagonal elements will be positive given that each areal unit has at least one neighbor. It is because the neighbor of a neighbor to an areal unit i includes areal unit i itself. The final result of Eq. 18 is complex and cannot be generalized. Thus, following Lesage and Pace (2008), the partial derivatives are denoted as $n \times n$ matrix $\mathbf{S_r}(\mathbf{W})$ for any r^{th} explanatory variable.

$$S_r(W) = \begin{bmatrix} S_r(W)_{11} & S_r(W)_{12} & \dots & S_r(W)_{1n} \\ S_r(W)_{21} & S_r(W)_{22} & \dots & S_r(W)_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_r(W)_{n1} & S_r(W)_{n2} & \dots & S_r(W)_{nn} \end{bmatrix} \tag{19}$$

As a result, Eq. 19 suggests that if x_{ir} changes, not only y_i changes but also y_j changes for all i and j. It is called global spillovers which arise due to the infinite series expansion of $(I_n - \rho W)^{-1}$. Global spillovers have a decay of influence for higher–order neighbors.

It is also the case that own partial derivative for the i^{th} areal unit (i.e., $S_r(W)_{ii}$ which measures the impact on y_i from a change in x_{ir}) usually does not equal to β_r as in OLS.[85] This impact includes the *feedback effects* (e.g., x_{ir} affects x_{jr}, and x_{jr} affects back x_{ir}) as well as longer paths (e.g., impacts go from x_{ir} to x_{jr} to x_{lr} and back to x_{ir}). Moreover, it is important to note that *global diffusion* exists in the error terms due to the infinite series expansion of $(I_n - \rho W)^{-1}$.

The partial derivatives matrix $\mathbf{S_r}(\mathbf{W})$ for SDEM can be constructed in the same manner. Since there is not a term such as $(I_n - \lambda W)^{-1}$ in front of the explanatory variables as in SDM, only local spillovers exist in SDEM (see Eq. 15). Note that a model with only local spillovers exhibits no feedback effects. SDEM also shows global diffusion in error terms as in SDM.

In matrix $\mathbf{S_r}(\mathbf{W})$, it is important to note that every diagonal element represents a *direct impact* and every off–diagonal element represents an *indirect impact*. Moreover, direct and indirect impacts are different for the various areal units in the sample. That means, if we have n areal units and k explanatory variables, we obtain k different $n \times n$ matrices of direct and indirect impacts. Therefore, presentation of these impacts is a serious problem.

To analyze the significant amount of information pertaining to the partial derivatives from these models, Lesage and Pace (2008) propose to report three different impact measures by using elements of matrix $\mathbf{S_r}(\mathbf{W})$: (1) *direct impact* measured by the average of diagonal elements; (2) *total impact* measured by the average of either row sums or column sums of all elements; and (3) *indirect impact* measured by the average of either row sums or column sums of off–diagonal elements, which is simply the difference between the *total impact* and *direct impact*. Since the indirect impact represents spatial spillovers in the data, sometimes it is used interchangeably in

[85] With three explanatory variables, Elhorst (2010) shows that all elements in matrix $\mathbf{S_r}(\mathbf{W})$ depends on the structure of weight matrix $\mathbf{W}$, β_r, θ_r, and ρ.

the literature. These impacts are given as

$$\bar{M}(r)_{direct} = n^{-1} tr(S_r(W)) \tag{20a}$$

$$\bar{M}(r)_{total} = n^{-1} \iota_n' S_r(W) \iota_n \tag{20b}$$

$$\bar{M}(r)_{indirect} = \bar{M}(r)_{total} - \bar{M}(r)_{direct} \tag{20c}$$

where ι_n is an $n \times 1$ vector of ones, and

$$S_r(W) = V(W)(I_n\beta_r + W\theta_r) \qquad \text{(for SDM)}$$

$$S_r(W) = (I_n\beta_r + W\theta_r) \qquad \text{(for SDEM)}$$

$$V(W) = (I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \ldots$$

Lesage and Pace (2008) also separate the total impact into two different impacts. The sum across the i^{th} row of matrix $\mathbf{S_r(W)}$ represents the total impact on individual observation y_i resulting from changing the r^{th} explanatory variable by the same amount across all n observations. There are n of these sums given the column vector $S_r(W)\iota_n$. Thus, an average of these total impacts is called *average total impact to an observation*. Summing down the j^{th} column of matrix $\mathbf{S_r(W)}$ gives the total impact over all y_i from changing the r^{th} explanatory variable by an amount in the j^{th} observation. There are n of these sums given the row vector $\iota_n' S_r(W)$. Thus, an average of these total impacts is called *average total impact from an observation*.

The estimated direct, indirect and total impacts of explanatory variables should eventually be used in hypothesis test for inference. However, one difficulty is that these impacts consist of different coefficient estimates according to complex mathematical formula, and the dispersion of these impacts depends on the dispersion of all coefficient estimates involved (e.g., see Eq. 18). For instance, consider the partial derivatives of Y with respect to the r^{th} explanatory variable for SDM as shown in Eq. 18. If the coefficient estimates of ρ, β_r, and θ_r happen to be significant, this does not automatically guarantee that these impacts of the r^{th} explanatory variable are also significant (Elhorst, 2010).

In order to draw inferences regarding the statistical significance of these impacts, Lesage (1997) suggests simulating the distribution of these impacts using the variance–covariance matrix implied by the ML estimates.[86] In the present study, the distribution of impacts are simulated 10000 times to calculate standard deviations, t–values, and p–values.[87]

[86] Specifically, a Bayesian Markov Chain Monte Carlo (MCMC) estimation method is used. It provides a large number of draws for the model parameters.

[87] Since calculating the exact measure of the matrix $(I_n - \lambda W)^{-1}$ is computationally inefficient, traces of the powers of matrix $\mathbf{W}$ is used to approximate the infinite series expansion shown in Eq. 17.

9 Results and Discussion

Table 15 presents the direct impact estimates of explanatory variables along with some diagnostic tests and model fitting measures for all models.[88] Table 16 and Table 17 show the indirect and total impact estimates respectively.

Leaving the coefficient estimates aside for a moment, what stands out in Table 15 is that all models suggest a statistically significant strong positive spatial autocorrelation at the 1% significance level. The highest spatial autocorrelation is observed in model $PM_{10}{}^{\dagger,c,\bullet,q1}$ with 0.794 and the lowest in model $SO_2{}^{\dagger,b,\bullet,q1}$ with 0.441. To further confirm these results, LR diagnostic test for the absence of spatial dependence is performed. The LR tests yield the same results and confirm the statistically significant spatial autocorrelation for all models. Moreover, to make sure there is no spatial autocorrelation left off in the residuals, Global Moran's I test is performed. In all models, the test statistic is close to zero and non–significant at any conventional levels of significance.

While interpreting the coefficient estimates one should be careful since some explanatory variables are in the natural logarithmic form and some are in raw form. For instance, in model $CO^{c,\circ,q1}$, the direct impact estimate of *log population density* should be interpreted as "while holding other variables fixed, 1% increase in *population density* is expected to change CO emissions by approximately -0.483%." It is the case when both the dependent and explanatory variables are in logarithmic form. However, the interpretation changes if the dependent variable is in the natural logarithmic form and the explanatory variable is not. For instance, in model $NH_3{}^{\dagger,b,\bullet,q1}$, the direct impact estimate of *Muslims* should be interpreted as "while holding other variables fixed, one percentage point increase in *Muslims* is expected to change NH_3 emissions by approximately -4.1% ($= 100\% \cdot -0.041$)." For the details of these approximations, see Appendix H.

In interpreting impact measures, the focus is on understanding which characteristics of areal units, specifically religion variables, exert statistically significant impacts on emissions. Therefore, statistically non–significant impact estimates are ignored in general. The present study uses emissions as a quantitative measure and a proxy for environmental performance in the interest of area although these two terms are negatively related. Thus, extra care is needed while interpreting the results.

9.1 Direct Impacts

First, the influence of explanatory variables is measured using direct impacts. Specifically, direct impacts reflect the average change in emissions within an areal unit from a change in an explanatory variable in the own areal unit. Table 15 shows the direct impact estimates.

[88] Impact estimates of intercept are omitted in all models due to the construction of spatial models with row–standardized weight style. See Section 8.4.1 for more information.

The direct impact estimates of *log income* are positive and statistically significant in all models. This outcome is unexpected especially for the models with only *log income* (i.e., $CO^{c,\circ,q1}$ and $VOC^{b,\bullet,q1}$). This result is contrary to the findings of Esty and Porter (2001) although the estimates are negligibly low. On the other hand, the direct impact estimates of *log income squared* are negative and statistically significant in all models. It suggests that for low–level income, pollution increases as income increases; however, after a point of high–level income, pollution decreases as income increases. The positive estimates of *log income* and the negative estimates of *log income squared* confirm the findings of Grossman and Krueger (1994) and our prior expectation (i.e., inverted U–shape relationship between *income* and pollution as suggested by Environmental Kuznets Curve theory).[89] Figure 18 presents the Environmental Kuznets Curves by the dependent variable of each spatial model, where *log income squared* is employed as an explanatory variable. Plots suggest that there is an inverted U–shape relationship between *income* and the pollution variable of each spatial model. As a result, we conclude that there is a U–shape relationship between *income* and environmental performance.

From Table 15, it can be seen that the direct impact estimates of *gas price* are negative and statistically significant in most models, except being positive in model ${PM_{10}}^{\dagger,c,\bullet,q1}$. Similarly, direct impact estimates of *gas tax/fee* are negative and statistically significant in almost all models, though the magnitudes are trivial compared to the estimates of *gas price*. When the robustness and magnitudes of our results are considered, it seems that an increase in *gas tax/fee* is effective in reducing most emission types, but it is a less efficient policy tool compared to an increase in *gas price*. As a result, we conclude that *gas price* and *gas tax/fee* are positively related to environmental performance. Thus, we confirm the findings of Esty and Porter (2001). On the other hand, *renewable energy consumption* seems to be ineffective in changing environmental performance since the direct impact estimates are mostly non–significant. It is an unexpected outcome since the usage of renewable energy sources is considered to be effective in reducing emissions as pointed out by Ohlan (2015).

As indicated before, the education variable used in each model varies. The education variable exhibits statistically significant direct impacts only where *bachelor's degree or more* is used as an explanatory variable. The non–significant estimates of *some college or more* might seem an opposing result when compared to the negative and significant estimates of *bachelor's degree or more*. However, it suggests a richer result. Since *bachelor's degree or more* is nested in *some college or more*, we can say education is effective in reducing emissions only after a certain level which is *bachelor's degree or more* according to our analysis. The highest direct impact estimate is in model ${SO_2}^{\dagger,b,\bullet,q2}$ where one percentage point increase in *bachelor's degree or more* leads to 2.2% decrease in emissions whereas the lowest is in model ${NH_3}^{\dagger,b,\bullet,q1}$ with 1% reduction. As a result, the direct impact estimates of education are as expected and we confirm

[89] In models with *log income squared*, one can calculate the approximate direct impact estimates of *log income* by using Eq. 30. Using the lower quantile of *log income* and applying Eq. 30 gives that 1% increase in *income* leads to −0.117%, 0.9%, 0.154%, 0.272%, and 1.68% change in emissions in models ${NH_3}^{\dagger,b,\bullet,q1}$, ${NO_x}^{\dagger,b,\circ,q1}$, ${PM_{10}}^{\dagger,c,\bullet,q1}$, ${PM_{2.5}}^{\dagger,c,\bullet,q1}$, and ${SO_2}^{\dagger,b,\bullet,q2}$ respectively. See Appendix I.2 for more information.

the previous literature that the better–educated individuals express pro–environment stance and increase environmental performance.

In all models, the direct impact estimates of *log population density* are negative and statistically significant at the 1% significance level. The change in emissions from 1% increase in *population density* varies from −0.743% in model $NH_3^{\dagger,b,\bullet,q1}$ to −0.251% in model $SO_2^{\dagger,b,\bullet,q2}$. Consequently, our results confirm the findings of Makido et al. (2012) and Gately et al. (2015) by providing spatial econometric evidence for a positive association between *population density* and environmental performance.

The weather events seem to be related to emissions as suggested by (EPA, 2011b). The direct impact estimates of *log mean daily precipitation* are positive and statistically significant only in models $CO^{c,\circ,q1}$, $PM_{2.5}^{\dagger,c,\bullet,q1}$, and $VOC^{b,\bullet,q1}$. Although the magnitudes of these estimates are negligibly low, we confirm that relative precipitation is associated with low environmental performance. The direct impact estimates of *log mean daily maximum heat index* indicate conflicting results across models. Specifically, the estimates are positive and statistically significant in models $NH_3^{\dagger,b,\bullet,q1}$ and $NO_x^{\dagger,b,\circ,q1}$ but negative in model $CO^{c,\circ,q1}$. These conflicting results might arise due to the composition of different CAPs and how they react within various climates.

Note that we do not have any prior expectations about the sign and magnitude of religion variable estimates. In fact, that is the question we try to answer. Generally speaking, the direct impact estimates of religion variables on emissions can be separated into three groups. In the first group, *Evangelical Protestants*, *Black Protestants*, *Mainline Protestants*, and *Orthodox Christians* exhibit positive and statistically significant estimates in more than half of the models. Among the Protestant religions, *Black Protestants* shows the highest positive direct impact estimates in model $SO_2^{\dagger,b,\bullet,q2}$ with 4.2% and *Evangelical Protestants* shows the lowest in model $PM_{10}^{\dagger,c,\bullet,q1}$ with 0.1%. Compared to the Protestant groups, *Orthodox Christians* consistently exhibits the largest direct impacts on emissions. For instance, in models $CO^{c,\circ,q1}$, $SO_2^{\dagger,b,\bullet,q2}$, and $VOC^{b,\bullet,q1}$, one percentage point increase in *Orthodox Christians* leads to 11.9%, 25.5%, and 11.8% increase in emissions respectively.

In the second group, *Catholics*, *Mormons*, and *Jews* show statistically significant positive direct impact estimates in less than half of the models. The magnitudes of impacts for *Jews* are substantial compared to other two religions. For instance, one percentage point rise in *Catholics* and *Mormons* are expected to increase emissions up to 0.7% in models $SO_2^{\dagger,b,\bullet,q2}$ and $NH_3^{\dagger,b,\bullet,q1}$ respectively whereas the same change in *Jews* leads to 3.9% increase in model $PM_{10}^{\dagger,c,\bullet,q1}$.

In the third group, *Muslims* and *Buddhists* are the only religious groups which show negative and statistically significant direct impact estimates. However, the results are not robust across models. Specifically, *Buddhists* exhibits the highest decrease in emissions across all religions and models (e.g., one percentage point increase in *Buddhists* would lead to a −4.6% change in emissions in model $PM_{10}^{\dagger,c,\bullet,q1}$). On the other hand, *Muslims* shows negative direct impact estimates in models $NH_3^{\dagger,b,\bullet,q1}$ and $PM_{10}^{\dagger,c,\bullet,q1}$, but positive direct impacts in models $CO^{c,\circ,q1}$

and $VOC^{b,\bullet,q1}$.

What stands out in Table 15 is that most of the direct impact estimates of religion variables pertaining to Judeo–Christian beliefs are positive and statistically significant. It seems that this result is robust across models for *Evangelical Protestants*, *Black Protestants*, *Mainline Protestants*, *Orthodox Christians*, and *Jews*. As a result, in direct impacts, there appears to be a negative association between Judeo–Christian beliefs and environmental performance.

9.2 Indirect Impacts

Second, the influence of explanatory variables is measured using indirect impacts which capture the average change in emissions of neighboring areal units (i.e., spatial spillovers) from a change in an explanatory variable of an areal unit. Table 16 presents the indirect impact estimates which help us to identify factors producing the largest and lowest spillovers. Negative indirect impacts could be considered as spatial benefits (i.e., positive externality) since they indicate impacts on neighboring areal units that lead to an increase in the environmental performance. On the other hand, positive indirect impacts would represent negative externality since they indicate that neighboring areal units suffer from a decrease in environmental performance.

As shown in Table 16, indirect impact estimates of *log income* and *log income squared* are statistically significant only in models $PM_{10}{}^{\dagger,c,\bullet,q1}$ and $VOC^{b,\bullet,q1}$. In these models, *income* exhibits negative externality on neighboring areal units. In general, these results suggest that spatial spillovers arising from *income* are not robust across models as it is the case in direct impacts. Therefore, it can be said that *income* does not exhibit robust indirect impacts on environmental performance.

The results show that *gas price* and *gas tax/fee* do not exhibit any impacts on the environmental performance of neighboring areal units since none of the indirect impact estimates are statistically significant. On the other hand, the indirect impact estimates of *renewable energy consumption* are positive and statistically significant in three models, though it is negative in model $SO_2{}^{\dagger,b,\bullet,q2}$. Although the magnitudes of these estimates are negligibly small, it is an unexpected outcome since the usage of renewable energy sources is considered to be effective in emission reduction, not accession.

The indirect impact estimates of *some college or more* are positive and statistically significant in models where it is used as an education variable. However, the estimates of *bachelor's degree or more* are negative and statistically significant, except in model $NH_3{}^{\dagger,b,\bullet,q1}$. These results might seem conflicting. However, a positive indirect impact for *some college or more* and a negative one for *bachelor's degree or more* tell the same story. The puzzle can be resolved if one considers the fact that an increase in *some college or more* leads to a relative decrease in *bachelor's degree or more* in most cases.[90] As a result, education exhibits negative and statistically significant

[90]Also, note the conclusion made in Section 9.1. That is, education is effective in reducing emissions only after a certain level (i.e., *bachelor's degree or more*).

indirect impacts (i.e., positive externality) on emissions of neighboring areal units. Therefore, we conclude that the better–educated individuals have positive impacts on the environmental performance of neighboring areal units.

Although the magnitudes are negligibly low, the indirect impact estimates of *log population density* are positive and statistically significant in most models. On the contrary, the estimate in model $SO_2^{\dagger,b,\bullet,q2}$ is negative and statistically significant. The possible reason for mixed results could be the source of emission types and how they are related to human activities. For instance, NO_x is mainly emitted from combustion in vehicles and power plants (Clark et al., 2014). One possible explanation of positive indirect impact estimates in most models might be the traffic flow between metropolitan areas (i.e., cities with high population density) and surrounding areas. Metropolitan areas often attract workers from surrounding areas which in turn causes an increase in vehicle usage and traffic emissions not only in metropolitan areas but also in the surrounding areas. However, traffic emissions increase more in the surrounding areas since medium and low–density cities do not have as effective public transportation or walking accessibility as metropolitan areas (Gately et al., 2015). As a result, we conclude that *population density* exhibits negative indirect impacts on environmental performance.

From Table 16, it can be seen that the indirect impact estimates of *log mean daily precipitation* is negative and statistically significant in most models, though the magnitudes of these estimates are not large across models. On the other hand, *log mean daily maximum heat index* exhibits positive and statistically significant indirect impact estimates, expect in model $NO_x^{\dagger,b,\circ,q1}$. Thus, it can be said that an increase in *log mean daily precipitation* and *mean daily maximum heat index* in the own areal unit lead to a reduction and rise respectively in environmental performance of neighboring areal units.

Among the religion variables, *Mainline Protestants*, *Catholics*, and *Orthodox Christians* exhibit the most interesting indirect impact estimates. Although the impacts of *Mainline Protestants* are statistically significant in most models, the signs are mixed. Specifically, *Mainline Protestants* exhibits positive externality in models $CO^{c,\circ,q1}$, $SO_2^{\dagger,b,\bullet,q2}$, and $VOC^{b,\bullet,q1}$ whereas negative externality in model $NH_3^{\dagger,b,\bullet,q1}$. The indirect impact estimates of *Catholics* are positive in all models and statistically significant in most models. The results are almost robust across models although the magnitudes of these estimates are small (e.g., the highest change in emissions of neighboring areal units is 1.2% in model $NH_3^{\dagger,b,\bullet,q1}$ from one percentage point increase in *Catholics* in the own areal unit). The most striking result for religion variables in Table 16 is that *Orthodox Christians* shows an extreme negative externality in model $SO_2^{\dagger,b,\bullet,q2}$ (i.e., one percentage point increase in *Orthodox Christians* in the own areal unit leads to 296.2% increase in emissions of neighboring areal units). Unfortunately, this is the only statistically significant result across models; and thus, making a general conclusion for *Orthodox Christians* would be invalid. As a result, we conclude that the only religious body that has spatial spillovers across all models is *Catholics*, which exhibits negative indirect impacts on environmental performance.

9.3 Total Impacts

Third, the influence of explanatory variables is measured using total impact estimates as given in Table 17. Since total impacts take both direct and indirect impacts into account, they allow us to draw an inference regarding what factors are important in affecting emissions in general. When considering total impacts, we should keep in mind that they reflect average impacts on the own areal unit plus average impacts on neighboring areal units. For instance, in model $CO^{c,\circ,q1}$, 1% increase in *mean daily precipitation* would lead to 0.174% increase in emissions in the own and neighboring areal units in average (i.e., increase in emissions in all areal units).

As shown in Table 17, the total impact estimates of *log income* are positive and statistically significant in models $PM_{10}{}^{\dagger,c,\bullet,q1}$, $PM_{2.5}{}^{\dagger,c,\bullet,q1}$, $SO_2{}^{\dagger,b,\bullet,q2}$, and $VOC^{b,\bullet,q1}$. However, estimates of *log income squared* are negative and statistically significant in models $PM_{10}{}^{\dagger,c,\bullet,q1}$, $PM_{2.5}{}^{\dagger,c,\bullet,q1}$, and $SO_2{}^{\dagger,b,\bullet,q2}$. These results indicate an inverted U–shape relationship between *income* and pollution as suggested by Environmental Kuznets Curve theory. However, this relationship can be observed for only some emission types. As a result, in total impacts, we conclude that there is a U–shape relationship between *income* and environmental performance.

From Table 17, it can be seen that the total impact estimates of *gas price* are negative and statistically significant in only three models whereas the estimates of *gas tax/fee* are negative and statistically significant in almost all models. The magnitudes and robustness of our results point out that increasing *gas tax/fee* seems to be effective in reducing most emission types; however, it is a less efficient policy tool compared to an increase in *gas price*. As a result, we conclude that *gas price* and *gas tax/fee* are positively associated with environmental performance, and confirm the findings of Esty and Porter (2001). On the other hand, the total impact estimates of *renewable energy consumption* are positive and statistically significant in four models, though it is negative in model $SO_2{}^{\dagger,b,\bullet,q2}$. Although the magnitudes of these estimates are negligibly small, it is an unexpected outcome since the usage of renewable energy sources is considered to be effective in emission reduction as pointed by Ohlan (2015).

The total impact estimates of education variables exhibit the similar pattern as seen in the direct and indirect impacts. All total impacts estimates of *some college or more* are positive and statistically significant. However, estimates of *bachelor's degree or more* are negative and statistically significant. These results confirm our previous claims that an increase in *some college or more* leads to a relative decrease in *bachelor's degree or more* in most cases, and education is effective in decreasing emissions after a certain level of education, which is *bachelor's degree or more*. As a result, education exhibits negative and statistically significant total impacts on emissions, except $NH_3{}^{\dagger,b,\bullet,q1}$ model. Therefore, we conclude that the better–educated individuals have positive total impacts on environmental performance.

The total impact estimates of *log population density* are negative and statistically significant in all models. This outcome confirms the findings of Makido et al. (2012) and Gately et al. (2015) once again. As a result, we conclude that *population density* exhibits positive total impacts on environmental performance.

From Table 17, it can be seen that the total impact estimates of *log mean daily precipitation* is negative and statistically significant in four models, except it is positive in models $\mathrm{CO}^{c,\circ,q1}$ and $\mathrm{SO_2}^{\dagger,b,\bullet,q2}$. The total impact estimates of *log mean daily maximum heat index* is positive and statistically significant in four models. As a result, our findings indicate that *mean daily precipitation* has positive total impacts on environmental performance whereas *mean daily maximum heat index* has negative total impacts.

Among the variables pertaining to religion, *Catholics*, *Orthodox Christians*, *Hindus*, and *Buddhists* exhibit the most interesting total impact estimates. The only religious body which shows positive total impact estimates in all models is *Catholics*. These estimates are statistically significant in most models, and the magnitudes are ranging from 1.6% in model $\mathrm{NH_3}^{\dagger,b,\bullet,q1}$ to 0.7% in models $\mathrm{PM_{2.5}}^{\dagger,c,\bullet,q1}$ and $\mathrm{VOC}^{b,\bullet,q1}$. Although the magnitudes of these estimates are small, the results are almost robust across models. As it is the case in the indirect impacts, *Orthodox Christians* shows an extremely large total impact estimate in model $\mathrm{SO_2}^{\dagger,b,\bullet,q2}$ (i.e., one percentage point increase in *Orthodox Christians* leads to 321.7% increase in emissions). Once again, this is the only statistically significant result across models; and thus, making a general conclusion for *Orthodox Christians* would be invalid. The only two religious bodies which exhibit negative total impact estimates in all models are *Hindus* and *Buddhists*. However, they are statistically significant only in one model for *Hindus* and two models for *Buddhists*, and the magnitudes are large. Specifically, one percentage point increase in *Hindus* leads to 106.1% decrease in emissions in model $\mathrm{SO_2}^{\dagger,b,\bullet,q2}$ and the same change in *Buddhists* leads to 17.1% and 52.6% decrease in models $\mathrm{NO}_x{}^{\dagger,b,\circ,q1}$ and $\mathrm{PM_{10}}^{\dagger,c,\bullet,q1}$ respectively. As a result, we conclude that the only religious body that indicates negative total impacts on environmental performance across all models is *Catholics*. On the other hand, *Hindus* and *Buddhists* are the only two religious bodies that have positive total impacts on environmental performance, though the significance of estimates is not robust across models.

Finally, Appendix I presents some ESDA thematic mapping for all base model spatial regressions used as a final model in interpreting.

10 Conclusion

As stated previously, the primary question that this study tries to answer is: “Does religion have any impact on the environmental performance of a county after controlling for the other important determinants of environmental performance?” Although the direct, indirect and total impact estimates of a given religious body are often not robust across all models, our results suggest clear conclusions about the impacts of some religious bodies on environmental performance. An important point is that our results does not suggest causality between religious bodies and environmental performance but instead suggests an association.

The thorough spatial analyses reveal that most of the religious bodies exhibit negative direct impacts on environmental performance. Among those, *Evangelical Protestants*, *Black Protestants*, *Mainline Protestants*, *Orthodox Christians*, and *Jews* have robust direct impacts across all models. The magnitudes of direct impact estimates of *Orthodox Christians* are large compared to all Protestant groups and *Jews*. Therefore, we argue that there appears to be a negative direct association between Judeo–Christian beliefs and environmental performance. Moreover, this negative association gains strength in the following order: *Evangelical Protestants*, *Mainline Protestants*, *Black Protestants*, *Jews* and *Orthodox Christians*.

For the indirect impacts, only *Catholics* shows statistically significant and robust impacts across all models. *Catholics* in the own areal unit appears to be negatively associated with environmental performance on neighboring areal units, though the magnitudes of these impacts are generally less than 1%.

If both the direct and indirect impacts are considered together (i.e., total impacts), once again *Catholics* is the only religious body that appears to be negatively associated with environmental performance across all models. On the other hand, *Hindus* and *Buddhists* are the only two religious bodies that seem to have a strong positive association with environmental performance. These associations are not robust across models; and thus, the validity of our results are still in question.

To sum up, taking into account the strong spatial autocorrelation in all variables suggested by formal tests and ESDA, our spatial regression analyses show that religion appears to be associated with environmental performance. Specifically, religions pertaining to Judeo-Christian beliefs such as *Evangelical Protestants*, *Black Protestants*, *Mainline Protestants*, *Catholics*, *Orthodox Christians*, and *Jews* exhibit a negative association with environmental performance, whereas *Hindus* and *Buddhists* are the only ones that show a positive association. Therefore, we find strong support for White's Thesis by providing spatial econometric evidence.

Table 1: Data and Shapefile Description Summary

Variable	Description	Source	# of Obs.
Environmental Performance	Emissions measured in tons per capita in 2011. The selected variable names are *CO*, NH_3, NO_x, PM_{10}, $PM_{2.5}$, SO_2 and *VOC*.	EPA (2011a,b)	3237
Religion	Adherents to religions as a percentage of total population in 2010. The selected variable names are *Adherents*, *Protestants*, *Evangelical Protestants*, *Black Protestants*, *Mainline Protestants*, *Catholics*, *Orthodox Christians*, *Mormons*, *Muslims*, *Jews*, *Hindus*, and *Buddhists*.	ARDA (2010) and Grammich et al. (2012)	3149
Income	Income per capita measured in 2010 prices. The selected variable name is *income*.	Census Bureau (2010a) and Social Explorer (2010a)	3221
Gas Price	Gas prices excluding taxes and fees per gallon measured in 2010 prices. The selected variable name is *gas price*.	EIA (2010a)	51
Gas Tax/Fee	Total gas taxes and fees per gallon measured in 2010 prices. The selected variable name is *gas tax/fee*.	API (2010) and TPC (2010)	51
Renewable Energy Consumption	Renewable energy consumption in British thermal units per capita in 2010. The selected variable name is *renewable energy consumption*.	EIA (2010b)	51

Table 1 (continued)

Variable	Description	Source	# of Obs.
Education	Highest educational attainment for population 25 years and over as a percentage of population 25 Years and over in 2010. The selected variable names are *some college or more*, *bachelor's degree or more*, and *doctorate degree*.	Census Bureau (2010a) and Social Explorer (2010a)	3211
Population Density	Population per square mile in 2010. The selected variable name is *population density*.	Census Bureau (2010d) and Social Explorer (2010b)	3211
Precipitation	Mean daily precipitation in mm in 2010. The selected variable name is *Mean daily precipitation*.	NLDAS (2010b)	3112
Temperature	Mean daily maximum heat index in °F in 2010. The selected variable name is *Mean daily maximum heat index*.	NLDAS (2010a)	3112
Population Centroid	Mean center of population in longitude and latitude decimal degrees. The selected variable name is *population centroid*.	Census Bureau (2010c)	3221
Shapefiles	TIGER/Line and Cartographic Boundary Shapefiles in 2010.	Census Bureau (2010e,b)	3221

Table 2: Summary Statistics

Variables	# of Obs.	Min.	Max.	Median	Mean	Std. Dev.	Skewness	Kurtosis
CO	3109	0.03	189.99	0.33	0.90	4.54	27.98	1029.03
Log CO	3109	−3.65	5.25	−1.10	−0.90	0.95	1.25	2.84
NH_3	3109	0.00	3.59	0.03	0.09	0.19	5.37	54.23
Log NH_3	3109	−9.11	1.28	−3.64	−3.67	1.73	−0.10	−0.47
NO_x	3109	0.01	23.29	0.07	0.15	0.50	32.70	1444.22
Log NO_x	3109	−4.94	3.15	−2.70	−2.52	0.95	0.97	1.27
PM_{10}	3109	0.00	19.38	0.15	0.36	0.69	10.22	211.36
Log PM_{10}	3109	−6.24	2.96	−1.89	−1.91	1.38	−0.10	−0.28
$PM_{2.5}$	3109	0.00	15.64	0.04	0.11	0.40	23.73	784.71
Log $PM_{2.5}$	3109	−6.89	2.75	−3.17	−3.17	1.29	0.20	0.07
SO_2	3109	0.00	8.74	0.00	0.06	0.29	14.67	324.29
Log SO_2	3109	−8.98	2.17	−5.44	−5.15	1.74	0.97	0.98
VOC	3109	0.01	58.32	0.08	0.31	1.61	23.95	746.05
Log VOC	3109	−4.68	4.07	−2.51	−2.27	1.13	1.21	2.17
Adherents (%)	3109	3.07	192.46	49.88	51.58	18.16	0.75	1.91
Protestants (%)	3109	0.00	138.33	35.03	36.34	18.93	0.56	0.44
Evangelical Protestants (%)	3109	0.00	130.87	18.95	23.26	16.32	1.05	1.24
Black Protestants (%)	3109	0.00	31.56	0.00	1.48	3.30	3.78	17.86
Mainline Protestants (%)	3109	0.00	83.54	8.59	11.60	10.10	2.17	6.36
Catholics (%)	3109	0.00	99.96	7.91	12.38	13.50	1.84	4.56
Orthodox Christians (%)	3109	0.00	13.29	0.00	0.06	0.31	25.99	1043.50

Table 2 (continued)

Variables	# of Obs.	Min.	Max.	Median	Mean	Std. Dev.	Skewness	Kurtosis
Mormons (%)	3109	0.00	100.79	0.61	2.24	8.58	7.62	62.48
Muslims (%)	3109	0.00	28.99	0.00	0.23	1.06	12.30	232.57
Jews (%)	3109	0.00	32.56	0.00	0.12	0.87	24.16	774.18
Hindus (%)	3109	0.00	8.91	0.00	0.06	0.35	12.98	228.80
Buddhists (%)	3109	0.00	10.22	0.00	0.07	0.36	15.91	372.82
Income	3109	7772.00	64381.00	21733.00	22450.77	5369.39	1.63	6.05
Log Income	3109	8.96	11.07	9.99	9.99	0.22	0.37	1.46
Gas Price	3109	2.23	2.53	2.28	2.30	0.06	1.45	2.02
Log Gas Price	3109	0.80	0.93	0.83	0.83	0.03	1.38	1.71
Gas Tax/Fee	3109	0.32	0.65	0.41	0.44	0.07	1.00	0.38
Log Gas Tax/Fee	3109	−1.13	−0.43	−0.88	−0.83	0.16	0.70	−0.27
Renewable Energy Consumption	3109	0.00	0.16	0.02	0.04	0.04	1.94	2.71
Log Renewable Energy Consumption	3109	−6.31	−1.83	−3.80	−3.60	0.71	0.97	0.16
Some College or More (%)	3109	17.95	89.27	46.95	47.37	10.94	0.30	−0.12
Bachelor's Degree or More (%)	3109	3.68	70.96	16.82	19.01	8.67	1.55	3.21
Doctorate Degree (%)	3109	0.00	16.49	0.47	0.69	0.86	5.23	54.02
Population Density	3109	0.12	69468.42	45.64	261.48	1733.23	26.77	921.91
Log Population Density	3109	−2.10	11.15	3.82	3.81	1.72	0.09	0.56
Mean Daily Precipitation	3109	0.38	8.94	2.70	2.62	0.82	0.57	4.92
Log Mean Daily Precipitation	3109	−0.96	2.19	0.99	0.91	0.36	−1.23	2.25
Mean Daily Max. Heat Index	3109	78.86	99.21	92.28	91.71	4.51	−0.59	−0.40
Log Mean Daily Max. Heat Index	3109	4.37	4.60	4.52	4.52	0.05	−0.69	−0.21

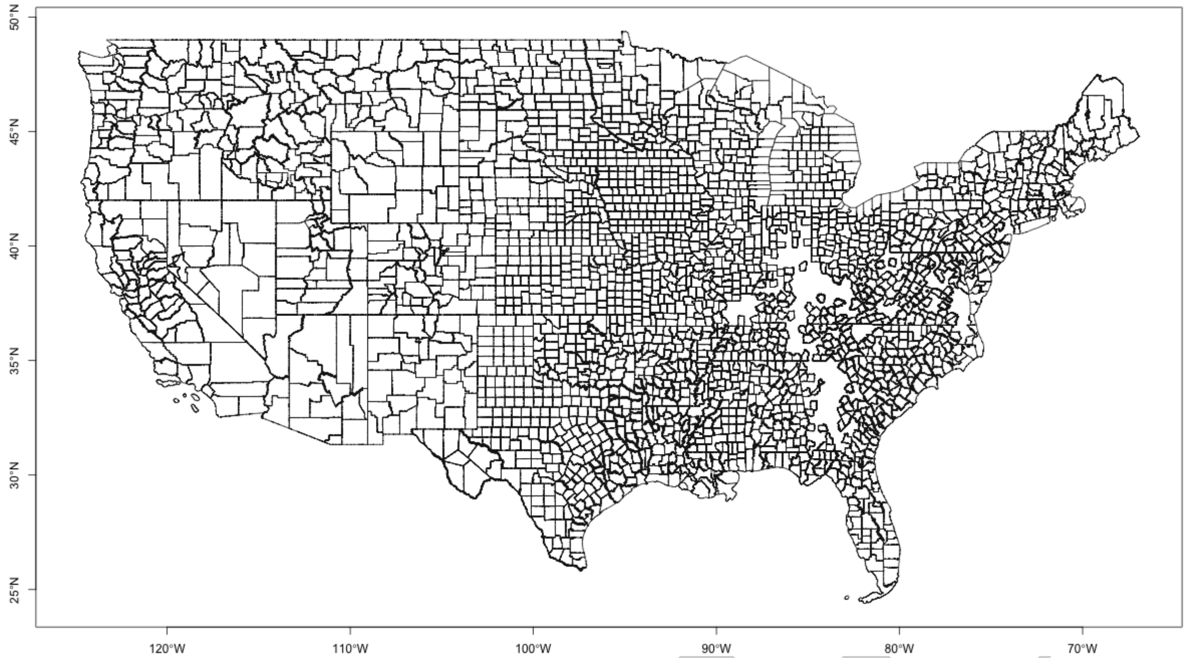


Figure 3: TIGER/Line Shapefile with County Borders Layer in Census 2010

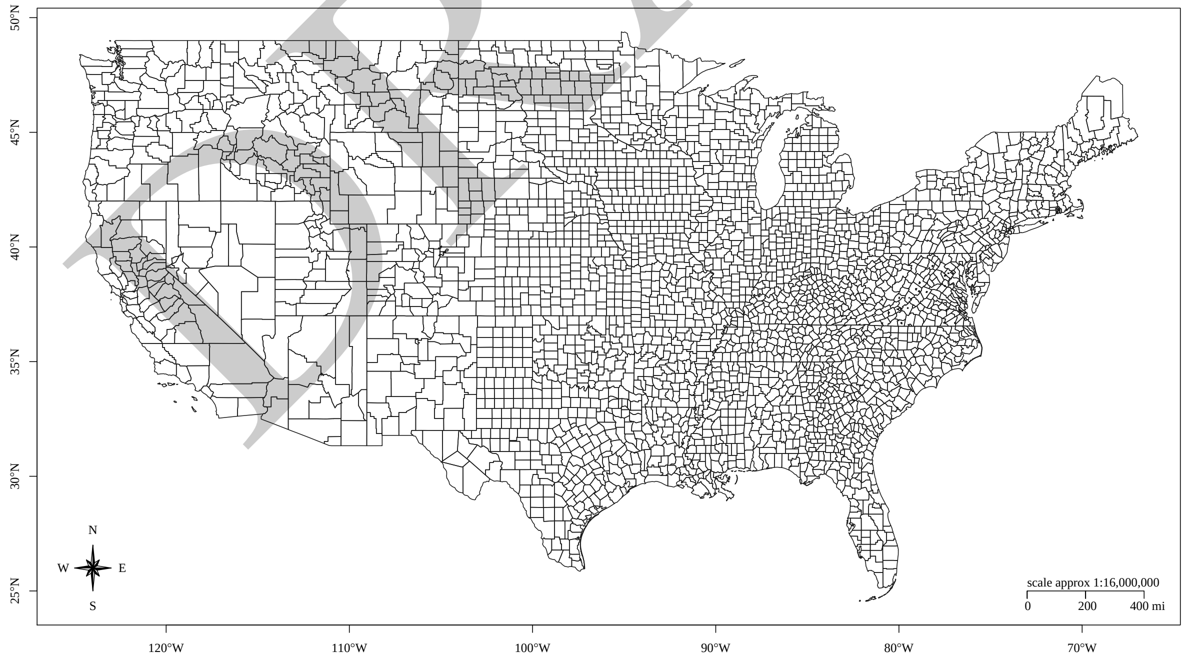


Figure 4: Cartographic Boundary Shapefile with County Borders Layer in Census 2010

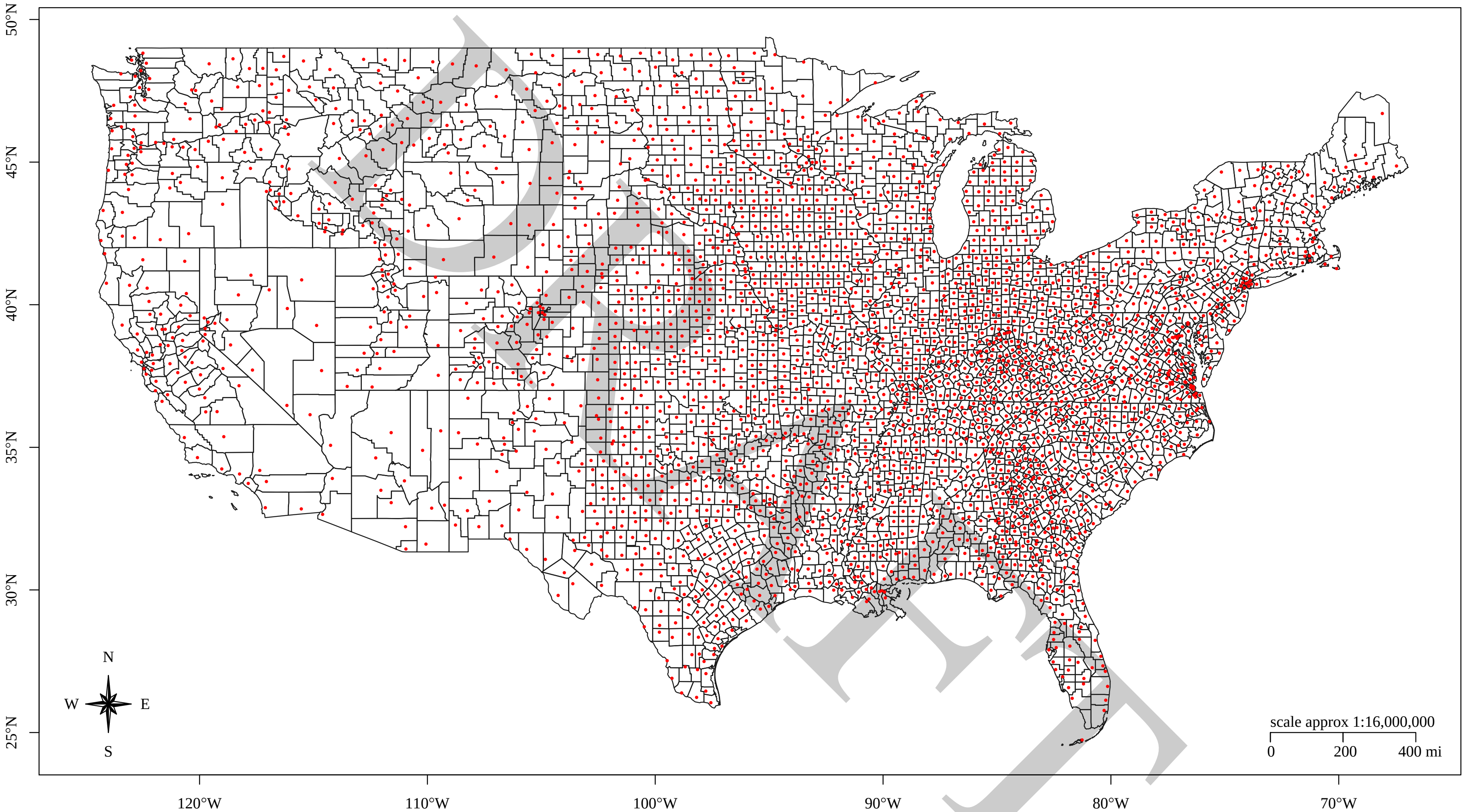
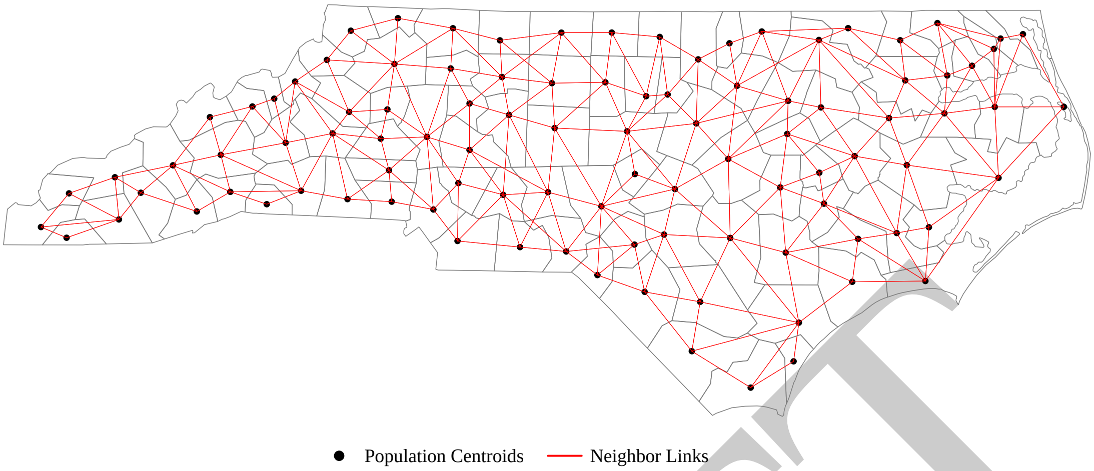
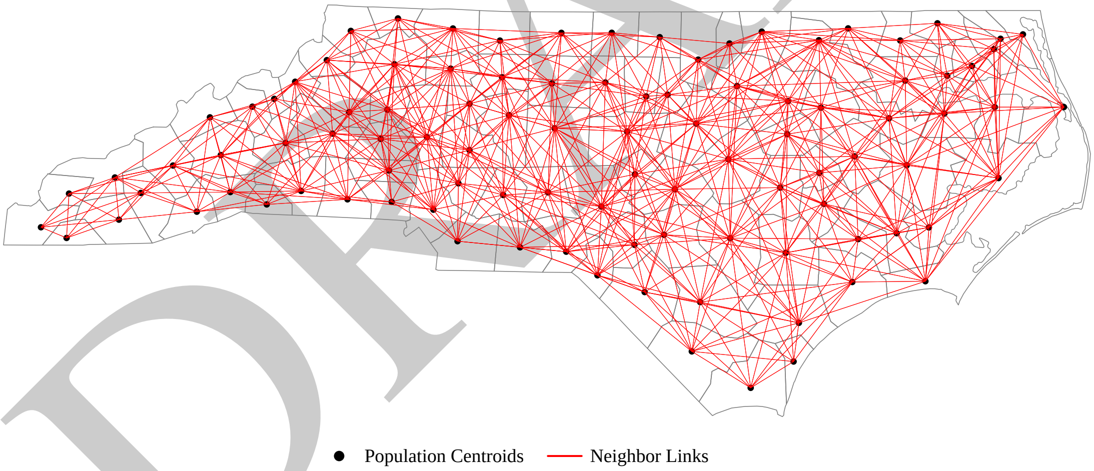


Figure 5: Population Centroids in Census 2010

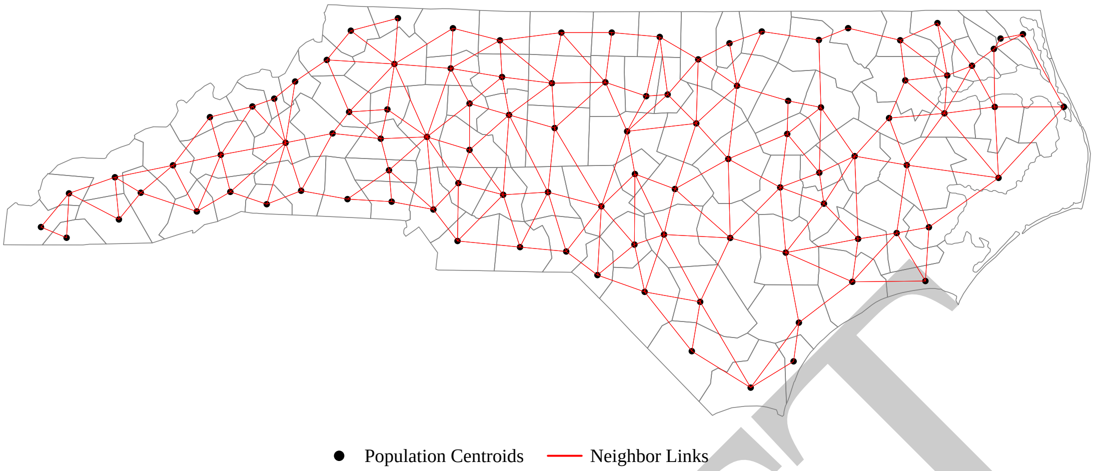


(a) 1st Order Queen Contiguity

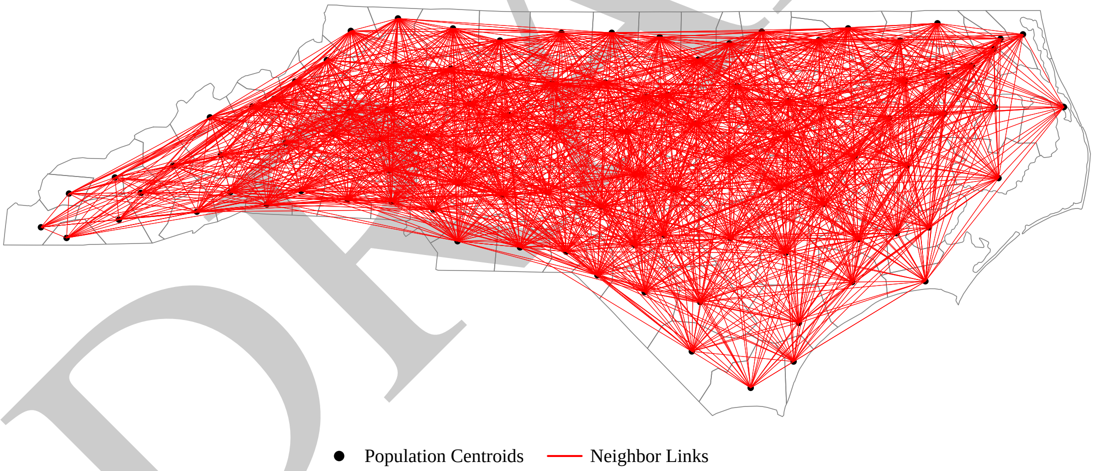


(b) 2nd Order Queen Contiguity

Figure 6: Neighbor Links by Contiguity Type

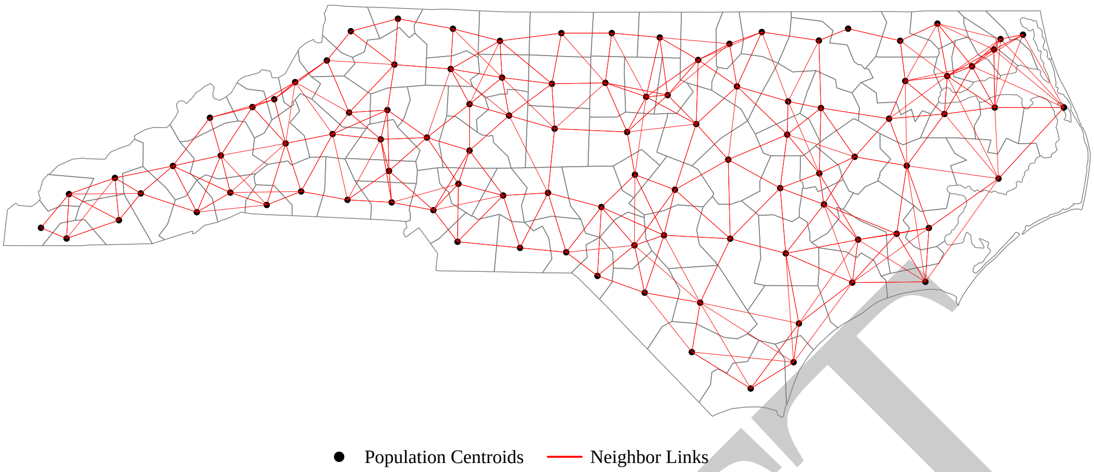


(c) Sphere of Influence Contiguity

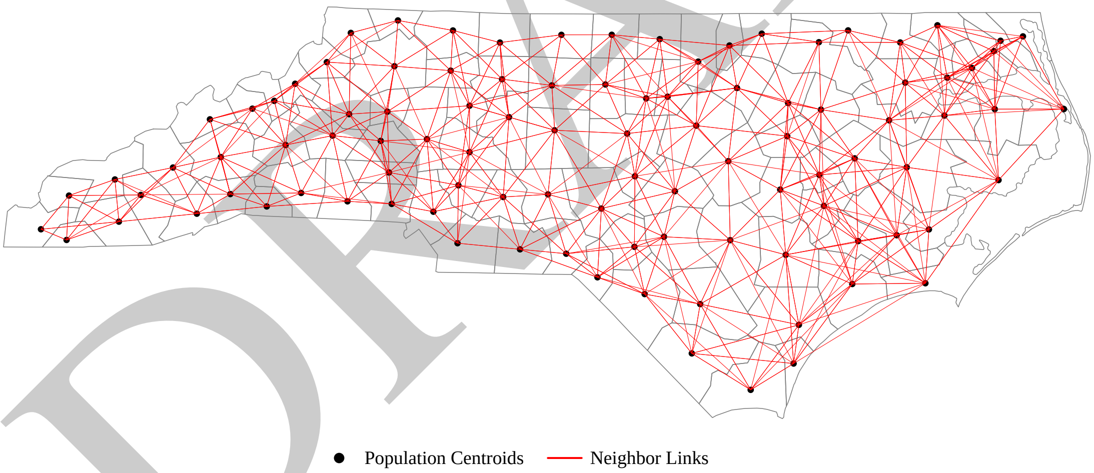


(d) Minimum Distance Contiguity

Figure 6 (continued)

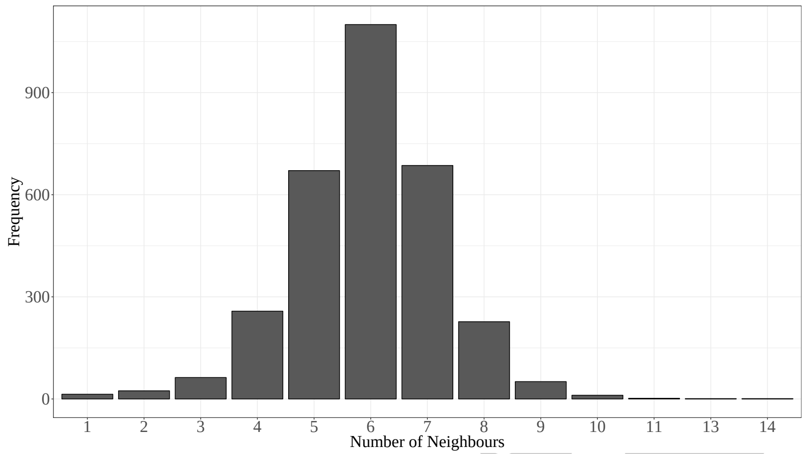


(e) 6 Nearest Neighbors Contiguity

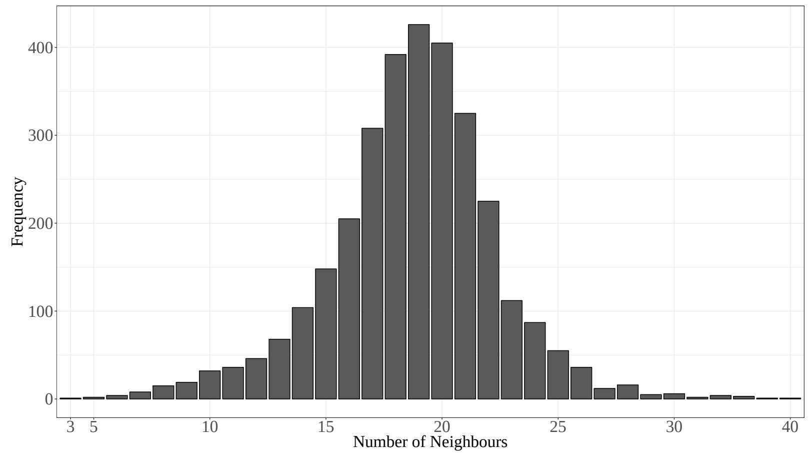


(f) 10 Nearest Neighbors Contiguity

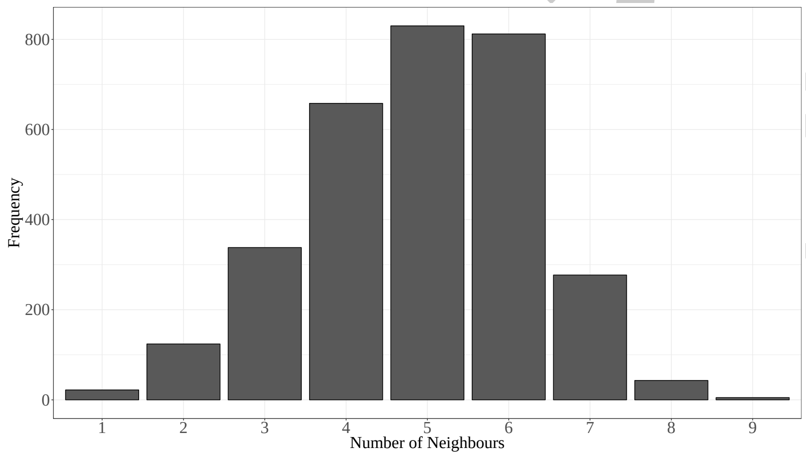
Figure 6 (continued)



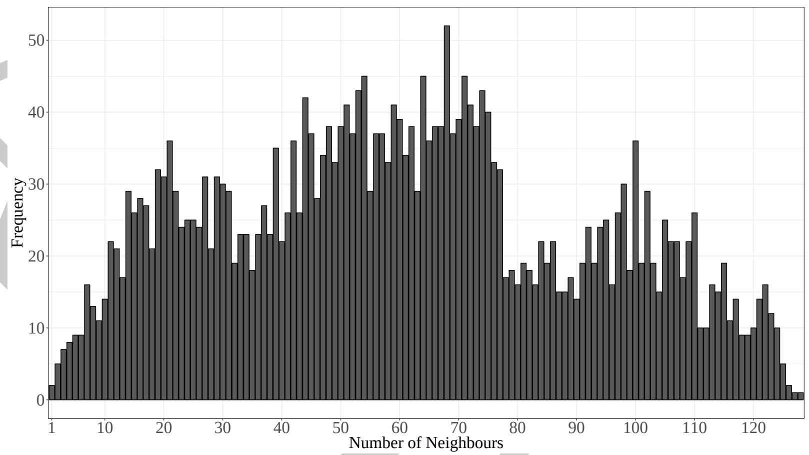
(a) 1st Order Queen Contiguity



(b) 2nd Order Queen Contiguity

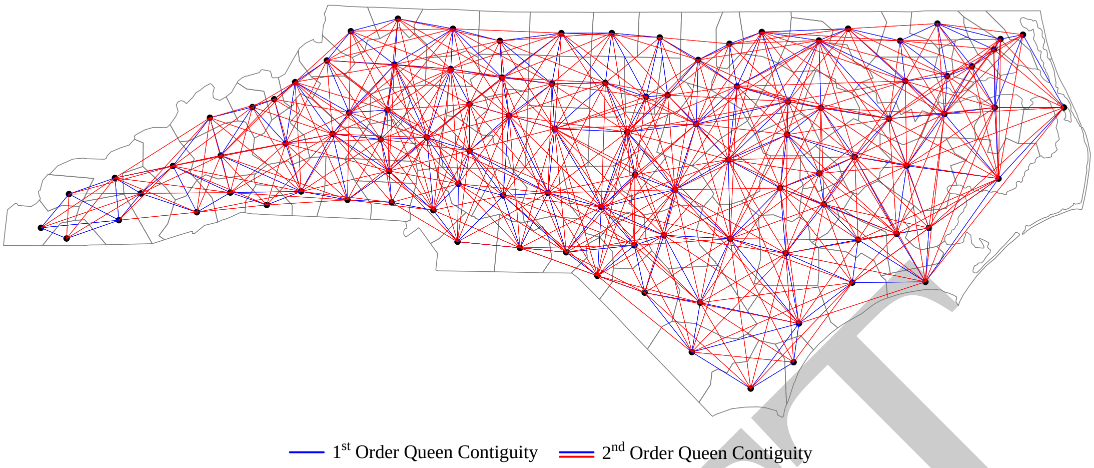


(c) Sphere of Influence Contiguity

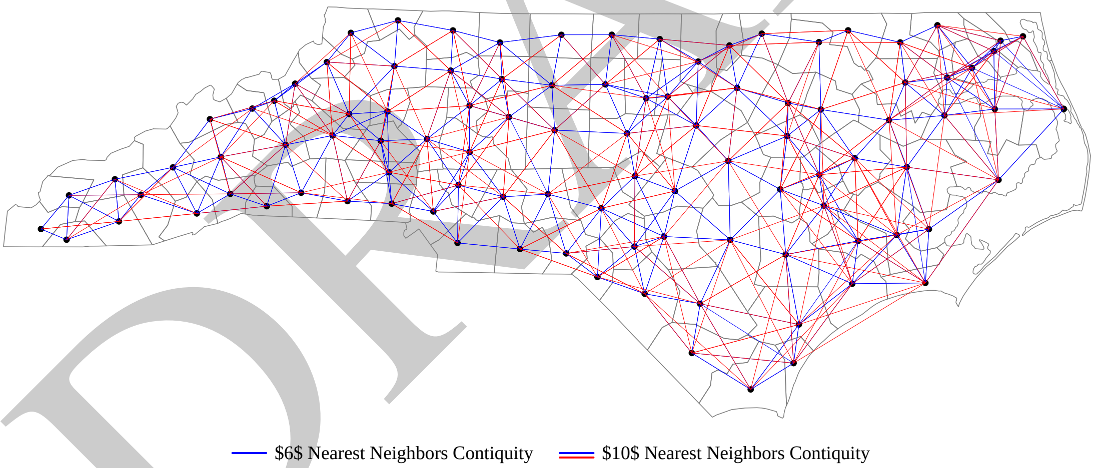


(d) Minimum Distance Contiguity

Figure 7: Frequency Distribution of Number of Neighbors by Contiguity Type



(a) 1st Order Queen Contiguity vs. 2nd Order Queen Contiguity



(b) 6 Nearest Neighbors Contiguity vs. 10 Nearest Neighbors Contiguity

Figure 8: Contiguity Comparisons

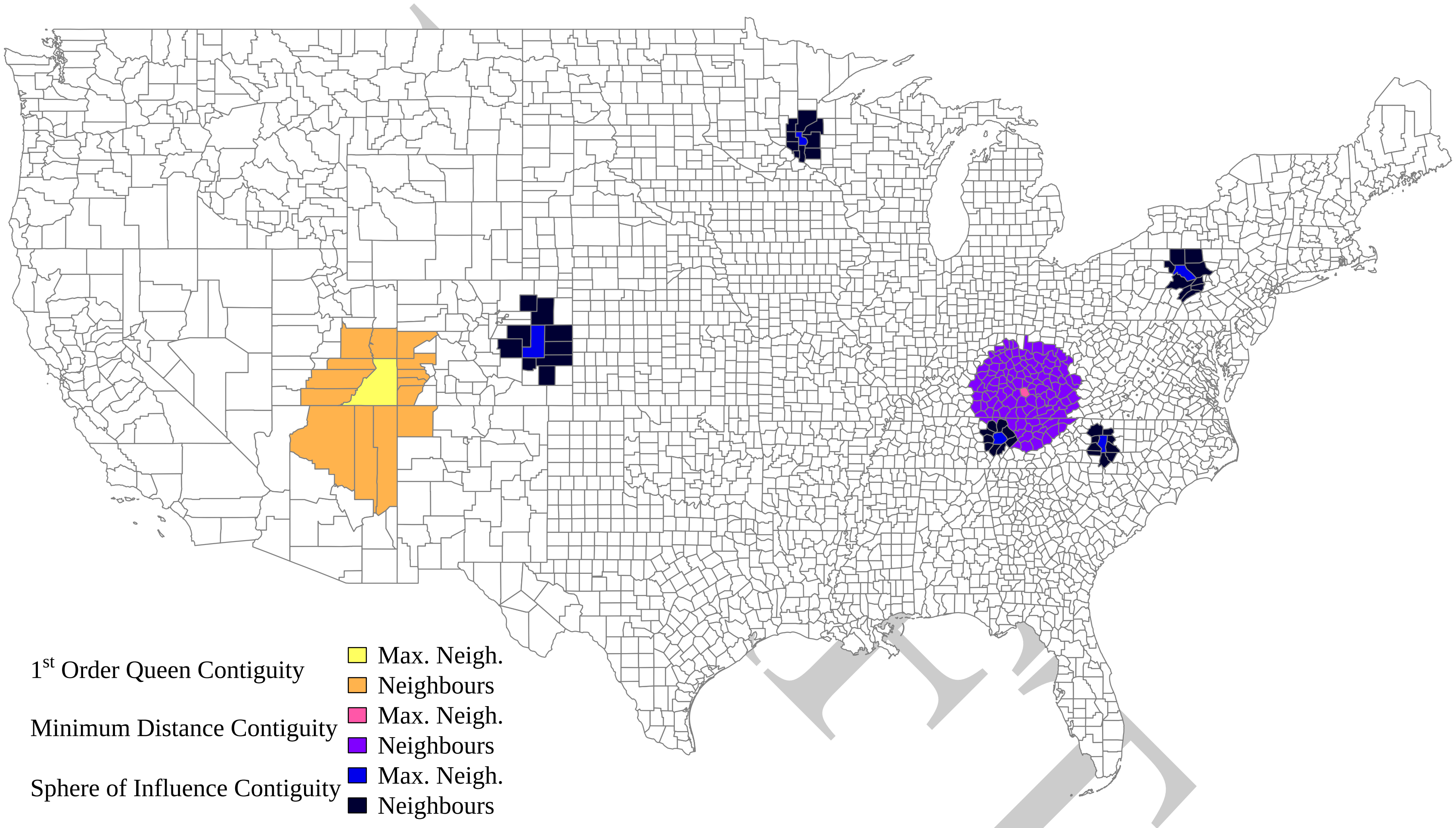


Figure 9: Counties with the Maximum Number of Neighbors vs. Their Neighbors by Contiguity Type

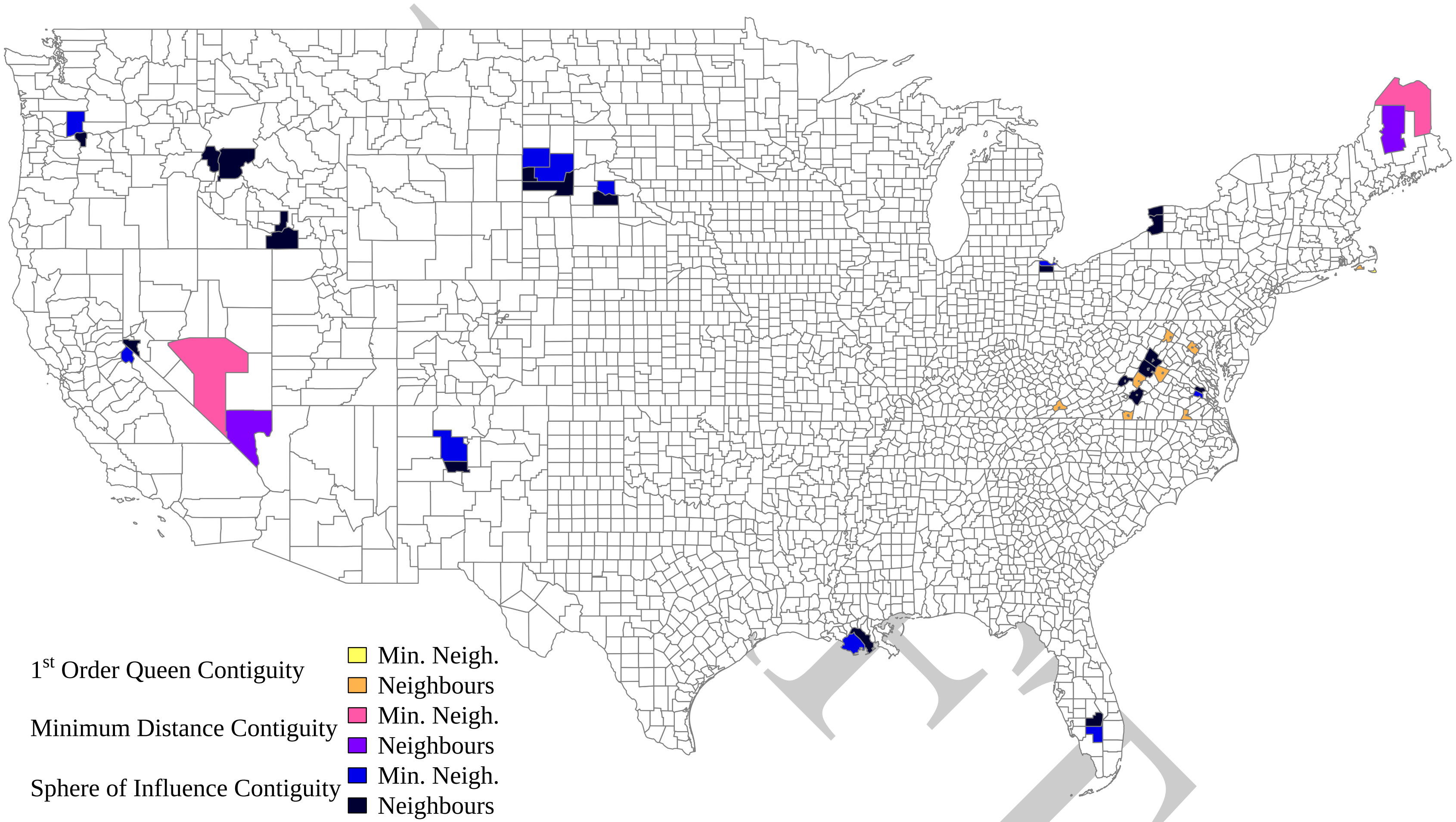
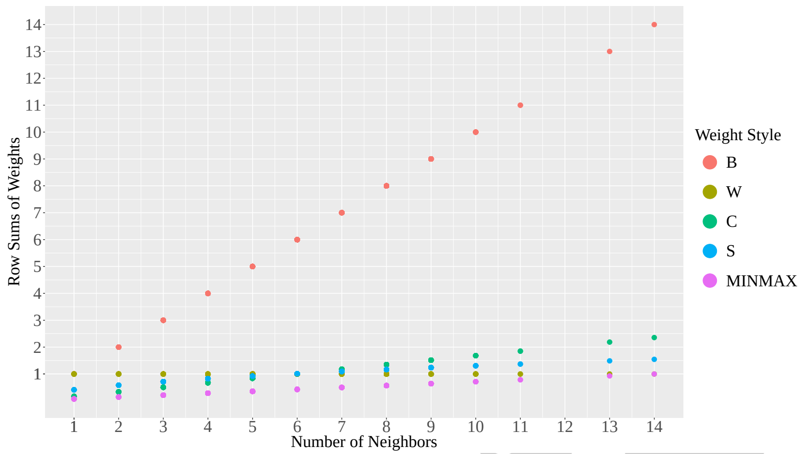
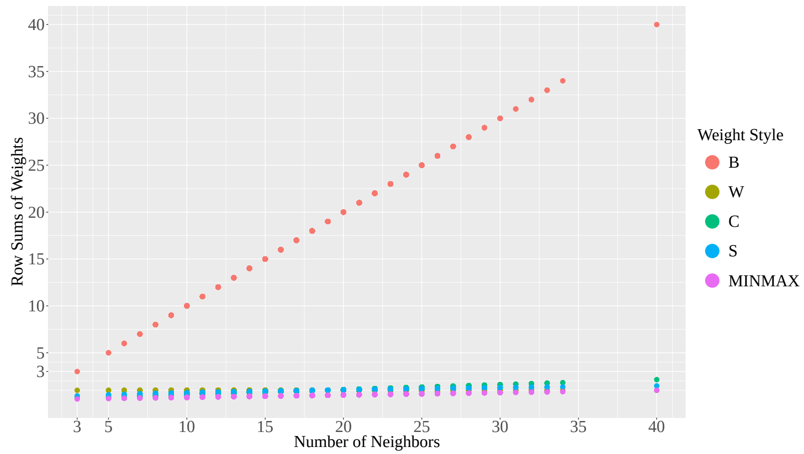


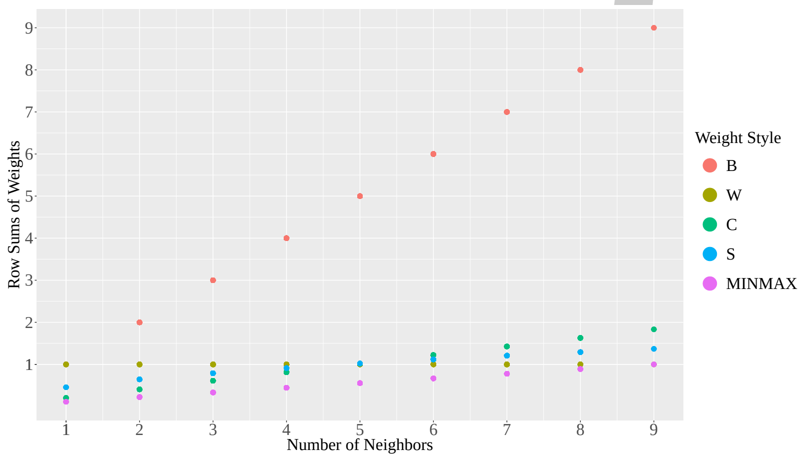
Figure 10: Counties with the Minimum Number of Neighbors vs. Their Neighbors by Contiguity Type



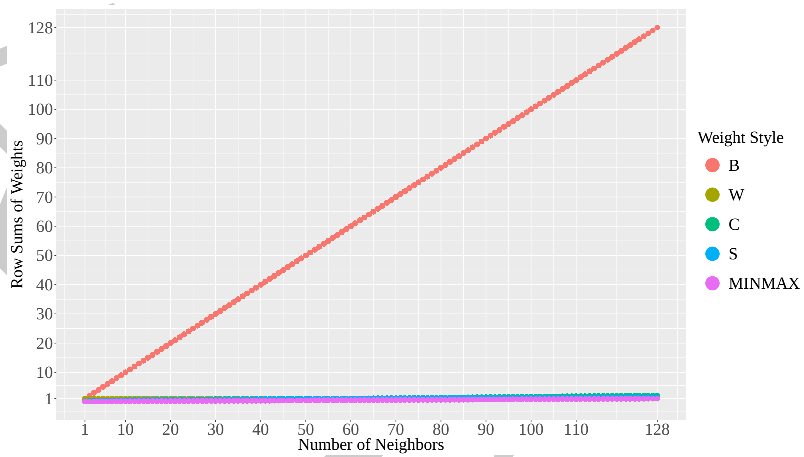
(a) 1st Order Queen Contiguity



(b) 2nd Order Queen Contiguity



(c) Sphere of Influence Contiguity



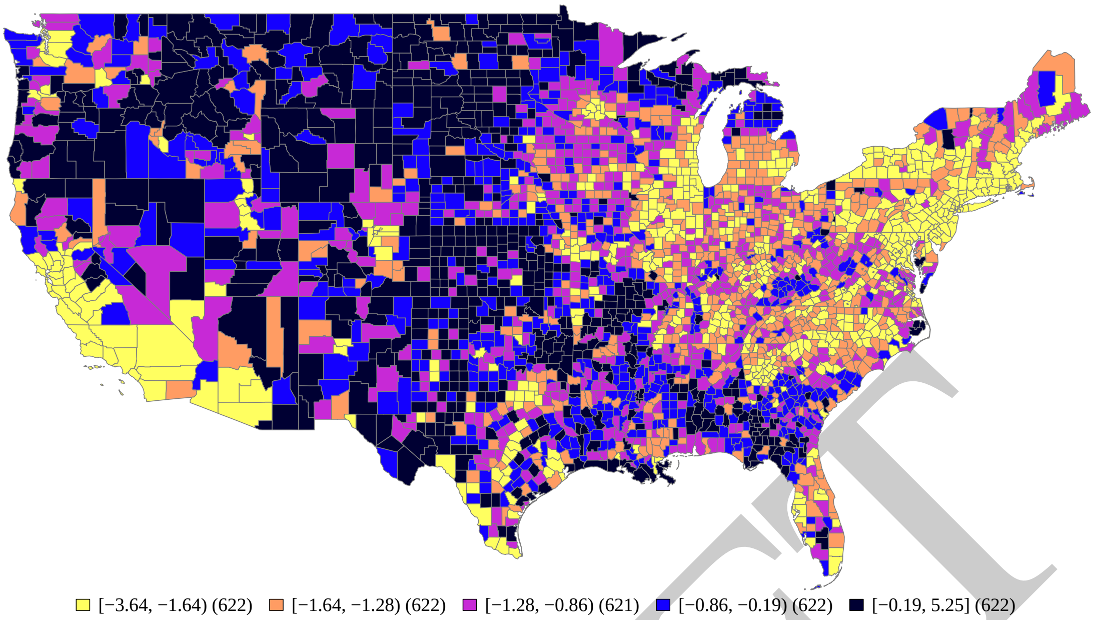
(d) Minimum Distance Contiguity

Figure 11: Number of Neighbors vs. Row Sums of Weights by Contiguity Type and Weight Style

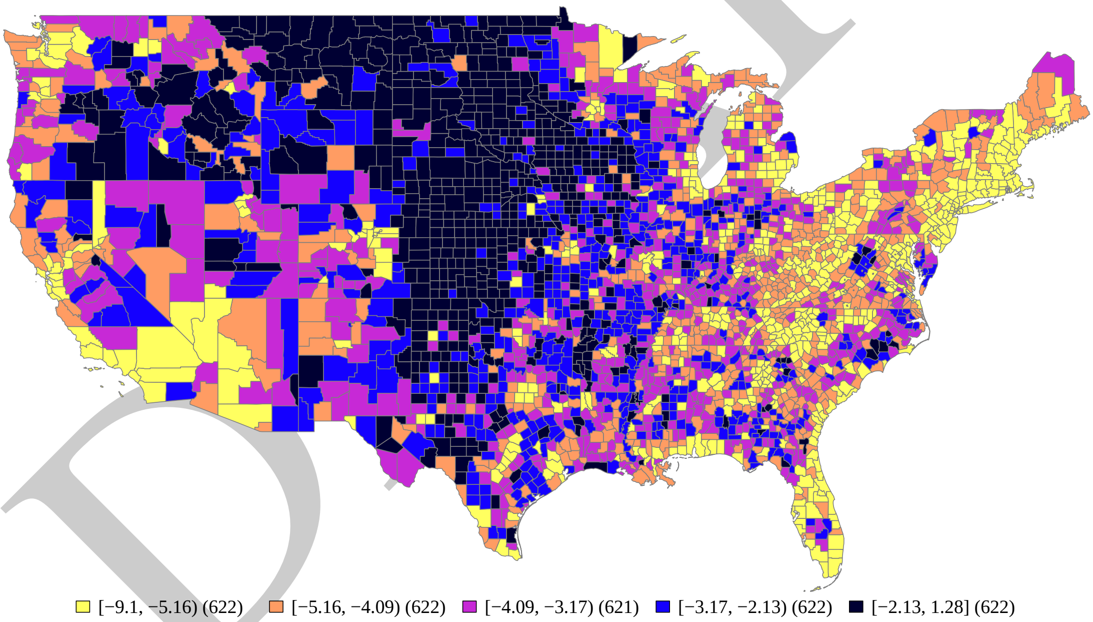
Table 3: Global Moran's I Test Statistics for Dependent Variables by Weight Matrix

	Dependent Variable:						
Weight Matrix	CO	NH_3	NO_x	PM_{10}	$PM_{2.5}$	SO_2	VOC
q1.b	0.551***	0.657***	0.515***	0.727***	0.652***	0.201***	0.629***
q1.w	0.558***	0.675***	0.523***	0.749***	0.669***	0.206***	0.628***
q1.c	0.551***	0.657***	0.515***	0.727***	0.652***	0.201***	0.629***
q1.s	0.554***	0.665***	0.518***	0.736***	0.659***	0.204***	0.628***
q1.minmax	0.551***	0.657***	0.515***	0.727***	0.652***	0.201***	0.629***
q2.b	0.476***	0.567***	0.451***	0.668***	0.587***	0.157***	0.549***
q2.w	0.476***	0.583***	0.457***	0.687***	0.599***	0.160***	0.547***
q2.c	0.476***	0.567***	0.451***	0.668***	0.587***	0.157***	0.549***
q2.s	0.475***	0.574***	0.454***	0.677***	0.592***	0.158***	0.548***
q2.minmax	0.476***	0.567***	0.451***	0.668***	0.587***	0.157***	0.549***
k6.b	0.545***	0.699***	0.528***	0.767***	0.678***	0.207***	0.617***
k6.w	0.545***	0.699***	0.528***	0.767***	0.678***	0.207***	0.617***
k6.c	0.545***	0.699***	0.528***	0.767***	0.678***	0.207***	0.617***
k6.s	0.545***	0.699***	0.528***	0.767***	0.678***	0.207***	0.617***
k6.minmax	0.545***	0.699***	0.528***	0.767***	0.678***	0.207***	0.617***
k10.b	0.514***	0.660***	0.503***	0.741***	0.650***	0.185***	0.587***
k10.w	0.514***	0.660***	0.503***	0.741***	0.650***	0.185***	0.587***
k10.c	0.514***	0.660***	0.503***	0.741***	0.650***	0.185***	0.587***
k10.s	0.514***	0.660***	0.503***	0.741***	0.650***	0.185***	0.587***
k10.minmax	0.514***	0.660***	0.503***	0.741***	0.650***	0.185***	0.587***
soi.b	0.551***	0.688***	0.533***	0.753***	0.675***	0.215***	0.632***
soi.w	0.573***	0.711***	0.544***	0.778***	0.695***	0.217***	0.648***
soi.c	0.551***	0.688***	0.533***	0.753***	0.675***	0.215***	0.632***
soi.s	0.561***	0.699***	0.539***	0.765***	0.685***	0.216***	0.640***
soi.minmax	0.551***	0.688***	0.533***	0.753***	0.675***	0.215***	0.632***

Notes: Normality assumption is used in all tests. All dependent variables are in the natural logarithmic form. *q1*, *q2*, *k6*, *k10*, and *soi* indicate 1st Order Queen, 2nd Order Queen, 6 Nearest Neighbors, 10 Nearest Neighbors, and Sphere of Influence contiguities respectively. *b*, *w*, *c*, *s*, and *minmax* indicate binary, row–standardized, global–standardized, variance–stabilizing, and minmax–normalized weight styles respectively. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

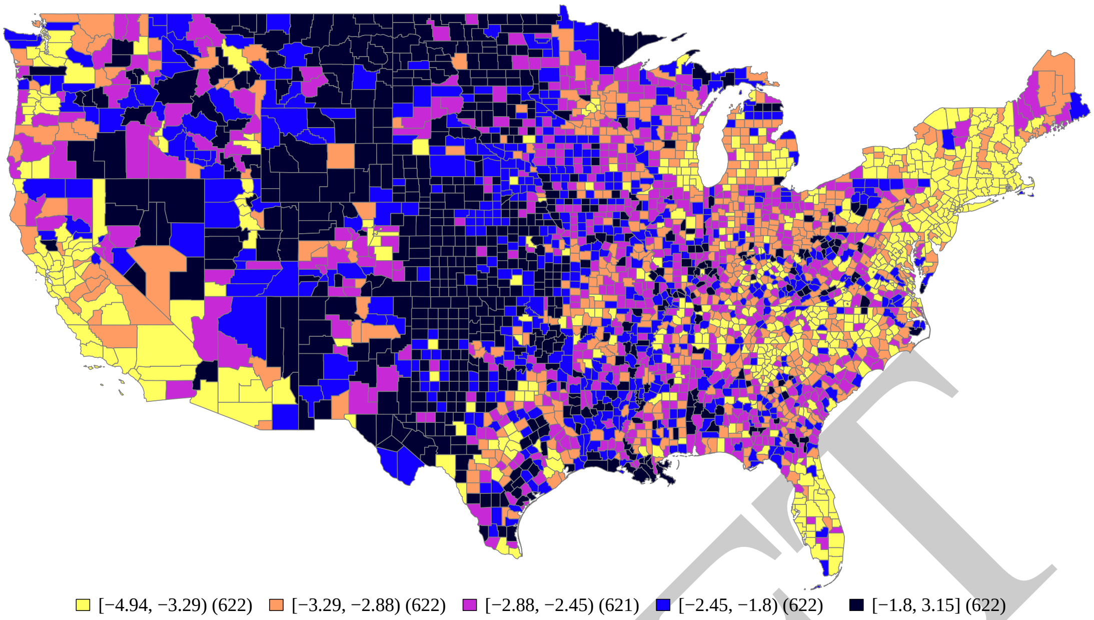


(a) Log CO

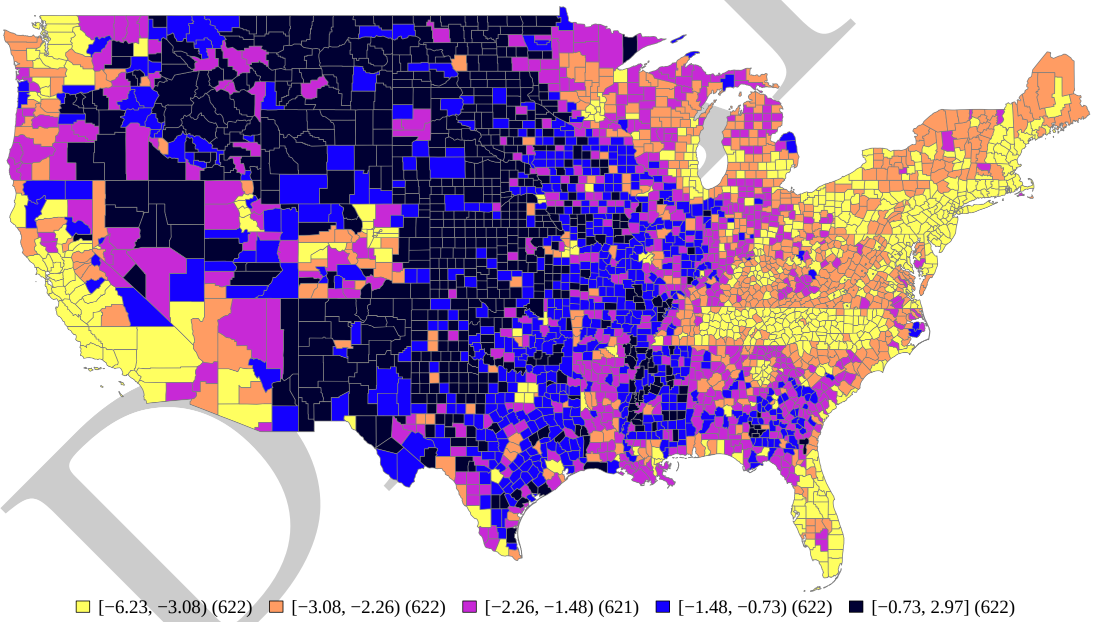


(b) Log NH_3

Figure 12: Thematic Maps of All Dependent Variables

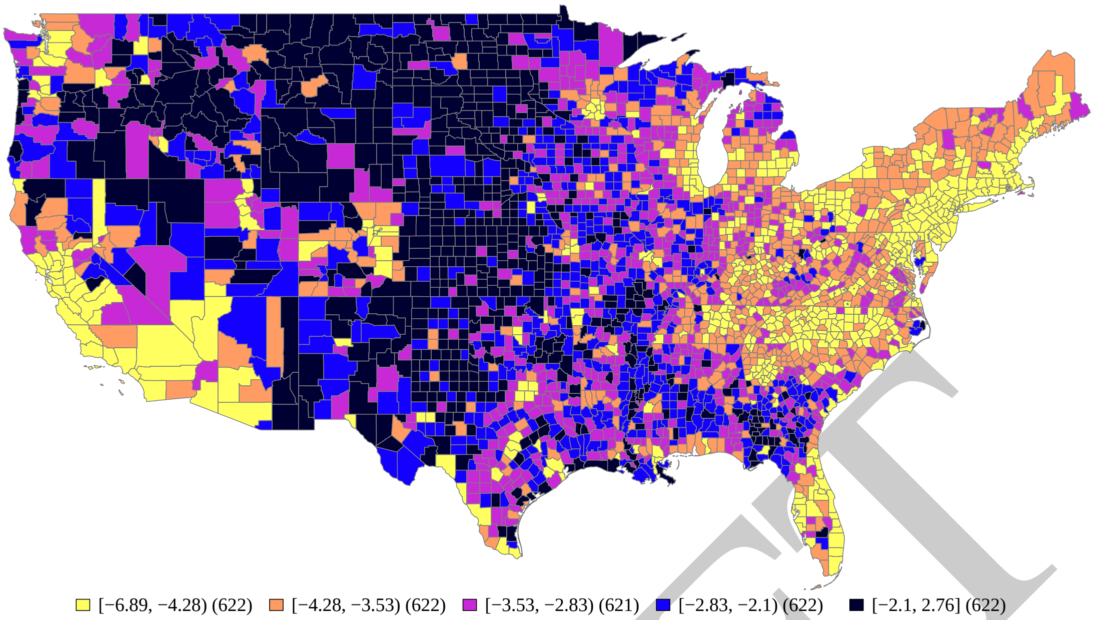


(c) Log NO_x

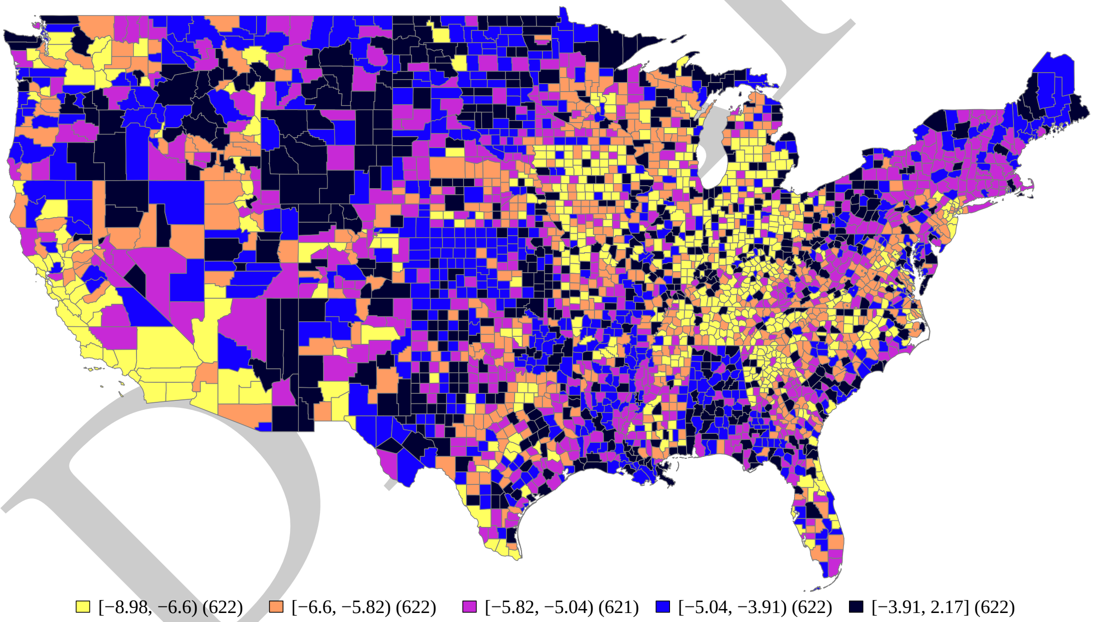


(d) Log PM_{10}

Figure 12 (continued)



(e) Log $PM_{2.5}$



(f) Log SO_2

Figure 12 (continued)

Table 4: LR Test Statistics for Base Model OLS by Education Variable

Base Model	Some College or More	Bachelor's Degree or More	Doctorate Degree
CO	0.004	0.330	0.046
NH_3	17.735***	7.409***	17.770***
NO_x	13.994***	2.744*	12.295***
PM_{10}	25.581***	22.545***	20.829***
$PM_{2.5}$	13.034***	6.776***	10.828***
SO_2	20.549***	10.924***	22.504***
VOC	0.179	0.637	0.109

Notes: LR tests are performed to decide including *log income squared* in a base model. Each base model is labeled with its dependent variable label. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

Table 5: AIC of Base Model OLS by Education Variable

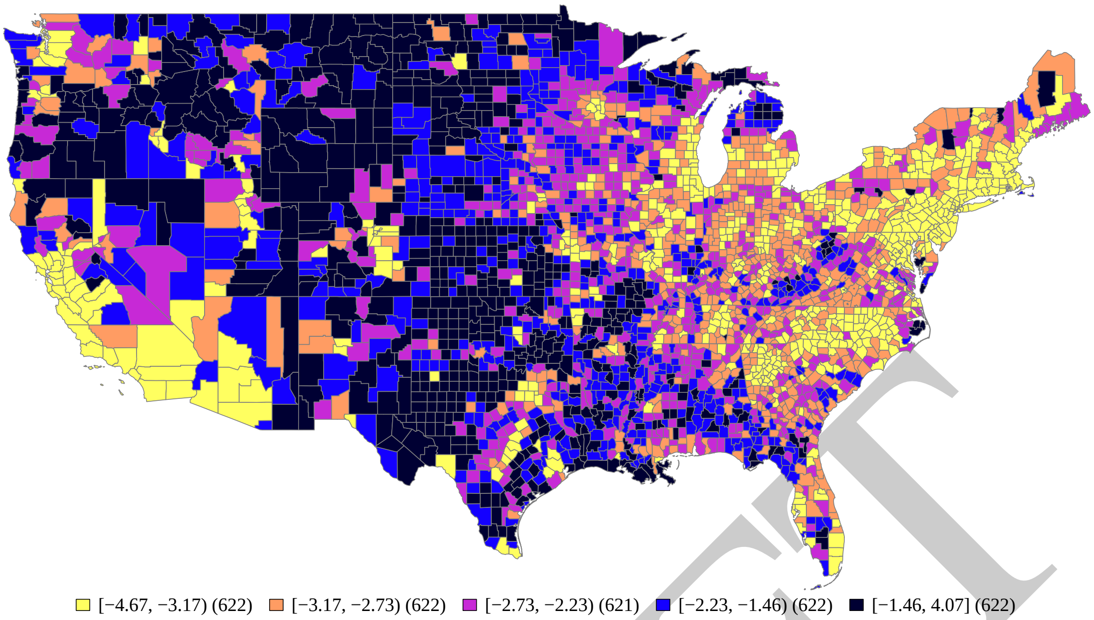
Base Model	Some College or More	Bachelor's Degree or More	Doctorate Degree
CO	5451.45*	5454.66	5455.53
$NH_3{}^{\dagger}$	7909.90	7896.66*	7916.50
$NO_x{}^{\dagger}$	6019.92	5999.62*	6032.64
$PM_{10}{}^{\dagger}$	5684.31*	5746.77	5747.55
$PM_{2.5}{}^{\dagger}$	5584.51*	5600.31	5602.26
$SO_2{}^{\dagger}$	11995.46	11988.91*	12002.82
VOC	6369.00	6355.27*	6369.05

Notes: Each base model is labeled with its label of the dependent variable. † indicates a base model including *log income squared*. * indicates the model with the minimum AIC among a group of base models with different education variables.

Table 6: BIC of Base Model OLS by Education Variable

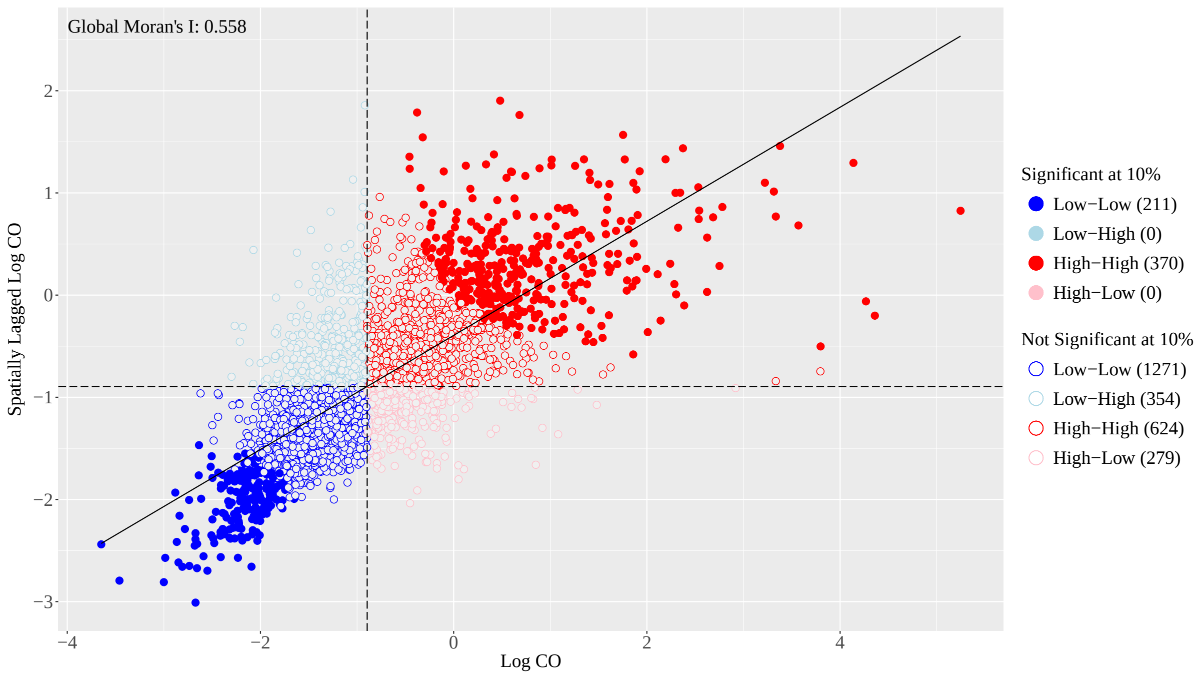
Base Model	Some College or More	Bachelor's Degree or More	Doctorate Degree
CO	5572.29*	5575.51	5576.37
$NH_3{}^{\dagger}$	8036.79	8023.54*	8043.38
$NO_x{}^{\dagger}$	6146.81	6126.51*	6159.53
$PM_{10}{}^{\dagger}$	5811.19*	5873.66	5874.43
$PM_{2.5}{}^{\dagger}$	5711.39*	5727.19	5729.14
$SO_2{}^{\dagger}$	12122.34	12115.79*	12129.70
VOC	6489.84	6476.11*	6489.89

Notes: Each base model is labeled with its label of the dependent variable. † indicates a base model including *log income squared*. * indicates the model with the minimum BIC among a group of base models with different education variables.

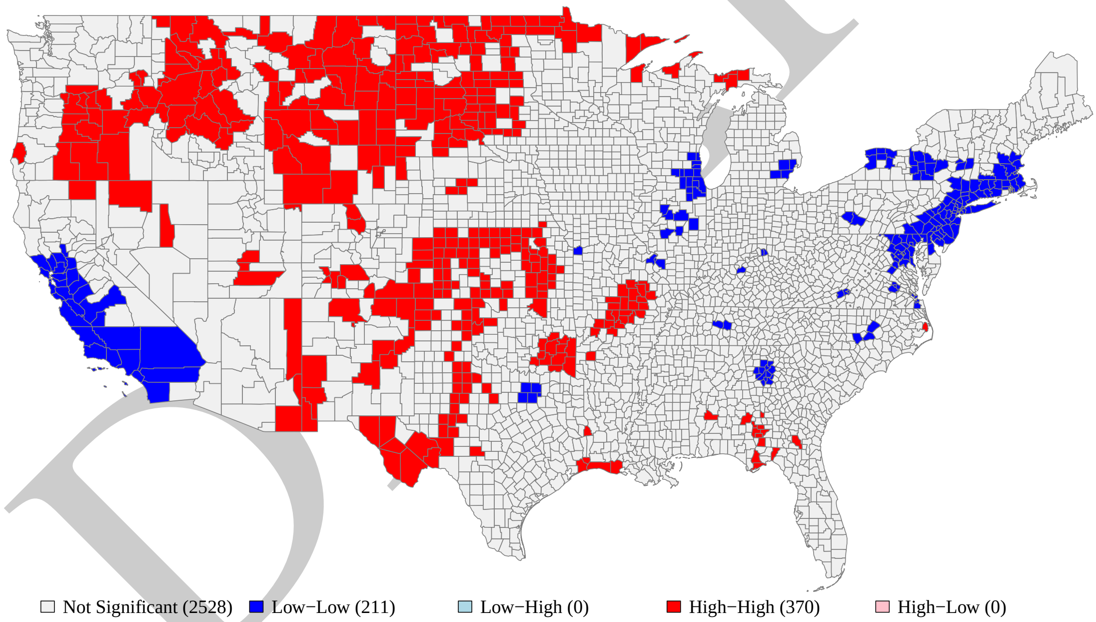


(g) Log VOC

Figure 12 (continued)

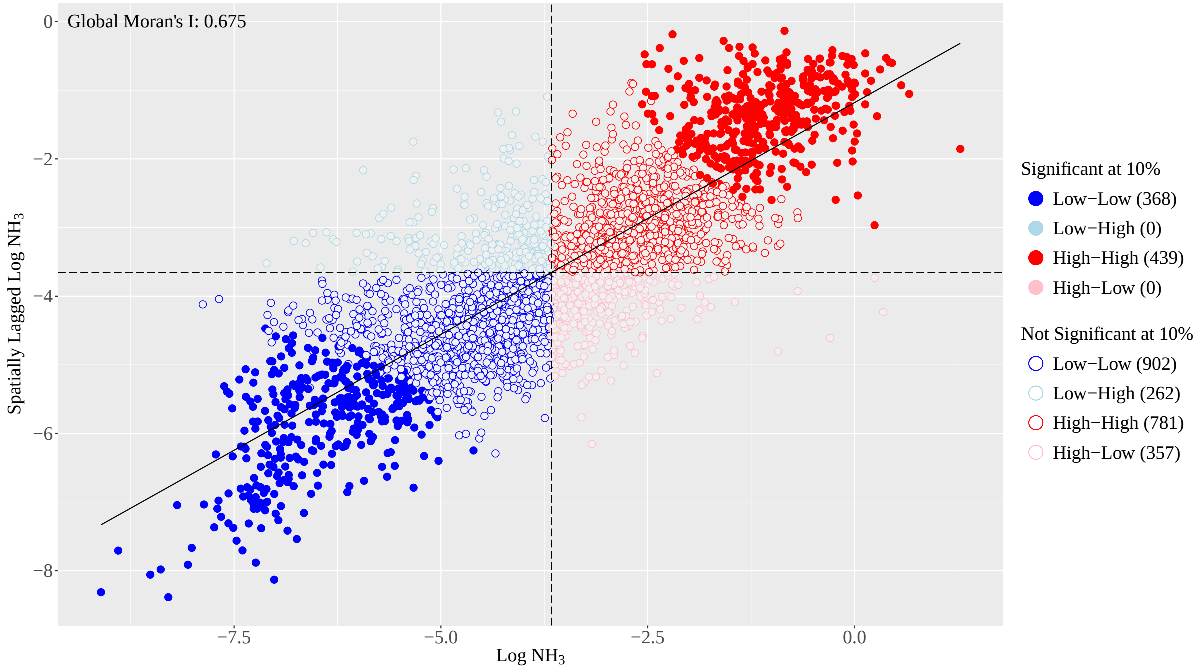


(a) LISA Moran Scatterplot for Log CO

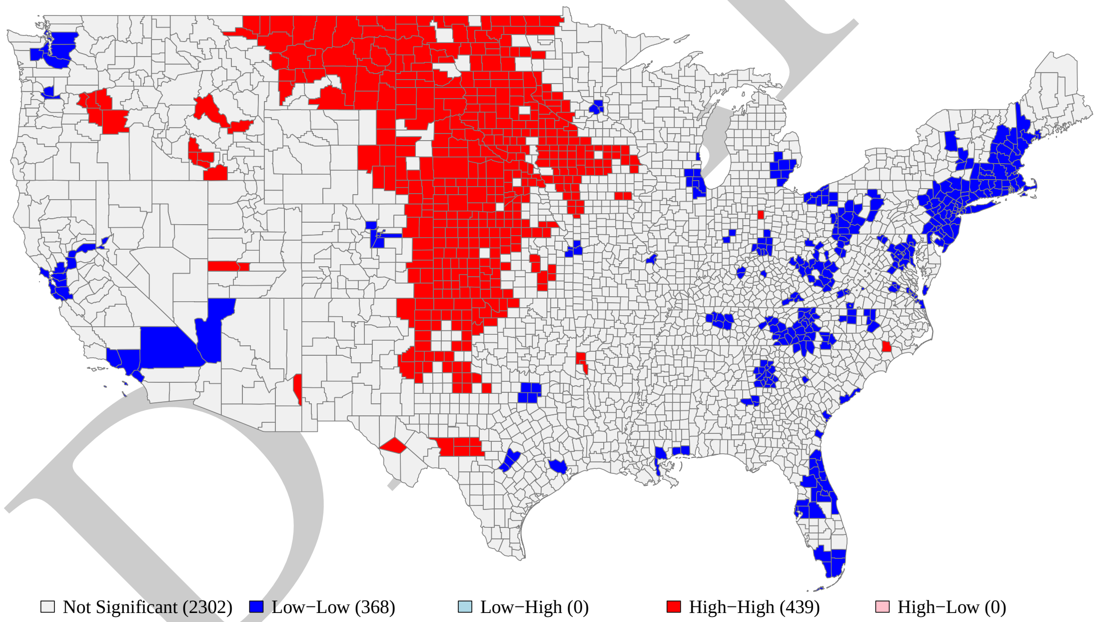


(b) LISA Cluster Map for Log CO

Figure 13: LISA Moran Scatterplot and Cluster Map by Dependent Variable

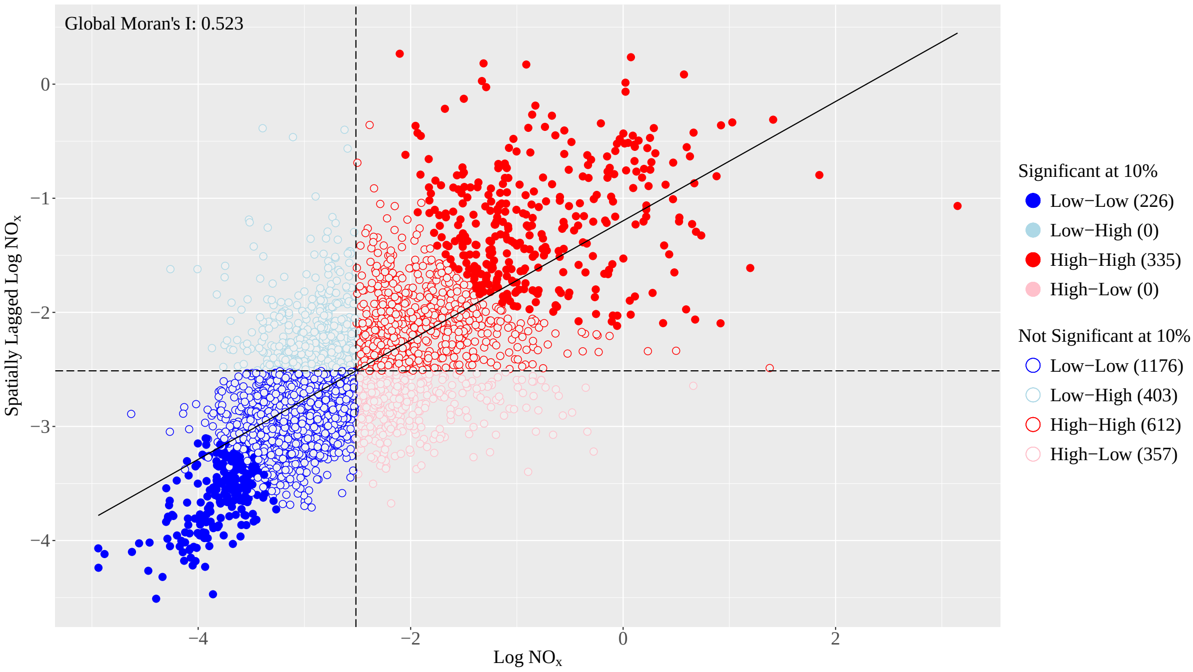


(c) LISA Moran Scatterplot for Log NH_3

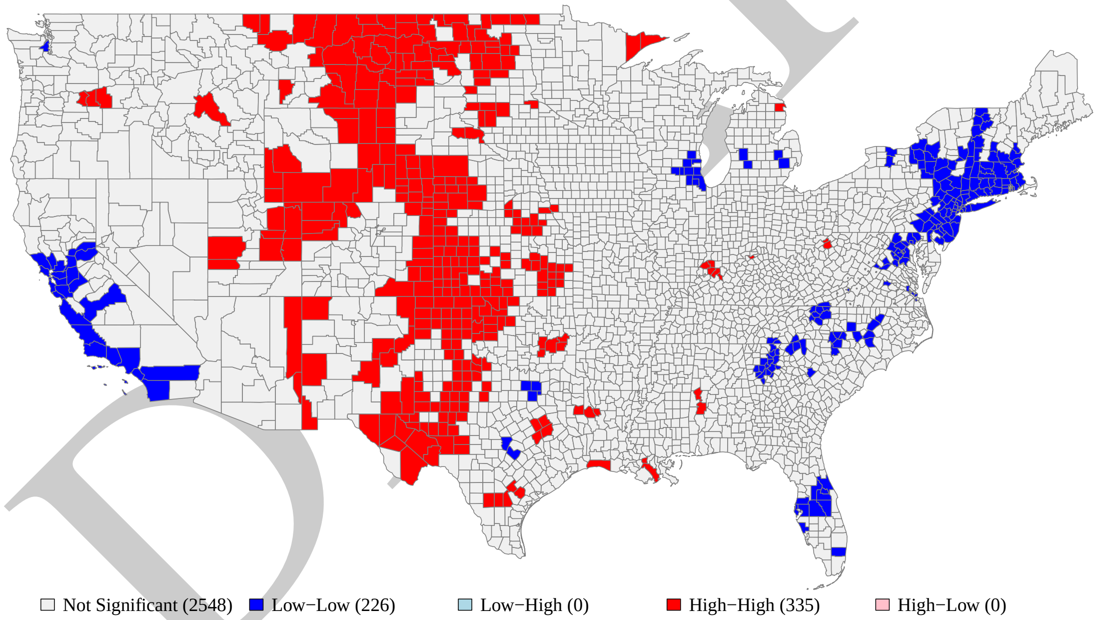


(d) LISA Cluster Map for Log NH_3

Figure 13 (continued)

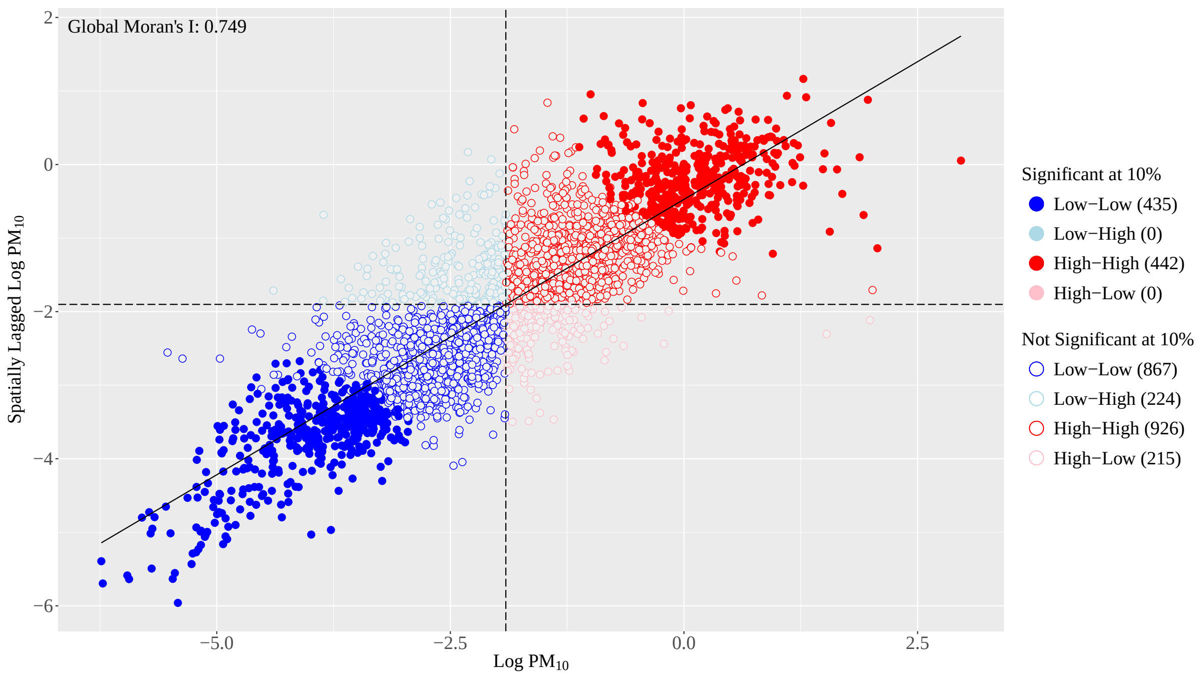


(e) LISA Moran Scatterplot for Log NO_x

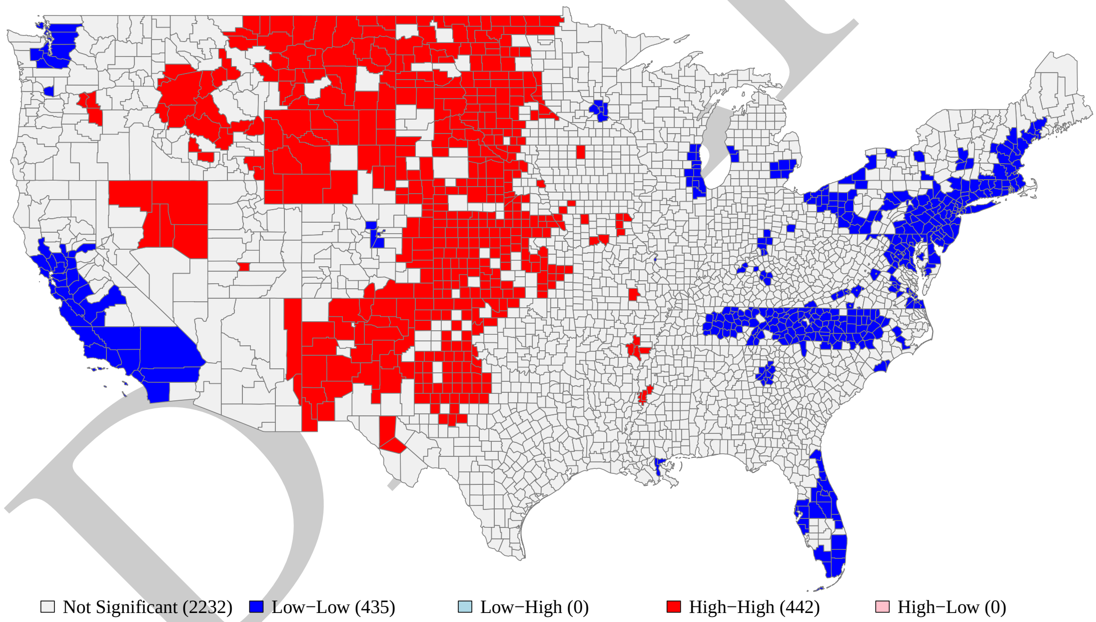


(f) LISA Cluster Map for Log NO_x

Figure 13 (continued)

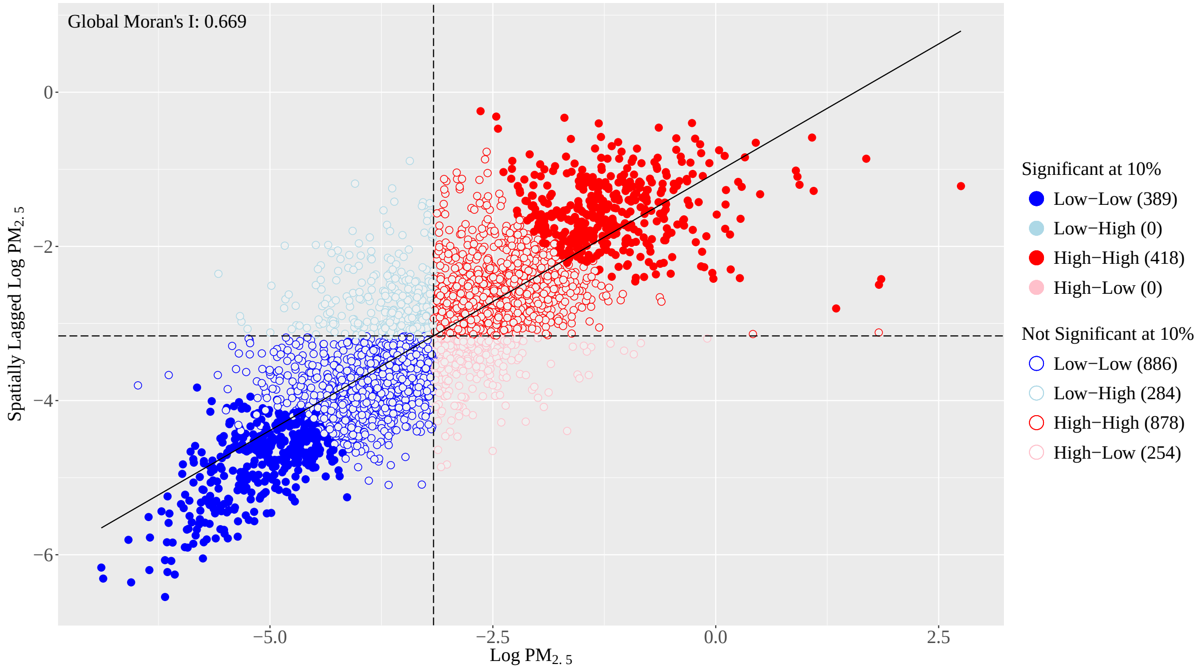


(g) LISA Moran Scatterplot for Log PM_{10}

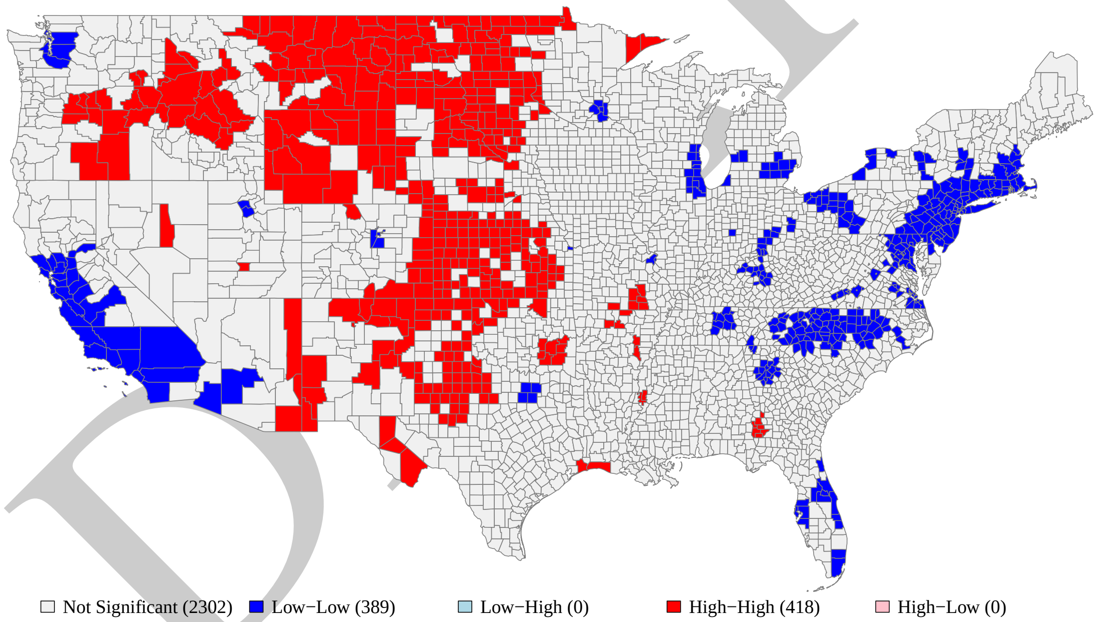


(h) LISA Cluster Map for Log PM_{10}

Figure 13 (continued)

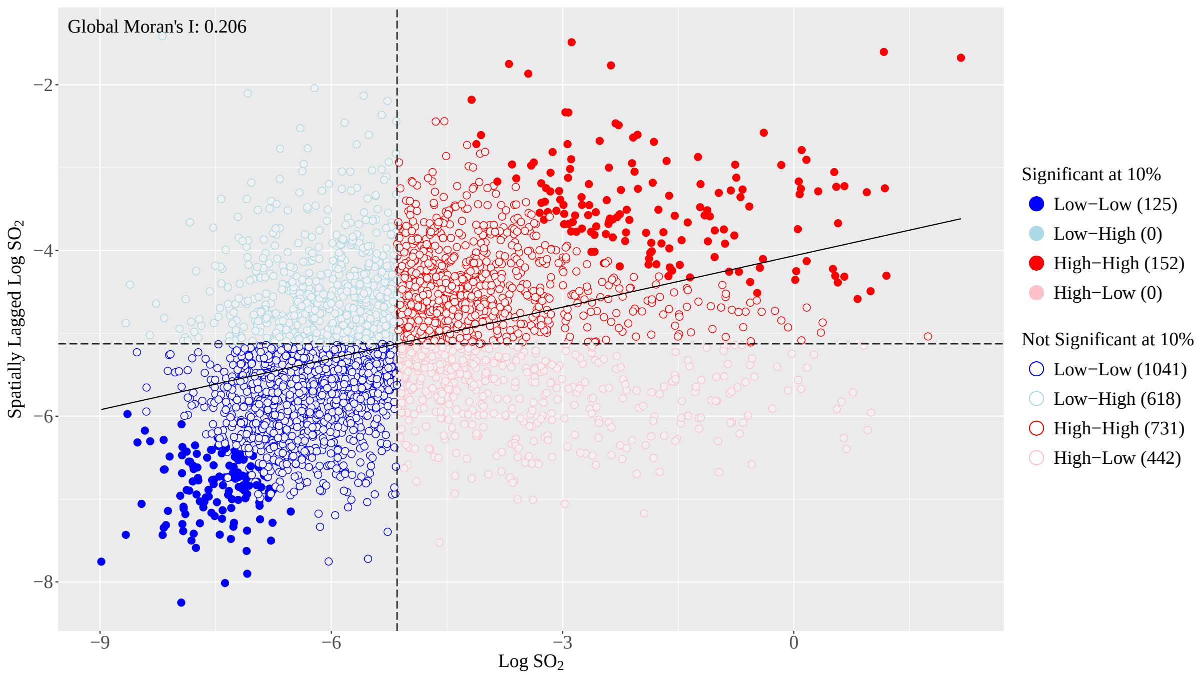


(i) LISA Moran Scatterplot for Log $PM_{2.5}$

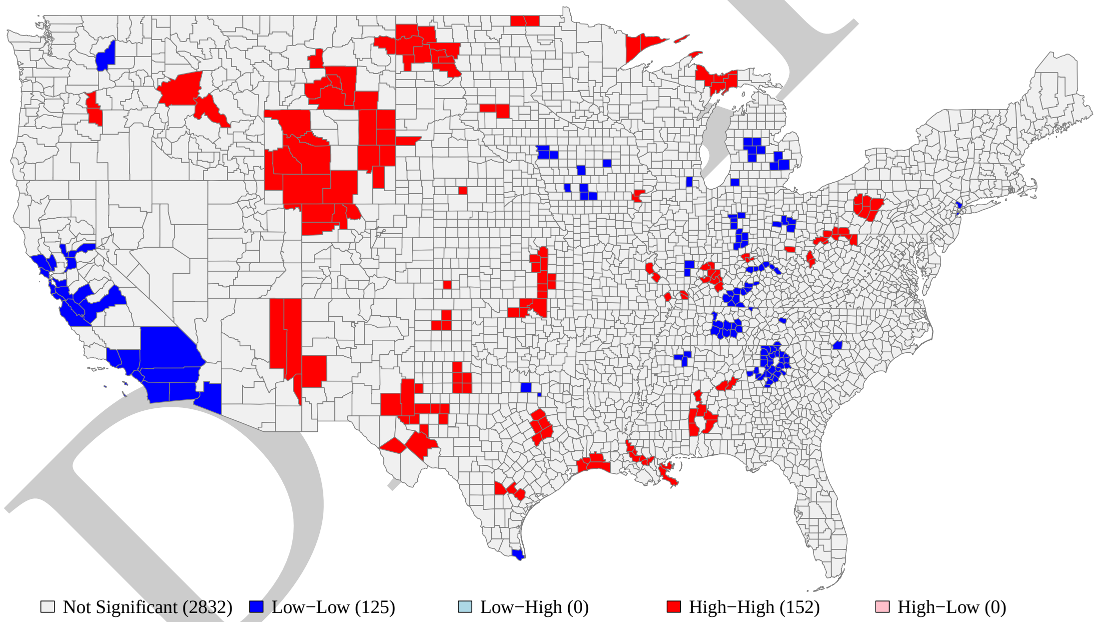


(j) LISA Cluster Map for Log $PM_{2.5}$

Figure 13 (continued)

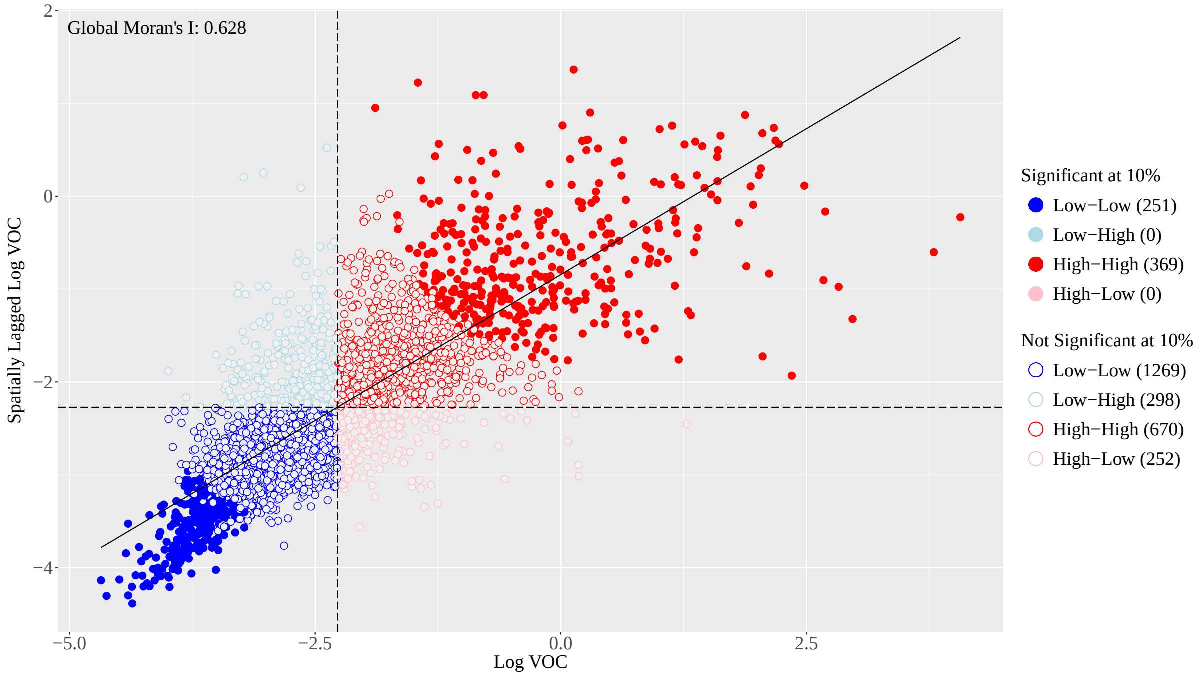


(k) LISA Moran Scatterplot for Log SO_2

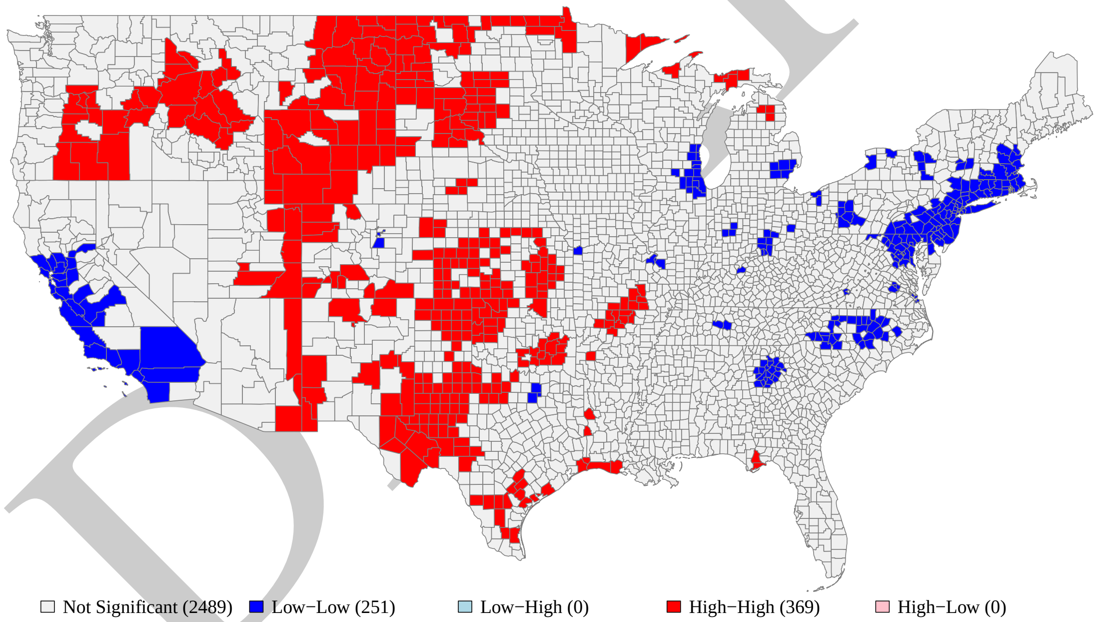


(l) LISA Cluster Map for Log SO_2

Figure 13 (continued)



(m) LISA Moran Scatterplot for Log VOC



(n) LISA Cluster Map for Log VOC

Figure 13 (continued)

Table 7: CPD Test *P*-Values for Base Model OLS by Pairs of Education Variables

Base Model	Hypothesis Tests	Some College or More / Bachelor's Degree or More	Some College or More / Doctorate Degree	Bachelor's Degree or More / Doctorate Degree
CO	H_0: M_1 vs. H_1: M_2	0.498	0.641	0.688
	H_0: M_2 vs. H_1: M_1	$< 0.001_{c}^{***}$	$< 0.001_{c}^{***}$	$< 0.001_{b}^{***}$
NH_3 †	H_0: M_1 vs. H_1: M_2	$< 0.001_{b}^{***}$	0.952	0.086_{d}^{*}
	H_0: M_2 vs. H_1: M_1	0.368	$< 0.001_{c}^{***}$	$< 0.001_{b}^{***}$
NO_x †	H_0: M_1 vs. H_1: M_2	$< 0.001_{b}^{***}$	0.002_{d}^{***}	0.558
	H_0: M_2 vs. H_1: M_1	0.806	$< 0.001_{c}^{***}$	$< 0.001_{b}^{***}$
PM_{10} †	H_0: M_1 vs. H_1: M_2	$< 0.001_{b}^{***}$	0.157	0.485
	H_0: M_2 vs. H_1: M_1	$< 0.001_{c}^{***}$	$< 0.001_{c}^{***}$	0.153
$PM_{2.5}$ †	H_0: M_1 vs. H_1: M_2	0.051_{b}^{*}	0.354	0.609
	H_0: M_2 vs. H_1: M_1	$< 0.001_{c}^{***}$	$< 0.001_{c}^{***}$	$< 0.001_{b}^{***}$
SO_2 †	H_0: M_1 vs. H_1: M_2	$< 0.001_{b}^{***}$	0.536	0.098_{d}^{*}
	H_0: M_2 vs. H_1: M_1	0.868	$< 0.001_{c}^{***}$	$< 0.001_{b}^{***}$
VOC	H_0: M_1 vs. H_1: M_2	$< 0.001_{b}^{***}$	0.150	0.193
	H_0: M_2 vs. H_1: M_1	0.092_{c}^{*}	0.128	$< 0.001_{b}^{***}$

Notes: Each base model is labeled with its label of the dependent variable. † indicates a base model including *log income squared*. $_c$, $_b$, and $_d$ indicate that the selected education variable for a base model is *log some college or more*, *log bachelor's degree or more*, and *log doctorate degree* respectively. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

Table 8: DMJ Test *P*-Values for Base Model OLS by Pairs of Education Variables

Base Model	Hypothesis Tests	Some College or More / Bachelor's Degree or More	Some College or More / Doctorate Degree	Bachelor's Degree or More / Doctorate Degree
CO	H_0: M_1 vs. H_1: M_2	$< 0.001_b^{***}$	0.430	0.288
	H_0: M_2 vs. H_1: M_1	$< 0.001_c^{***}$	0.031_c^{**}	0.158
NH_3†	H_0: M_1 vs. H_1: M_2	$< 0.001_b^{***}$	0.953	0.015_d^{**}
	H_0: M_2 vs. H_1: M_1	0.323	0.010_c^{**}	$< 0.001_b^{***}$
NO_x†	H_0: M_1 vs. H_1: M_2	$< 0.001_b^{***}$	0.032_d^{**}	0.540
	H_0: M_2 vs. H_1: M_1	0.809	$< 0.001_c^{***}$	$< 0.001_b^{***}$
PM_{10}†	H_0: M_1 vs. H_1: M_2	$< 0.001_b^{***}$	0.061_d^{*}	0.552
	H_0: M_2 vs. H_1: M_1	$< 0.001_c^{***}$	$< 0.001_c^{***}$	0.289
$PM_{2.5}$†	H_0: M_1 vs. H_1: M_2	$< 0.001_b^{***}$	0.141	0.055_d^{*}
	H_0: M_2 vs. H_1: M_1	$< 0.001_c^{***}$	$< 0.001_c^{***}$	0.018_b^{**}
SO_2†	H_0: M_1 vs. H_1: M_2	0.011_b^{**}	0.138	$< 0.001_d^{***}$
	H_0: M_2 vs. H_1: M_1	0.866	0.002_c^{***}	$< 0.001_b^{***}$
VOC	H_0: M_1 vs. H_1: M_2	$< 0.001_b^{***}$	0.384	0.088_d^{*}
	H_0: M_2 vs. H_1: M_1	0.023_c^{**}	0.367	$< 0.001_b^{***}$

Notes: Each base model is labeled with its label of the dependent variable. † indicates a base model including *log income squared*. $_c$, $_b$, and $_d$ indicate that the selected education variable for a base model is *log some college or more*, *log bachelor's degree or more*, and *log doctorate degree* respectively. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

Table 9: Base Model OLS Results

	Base Model OLS:						
	CO^{c}	$NH_3^{\dagger,b}$	$NO_x^{\dagger,b}$	$PM_{10}^{\dagger,c}$	$PM_{2.5}^{\dagger,c}$	$SO_2^{\dagger,b}$	VOC^{b}
Intercept	-0.205	-90.252^{***}	-44.217^{***}	-101.355^{***}	-70.635^{***}	-145.076^{***}	-9.328^{***}
	(1.921)	(20.078)	(14.799)	(13.418)	(13.204)	(38.774)	(2.209)
Evangelical Protestants (%)	-0.001	0.006^{***}	0.001	0.001	0.001	-0.004	0.002^{*}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Black Protestants (%)	0.021^{***}	-0.030^{***}	0.008^{**}	-0.010^{***}	0.010^{***}	0.058^{***}	0.008^{**}
	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.010)	(0.004)
Mainline Protestants (%)	-0.007^{***}	0.033^{***}	0.006^{***}	0.011^{***}	0.007^{***}	-0.005	-0.007^{***}
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)
Catholics (%)	0.000	0.008^{***}	0.002^{**}	0.003^{***}	0.003^{***}	0.010^{***}	0.004^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)
Orthodox Christians (%)	0.148^{***}	-0.031	0.109^{***}	0.049	0.103^{***}	0.229^{**}	0.139^{***}
	(0.037)	(0.055)	(0.041)	(0.039)	(0.038)	(0.107)	(0.043)
Mormons (%)	-0.003^{**}	0.013^{***}	0.004^{**}	0.007^{***}	0.000	0.002	-0.001
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)
Muslims (%)	0.022^{**}	-0.049^{***}	0.022^{**}	-0.028^{***}	-0.004	0.015	0.055^{***}
	(0.010)	(0.015)	(0.011)	(0.011)	(0.010)	(0.029)	(0.012)
Jews (%)	0.033^{**}	0.005	0.005	0.027^{*}	0.027^{**}	0.004	0.039^{***}
	(0.013)	(0.020)	(0.015)	(0.014)	(0.014)	(0.038)	(0.015)

Table 9 (continued)

	Base Model OLS:						
	CO^{c}	$NH_3{}^{\dagger,b}$	$NO_x{}^{\dagger,b}$	$PM_{10}{}^{\dagger,c}$	$PM_{2.5}{}^{\dagger,c}$	$SO_2{}^{\dagger,b}$	VOC^{b}
Hindus (%)	−0.013	−0.053	0.000	−0.021	−0.030	−0.052	−0.020
	(0.031)	(0.046)	(0.034)	(0.032)	(0.032)	(0.088)	(0.036)
Buddhists (%)	−0.008	0.035	−0.067**	−0.043	−0.026	−0.148*	0.004
	(0.030)	(0.044)	(0.033)	(0.031)	(0.031)	(0.086)	(0.035)
Log Income	0.302***	10.902***	5.782**	13.023***	9.323***	26.625***	0.768***
	(0.081)	(3.948)	(2.910)	(2.637)	(2.595)	(7.625)	(0.091)
Log Income Squared		−0.538***	−0.241*	−0.666***	−0.467***	−1.262***	
		(0.198)	(0.146)	(0.132)	(0.130)	(0.383)	
Log Gas Price	−1.859***	−1.217	−2.650***	0.726	0.480	−1.697	−2.751***
	(0.594)	(0.887)	(0.654)	(0.619)	(0.610)	(1.713)	(0.688)
Log Gas Tax/Fee	−0.471***	0.703***	−0.266***	−0.558***	−0.382***	−0.381	−0.620***
	(0.084)	(0.127)	(0.094)	(0.089)	(0.088)	(0.246)	(0.098)
Log Renewable Energy Consumption	0.055***	0.366***	0.016	0.129***	0.113***	−0.029	0.016
	(0.019)	(0.028)	(0.021)	(0.020)	(0.020)	(0.054)	(0.022)
Some College or More (%)	0.003**			0.014***	0.007***		
	(0.002)			(0.002)	(0.002)		
Bachelor's Degree or More (%)		−0.015***	−0.017***			−0.023***	−0.010***
		(0.003)	(0.002)			(0.006)	(0.002)

Table 9 (continued)

	Base Model OLS:						
	CO^{c}	$NH_3{}^{\dagger,b}$	$NO_x{}^{\dagger,b}$	$PM_{10}{}^{\dagger,c}$	$PM_{2.5}{}^{\dagger,c}$	$SO_2{}^{\dagger,b}$	VOC^{b}
Log Population Density	−0.482***	−0.686***	−0.369***	−0.656***	−0.665***	−0.297***	−0.555***
	(0.009)	(0.014)	(0.010)	(0.009)	(0.009)	(0.027)	(0.011)
Log Mean Daily Precipitation	0.175***	−0.318***	−0.387***	−0.330***	0.004	−0.001	−0.029
	(0.038)	(0.055)	(0.041)	(0.039)	(0.039)	(0.107)	(0.043)
Log Mean Daily Max. Heat Index	−0.169	8.175***	2.657***	8.226***	5.024***	0.546	0.768**
	(0.334)	(0.499)	(0.368)	(0.347)	(0.341)	(0.964)	(0.389)
Observations	3109	3109	3109	3109	3109	3109	3109
Residual Std. Error	0.580	0.859	0.633	0.602	0.592	1.658	0.670
R^2	0.630	0.756	0.560	0.811	0.792	0.097	0.649
Adjusted R^2	0.628	0.754	0.557	0.810	0.791	0.091	0.647
AIC	5451.452	7896.658	5999.623	5684.306	5584.507	11988.909	6355.265
BIC	5572.293	8023.541	6126.506	5811.190	5711.390	12115.792	6476.106
Log Likelihood	−2705.726	−3927.329	−2978.811	−2821.153	−2771.253	−5973.454	−3157.633
F Statistic	292.635***	503.629***	206.836***	697.194***	619.232***	17.467***	317.604***
Breusch-Pagan Test	208.465***	113.814***	163.645***	220.598***	171.725***	108.238***	417.296***

Notes: Each base model is labeled with its label of the dependent variable. †, c, and b indicate a base model including *log income squared*, *log some college or more*, and *log bachelor's degree or more* respectively. Standard errors are in parenthesis. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

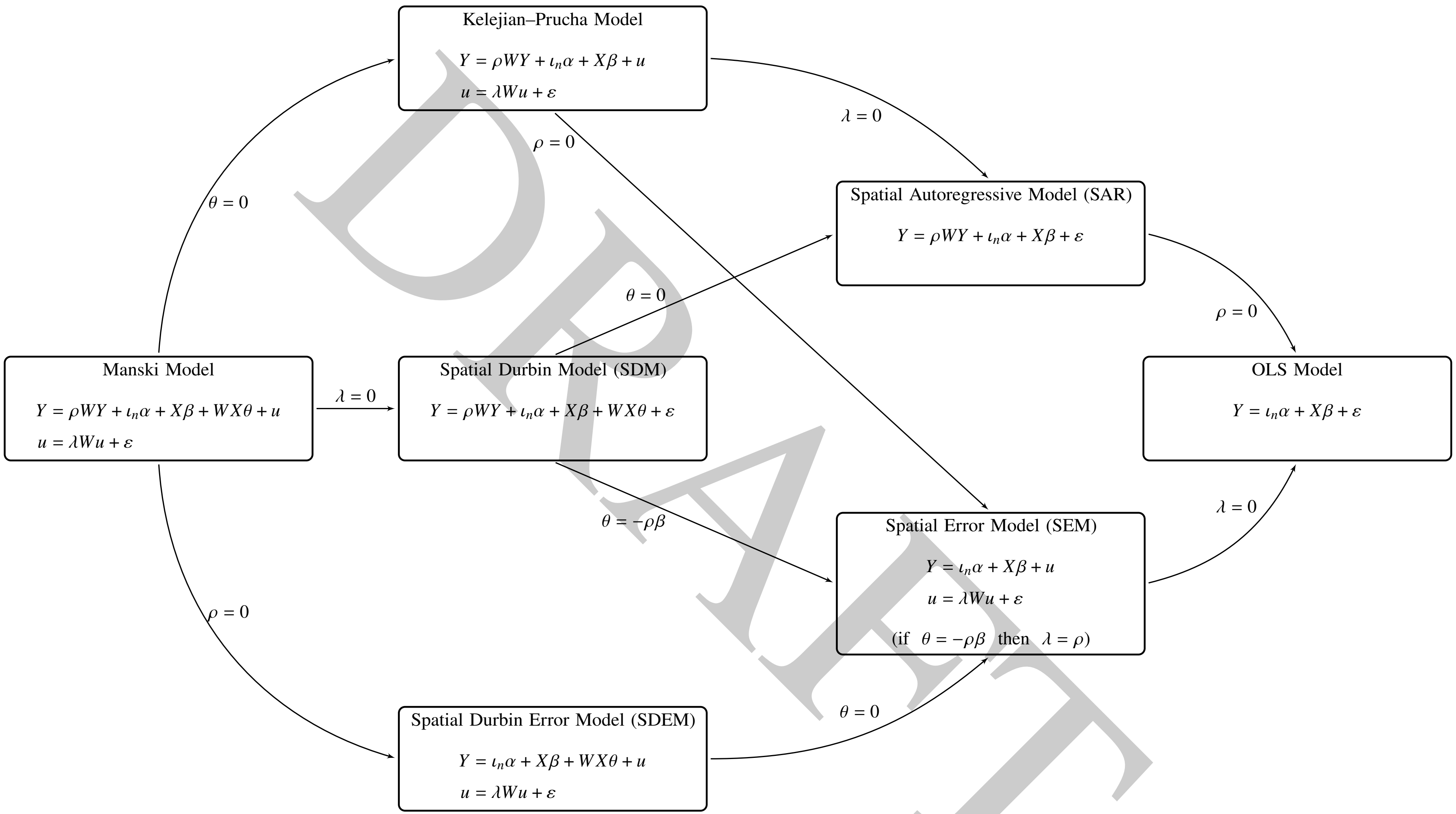


Figure 14: The Relationship Between Different Spatial Models for Cross–Section Areal Data

Table 10: LM Test Statistics for Base Model OLS by Weight Matrix

		Weight Matrix:			
Base Model	LM Test	q1.w	q2.w	k10.w	soi.w
CO^c	LM_ρ	669.22^{***}	781.92^{***}	603.52^{***}	503.71^{***}
	LM_λ	1163.96^{***}	2024.78^{***}	1388.70^{***}	841.24^{***}
	LM^*_ρ	2.08	0.13	0.11	0.51
	LM^*_λ	496.82^{***}	1243.00^{***}	785.28^{***}	338.05^{***}
${NH_3}^{\dagger,b}$	LM_ρ	865.28^{***}	680.89^{***}	490.71^{***}	821.71^{***}
	LM_λ	1805.82^{***}	2740.75^{***}	2401.75^{***}	1505.45^{***}
	LM^*_ρ	0.22	0.04	2.78^{*}	0.46
	LM^*_λ	940.76^{***}	2059.90^{***}	1913.82^{***}	684.20^{***}
${NO_x}^{\dagger,b}$	LM_ρ	461.02^{***}	500.41^{***}	302.22^{***}	383.17^{***}
	LM_λ	605.84^{***}	827.84^{***}	746.66^{***}	477.05^{***}
	LM^*_ρ	2.60	17.41^{***}	1.63	4.53^{**}
	LM^*_λ	147.41^{***}	344.84^{***}	446.08^{***}	98.41^{***}
${PM_{10}}^{\dagger,c}$	LM_ρ	1399.81^{***}	1754.14^{***}	1152.65^{***}	1252.74^{***}
	LM_λ	2976.54^{***}	6342.10^{***}	4199.04^{***}	2337.39^{***}
	LM^*_ρ	5.00^{**}	50.77^{***}	13.64^{***}	13.71^{***}
	LM^*_λ	1581.73^{***}	4638.72^{***}	3060.03^{***}	1098.37^{***}
${PM_{2.5}}^{\dagger,c}$	LM_ρ	695.69^{***}	856.25^{***}	401.22^{***}	578.04^{***}
	LM_λ	1632.45^{***}	3114.28^{***}	2020.18^{***}	1232.03^{***}
	LM^*_ρ	0.25	13.48^{***}	0.43	0.10
	LM^*_λ	937.01^{***}	2271.51^{***}	1619.39^{***}	654.08^{***}
${SO_2}^{\dagger,b}$	LM_ρ	212.06^{***}	338.68^{***}	229.11^{***}	183.21^{***}
	LM_λ	188.12^{***}	305.20^{***}	251.09^{***}	160.65^{***}
	LM^*_ρ	29.26^{***}	33.49^{***}	18.71^{***}	28.85^{***}
	LM^*_λ	5.31^{**}	0.00	40.69^{***}	6.30^{**}
VOC^b	LM_ρ	1143.36^{***}	1366.54^{***}	909.25^{***}	940.45^{***}
	LM_λ	1608.12^{***}	2683.02^{***}	1996.91^{***}	1264.99^{***}
	LM^*_ρ	22.40^{***}	65.99^{***}	30.30^{***}	23.73^{***}
	LM^*_λ	487.16^{***}	1382.47^{***}	1117.97^{***}	348.27^{***}

Notes: Each base model is labeled with its label of the dependent variable. †, ^{c}, and ^{b} indicate a base model including *log income squared*, *log some college or more*, and *log bachelor's degree or more* respectively. *q1*, *q2*, *k10*, and *soi* indicate 1st Order Queen, 2nd Order Queen, 10 Nearest Neighbors, and Sphere of Influence contiguities respectively. *w* indicates row–standardized weight style. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

Table 11: LR Test Statistics for Base Model OLS vs. Spatial Models by Weight Matrix

Base Model	LR Test	*Weight Matrix:* q1.w	q2.w	k10.w	soi.w
CO^c	H_0: OLS vs. H_1: SAR	523.64***	489.02***	472.34***	426.16***
	H_0: OLS vs. H_1: SEM	864.58***	799.68***	769.56***	668.41***
	H_0: OLS vs. H_1: SDM	919.91***	823.25***	813.88***	738.44***
	H_0: OLS vs. H_1: SDEM	922.98***	825.02***	806.55***	742.12***
${NH_3}^{\dagger,b}$	H_0: OLS vs. H_1: SAR	740.26***	513.09***	665.57***	738.97***
	H_0: OLS vs. H_1: SEM	1451.13***	1176.09***	1337.82***	1292.93***
	H_0: OLS vs. H_1: SDM	1555.42***	1237.25***	1427.31***	1422.45***
	H_0: OLS vs. H_1: SDEM	1546.70***	1242.71***	1434.31***	1412.65***
${NO_x}^{\dagger,b}$	H_0: OLS vs. H_1: SAR	363.44***	309.63***	344.64***	325.51***
	H_0: OLS vs. H_1: SEM	473.30***	408.90***	437.68***	399.00***
	H_0: OLS vs. H_1: SDM	523.32***	449.88***	476.37***	453.75***
	H_0: OLS vs. H_1: SDEM	523.59***	446.34***	472.71***	451.47***
${PM_{10}}^{\dagger,c}$	H_0: OLS vs. H_1: SAR	1185.77***	1262.01***	1219.90***	1125.02***
	H_0: OLS vs. H_1: SEM	2266.56***	2187.58***	2151.71***	1963.09***
	H_0: OLS vs. H_1: SDM	2387.05***	2223.71***	2221.14***	2086.35***
	H_0: OLS vs. H_1: SDEM	2364.45***	2216.70***	2204.32***	2057.09***
${PM_{2.5}}^{\dagger,c}$	H_0: OLS vs. H_1: SAR	570.59***	603.78***	559.22***	506.74***
	H_0: OLS vs. H_1: SEM	1170.55***	1127.91***	1059.80***	963.95***
	H_0: OLS vs. H_1: SDM	1244.40***	1153.93***	1115.57***	1035.90***
	H_0: OLS vs. H_1: SDEM	1238.39***	1148.25***	1100.72***	1029.44***
${SO_2}^{\dagger,b}$	H_0: OLS vs. H_1: SAR	172.61***	192.63***	185.41***	160.36***
	H_0: OLS vs. H_1: SEM	160.01***	184.75***	173.88***	146.79***
	H_0: OLS vs. H_1: SDM	225.41***	241.13***	233.86***	212.92***
	H_0: OLS vs. H_1: SDEM	220.37***	231.35***	230.22***	206.54***
VOC^b	H_0: OLS vs. H_1: SAR	875.90***	794.64***	807.36***	787.65***
	H_0: OLS vs. H_1: SEM	1208.87***	1059.89***	1070.01***	1015.92***
	H_0: OLS vs. H_1: SDM	1259.57***	1097.62***	1117.69***	1106.48***
	H_0: OLS vs. H_1: SDEM	1253.67***	1099.29***	1102.99***	1108.51***

Notes: Each base model is labeled with its label of the dependent variable. †, ^{c}, and ^{b} indicate a base model including *log income squared*, *log some college or more*, and *log bachelor's degree or more* respectively. *q1*, *q2*, *k10*, and *soi* indicate 1st Order Queen, 2nd Order Queen, 10 Nearest Neighbors, and Sphere of Influence contiguities respectively. *w* indicates row–standardized weight style. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

Table 12: LR Test Statistics for Nested Spatial Models by Weight Matrix

		Weight Matrix:			
Base Model	LR Test	q1.w	q2.w	k10.w	soi.w
CO^c	H_0: SAR vs. H_1: SDM	396.26***	334.23***	341.54***	312.28***
	H_0: SEM vs. H_1: SDM	55.33***	23.57	44.32***	70.03***
	H_0: SEM vs. H_1: SDEM	58.40***	25.34	36.99***	73.71***
$NH_3{}^{\dagger,b}$	H_0: SAR vs. H_1: SDM	815.16***	724.16***	761.75***	683.48***
	H_0: SEM vs. H_1: SDM	104.29***	61.16***	89.50***	129.51***
	H_0: SEM vs. H_1: SDEM	95.56***	66.62***	96.49***	119.71***
$NO_x{}^{\dagger,b}$	H_0: SAR vs. H_1: SDM	159.87***	140.25***	131.74***	128.23***
	H_0: SEM vs. H_1: SDM	50.01***	40.98***	38.70***	54.74***
	H_0: SEM vs. H_1: SDEM	50.29***	37.43***	35.03**	52.46***
$PM_{10}{}^{\dagger,c}$	H_0: SAR vs. H_1: SDM	1201.27***	961.70***	1001.25***	961.32***
	H_0: SEM vs. H_1: SDM	120.48***	36.13**	69.44***	123.26***
	H_0: SEM vs. H_1: SDEM	97.89***	29.11*	52.61***	94.00***
$PM_{2.5}{}^{\dagger,c}$	H_0: SAR vs. H_1: SDM	673.81***	550.14***	556.35***	529.16***
	H_0: SEM vs. H_1: SDM	73.86***	26.02	55.77***	71.94***
	H_0: SEM vs. H_1: SDEM	67.84***	20.35	40.92***	65.49***
$SO_2{}^{\dagger,b}$	H_0: SAR vs. H_1: SDM	52.80***	48.50***	48.45***	52.56***
	H_0: SEM vs. H_1: SDM	65.40***	56.38***	59.99***	66.13***
	H_0: SEM vs. H_1: SDEM	60.36***	46.60***	56.34***	59.75***
VOC^b	H_0: SAR vs. H_1: SDM	383.67***	302.98***	310.34***	318.83***
	H_0: SEM vs. H_1: SDM	50.70***	37.73***	47.68***	90.55***
	H_0: SEM vs. H_1: SDEM	44.80***	39.40***	32.97**	92.58***

Notes: Each base model is labeled with its label of the dependent variable. †, ^{c}, and ^{b} indicate a base model including *log income squared*, *log some college or more*, and *log bachelor's degree or more* respectively. *q1*, *q2*, *k10*, and *soi* indicate 1st Order Queen, 2nd Order Queen, 10 Nearest Neighbors, and Sphere of Influence contiguities respectively. *w* indicates row–standardized weight style. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

Table 13: AIC of Base Model OLS and Spatial Models by Weight Matrix

			Weight Matrix:			
Base Model	OLS	Spatial Model	q1.w	q2.w	k10.w	soi.w
CO^{c}	5451.0	SAR	4929.8	4964.4	4981.1	5027.3
		SEM	4588.9	4653.8	4683.9	4785.0
		SDM	4569.5	4666.2	4675.6	4751.0
		SDEM	4566.5*	4664.4	4682.9	4747.3
$NH_3{}^{\dagger,b}$	7897.0	SAR	7158.4	7385.6	7233.1	7159.7
		SEM	6447.5	6722.6	6560.8	6605.7
		SDM	6381.2*	6699.4	6509.3	6514.2
		SDEM	6390.0	6693.9	6502.4	6524.0
$NO_x{}^{\dagger,b}$	6000.0	SAR	5638.2	5692.0	5657.0	5676.1
		SEM	5528.3	5592.7	5563.9	5602.6
		SDM	5516.3	5589.7	5563.2	5585.9
		SDEM	5516.0*	5593.3	5566.9	5588.2
$PM_{10}{}^{\dagger,c}$	5684.0	SAR	4500.5	4424.3	4466.4	4561.3
		SEM	3419.7	3498.7	3534.6	3723.2
		SDM	3337.3*	3500.6	3503.2	3638.0
		SDEM	3359.9	3507.6	3520.0	3667.2
$PM_{2.5}{}^{\dagger,c}$	5585.0	SAR	5015.9	4982.7	5027.3	5079.8
		SEM	4416.0	4458.6	4526.7	4622.6
		SDM	4380.1*	4470.6	4508.9	4588.6
		SDEM	4386.1	4476.3	4523.8	4595.1
$SO_2{}^{\dagger,b}$	11989.0	SAR	11818.3	11798.3	11805.5	11830.5
		SEM	11830.9	11806.2	11817.0	11844.1
		SDM	11803.5	11787.8*	11795.0	11816.0
		SDEM	11808.5	11797.6	11798.7	11822.4
VOC^{b}	6355.0	SAR	5481.4	5562.6	5549.9	5569.6
		SEM	5148.4	5297.4	5287.3	5341.3
		SDM	5133.7*	5295.6	5275.6	5286.8
		SDEM	5139.6	5294.0	5290.3	5284.8

Notes: Each base model is labeled with its label of the dependent variable. †, ^{c}, and ^{b} indicate a base model including *log income squared*, *log some college or more*, and *log bachelor's degree or more* respectively. *q1*, *q2*, *k10*, and *soi* indicate 1st Order Queen, 2nd Order Queen, 10 Nearest Neighbors, and Sphere of Influence contiguities respectively. *w* indicates row–standardized weight style. * indicates the model with the minimum AIC among a group of OLS and spatial models with different weight matrices by each base model.

Table 14: Log Likelihood of Base Model OLS and Spatial Models by Weight Matrix

			Weight Matrix:			
Base Model	OLS	Spatial Model	q1.w	q2.w	k10.w	soi.w
CO^{c}	−2706.0	SAR	−2443.9	−2461.2	−2469.6	−2492.6
		SEM	−2273.4	−2305.9	−2320.9	−2371.5
		SDM	−2245.8	−2294.1	−2298.8	−2336.5
		SDEM	−2244.2*	−2293.2	−2302.4	−2334.7
$NH_3{}^{\dagger,b}$	−3927.0	SAR	−3557.2	−3670.8	−3594.5	−3557.8
		SEM	−3201.8	−3339.3	−3258.4	−3280.9
		SDM	−3149.6*	−3308.7	−3213.7	−3216.1
		SDEM	−3154.0	−3306.0	−3210.2	−3221.0
$NO_x{}^{\dagger,b}$	−2979.0	SAR	−2797.1	−2824.0	−2806.5	−2816.1
		SEM	−2742.2	−2774.4	−2760.0	−2779.3
		SDM	−2717.2	−2753.9	−2740.6	−2751.9
		SDEM	−2717.0*	−2755.6	−2742.5	−2753.1
$PM_{10}{}^{\dagger,c}$	−2821.0	SAR	−2228.3	−2190.1	−2211.2	−2258.6
		SEM	−1687.9	−1727.4	−1745.3	−1839.6
		SDM	−1627.6*	−1709.3	−1710.6	−1778.0
		SDEM	−1638.9	−1712.8	−1719.0	−1792.6
$PM_{2.5}{}^{\dagger,c}$	−2771.0	SAR	−2486.0	−2469.4	−2491.6	−2517.9
		SEM	−2186.0	−2207.3	−2241.4	−2289.3
		SDM	−2149.1*	−2194.3	−2213.5	−2253.3
		SDEM	−2152.1	−2197.1	−2220.9	−2256.5
$SO_2{}^{\dagger,b}$	−5973.0	SAR	−5887.1	−5877.1	−5880.7	−5893.3
		SEM	−5893.4	−5881.1	−5886.5	−5900.1
		SDM	−5860.7	−5852.9*	−5856.5	−5867.0
		SDEM	−5863.3	−5857.8	−5858.3	−5870.2
VOC^{b}	−3158.0	SAR	−2719.7	−2760.3	−2754.0	−2763.8
		SEM	−2553.2	−2627.7	−2622.6	−2649.7
		SDM	−2527.8*	−2608.8	−2598.8	−2604.4
		SDEM	−2530.8	−2608.0	−2606.1	−2603.4

Notes: Each base model is labeled with its label of the dependent variable. †, ^{c}, and ^{b} indicate a base model including *log income squared*, *log some college or more*, and *log bachelor's degree or more* respectively. *q1*, *q2*, *k10*, and *soi* indicate 1st Order Queen, 2nd Order Queen, 10 Nearest Neighbors, and Sphere of Influence contiguities respectively. *w* indicates row–standardized weight style. * indicates the model with the maximum log likelihood among a group of OLS and spatial models with different weight matrices by each base model.

Table 15: Base Model Spatial Regression Results – Direct Impacts

	Base Model Spatial Regression:						
	$CO^{c,\circ,q1}$	$NH_3^{\dagger,b,\bullet,q1}$	$NO_x^{\dagger,b,\circ,q1}$	$PM_{10}^{\dagger,c,\bullet,q1}$	$PM_{2.5}^{\dagger,c,\bullet,q1}$	$SO_2^{\dagger,b,\bullet,q2}$	$VOC^{b,\bullet,q1}$
Evangelical Protestants (%)	0.001	0.007***	0.003***	0.001*	0.002**	0.003	0.001
Black Protestants (%)	0.013***	0.000	0.019***	−0.005	0.007*	0.042***	0.013***
Mainline Protestants (%)	0.001	0.014***	0.005***	0.010***	0.008***	−0.001	0.004***
Catholics (%)	−0.001	0.004***	0.001	0.000	0.000	0.007**	0.000
Orthodox Christians (%)	0.119***	−0.060	0.094**	0.041	0.083**	0.255**	0.118***
Mormons (%)	0.003	0.007*	0.002	0.000	0.003	0.011	0.005
Muslims (%)	0.027***	−0.041***	0.016	−0.017*	0.001	0.008	0.052***
Jews (%)	0.025**	0.008	0.004	0.039***	0.026**	−0.035	0.023
Hindus (%)	−0.009	−0.062	0.006	−0.013	−0.018	−0.065	−0.010
Buddhists (%)	−0.012	0.012	−0.054	−0.046*	−0.036	−0.069	−0.007
Log Income	0.324***	9.263***	5.629*	7.465***	8.450***	19.060**	0.586***
Log Income Squared		−0.476***	−0.240	−0.371***	−0.415***	−0.882**	
Log Gas Price	−2.949*	0.311	−5.061***	4.456***	−0.378	−7.277*	−1.820
Log Gas Tax/Fee	−0.482***	0.004	−0.453**	−0.671***	−0.544***	−0.338	−0.385*
Log Renewable Energy Consumption	−0.032	0.050	0.102*	−0.012	0.015	0.396***	−0.021
Some College or More (%)	0.000			−0.001	−0.002		
Bachelor's Degree or More (%)		−0.010***	−0.013***			−0.022***	−0.011***
Log Population Density	−0.483***	−0.743***	−0.393***	−0.595***	−0.630***	−0.251***	−0.513***
Log Mean Daily Precipitation	0.280***	−0.113	−0.135	−0.027	0.257***	0.078	0.215***
Log Mean Daily Max. Heat Index	−2.750**	6.650***	7.341***	0.806	−0.254	1.507	−1.489

Table 15 (continued)

	Base Model Spatial Regression:						
	$CO^{c,\circ,q1}$	$NH_3{}^{\dagger,b,\bullet,q1}$	$NO_x{}^{\dagger,b,\circ,q1}$	$PM_{10}{}^{\dagger,c,\bullet,q1}$	$PM_{2.5}{}^{\dagger,c,\bullet,q1}$	$SO_2{}^{\dagger,b,\bullet,q2}$	$VOC^{b,\bullet,q1}$
Spatial Dependence Param: $\rho^{\bullet}$		0.724***		0.794***	0.669***	0.441***	0.684***
Spatial Dependence Param: λ°	0.620***		0.518***				
Observations	3109	3109	3109	3109	3109	3109	3109
Residual Std. Error	0.478	0.628	0.564	0.378	0.460	1.580	0.518
AIC	4566.470	6381.236	5516.029	3337.259	4380.102	11787.779	5133.694
Log Likelihood	−2244.235	−3149.618	−2717.015	−1627.630	−2149.051	−5852.889	−2527.847
LR Test Statistic for $\rho^{\bullet}$		1347.130***		2122.741***	1106.853***	121.787***	1126.268***
LR Test Statistic for λ°	819.262***		451.437***				
Global Moran's I of Residuals	−0.023	−0.016	−0.010	−0.046	−0.030	−0.008	−0.021

Notes: Each base model is labeled with its label of the dependent variable. †, ^{c}, and ^{b} indicate a base model including *log income squared*, *log some college or more*, and *log bachelor's degree or more* respectively. ^{q1} and ^{q2} indicate a model estimated with 1st Order Queen and 2nd Order Queen contiguities respectively. $^{\circ}$ and $^{\bullet}$ indicate a model estimated with Spatial Durbin Error Model (SDEM) and Spatial Durbin Model (SDM) respectively. In all models row–standardized weight style is used. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

Table 16: Base Model Spatial Regression Results – Indirect Impacts

	Base Model Spatial Regression:						
	$CO^{c,\circ,q1}$	$NH_3{}^{\dagger,b,\bullet,q1}$	$NO_x{}^{\dagger,b,\circ,q1}$	$PM_{10}{}^{\dagger,c,\bullet,q1}$	$PM_{2.5}{}^{\dagger,c,\bullet,q1}$	$SO_2{}^{\dagger,b,\bullet,q2}$	$VOC^{b,\bullet,q1}$
Evangelical Protestants (%)	−0.003	0.008	−0.004*	0.001	−0.002	−0.024***	0.000
Black Protestants (%)	0.009	−0.058***	−0.014	−0.015	0.002	0.072**	−0.011
Mainline Protestants (%)	−0.012***	0.035***	−0.002	0.008	−0.001	−0.023**	−0.023***
Catholics (%)	0.003	0.012**	0.002	0.008**	0.007**	0.005	0.007*
Orthodox Christians (%)	0.053	−0.406	0.100	−0.359	−0.083	2.962***	0.150
Mormons (%)	−0.007	0.018**	0.002	0.009	−0.004	−0.023	−0.007
Muslims (%)	0.020	−0.054	−0.008	−0.059	0.019	0.022	0.110*
Jews (%)	0.026	−0.098	0.011	0.121	0.084	−0.141	0.019
Hindus (%)	−0.102	−0.166	−0.140	−0.105	−0.191	−0.995	−0.216
Buddhists (%)	0.005	−0.019	−0.118	−0.480***	−0.211	−0.772	0.123
Log Income	−0.316	10.749	−0.390	26.380*	14.453	55.662	1.098***
Log Income Squared		−0.557	0.041	−1.426**	−0.787	−2.681	
Log Gas Price	0.296	0.576	2.013	−1.409	2.004	6.528	−2.493
Log Gas Tax/Fee	−0.031	0.599	0.118	−0.235	−0.050	−1.013	−0.426
Log Renewable Energy Consumption	0.114*	0.409***	−0.064	0.155*	0.104	−0.701***	0.121
Some College or More (%)	0.011***			0.051***	0.031***		
Bachelor's Degree or More (%)		0.018	−0.014**			−0.077**	−0.023**
Log Population Density	0.025	0.300***	0.051**	0.154***	0.075**	−0.353***	−0.054
Log Mean Daily Precipitation	−0.106	−0.644***	−0.310**	−0.702***	−0.465***	0.535	−0.219
Log Mean Daily Max. Heat Index	2.890**	2.112	−4.654***	8.564***	6.028***	−2.397	2.530

Notes: Notes in the previous table apply.

Table 17: Base Model Spatial Regression Results – Total Impacts

	Base Model Spatial Regression:						
	$CO^{c,\circ,q1}$	$NH_3^{\dagger,b,\bullet,q1}$	$NO_x^{\dagger,b,\circ,q1}$	$PM_{10}^{\dagger,c,\bullet,q1}$	$PM_{2.5}^{\dagger,c,\bullet,q1}$	$SO_2^{\dagger,b,\bullet,q2}$	$VOC^{b,\bullet,q1}$
Evangelical Protestants (%)	−0.003	0.015***	−0.001	0.003	0.001	−0.021***	0.000
Black Protestants (%)	0.022**	−0.058***	0.006	−0.020	0.009	0.115***	0.003
Mainline Protestants (%)	−0.011***	0.049***	0.003	0.018***	0.008*	−0.024**	−0.019***
Catholics (%)	0.002	0.016***	0.004	0.008*	0.007**	0.011	0.007*
Orthodox Christians (%)	0.172	−0.466	0.194	−0.317	0.000	3.217***	0.268
Mormons (%)	−0.004	0.025***	0.005	0.008	−0.002	−0.012	−0.002
Muslims (%)	0.048	−0.095	0.008	−0.076	0.020	0.030	0.161**
Jews (%)	0.050	−0.089	0.015	0.159	0.110	−0.176	0.042
Hindus (%)	−0.111	−0.228	−0.133	−0.118	−0.210	−1.061*	−0.226
Buddhists (%)	−0.007	−0.007	−0.171*	−0.526***	−0.247	−0.841	0.115
Log Income	0.008	20.011	5.239	33.844**	22.903**	74.723**	1.684***
Log Income Squared		−1.034	−0.199	−1.797**	−1.202**	−3.563*	
Log Gas Price	−2.654**	0.887	−3.048**	3.047	1.626	−0.749	−4.313**
Log Gas Tax/Fee	−0.513***	0.603	−0.335*	−0.905***	−0.594**	−1.351**	−0.810***
Log Renewable Energy Consumption	0.082*	0.459***	0.038	0.143*	0.119**	−0.305**	0.100
Some College or More (%)	0.011***			0.050***	0.029***		
Bachelor's Degree or More (%)		0.008	−0.027***			−0.098***	−0.033***
Log Population Density	−0.458***	−0.443***	−0.343***	−0.441***	−0.555***	−0.604***	−0.567***
Log Mean Daily Precipitation	0.174**	−0.757***	−0.445***	−0.730***	−0.208*	0.613**	−0.005
Log Mean Daily Max. Heat Index	0.140	8.762***	2.687***	9.370***	5.775***	−0.891	1.041

Notes: Notes in the previous table apply.

References

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Vol. 4. Springer Science & Business Media.

— (1995). “Local Indicators of Spatial Association-LISA”. *Geographical Analysis* 27.2, pp. 93–115.

— (1999). “Interactive Techniques and Exploratory Spatial Data Analysis”. *Geographical Information Systems: Principles, Techniques, Management and Applications, Eds., P. Longley, M. Goodchild, D. Maguire, and D. Rhind. Cambridge: Geoinformation Int*.

— (2001). “Spatial Econometrics”. *A Companion to Theoretical Econometrics* 310330.

Anselin, L. and S. Bao (1997). “Exploratory Spatial Data Analysis Linking Spacestat and Arcview”. In: *Recent developments in spatial analysis*. Springer, pp. 35–59.

Anselin, L., A.K. Bera, R. Florax, and M.J. Yoon (1996a). “Simple Diagnostic Tests for Spatial Dependence”. *Regional Science and Urban Economics* 26.1, pp. 77–104.

Anselin, L., M. Fisher, H. Scholten, and D. Unwin (1996b). “Spatial Analytical Perspectives on GIS”. *The Moran Scatterplot as an Esda Tool to Assess Local Instability in Spatial Association. Taylor and Francis: London*, pp. 111–125.

Anselin, L. and R. Florax (2012). *New Directions in Spatial Econometrics*. Springer Science & Business Media.

Anselin, L. and S.J. Rey (2010). “Perspectives on Spatial Data Analysis”. In: *Perspectives on Spatial Data Analysis*. Springer, pp. 1–20.

API (2010). *American Petroleum Institute (API), Motor Fuel Taxes*.

ARDA (2010). *U.S. Religion Census: Religious Congregations and Membership Study, 2010 (County File)*. URL: `https://goo.gl/1JTUY4` (visited on Jan. 30, 2017).

Assunção, R. and E. Krainski (2009). “Neighborhood Dependence in Bayesian Spatial Models”. *Biometrical Journal* 51.5, pp. 851–869.

Baller, R.D., L. Anselin, S.F. Messner, G. Deane, and D.F. Hawkins (2001). “Structural Covariates of U.S. County Homicide Rates: Incorporating Spatial Effects”. *Criminology* 39.3, pp. 561–588.

Barrett, D.B. and T.M. Johnson (2001). *World Christian Trends, AD 30-AD 2200: Interpreting the Annual Christian Megacensus*. Vol. 1. William Carey Library.

Barro, R. and R. Mccleary (2003). *Religion and Economic Growth*. Tech. rep.

Baumol, W.J. and W.E. Oates (1988). *The Theory of Environmental Policy*. Cambridge University Press.

Bavaud, F. (2010). “Models for Spatial Weights: A Systematic Look”. *Geographical Analysis* 30.2, pp. 153–171.

Bivand, R.S., W.G. Müller, and M. Reder (2009). “Power Calculations for Global and Local Moran’s I”. *Computational Statistics & Data Analysis* 53.8, pp. 2859–2872.

Bivand, R.S., E.J. Pebesma, and V. Gómez-Rubio (2008). *Applied Spatial Data Analysis with R*. Springer Science & Business Media.

Breusch, T.S. and A.R. Pagan (1979). “A Simple Test for Heteroscedasticity and Random Coefficient Variation”. *Econometrica* 47.5, p. 1287.

Britannica (2010). *Religion: Year In Review 2010, Worldwide Adherents of All Religions*. URL: `https://goo.gl/GTcw6o` (visited on Jan. 30, 2017).

Catton, W.R. and R.E. Dunlap (1978). “Environmental Sociology: A New Paradigm”. *The American Sociologist*, pp. 41–49.

Census Bureau, U.S. (2010a). *American Community Survey 2006-2010 (5-Year Estimates), Selected County Characteristics: United States*.

— (2010b). *U.S. Census 2010, Cartographic Boundary Shapefiles - Counties*. URL: `https://goo.gl/HYRYSf` (visited on Jan. 30, 2017).

— (2010c). *U.S. Census 2010, Centers of Population by County: 2010*. URL: `https://goo.gl/jUIfP6` (visited on Jan. 30, 2017).

— (2010d). *U.S. Census 2010, Selected County Characteristics: United States*.

— (2010e). *U.S. Census 2010, TIGER/Line® Shapefiles*. URL: `https://goo.gl/5Y168r` (visited on Jan. 30, 2017).

Clark, L.P., D.B. Millet, and J.D. Marshall (2014). “National Patterns in Environmental Injustice and Inequality: Outdoor NO2 Air Pollution in the United States”. *PLoS ONE* 9.4. Ed. by Y. Zhang, e94431.

Cliff, A.D. and J.K. Ord (1972). “Testing for Spatial Autocorrelation among Regression Residuals”. *Geographical Analysis* 4.3, pp. 267–284.

— (1981). *Spatial Processes: Models & Applications*. Taylor & Francis.

Cox, D.R. (1961). “Tests of Separate Families of Hypotheses”. In: *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*. Vol. 1, pp. 105–123.

— (1962). “Further Results on Tests of Separate Families of Hypotheses”. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 406–424.

Cressie, N. (1993). “Statistics for Spatial Data: Wiley Series in Probability and Statistics”. *Wiley-Interscience, New York* 15, pp. 105–209.

Davidson, R. and J.G. MacKinnon (1981). “Several Tests for Model Specification in the Presence of Alternative Hypotheses”. *Econometrica* 49.3, p. 781.

Dulal, H.B., R. Foa, and S. Knowles (2011). “Social Capital and Cross-Country Environmental Performance”. *The Journal of Environment & Development* 20.2, pp. 121–144.

Dunlap, R.E., K.D.V. Liere, A.G. Mertig, and R.E. Jones (2000). "New Trends in Measuring Environmental Attitudes: Measuring Endorsement of the New Ecological Paradigm: A Revised NEP Scale". *Journal of Social Issues* 56.3, pp. 425–442.

Eckberg, D.L. and T.J. Blocker (1989). "Varieties of Religious Involvement and Environmental Concerns: Testing the Lynn White Thesis". *Journal for the Scientific Study of Religion* 28.4, p. 509.

— (1996). "Christianity, Environmentalism, and the Theoretical Problem of Fundamentalism". *Journal for the Scientific Study of Religion* 35.4, p. 343.

EIA (2010a). *U.S. Energy Information Administration (EIA), Gasoline Prices by Formulation, Grade, Sales Type*. URL: `https://goo.gl/iM6pVF` (visited on Jan. 30, 2017).

— (2010b). *U.S. Energy Information Administration (EIA), State Energy Data System (SEDS): 1960-2015 (complete)*. URL: `https://goo.gl/7CxVSw` (visited on Jan. 30, 2017).

Eilam, E. and T. Trop (2012). "Environmental Attitudes and Environmental Behavior—which is the Horse and which is the Cart?" *Sustainability* 4.12, pp. 2210–2246.

Elhorst, J.P. (2010). "Applied Spatial Econometrics: Raising the Bar". *Spatial Economic Analysis* 5.1, pp. 9–28.

EPA (2011a). *2011 National Emissions Inventory (NEI) Data*. URL: `https://goo.gl/jspqK5` (visited on Jan. 30, 2017).

— (2011b). *2011 National Emissions Inventory (NEI), Technical Support Document*. URL: `https://goo.gl/fSE3GH` (visited on Jan. 30, 2017).

Esty, D.C. and M.E. Porter (2001). "Ranking National Environmental Regulation and Performance: A Leading Indicator of Future Competitiveness?" *The Global Competitiveness Report* 2002, pp. 78–100.

Fischer, M.M. and A. Getis (2009). *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications*. Springer Science & Business Media.

Fisher, W.D. (1958). "On Grouping for Maximum Homogeneity". *Journal of the American Statistical Association* 53.284, pp. 789–798.

Florax, R.J., H. Folmer, and S.J. Rey (2003). "Specification Searches in Spatial Econometrics: The Relevance of Hendry's Methodology". *Regional Science and Urban Economics* 33.5, pp. 557–579.

Fotheringham, A.S., C. Brunsdon, and M. Charlton (2003). *Geographically Weighted Regression*. John Wiley & Sons, Limited.

Gately, C.K., L.R. Hutyra, and I.S. Wing (2015). "Cities, Traffic, and CO2 : A Multidecadal Assessment of Trends, Drivers, and Scaling Relationships". *Proceedings of the National Academy of Sciences* 112.16, pp. 4999–5004.

Ghali, M., J.M. Krieg, and K.S. Rao (2011). "A Bayesian Extension of the J-Test for Non-Nested Hypotheses". *Journal of Quantitative Economics* 9.1, pp. 53–72.

Godfrey, L.G. and M.H. Pesaran (1983). "Tests of Non-Nested Regression Models". *Journal of Econometrics* 21.1, pp. 133–154.

Gore, A. (2006). *An Inconvenient Truth: The Planetary Emergency of Global Warming and What We Can Do about it*. Rodale.

Grammich, C., K. Hadaway, R. Houseal, D.E. Jones, A. Krindatch, R. Stanley, and R.H. Taylor (2012). *2010 U.S. Religion Census: Religious Congregations & Membership Study*.

Greeley, A. (1993). "Religion and Attitudes toward the Environment". *Journal for the Scientific Study of Religion* 32.1, p. 19.

Greene, W.H. (2003). *Econometric Analysis*. Pearson Education India.

Grossman, G. and A. Krueger (1994). *Economic Growth and the Environment*. Tech. rep.

Gudipudi, R., T. Fluschnik, A.G.C. Ros, C. Walther, and J.P. Kropp (2016). "City Density and CO2 Efficiency". *Energy Policy* 91, pp. 352–361.

Hand, C.M. and K.D.V. Liere (1984). "Religion, Mastery-over-Nature, and Environmental Concern". *Social Forces* 63.2, p. 555.

Hayes, B.G. and M. Marangudakis (2001). "Religion and Attitudes Towards Nature in Britain". *British Journal of Sociology* 52.1, pp. 139–155.

Hoeting, J.A., R.A. Davis, A.A. Merton, and S.E. Thompson (2006). "Model Selection for Geostatistical Models". *Ecological Applications* 16.1, pp. 87–98.

Hsu, A., N. Alexandre, S. Cohen, P. Jao, E. Khusainova, and D. Mosteller (2016). *2016 Environmental Performance Index. New Haven, CT: Yale University*.

Işık, O. and M.M. Pınarcıoğlu (2007). "Geographies of a Silent Transition: A Geographically Weighted Regression Approach to Regional Fertility Differences in Turkey". *European Journal of Population / Revue Européenne De Démographie* 22.4, pp. 399–421.

Jones, C. and D.M. Kammen (2014). "Spatial Distribution of U.S. Household Carbon Footprints Reveals Suburbanization Undermines Greenhouse Gas Benefits of Urban Population Density". *Environmental Science & Technology* 48.2, pp. 895–902.

Kanagy, C.L. and H.M. Nelsen (1995). "Religion and Environmental Concern: Challenging the Dominant Assumptions". *Review of Religious Research* 37.1, p. 33.

Kelejian, H.H. and I.R. Prucha (1998). "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances". *The Journal of Real Estate Finance and Economics* 17.1, pp. 99–121.

— (2010). "Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances". *Journal of Econometrics* 157.1, pp. 53–67.

Kissling, W.D. and G. Carl (2007). "Spatial Autocorrelation and the Selection of Simultaneous Autoregressive Models". *Global Ecology and Biogeography* 0.0, 070618060123007–???

Koenker, R. (1981). "A Note on Studentizing a Test for Heteroscedasticity". *Journal of Econometrics* 17.1, pp. 107–112.

Lamla, M.J. (2009). "Long-Run Determinants of Pollution: A Robustness Analysis". *Ecological Economics* 69.1, pp. 135–144.

Land, K.C., G. Deane, and J.R. Blau (1991). "Religious Pluralism and Church Membership: A Spatial Diffusion Model". *American Sociological Review* 56.2, p. 237.

Lee, L.-F. (2007). "Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors and Fixed Effects". *Journal of Econometrics* 140.2, pp. 333–374.

Leenders, R.T. (2002). "Modeling Social Influence through Network Autocorrelation: Constructing the Weight Matrix". *Social Networks* 24.1, pp. 21–47.

Lesage, J.P. (1997). "Bayesian Estimation of Spatial Autoregressive Models". *International Regional Science Review* 20.1-2, pp. 113–129.

Lesage, J.P. and R.K. Pace (2008). "An Introduction to Spatial Econometrics". *Revue Déconomie Industrielle* 3, pp. 19–44.

Macionis, J.J. (2011). *Society: The Basics*. Pearson.

Makido, Y., S. Dhakal, and Y. Yamagata (2012). "Relationship between Urban Form and CO2 Emissions: Evidence from Fifty Japanese Cities". *Urban Climate* 2, pp. 55–67.

Manski, C.F. (1993). "Identification of Endogenous Social Effects: The Reflection Problem". *Review of Economic Studies* 60.3, p. 531.

Moran, P.A.P. (1950). "Notes on Continuous Stochastic Phenomena". *Biometrika* 37.1/2, p. 17.

Nasr, S.H. and W.C. Chittick (2007). *The Essential Seyyed Hossein Nasr*. World Wisdom, Inc.

NLDAS (2010a). *North America Land Data Assimilation System (NLDAS) Daily Air Temperatures and Heat Index, years 1979-2011 on CDC WONDER Online Database, released 2012*. URL: https://goo.gl/NDwzvv (visited on Jan. 30, 2017).

— (2010b). *North America Land Data Assimilation System (NLDAS) Daily Precipitation years 1979-2011 on CDC WONDER Online Database, released 2012*. URL: https://goo.gl/WVcFZ1 (visited on Jan. 30, 2017).

Ohlan, R. (2015). "The Impact of Population Density, Energy Consumption, Economic Growth and Trade Openness on CO2 Emissions in India". *Nat Hazards* 79.2, pp. 1409–1428.

Openshaw, S. (1984). "The Modifiable Areal Unit Problem". In: Geo Abstracts University of East Anglia.

Ord, J.K. (1975). "Estimation Methods for Models of Spatial Interaction". *Journal of the American Statistical Association* 70.349, pp. 120–126.

Papyrakis, E. (2012). "Environmental Performance in Socially Fragmented Countries". *Environmental and Resource Economics* 55.1, pp. 119–140.

Pesaran, M.H. (1974). "On the General Problem of Model Selection". *The Review of Economic Studies* 41.2, p. 153.

Pesaran, M.H. and A.S. Deaton (1978). "Testing Non-Nested Nonlinear Regression Models". *Econometrica* 46.3, p. 677.

Rinpoche, L.N. (1986). *The Assisi Declarations: Messages on Man and Nature from Buddhism, Christianity, Hinduism, Islam and Judaism*. WWF.

Sapienza, P., L. Zingales, and L. Guiso (2006). *Does Culture Affect Economic Outcomes?* Tech. rep.

Schultz, P.W., L. Zelezny, and N.J. Dalrymple (2000). "A Multinational Perspective on the Relation between Judeo-Christian Religious Beliefs and Attitudes of Environmental Concern". *Environment and Behavior* 32.4, pp. 576–591.

Seldadyo, H., J.P. Elhorst, and J.D. Haan (2010). "Geography and Governance: Does Space Matter?" *Papers in Regional Science* 89.3, pp. 625–640.

Social Explorer (2010a). *American Community Survey 2006-2010 (5-Year Estimates), Selected County Characteristics: United States, Prepared by Social Explorer*. URL: `https://goo.gl/SY0c3c` (visited on Jan. 30, 2017).

— (2010b). *U.S. Census 2010, Selected County Characteristics: United States, Prepared by Social Explorer*. URL: `https://goo.gl/otNgIP` (visited on Jan. 30, 2017).

Tabachnick, B.G., L.S. Fidell, and S.J. Osterlind (2001). "Using Multivariate Statistics".

Tiefelsdorf, M., D.A. Griffith, and B. Boots (1999). "A Variance-Stabilizing Coding Scheme for Spatial Link Matrices". *Environment and Planning a* 31.1, pp. 165–180.

Tiefelsdorf, M. (1998). "Some Practical Applications of Moran's I's Exact Conditional Distribution". *Papers in Regional Science* 77.2, pp. 101–129.

— (2002). "The Saddlepoint Approximation of Moran's I's and Local Moran's Ii's Reference Distributions and Their Numerical Evaluation". *Geographical Analysis* 34.3, pp. 187–206.

Tobler, W.R. (1970). "A Computer Movie Simulating Urban Growth in the Detroit Region". *Economic Geography* 46, p. 234.

TPC (2010). *Tax Policy Center (TPC), State Taxes on Gasoline and Diesel, 2010-2011*. URL: `https://goo.gl/tmYkWt` (visited on Jan. 30, 2017).

Vonk, M. (2012). "Sustainability, Values and Quality of Life What We Can Learn from Christian Communities". *Philosophia Reformata* 77.2, pp. 114–134.

West, S.G., J.F. Finch, and P.J. Curran (1995). "Structural Equation Models with Nonnormal Variables: Problems and Remedies".

White, L. (1967). "The Historical Roots of Our Ecologic Crisis". *American Association for the Advancement of Science* 155.3767, pp. 1203–1207.

Wolkomir, M.J., M. Futreal, E. Woodrum, and T. Hoban (1997). "Substantive Religious Belief and Environmentalism". *Social Science Quarterly*, pp. 96–108.

Woodrum, E. and M.J. Wolkomir (1997). "Religious Effects on Environmentalism". *Sociological Spectrum* 17.2, pp. 223–234.

Zhukov, Y.M. and B.M. Stewart (2012). "Choosing Your Neighbors: Networks of Diffusion in International Relations". *International Studies Quarterly : A Publication of the International Studies Association* 57.2, pp. 271–287.

Appendix

A Areal Unit Selection

In the spatial econometrics literature, data with spatial dimension is categorized into three groups: spatial point process, geostatistical data, and areal data (Bivand et al., 2008). The present study uses areal data in which observations are distributed into prespecified areal units. The prespecified areal units selected for this study is U.S counties in 2010. County is the arbitrary unit of analysis and might raise Modifiable Areal Unit Problem (MAUP)[91] (Openshaw, 1984). The MAUP is related to the fact that statistical measures for cross–sectional data are sensitive to the way how areal units are aggregated. Specifically, the level of aggregation and spatial arrangements of areal units affect the magnitude of various measures such as spatial autocorrelation coefficients and parameters in spatial regressions. For instance, counties may be too large as an areal unit of spatial analysis (e.g., for detecting the spatial autocorrelation) and the unobserved heterogeneity may create the MAUP. The problem of aggregation can also work in the opposite direction. If the spatial autocorrelation in data is a regional phenomenon, then aggregating data by counties will produce spatial autocorrelation not due to spatial interaction, but due to the counties with common local observations in the same region. In such a case, counties are too small as an areal unit.

The selection of areal unit should ideally be determined by theoretical considerations. However, data availability imposes severe constraints in practice as it is the case in this study. Thus, the present study uses counties as an areal unit for several reasons. First, the sample size is significantly augmented relative to using states or regions as an areal unit. Second, county is the smallest areal unit for performing spatial analysis with the necessary data available. Lastly, a county–level analysis can encompass the entire contiguous U.S. by including not only the urban areas but also the rural regions.

B Contiguity Matrix

In the literature, contiguities from the adjacency category are extensively used since they make substantive sense for the units with common borders being neighbors. In 1$^{\text{st}}$ Order Queen contiguity (i.e., the union of Rook and Bishop contiguities), two areal units are neighbors if they share any part of a shared border or vertex. This study also uses a higher order of Queen contiguity (i.e., 2$^{\text{nd}}$ Order Queen contiguity) in which second–order neighbor is a first order neighbor of a first order neighbor.

From interpoint distance contiguity category, Minimum Distance contiguity is used which

[91]For more information about MAUP and thorough literature review of empirical studies, see Anselin (1988) and Fotheringham et al. (2003).

ensures that every areal unit has at least one neighbor. Here, the neighbors of areal unit i are defined through interpoint distance. Interpoint distance specification defines $c_{ij} = 1$ if point j is located within a specified search distance of point i. In this case, the lower bound of the distance is equal to zero, and the upper bound is equal to the maximum first nearest neighbor distance. Thus, the most isolated unit in data will have one neighbor, and the rest of the units will have as many neighbors as can be found within the defined interpoint distance (Zhukov and Stewart, 2012). Minimum Distance contiguity is prone to capturing an excessive number of neighbors in regions with a high density of units such as south–east and north–east of U.S.

Another variety of interpoint distance contiguity employed is K Nearest Neighbors (KNN) contiguity. By estimating search distances individually for each areal unit, this method ensures that all units have the same number of neighbors while avoiding much of the noise associated with Minimum Distance contiguity. The disadvantage of KNN is that it produces asymmetric neighbors (i.e., unit i is a neighbor of unit j, but not vice–versa), and it may not reflect the true level of contiguity. The present study employs 6NN and 10NN contiguities. Since U.S. counties have five-to-six contiguous neighbors on average, 6NN contiguity is selected.[92] Moreover, using 10NN contiguity yields a ring around each county by roughly covering the first and second order contiguous neighbors.

The final contiguity employed is Sphere of Influence contiguity which uses graph–based methods. It is best suited for irregularly located areal units such as U.S. counties. Sphere of Influence contiguity is commonly used in the field of international relations to measure a country's level and range of cultural, economic, military, and politic influence over other nations. Like Minimum Distance contiguity, the Sphere of Influence contiguity ensures that all units have at least one neighbor. On the other hand, unlike KNN contiguities, the number of links per areal unit can vary, and also relatively long distance links are avoided. One disadvantage of Sphere of Influence contiguity is that it uses Euclidean distance but not great circle distance[93] as Minimum Distance and KNN contiguities. Moreover, inferences made from Sphere of Influence contiguity are not as immediately intuitive as the other approaches. For detailed information, see Zhukov and Stewart (2012).

Constructing Minimum Distance, 6NN, 10NN, and Sphere of Influence contiguities require specifying interpoints for distance calculation[94], where the interpoints may be geographical centroids, population centroids or any other theoretically–relevant set of coordinates. Since the population is clustered on a point very close to borders in some counties (e.g., counties in Arizona, New Mexico, and southern California), population centroids are used instead of geographical centroids.

[92]Average number of neighbors in 1st Order Queen Contiguity is 5.94.

[93]The distance along Earth's surface. See Seldadyo et al. (2010) for calculations.

[94]Great circle distance is used for constructing Minimum Distance, 6NN, and 10NN contiguities.

C Weight Matrix

The simplest transformation of a binary contiguity matrix **C** is no transformation at all. The matrix **C** can be left untouched and becomes the final weight matrix **W**. This approach is called binary weight style (B) which emphasizes strongly connected areal units and weights up areal units with many neighbors compared to those with few.

The matrix **C** is often row–standardized (W) such that each element in a row is divided by the sum of the elements in the row, and each row sums to unity:

$$w_{ij} = \frac{c_{ij}}{\sum_{j=1}^{n} c_{ij}} \quad \text{and} \quad \sum_{j=1}^{n} w_{ij} = 1 \quad \text{where} \quad i, j = 1, 2, \ldots, n$$

As row–standardization implies that spatially lagged observed values are averages over the sets of neighbors for areal unit i, it is very intuitive and commonly used in the literature. The weight style W emphasizes weekly connected areal units: the fewer the neighbors of areal unit i, the stronger their individual influence.

Global–standardized weight style (C), which emphasizes strongly connected areal units, is very similar to row–standardization, except it is standardized such that all elements in matrix **C** sum to n. Variance–stabilizing weight style (S), proposed by Tiefelsdorf et al. (1999), sums over all links to n and tries to create a balance between weight style W and C. The minmax–normalized weight style (MINMAX), based on Kelejian and Prucha (2010), divides the elements of matrix **C** by the minimum of the maximum row sums and maximum column sums. It is similar to the weight style C.

D Moran's I Test

The normality assumption assumes that observations of the variable of interest follow an iid Gaussian distribution. In the randomization, values are randomly assigned to areal units, then the Moran's I statistic is computed. This procedure is repeated several times to establish a simulated distribution. Then, the observed value of Moran's I statistic is compared to the simulated distribution to see how likely it is that the observed values could be considered from a random draw.[95] In the Monte Carlo simulation, observed values are randomly assigned to areal units. Then, the Moran's I statistic is computed by using n random permutations of the variable of interest to establish the rank of the observed Moran's I statistic in relation to the n simulated values.[96] Anselin and Florax (2012) state that it performs quite well even in the small sample

[95] The randomization differs from the normality assumption by introducing a correction term on the kurtosis of the variable of interest. As the variable departs from normality, the randomization compensates it by increasing the variance and decreasing the standard deviation.

[96] Since the rank of the observed statistic is computed relative to the reference distribution of Moran's I statistic for the permuted data, pseudo p–values are used in inference. The pseudo p–value of a statistic is the ratio of the

sizes.

Some other versions of Moran's I tests (e.g., computationally intensive Exact (Tiefelsdorf, 1998) and Saddlepoint approximation (Tiefelsdorf, 2002) methods) can also be used for spatial autocorrelation testing. However, as stated in Bivand et al. (2009), these tests make little difference to outcomes for spatial autocorrelation tests when the number of areal units is not small. Therefore, these methods for Moran's I tests are not considered.

D.1 Theoretical Moments of Global Moran's I Statistic

Under the null hypothesis of no spatial autocorrelation, the first moment of Global Moran's I statistics (i.e., mean) does not change with normality assumption and randomization. It depends only on the number of areal units n. The expected value, which converges to zero as n increases, is

$$E(I) = \frac{-1}{n-1}$$

However, the second moment (i.e., variance) changes with the normality assumption and randomization (Cliff and Ord, 1972, 1981). When the normality is assumed, the variance only depends on the characteristics of weight matrix $\mathbf{W}$, and it does not get changed by the variable of interest. On the other hand, in the randomization, variance also depends on the variable of interest. Specifically, the variance with the normality assumption is

$$V(I) = \frac{n^2 S_1 - nS_2 + 3S_0^2}{(n-1)(n+1)S_0^2} - E(I)^2$$

and with the randomization, it is

$$V(I) = \frac{n\left[(n_2 - 3n + 3)S_1 - nS_2 + 3S_0^2\right] - S_3\left[(n^2 - n)S_1 - 2nS_2 + 6S_0^2\right]}{(n-1)(n-2)(n-3)S_0^2} - E(I)^2$$

where

$$S_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \qquad S_1 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(w_{ij} + w_{ji})^2$$

$$S_2 = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} w_{ij} + \sum_{j=1}^{n} w_{ji}\right)^2 \qquad S_3 = \frac{n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^2}$$

number of values greater than or equal to observed statistic plus one and the number of simulations.

Under the null hypothesis of no spatial autocorrelation, Global Moran's I statistic is asymptotically normal. So, the statistic I becomes asymptotically standard normal, such that

$$I^* = \frac{I - E(I)}{\sqrt{V(I)}}$$

and the standardized value of the statistic I^* can be tested against a standard normal distribution.

E Local Moran's I Test

As with the Global Moran's I, hypothesis test for Local Moran's I can be performed with normality assumption and randomization. The normality assumption can be problematic since the number of neighbors of each areal unit is often very small. Moreover, randomization might give incorrect results since the mean of Local Moran's I keeps changing for one specific location during the permutation, which is not the case for global one. Therefore, following Anselin (1995), Local Moran's I statistic is generated under conditional randomization in the present study. Also, since Local Moran's I of every areal unit correlates with one another due to overlapping neighbors, Bonferroni adjustment (i.e., p–value adjustment for multiple tests) is used to acquire robust testing results.

F Exploratory Spatial Data Analysis

Exploratory Spatial Data Analysis (ESDA) is a critical first step for visualizing spatial patterns in the raw data, and for identifying spatial outliers, clusters, and hot spots. Moreover, ESDA can be used as a phase of analysis in which hypotheses proposed and spatial models suggested. For more information about ESDA techniques, see Anselin (1999), Fischer and Getis (2009), and Anselin and Rey (2010).

G Data Classification Methods

In the thematic mapping, all dependent and explanatory variables are classified either by quantile or Fisher data classification methods. Intervals for all *emission* variables, *population density*, and *mean daily precipitation* are determined by using quantile data classification method, which classifies data into a certain number of categories with an equal number of units in each category. However, the Fisher classification method is employed to better visualize the differences across counties for all religion and education variables, and *income*, *mean daily maximum heat index*, *gas price*, *gas tax/fee*, and *renewable energy consumption*. The Fisher classification, also known as natural breaks classification, uses the algorithm proposed by Fisher (1958). It tries to reduce the variance within classes and maximize the variance between classes. In the thematic mapping of the subsequent sections, Fisher classification method is used where it is necessary to classify groups more clearly. Otherwise, quantile method is employed.

H Interpretation of Log-Log and Log-Linear Models

Consider the simple linear regression with a constant term and two explanatory variables such as

$$y = \beta_0 + \beta_1 x_1 + \beta_1 x_2 \tag{21}$$

The standard interpretation of a regression parameter β_1 is that one unit change in the explanatory variable x_1 results in β_1 units change in the dependent variable y while holding all other variables fixed. That is

$$\frac{\Delta y}{\Delta x_1} = \beta_1 \tag{22}$$

However, interpreting parameter estimates is not always straightforward in a linear regression when some or all variables are in the natural logarithmic form. For instance, consider that the dependent and the first explanatory variables in Eq. 21 are transformed with natural logarithm. Then, the new regression equation will be

$$\ln y = \beta_0 + \beta_1 \ln x_1 + \beta_1 x_2 \tag{23}$$

In Eq. 23, if we follow the standard interpretation for parameter estimates β_1 and β_2, we will get

$$\frac{\Delta \ln y}{\Delta \ln x_1} = \beta_1 \tag{24}$$

$$\frac{\Delta \ln y}{\Delta x_2} = \beta_2 \tag{25}$$

As seen from Eq. 24 and Eq. 25, β_1 only gives us the change in $\ln y$ when there is a change in $\ln x_1$, and β_2 gives the change in $\ln y$ when there is a change in x_2. However, in most of the empirical studies, researchers are interested in the change when the original scale of a variable is used not the logarithmic scale.

To acquire an interpretation in original scale, we take advantage of the following logarithmic approximation (i.e., the difference in logs can be used to approximate proportionate changes). Let v_0 and v_1 be positive values. Then, for small changes in v, it can be shown that

$$\ln v_1 - \ln v_0 \approx (v_1 - v_0)/v_0 \approx \Delta v / v_0 \tag{26}$$

Multiplying Eq. 26 by 100 and using $\Delta \ln v = \ln v_1 - \ln v_0$ gives

$$100 \cdot \Delta \ln v \approx \%\Delta v \tag{27}$$

Plugging Eq. 27 into Eq. 24 and Eq. 25, and simplifying gives

$$\%\Delta y = \%\beta_1 \Delta x_1 \tag{28}$$

$$\%\Delta y = (100 \cdot \beta_2)\Delta x_2 \tag{29}$$

As a result, when both the dependent and explanatory variables are in the natural logarithmic form, the interpretation is that 1% change in x_1 is expected to change y by approximately $\%\beta_1$ while holding other variables fixed (see Eq. 28). On the other hand, when the dependent variable is in the natural logarithmic form but independent variable is not, then the interpretation is that 1 point change in x_2 is expected to change y by approximately $\%100 \cdot \beta_2$ while holding other variables fixed (see Eq. 29).

I ESDA for Base Model Spatial Regressions

This section presents some ESDA thematic mapping for all base model spatial regressions used in interpretation.

I.1 Correlation

As a part of ESDA, Figure 19 illustrates the spatial correlation of dependent variables for all counties with Wake County, NC by base model spatial regression while spatial dependence is taken into account.[97] The variance–covariance matrix of the dependent variable and disturbances for SDM and SDEM is defined respectively as

$$\hat{\Omega} = \hat{\sigma}^2\left[(I_n - \hat{\rho}W)'(I_n - \hat{\rho}W)\right]^{-1} \qquad \text{(for SDM)}$$

$$\hat{\Omega} = \hat{\sigma}^2\left[(I_n - \hat{\lambda}W)'(I_n - \hat{\lambda}W)\right]^{-1} \qquad \text{(for SDEM)}$$

I.2 Impacts From Measure

Figure 20 through Figure 37 illustrates the impacts on each areal unit from changing an explanatory variable by an amount in Wake County, NC by base model spatial regression. Specifically, these figures show the direct impact on Wake County and also the indirect impacts on all other counties from changing an explanatory variable in Wake County. These impacts are called *impacts from* in this study. The importance of *impacts from* is that they allow us to examine how a change in one areal unit spreads (i.e., spatial spillover) over other areal units.

[97]For thematic mapping, only the counties in North Carolina and Virginia are used since the spatial correlation of dependent variables for other counties with Wake County, NC becomes negligibly small in all base model spatial regressions.

The calculation of *impacts from* is very similar to the computation of *average total impact from an observation* explained in Section 8.7. Here, instead of taking an average of each column, only the column associated with Wake County is used. Therefore, to acquire *impacts from* of Wake County, the $n \times n$ matrix $\mathbf{S_r}(\mathbf{W})$ is calculated for each explanatory variable in all models, and then the column associated with Wake County is extracted.[98]

In these figures, the previously explained interpretation techniques are used while calculating the *impacts from*. For instance, impacts estimates are multiplied by 100 in Figure 20 but not in Figure 37. Also, in some base model spatial regressions, calculating *impacts from* of *income* needs special care since both the *log income* and *log income squared* are included into the model. In such models, the *impacts from* of *income* is calculated through an analogy from linear regressions.

In linear regressions, if both the linear and quadratic form of an explanatory variable are included in model, change in the dependent variable y with respect to change in the explanatory variable x is approximated by

$$\frac{\Delta y}{\Delta x} = \beta_1 + 2\beta_2 x \tag{30}$$

where β_1 and β_2 are the coefficient estimates of linear and quadratic form explanatory variables respectively. As it can be seen from Eq. 30, this approximation depends on a selected measure of x which can be sample mean, median, and lower or upper quartiles. By using the same approximation, the *impacts from* of *income* is calculated as

$$S_1(W) + 2S_2(W)\Theta_1 \tag{31}$$

where $S_1(W)$ and $S_2(W)$ are *impacts from* of *log income* and *log income squared* respectively; and Θ_1 is the lower quartile of *log income*.

In Figure 34, *impacts from* estimates of education variable are calculated by using the appropriate education variable (i.e., either *log some college or more* or *log bachelor's degree or more*) for each base model spatial regression.

J R Version Information

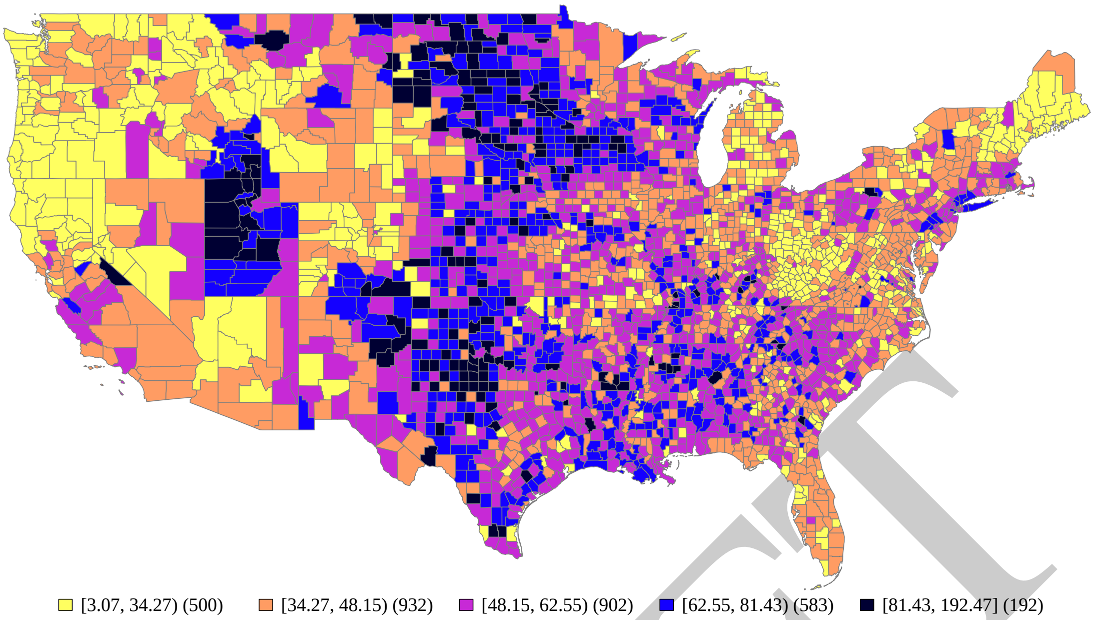
- R version 3.3.3 (2017-03-06), `x86_64-apple-darwin13.4.0`
- Locale: `en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8`
- Base packages: base, datasets, graphics, grDevices, grid, methods, stats, utils
- Other packages: acs 2.1.2, boot 1.3-20, car 2.1-6, choroplethr 3.6.1, choroplethrMaps 1.0.1, classInt 0.1-24, cowplot 0.9.2, devtools 1.13.4, dotCall64 0.9-5,

[98]For thematic mapping, only the counties in North Carolina and Virginia are used since the impacts fall on other counties become negligibly small in all base model spatial regressions.

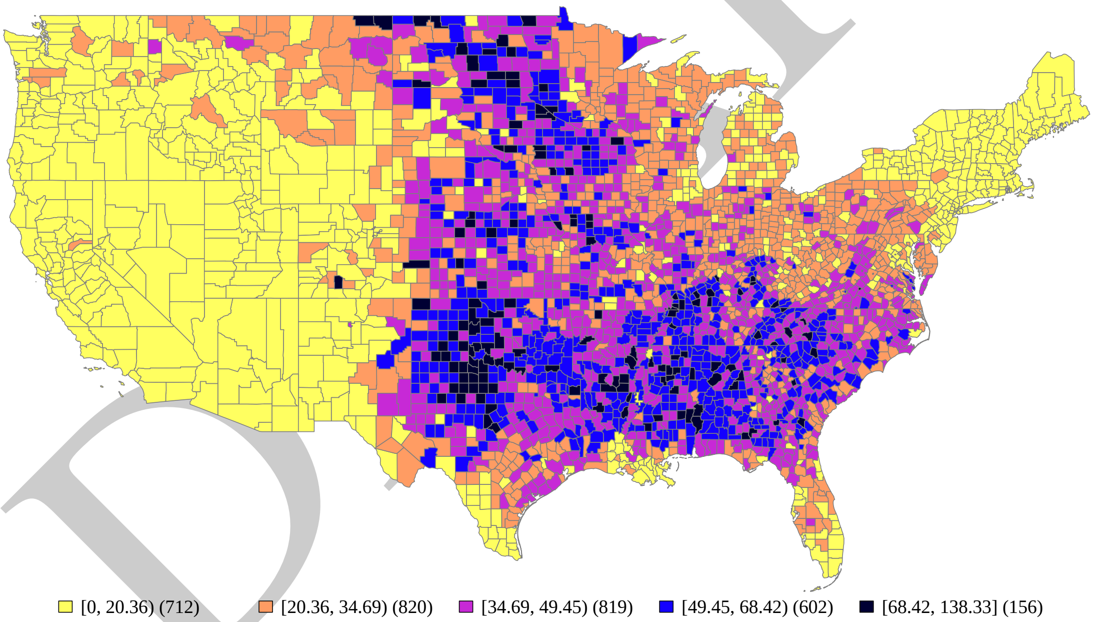
dplyr 0.7.4, fields 9.0, Formula 1.2-2, geoR 1.7-5.2, ggplot2 2.2.1.9000, gridExtra 2.3, gstat 1.1-5, gtable 0.2.0, gvlma 1.0.0.2, Hmisc 4.1-1, knitr 1.19, latex2exp 0.4.0, lattice 0.20-35, latticeExtra 0.6-28, lmtest 0.9-35, magrittr 1.5, maps 3.2.0, maptools 0.9-2, Matrix 1.2-11, moments 0.14, NCmisc 1.1.5, normtest 1.1, nortest 1.0-4, pastecs 1.3-18, pgirmess 1.6.7, plyr 1.8.4, png 0.1-7, proto 1.0.0, RANN 2.5.1, raster 2.6-7, RColorBrewer 1.1-2, reshape 0.8.7, reshape2 1.4.3, rgdal 1.2-16, sandwich 2.4-0, sp 1.2-6, spam 2.1-2, spData 0.2.7.0, spdep 0.7-4, spgwr 0.6-32, stargazer 5.2.1, stringi 1.1.6, stringr 1.2.0, survival 2.41-3, tidyr 0.7.2, tikzDevice 0.10-1, tmap 1.11, XLConnect 0.2-14, XLConnectJars 0.2-14, XML 3.98-1.9, xtable 1.8-2, zoo 1.8-1

- Loaded via a namespace (and not attached): acepack 1.4.1, assertthat 0.2.0, backports 1.1.2, base64enc 0.1-3, bindr 0.1, bindrcpp 0.2, bitops 1.0-6, checkmate 1.8.5, class 7.3-14, cluster 2.0.6, coda 0.19-1, codetools 0.2-15, colorspace 1.3-2, crosstalk 1.0.0, curl 3.1, data.table 1.10.4-3, DBI 0.7, deldir 0.1-14, dichromat 2.0-0, digest 0.6.15, e1071 1.6-8, expm 0.999-2, filehash 2.4-1, FNN 1.1, foreach 1.4.4, foreign 0.8-69, gdalUtils 2.0.1.7, gdata 2.18.0, geojsonlint 0.2.0, geosphere 1.5-7, ggmap 2.6.1, glue 1.2.0, gmodels 2.16.2, gtools 3.5.0, htmlTable 1.11.2, htmltools 0.3.6, htmlwidgets 1.0, httpuv 1.3.5, httr 1.3.1, intervals 0.15.1, iterators 1.0.9, jpeg 0.1-8, jsonlite 1.5, jsonvalidate 1.0.0, KernSmooth 2.23-15, lazyeval 0.2.1, leaflet 1.1.0, LearnBayes 2.15, lme4 1.1-15, mapproj 1.2-5, mapview 2.2.0, MASS 7.3-47, MatrixModels 0.4-1, memoise 1.1.0, mgcv 1.8-22, mime 0.5, minqa 1.2.4, munsell 0.4.3, nlme 3.1-131, nloptr 1.0.4, nnet 7.3-12, osmar 1.1-7, parallel 3.3.3, pbkrtest 0.4-7, pkgconfig 2.0.1, proftools 0.99-2, purrr 0.2.4, quantreg 5.34, R.methodsS3 1.7.1, R.oo 1.21.0, R.utils 2.6.0, R6 2.2.2, RandomFields 3.1.50, RandomFieldsUtils 0.3.25, rappdirs 0.3.1, Rcpp 0.12.14, RCurl 1.95-4.10, rgeos 0.3-26, RgoogleMaps 1.4.1, rJava 0.9-9, rjson 0.2.15, rlang 0.1.6, rmapshaper 0.3.0, rpart 4.1-11, rstudioapi 0.7, satellite 1.0.1, scales 0.5.0.9000, sf 0.5-5, shiny 1.0.5, spacetime 1.2-1, SparseM 1.77, splancs 2.01-40, splines 3.3.3, stats4 3.3.3, tcltk 3.3.3, tibble 1.3.4, tigris 0.6.2, tmaptools 1.2-2, tools 3.3.3, udunits2 0.13, units 0.5-1, uuid 0.1-2, V8 1.5, viridisLite 0.3.0, WDI 2.4, webshot 0.5.0, withr 2.1.1, xts 0.10-1, yaml 2.1.16

K Additional Tables and Figures

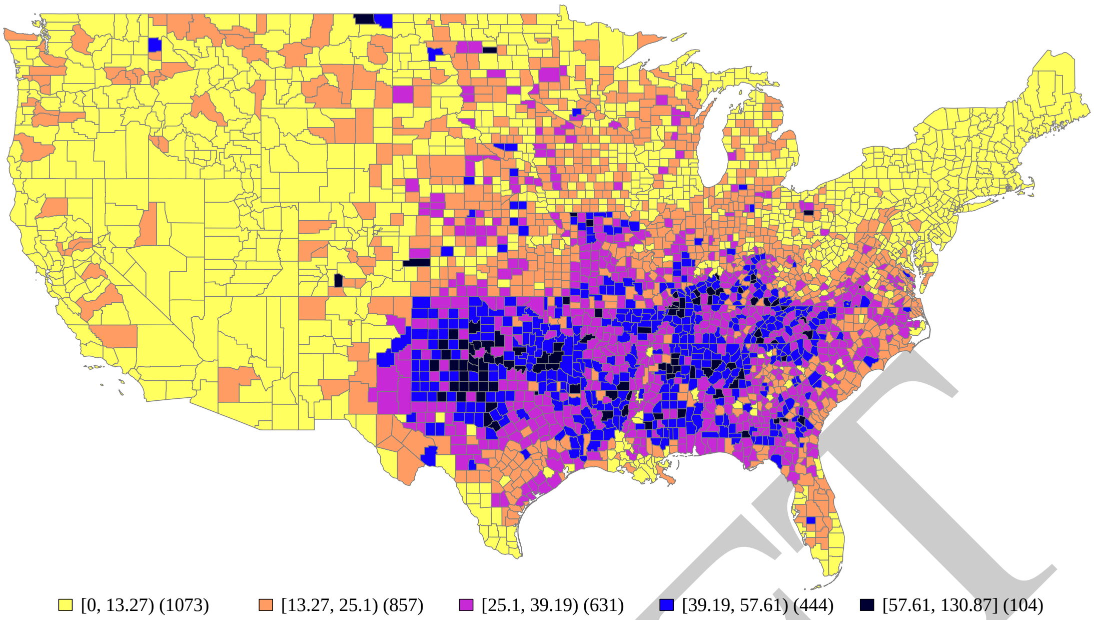


(a) Adherents (%)

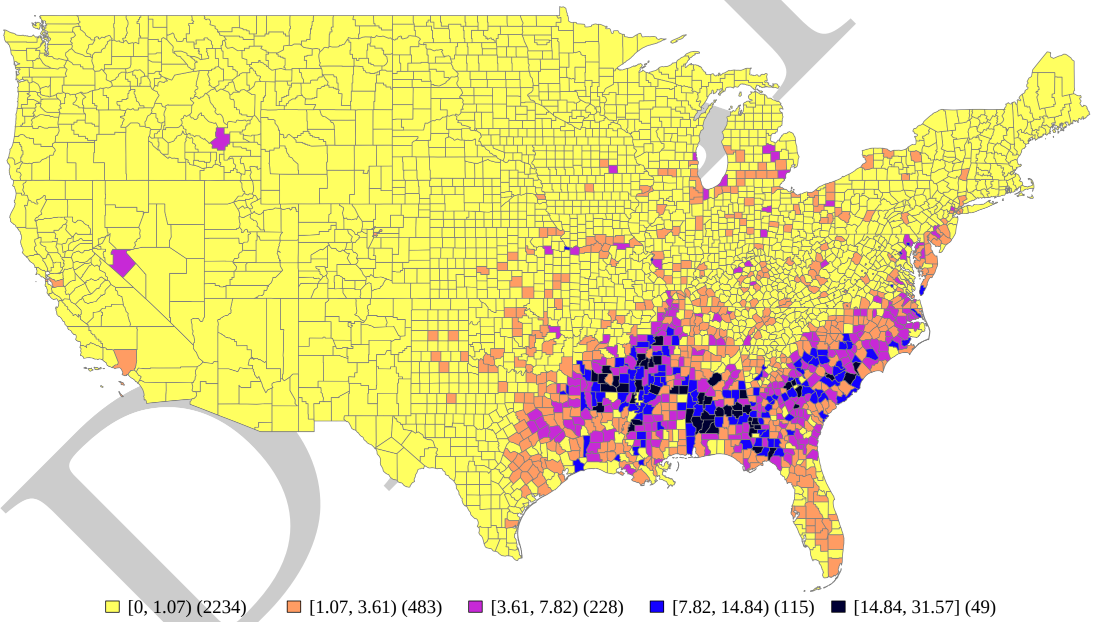


(b) Protestants (%)

Figure 15: Thematic Maps of Other Variables

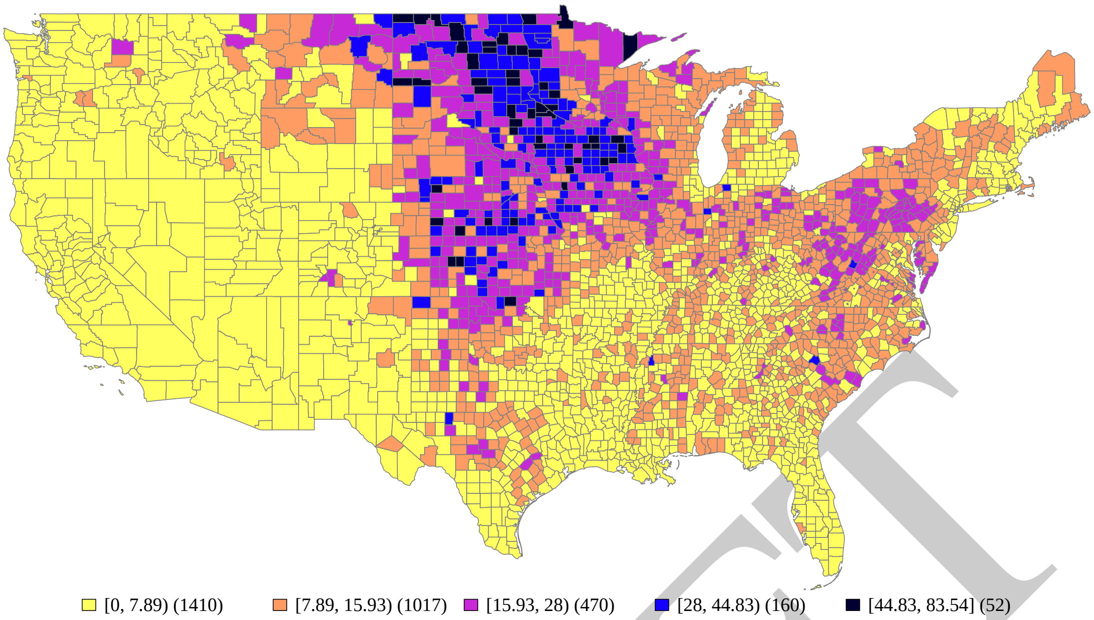


(c) Evangelical Protestants (%)

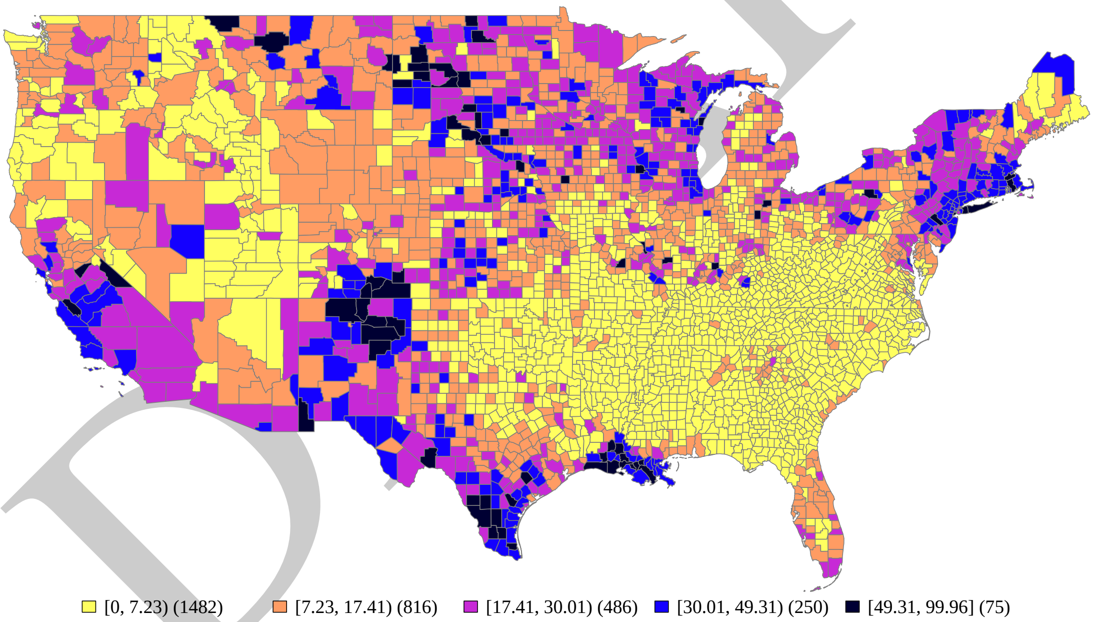


(d) Black Protestants (%)

Figure 15 (continued)

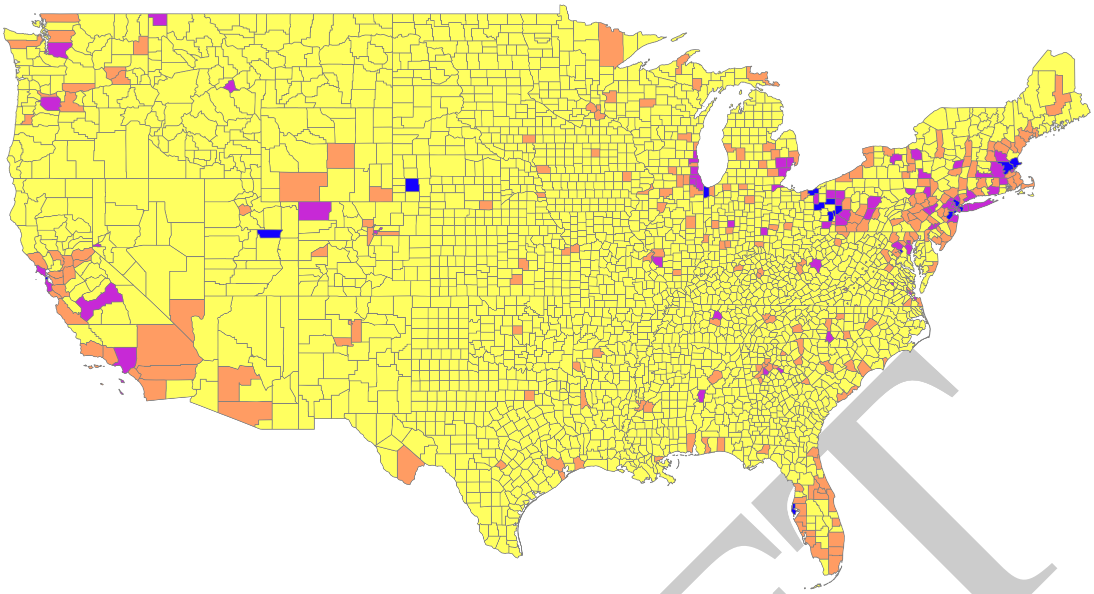


(e) Mainline Protestants (%)



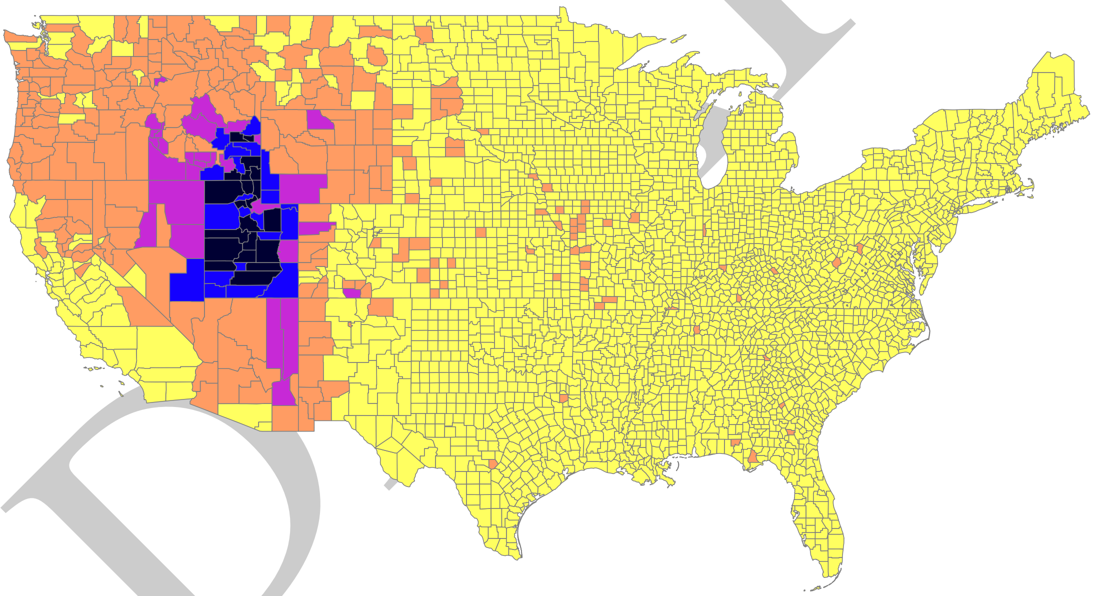
(f) Catholics (%)

Figure 15 (continued)



[0, 0.16) (2771) [0.16, 0.58) (245) [0.58, 1.27) (73) [1.27, 8.06) (19) [8.06, 13.3] (1)

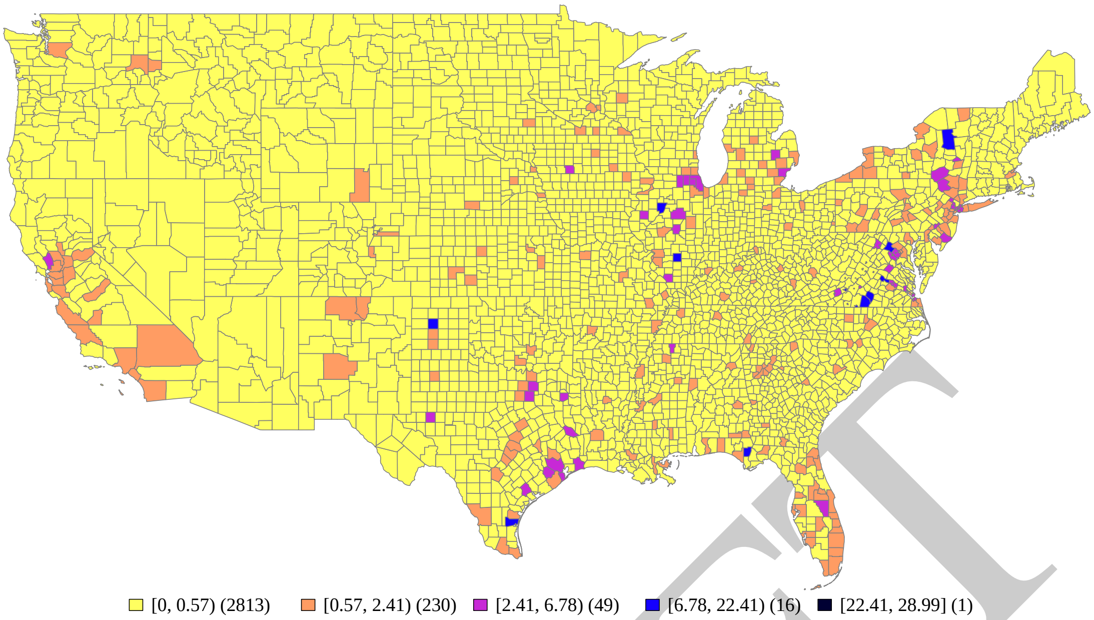
(g) Orthodox Christians (%)



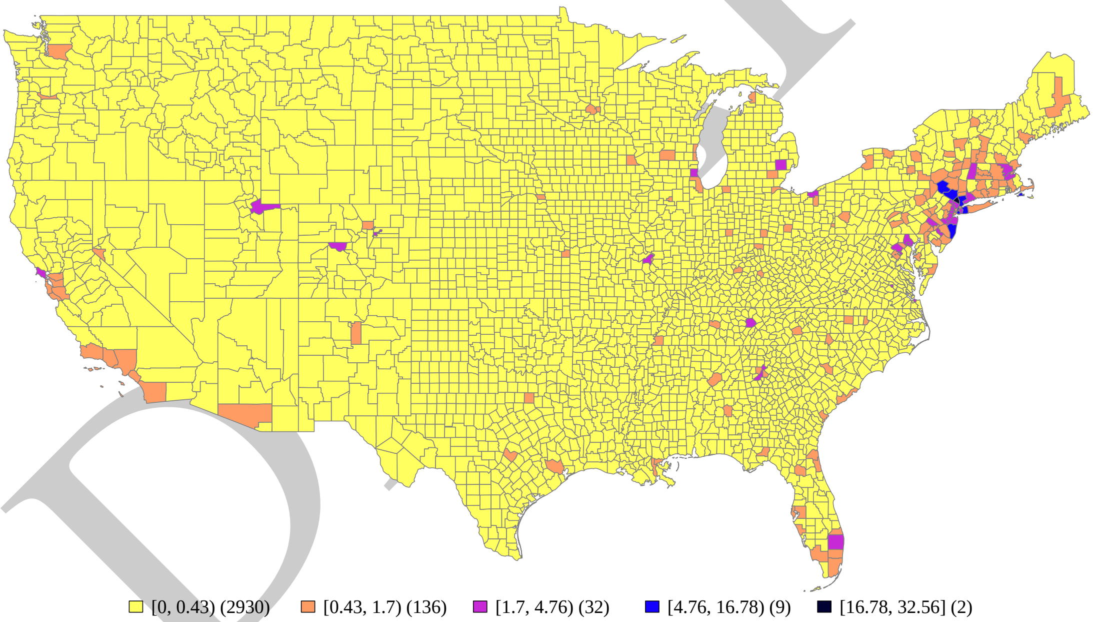
[0, 3.22) (2780) [3.22, 15.07) (257) [15.07, 41.26) (30) [41.26, 70.7) (21) [70.7, 100.79] (21)

(h) Mormons (%)

Figure 15 (continued)

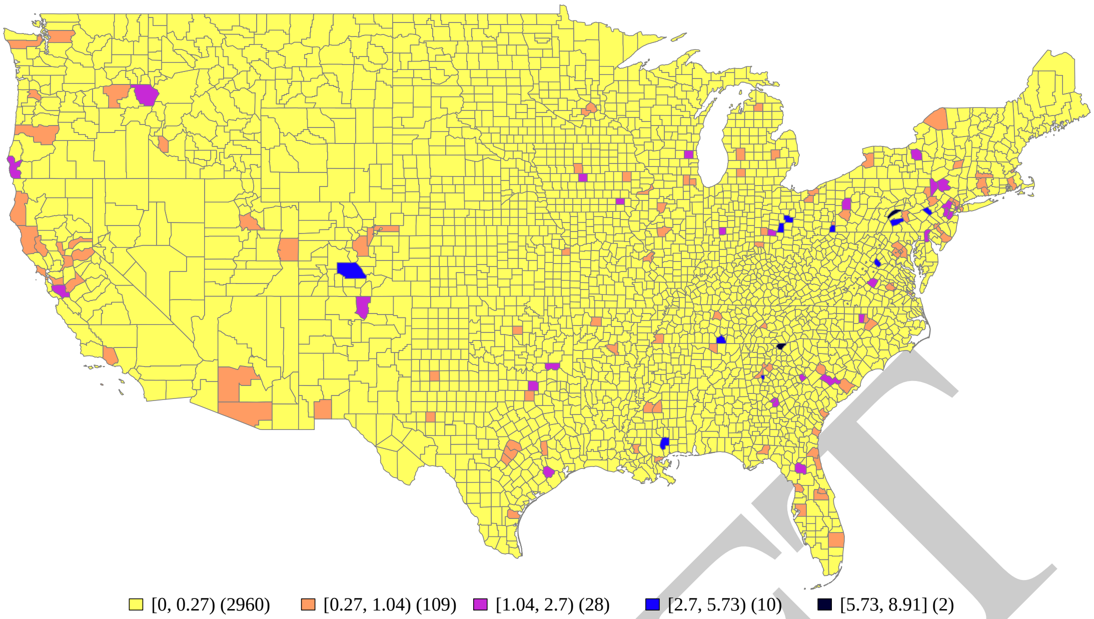


(i) Muslims (%)

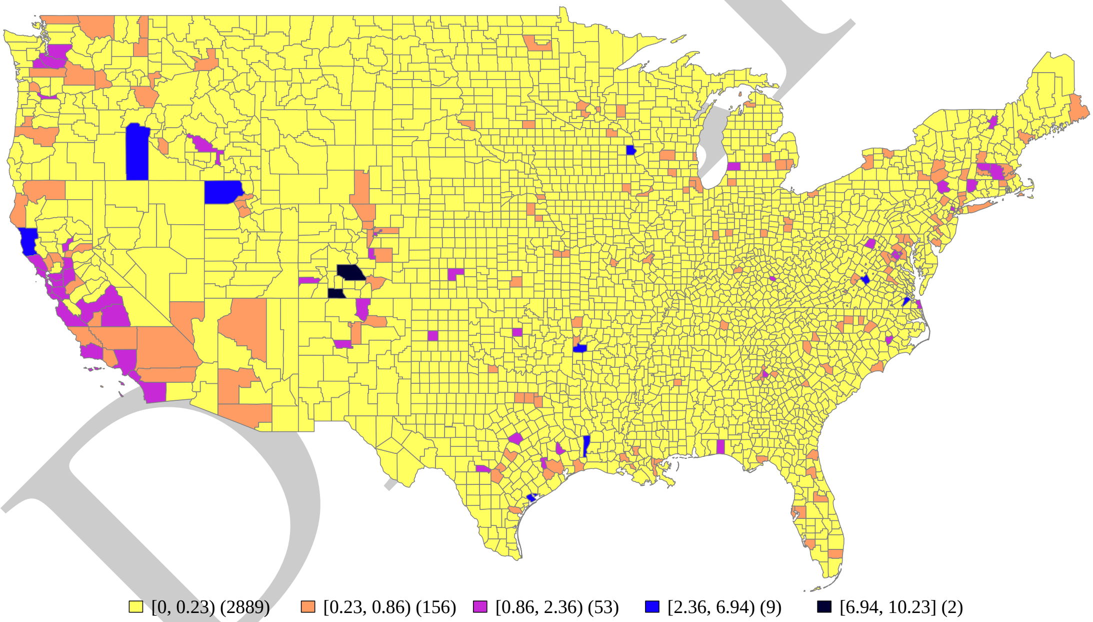


(j) Jews (%)

Figure 15 (continued)

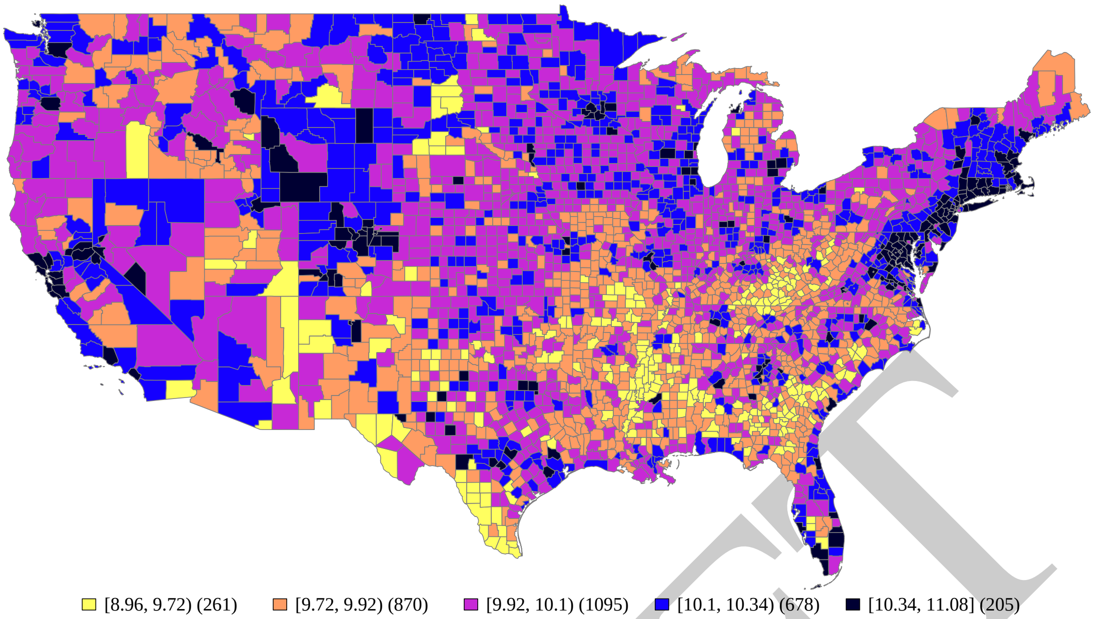


(k) Hindus (%)

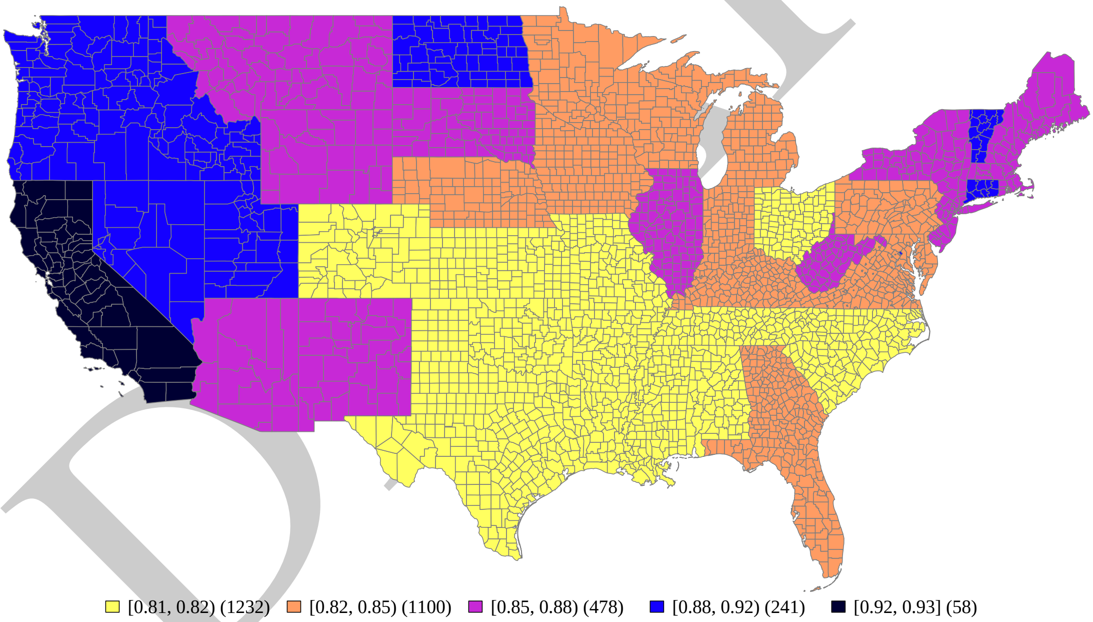


(l) Buddhists (%)

Figure 15 (continued)

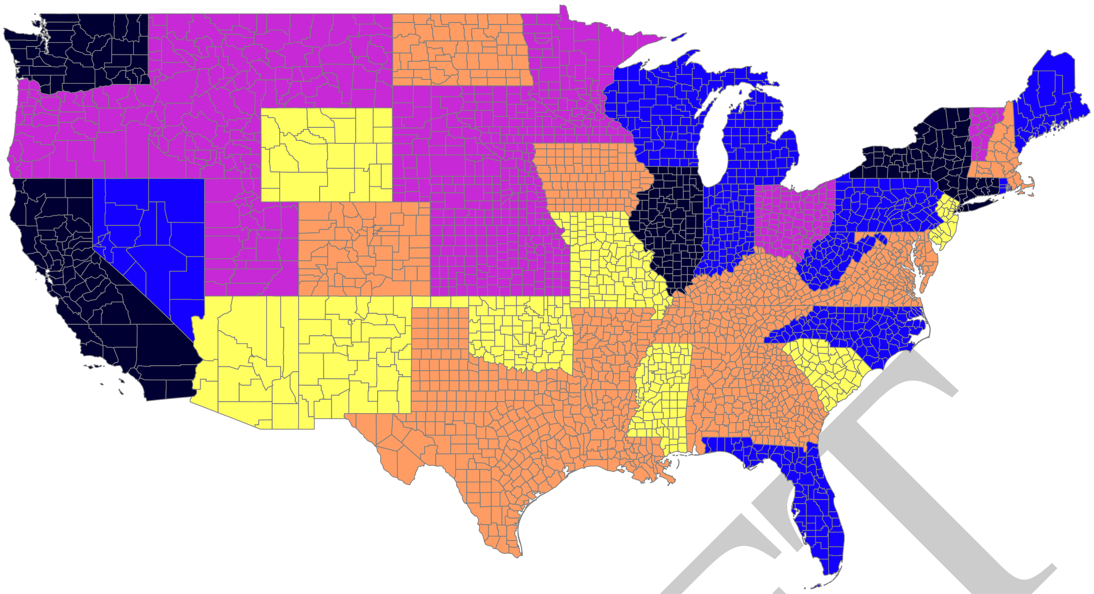


(m) Log Income

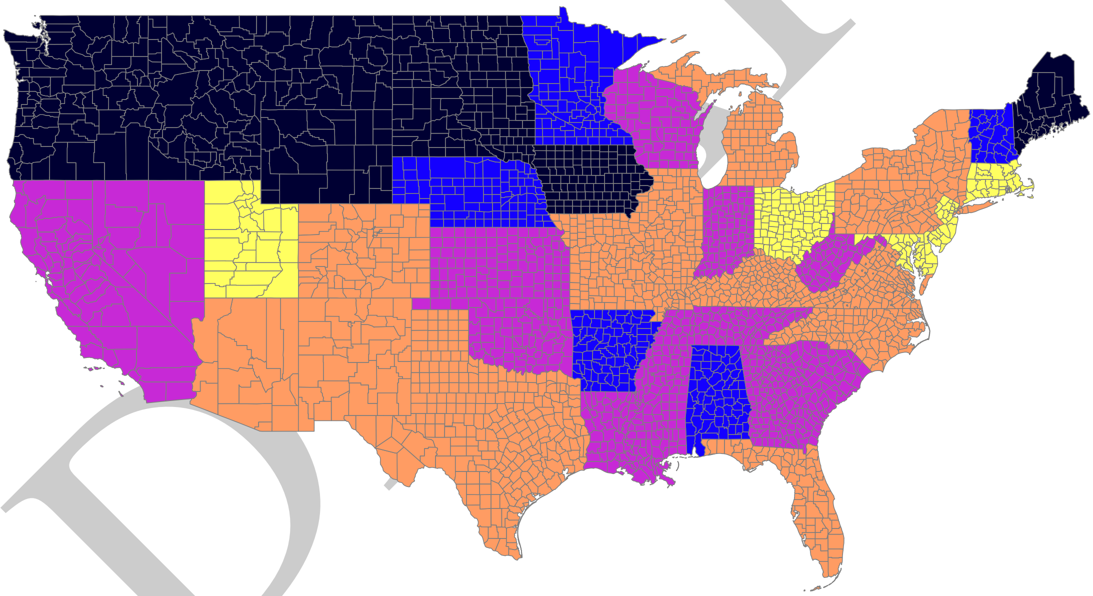


(n) Log Gas Price

Figure 15 (continued)

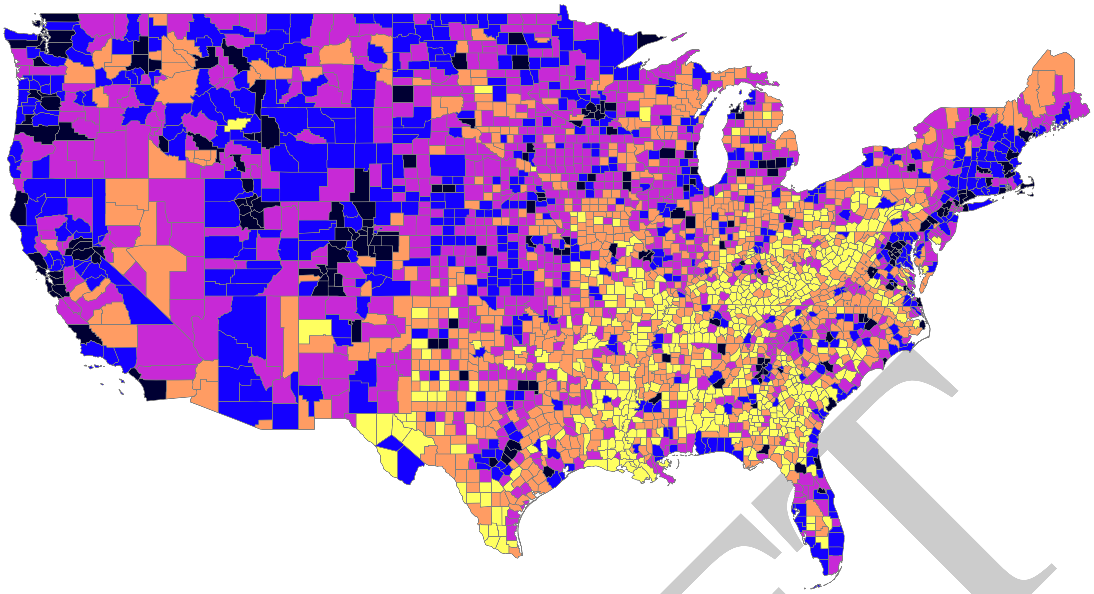


(o) Log Gas Tax/Fee



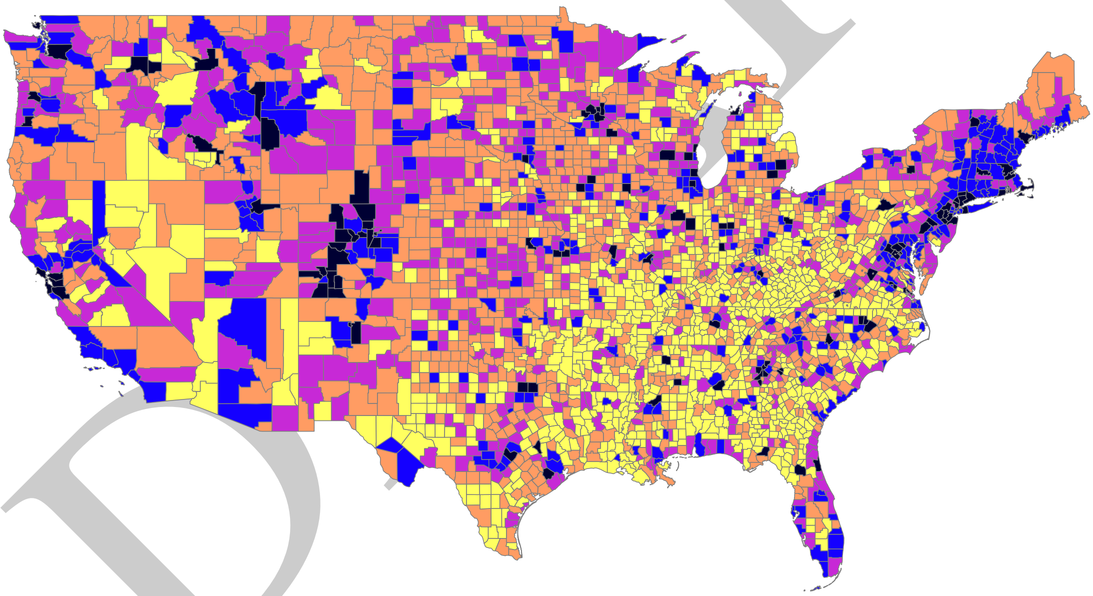
(p) Log Renewable Energy Consumption

Figure 15 (continued)



[17.95, 36.94) (564) [36.94, 45.84) (884) [45.84, 54.29) (864) [54.29, 64.2) (584) [64.2, 89.27] (213)

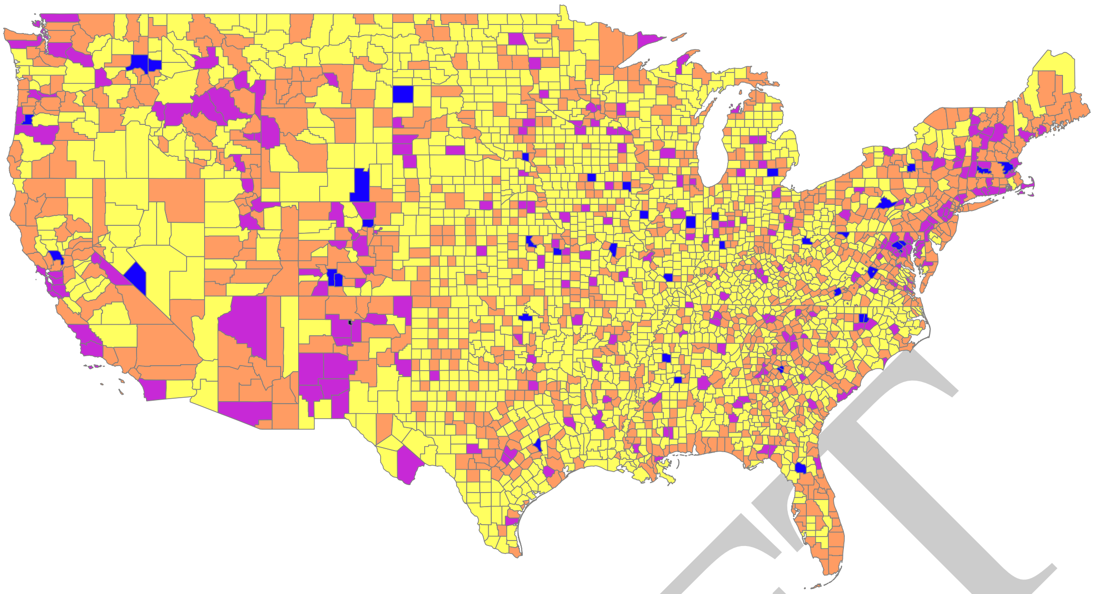
(q) Some College Degree or More (%)



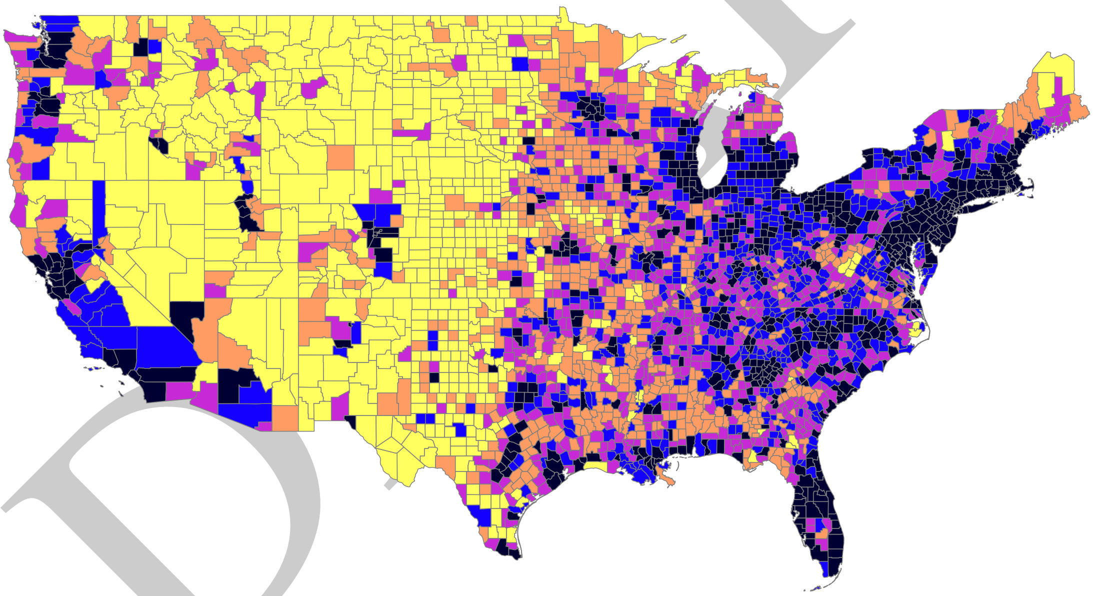
[3.68, 13.91) (957) [13.91, 19.51) (1048) [19.51, 26.64) (601) [26.64, 37.94) (365) [37.94, 70.96] (138)

(r) Bachelor's Degree or More (%)

Figure 15 (continued)

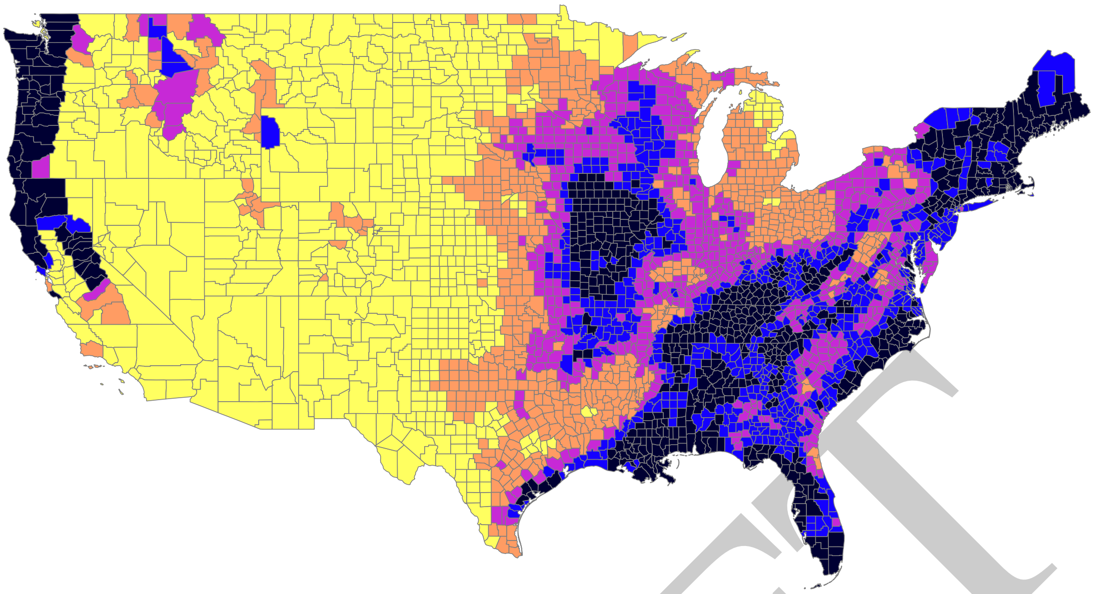


(s) Doctorate Degree (%)

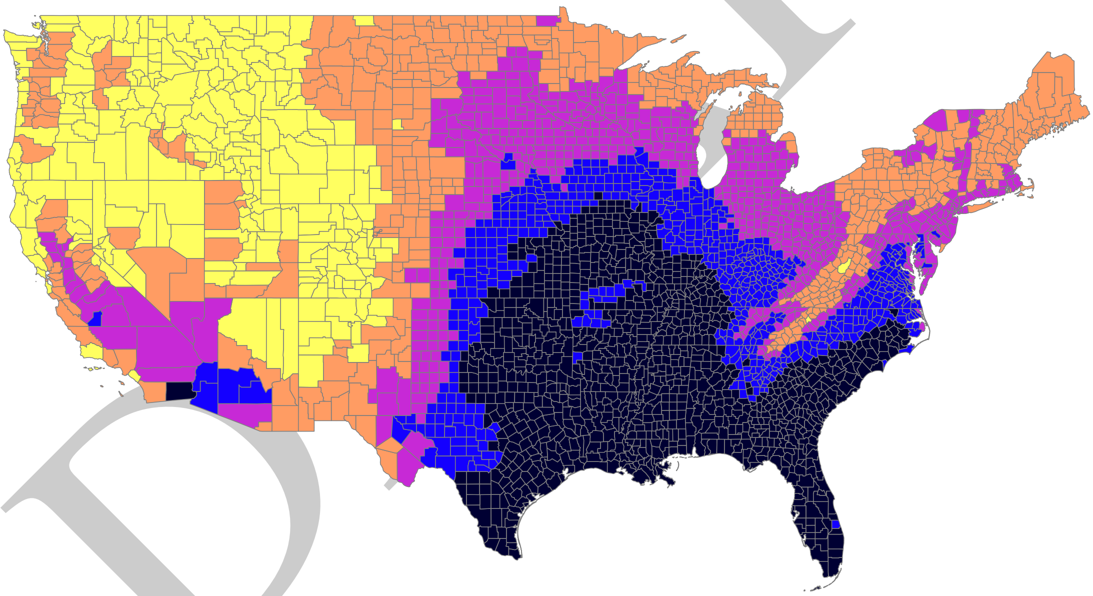


(t) Log Population Density

Figure 15 (continued)

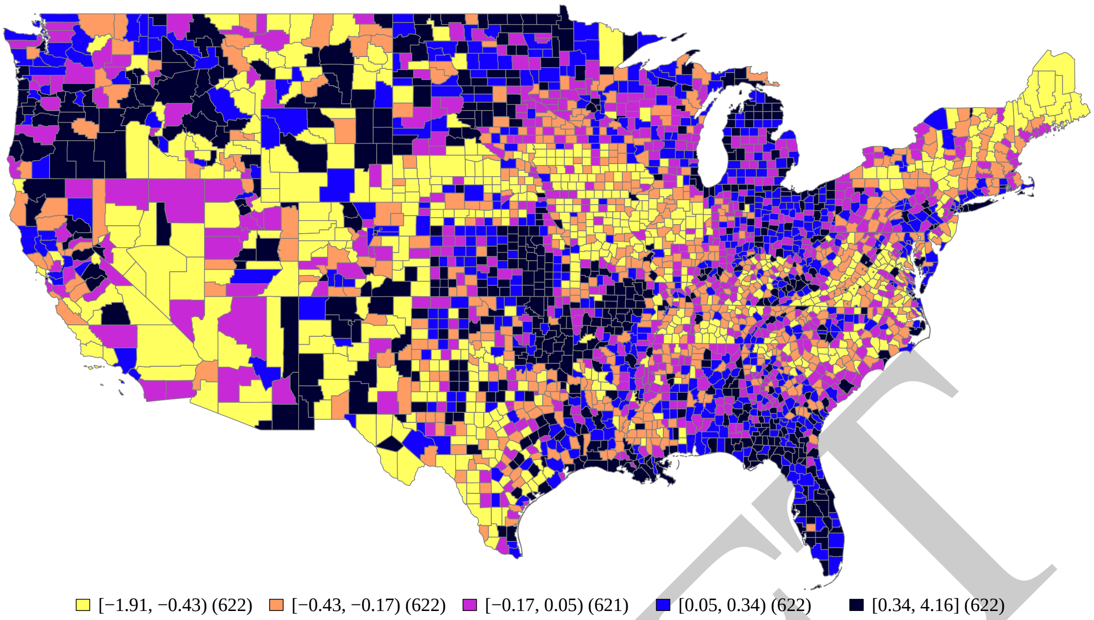


(u) Log Mean Daily Precipitation

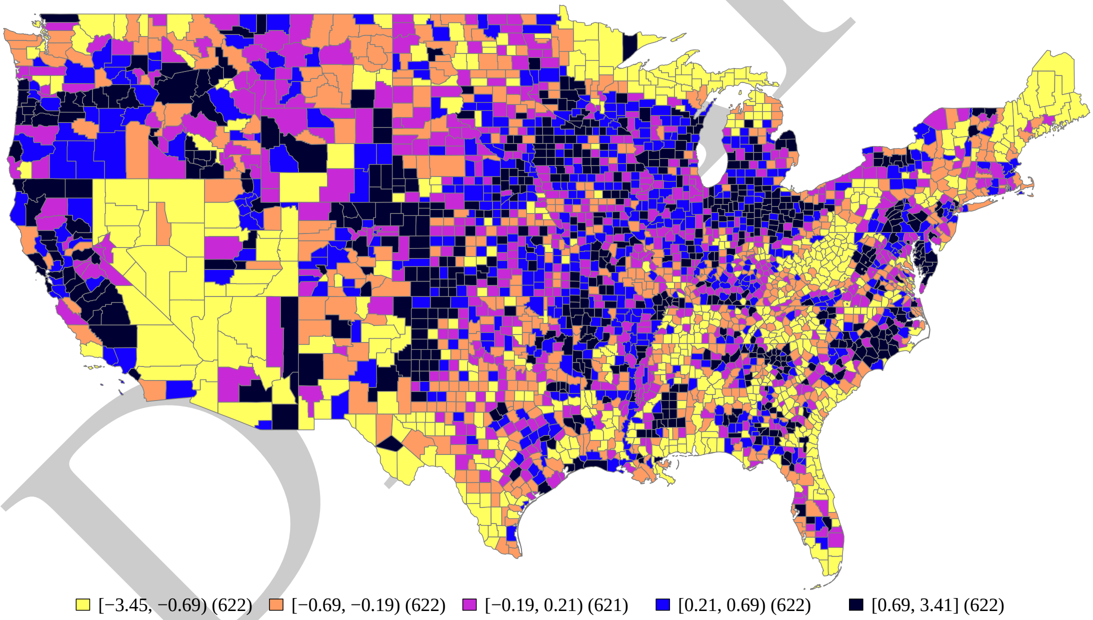


(v) Log Mean Daily Maximum Heat Index

Figure 15 (continued)

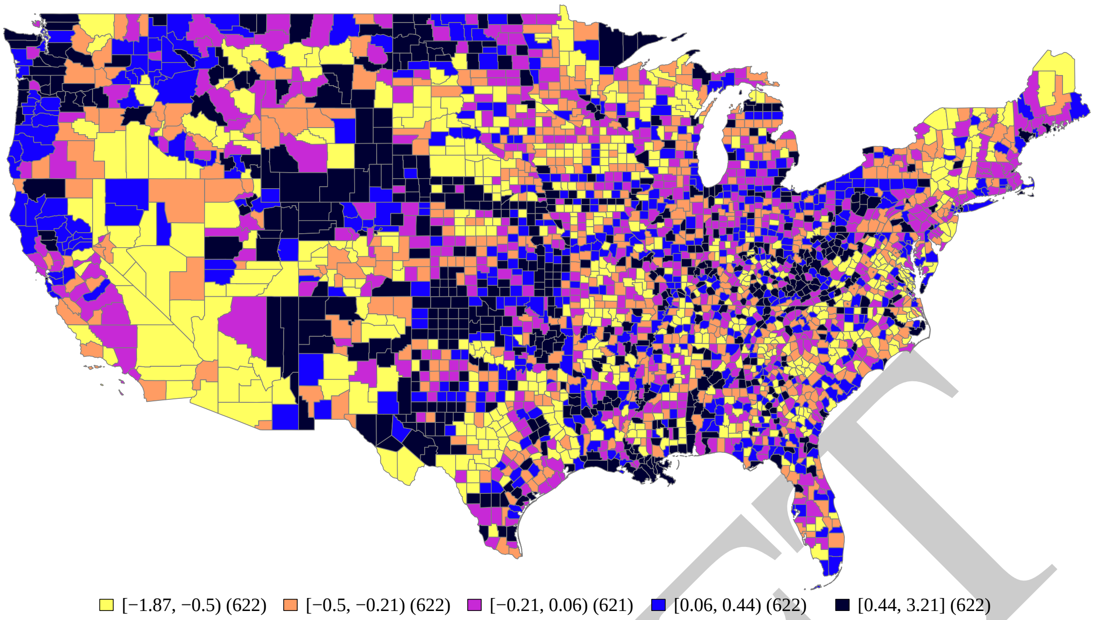


(a) CO[c]

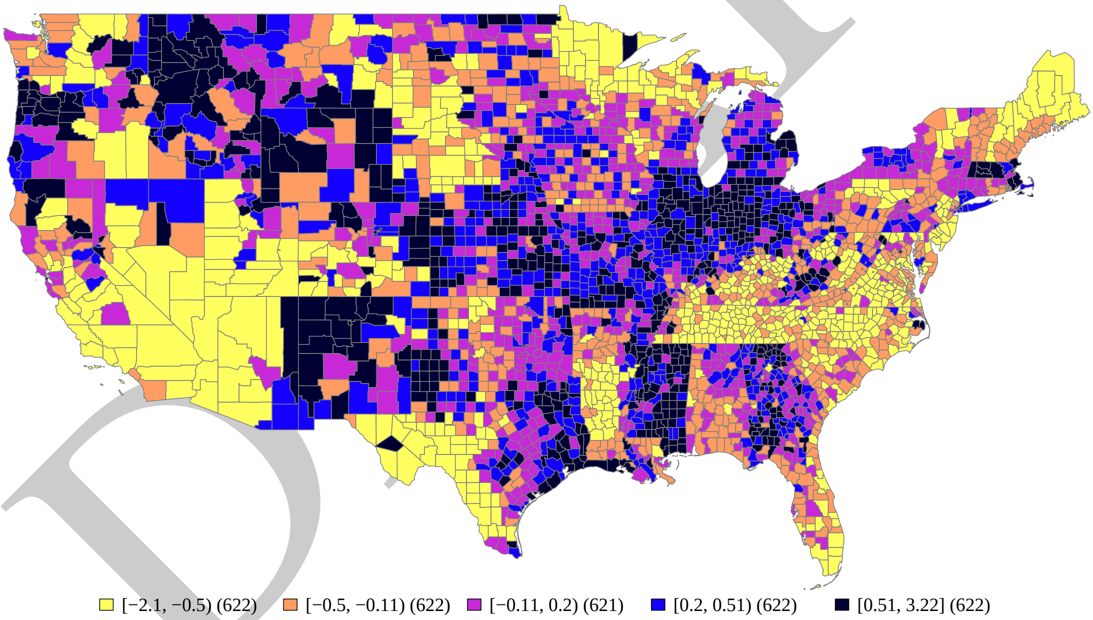


(b) NH_3[†,b]

Figure 16: Base Model OLS Residuals

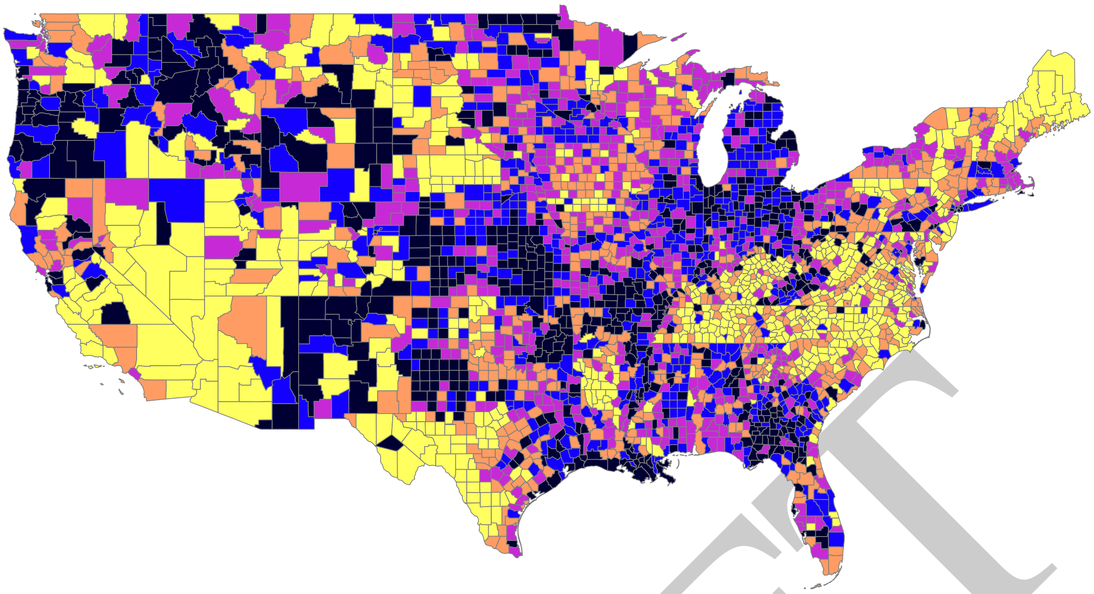


(c) $NO_x{}^{\dagger,b}$



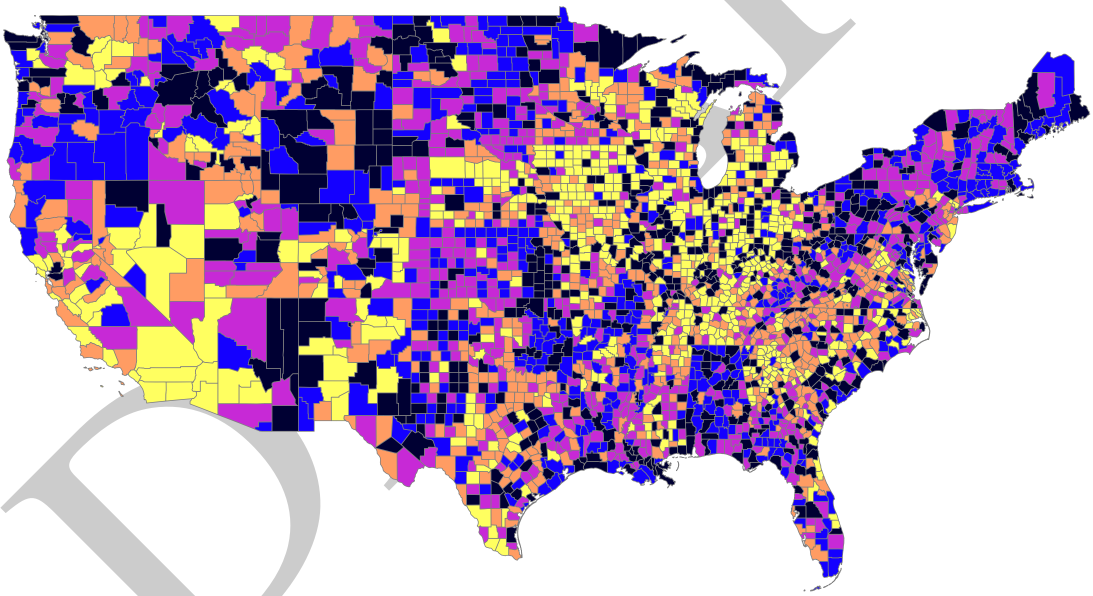
(d) $PM_{10}{}^{\dagger,c}$

Figure 16 (continued)



[−2.39, −0.45) (622) [−0.45, −0.12) (622) [−0.12, 0.13) (621) [0.13, 0.41) (622) [0.41, 4.21] (622)

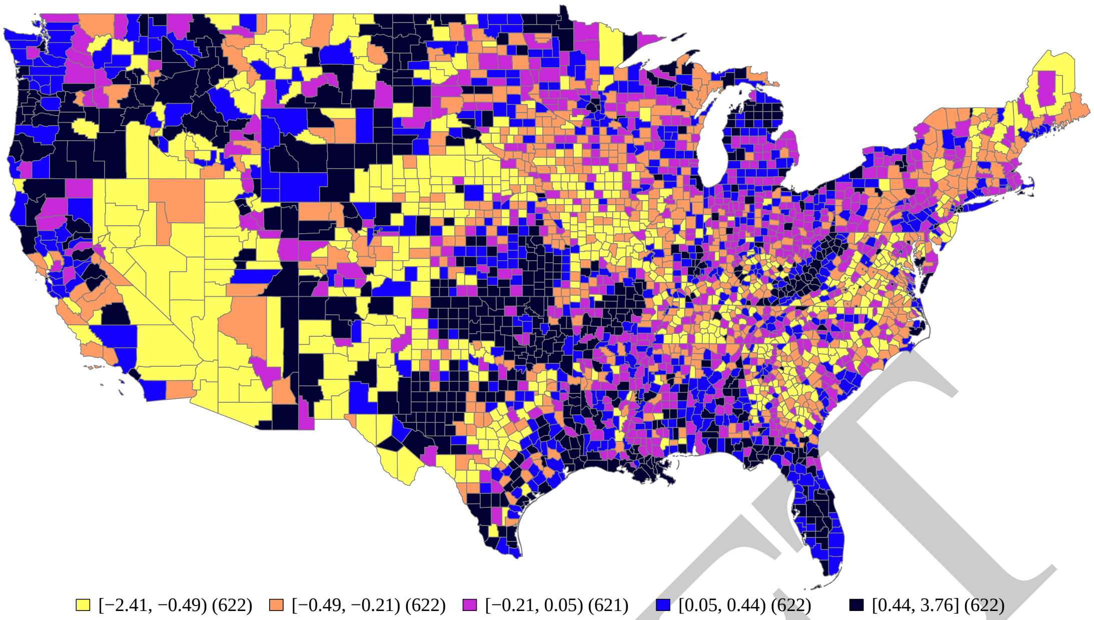
(e) $PM_{2.5}^{\dagger,c}$



[−3.55, −1.35) (622) [−1.35, −0.66) (622) [−0.66, 0.07) (621) [0.07, 1.11) (622) [1.11, 7.08] (622)

(f) $SO_2^{\dagger,b}$

Figure 16 (continued)



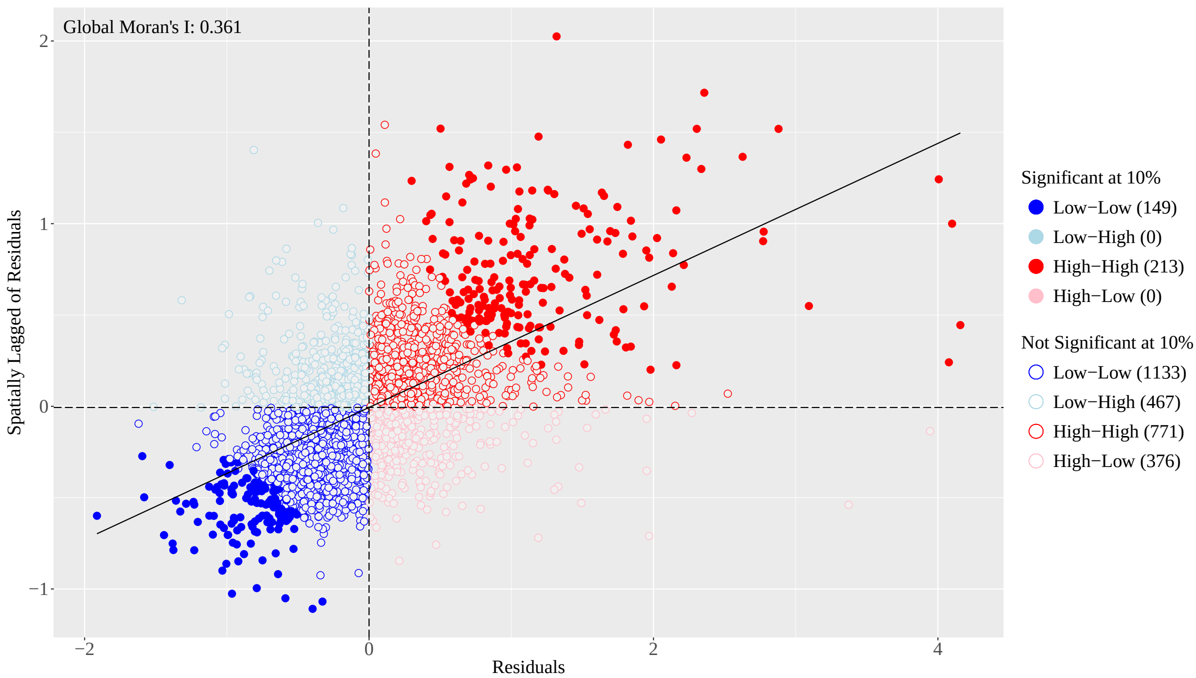
(g) VOC^c

Figure 16 (continued)

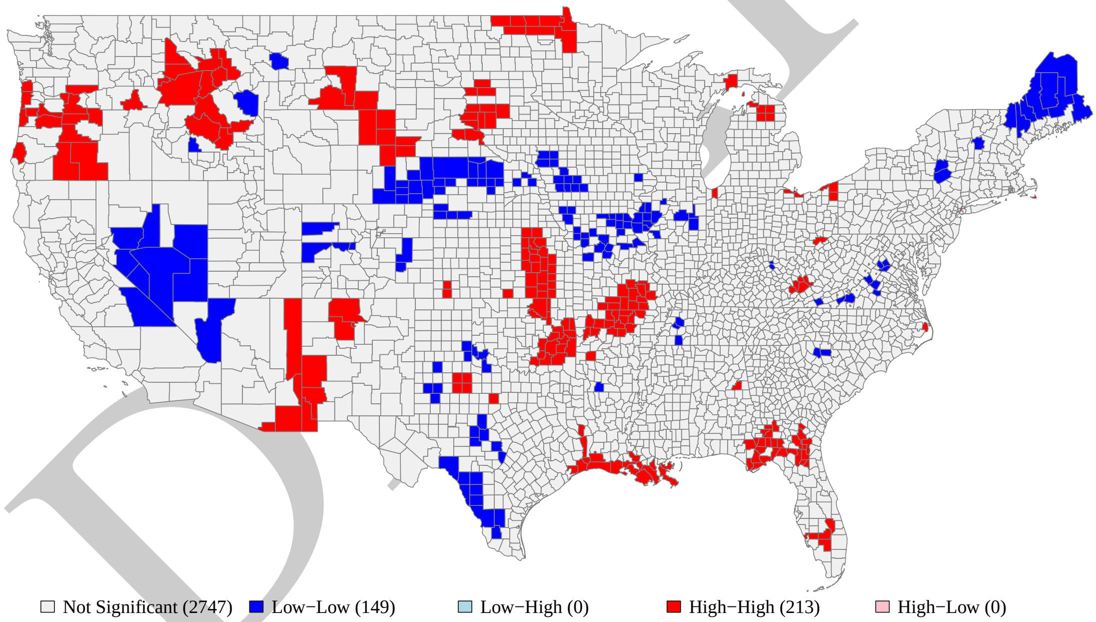
Table 18: Global Moran's I Test Statistics for Base Model OLS Residuals by Weight Matrix

	Base Model OLS:						
Weight Matrix	CO^{c}	$NH_3^{\dagger,b}$	$NO_x^{\dagger,b}$	$PM_{10}^{\dagger,c}$	$PM_{2.5}^{\dagger,c}$	$SO_2^{\dagger,b}$	VOC^{b}
q1.b	0.361***	0.443***	0.256***	0.569***	0.427***	0.142***	0.430***
q1.w	0.361***	0.450***	0.261***	0.578***	0.428***	0.145***	0.425***
q1.c	0.361***	0.443***	0.256***	0.569***	0.427***	0.142***	0.430***
q1.s	0.362***	0.447***	0.258***	0.573***	0.428***	0.144***	0.429***
q1.minmax	0.361***	0.443***	0.256***	0.569***	0.427***	0.142***	0.430***
q2.b	0.267***	0.303***	0.173***	0.469***	0.332***	0.102***	0.315***
q2.w	0.268***	0.311***	0.171***	0.474***	0.332***	0.104***	0.308***
q2.c	0.267***	0.303***	0.173***	0.469***	0.332***	0.102***	0.315***
q2.s	0.268***	0.307***	0.172***	0.471***	0.332***	0.103***	0.312***
q2.minmax	0.267***	0.303***	0.173***	0.469***	0.332***	0.102***	0.315***
k6.b	0.321***	0.438***	0.244***	0.550***	0.390***	0.144***	0.388***
k6.w	0.321***	0.438***	0.244***	0.550***	0.390***	0.144***	0.388***
k6.c	0.321***	0.438***	0.244***	0.550***	0.390***	0.144***	0.388***
k6.s	0.321***	0.438***	0.244***	0.550***	0.390***	0.144***	0.388***
k6.minmax	0.321***	0.438***	0.244***	0.550***	0.390***	0.144***	0.388***
k10.b	0.290***	0.381***	0.212***	0.504***	0.349***	0.123***	0.347***
k10.w	0.290***	0.381***	0.212***	0.504***	0.349***	0.123***	0.347***
k10.c	0.290***	0.381***	0.212***	0.504***	0.349***	0.123***	0.347***
k10.s	0.290***	0.381***	0.212***	0.504***	0.349***	0.123***	0.347***
k10.minmax	0.290***	0.381***	0.212***	0.504***	0.349***	0.123***	0.347***
soi.b	0.342***	0.455***	0.252***	0.560***	0.405***	0.151***	0.413***
soi.w	0.345***	0.462***	0.260***	0.576***	0.418***	0.151***	0.424***
soi.c	0.342***	0.455***	0.252***	0.560***	0.405***	0.151***	0.413***
soi.s	0.344***	0.458***	0.256***	0.567***	0.411***	0.151***	0.418***
soi.minmax	0.342***	0.455***	0.252***	0.560***	0.405***	0.151***	0.413***

Notes: Normality assumption is used in all tests. Each base model is labeled with its label of dependent variable. †, ^{c}, and ^{b} indicate a base model including *log income squared*, *log some college or more*, and *log bachelor's degree or more* respectively. *q1*, *q2*, *k6*, *k10*, and *soi* indicate 1st Order Queen, 2nd Order Queen, 6 Nearest Neighbors, 10 Nearest Neighbors, and Sphere of Influence contiguities respectively. *b*, *w*, *c*, *s*, and *minmax* indicate binary, row–standardized, global–standardized, variance–stabilizing, and minmax–normalized weight styles respectively. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels respectively.

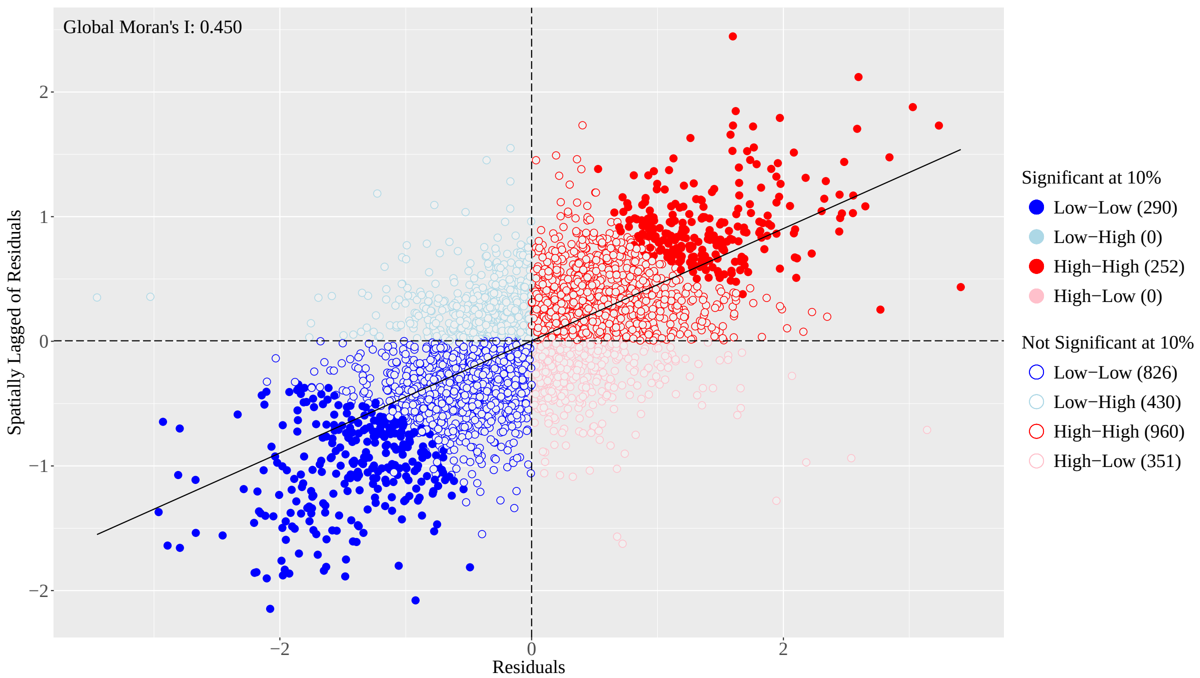


(a) LISA Moran Scatterplot for CO^c

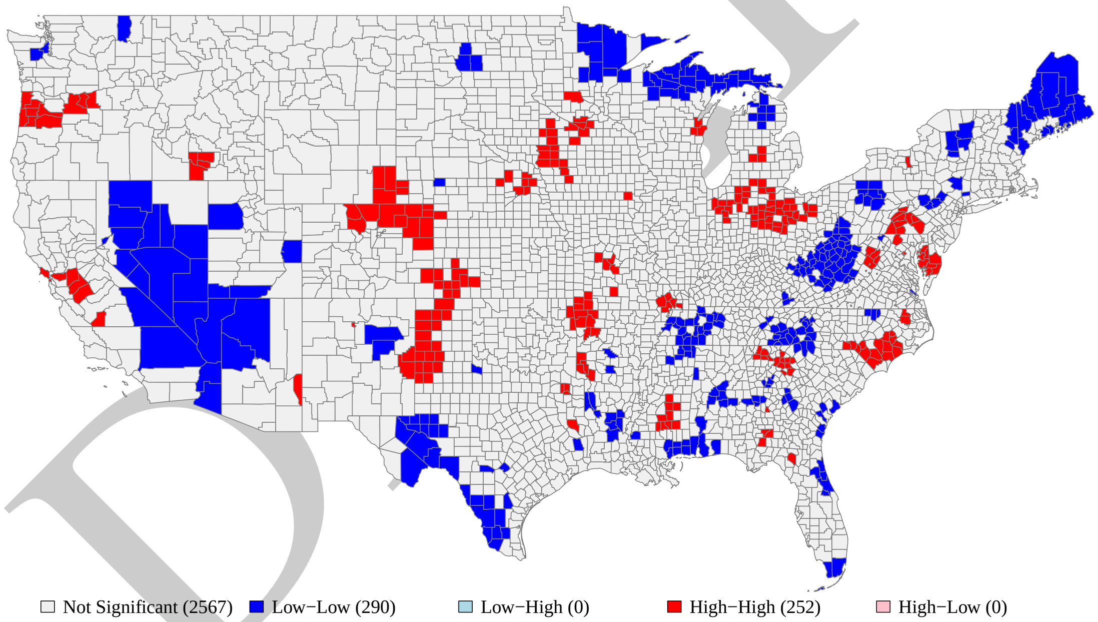


(b) LISA Cluster Map for CO^c

Figure 17: LISA Moran Scatterplot and Cluster Map by Base Model OLS Residuals

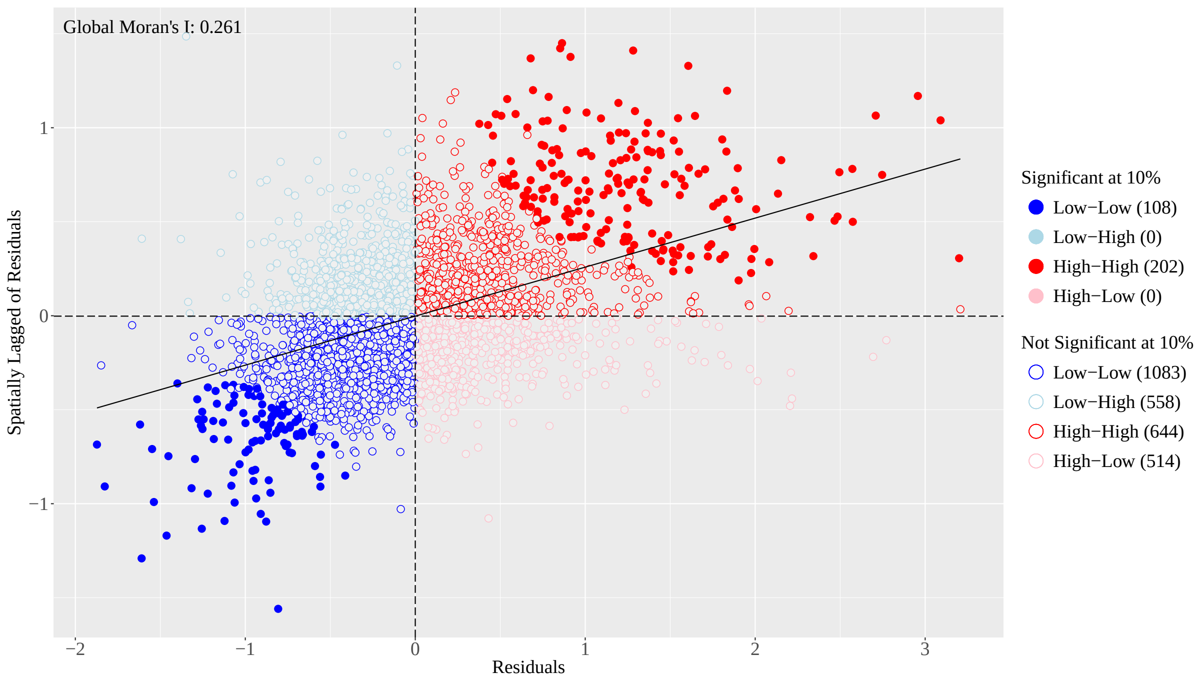


(c) LISA Moran Scatterplot for $NH_3{}^{\dagger,b}$

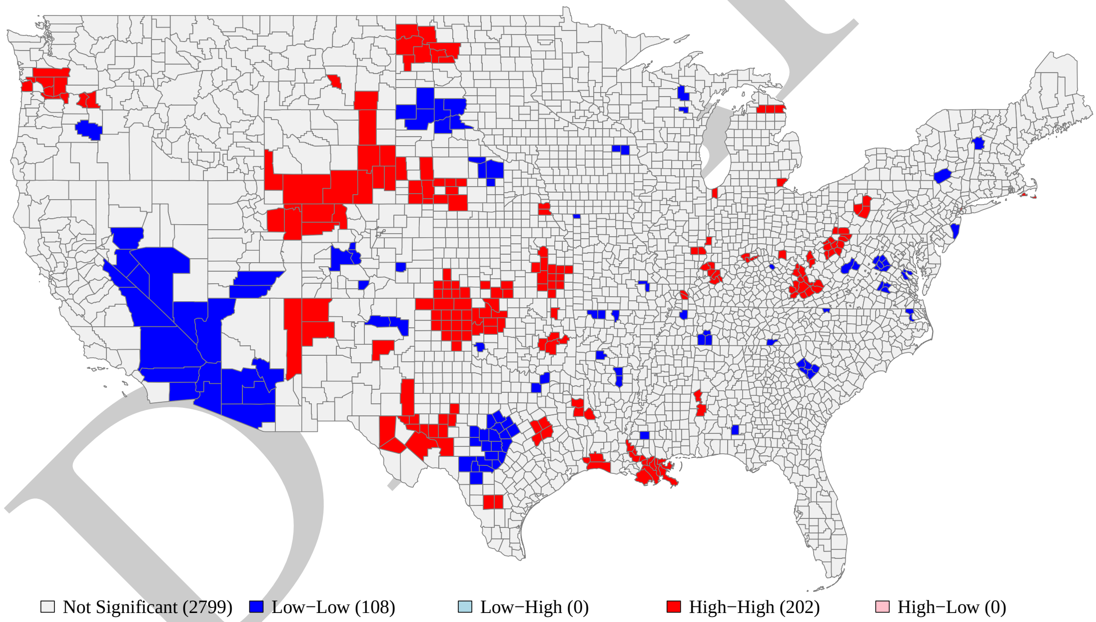


(d) LISA Cluster Map for $NH_3{}^{\dagger,b}$

Figure 17 (continued)

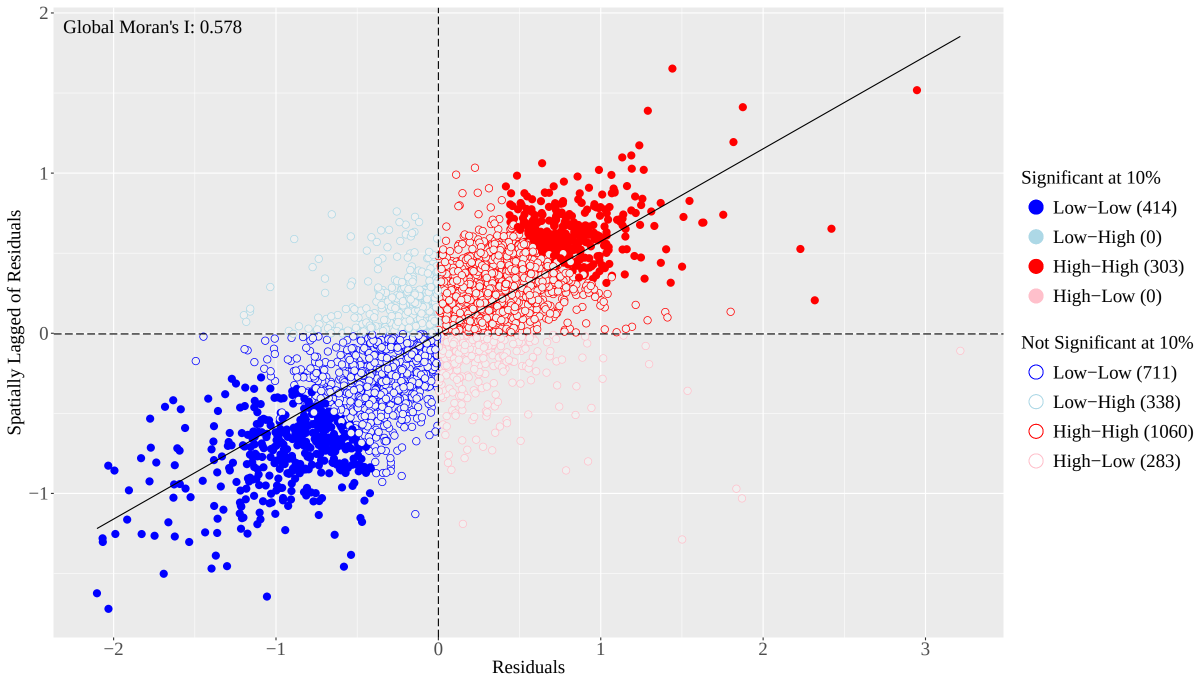


(e) LISA Moran Scatterplot for $NO_x{}^{\dagger,b}$

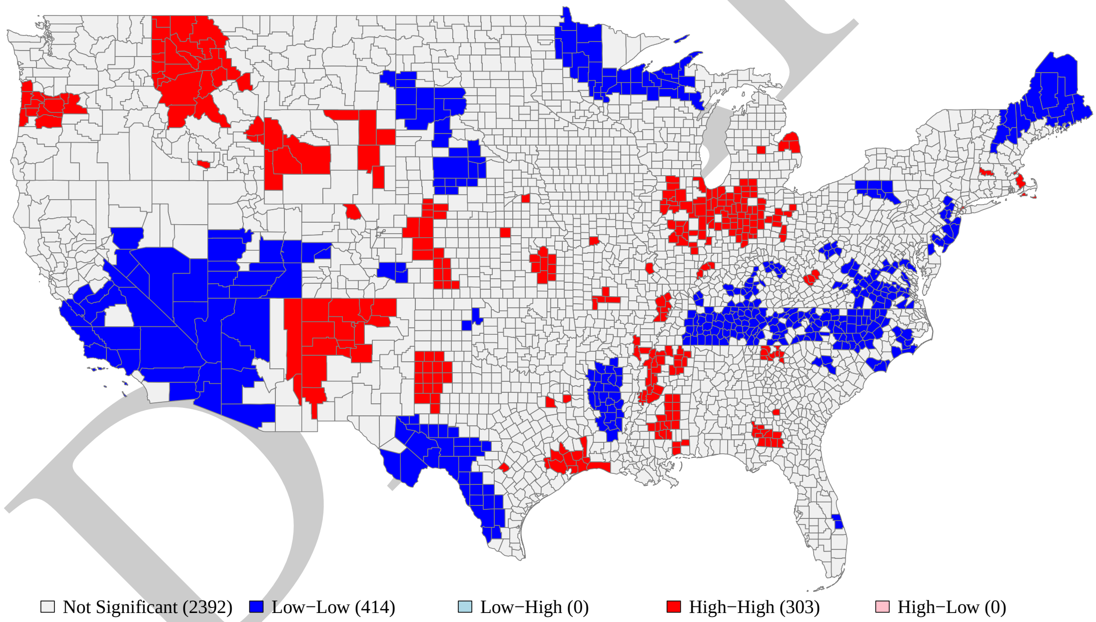


(f) LISA Cluster Map for $NO_x{}^{\dagger,b}$

Figure 17 (continued)

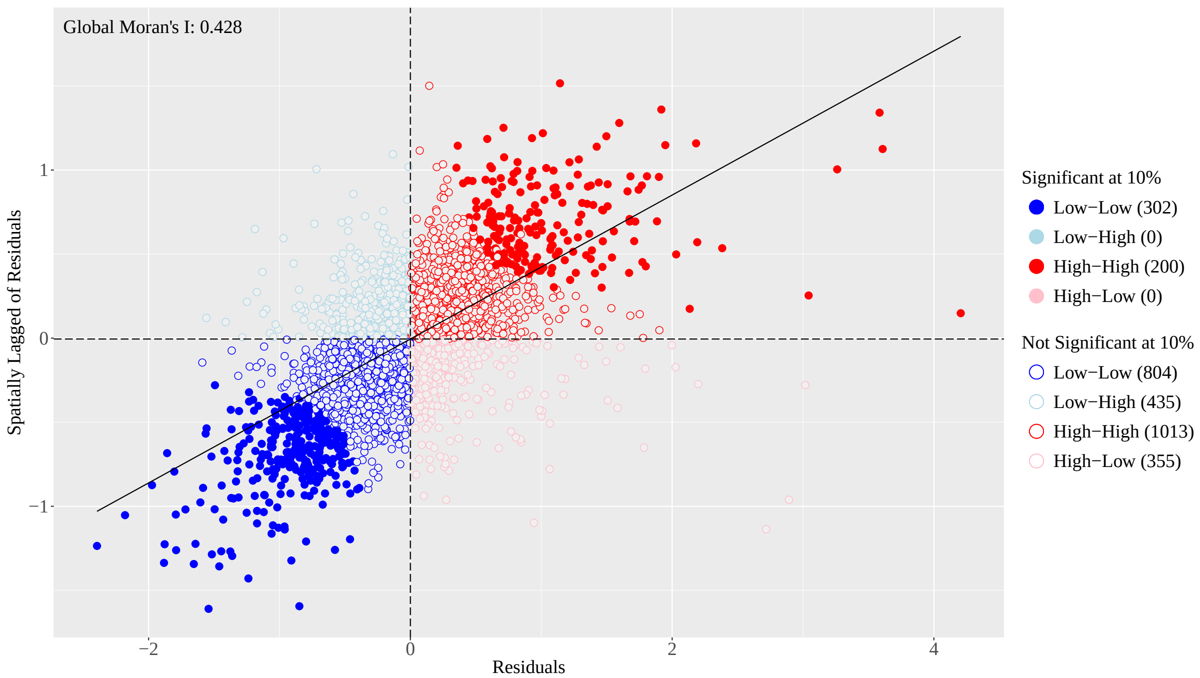


(g) LISA Moran Scatterplot for $\text{PM}_{10}{}^{\dagger,c}$

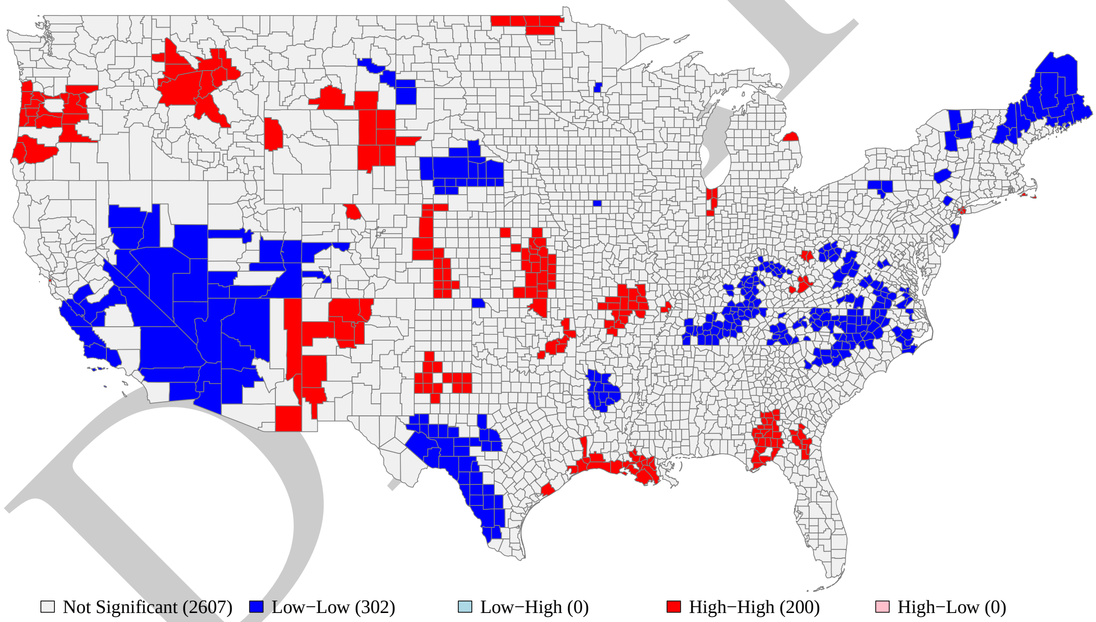


(h) LISA Cluster Map for $\text{PM}_{10}{}^{\dagger,c}$

Figure 17 (continued)

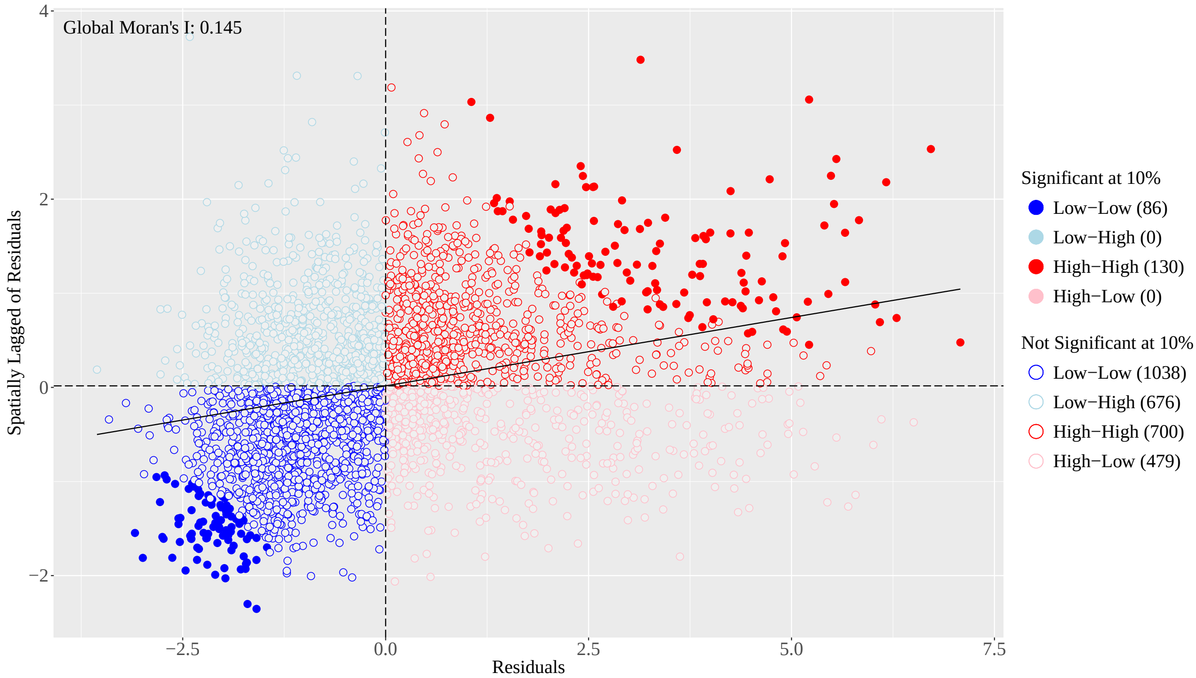


(i) LISA Moran Scatterplot for $PM_{2.5}{}^{\dagger,c}$

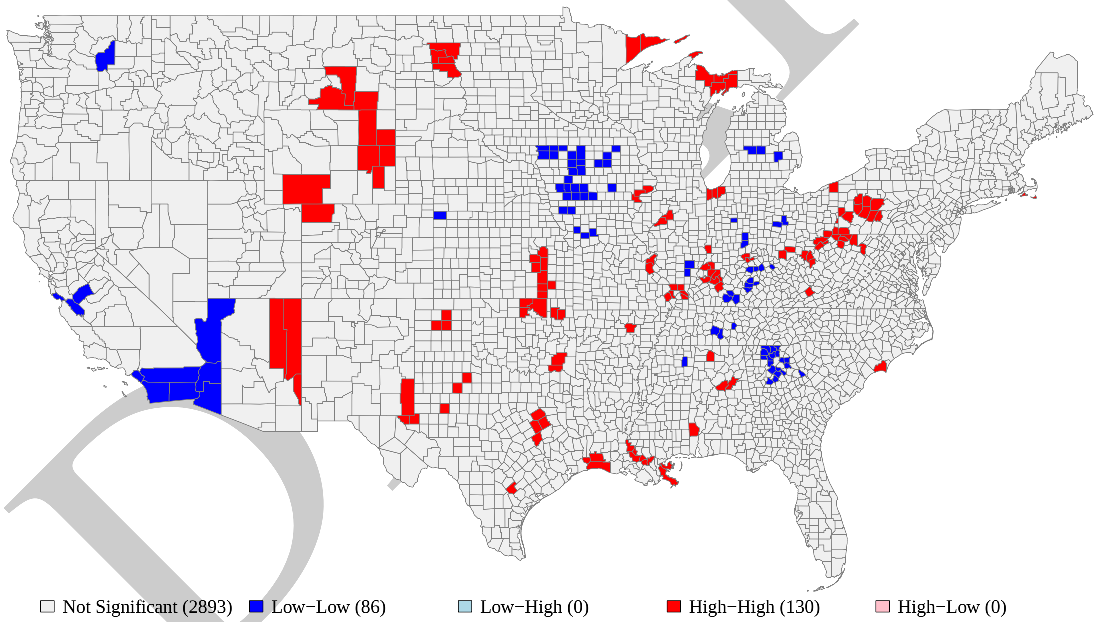


(j) LISA Cluster Map for $PM_{2.5}{}^{\dagger,c}$

Figure 17 (continued)



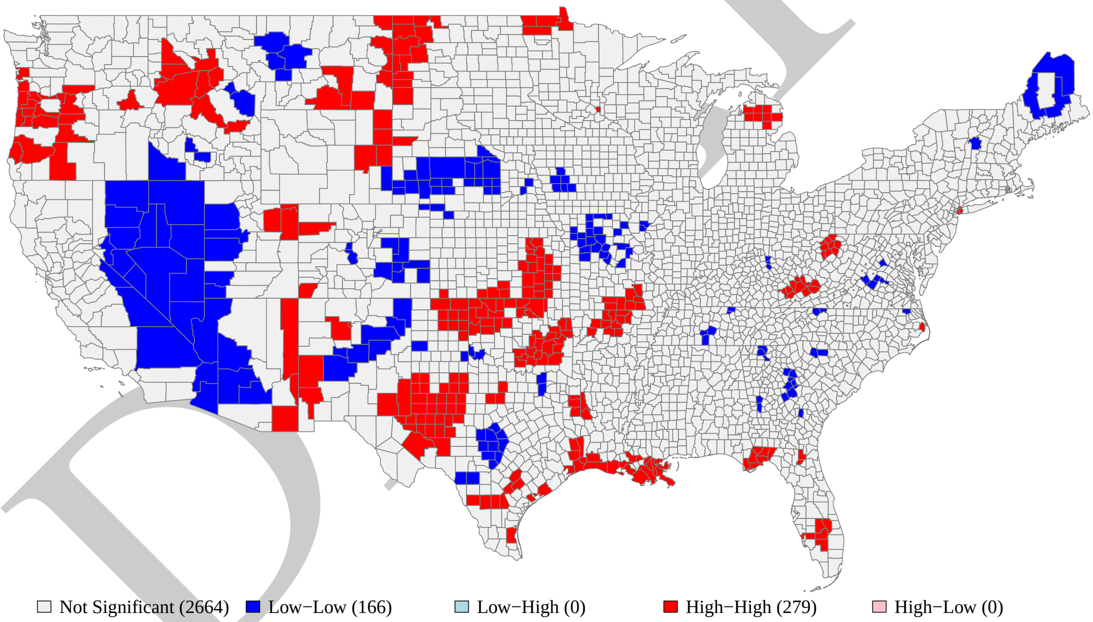
(k) LISA Moran Scatterplot for $SO_2{}^{\dagger,b}$



(l) LISA Cluster Map for $SO_2{}^{\dagger,b}$

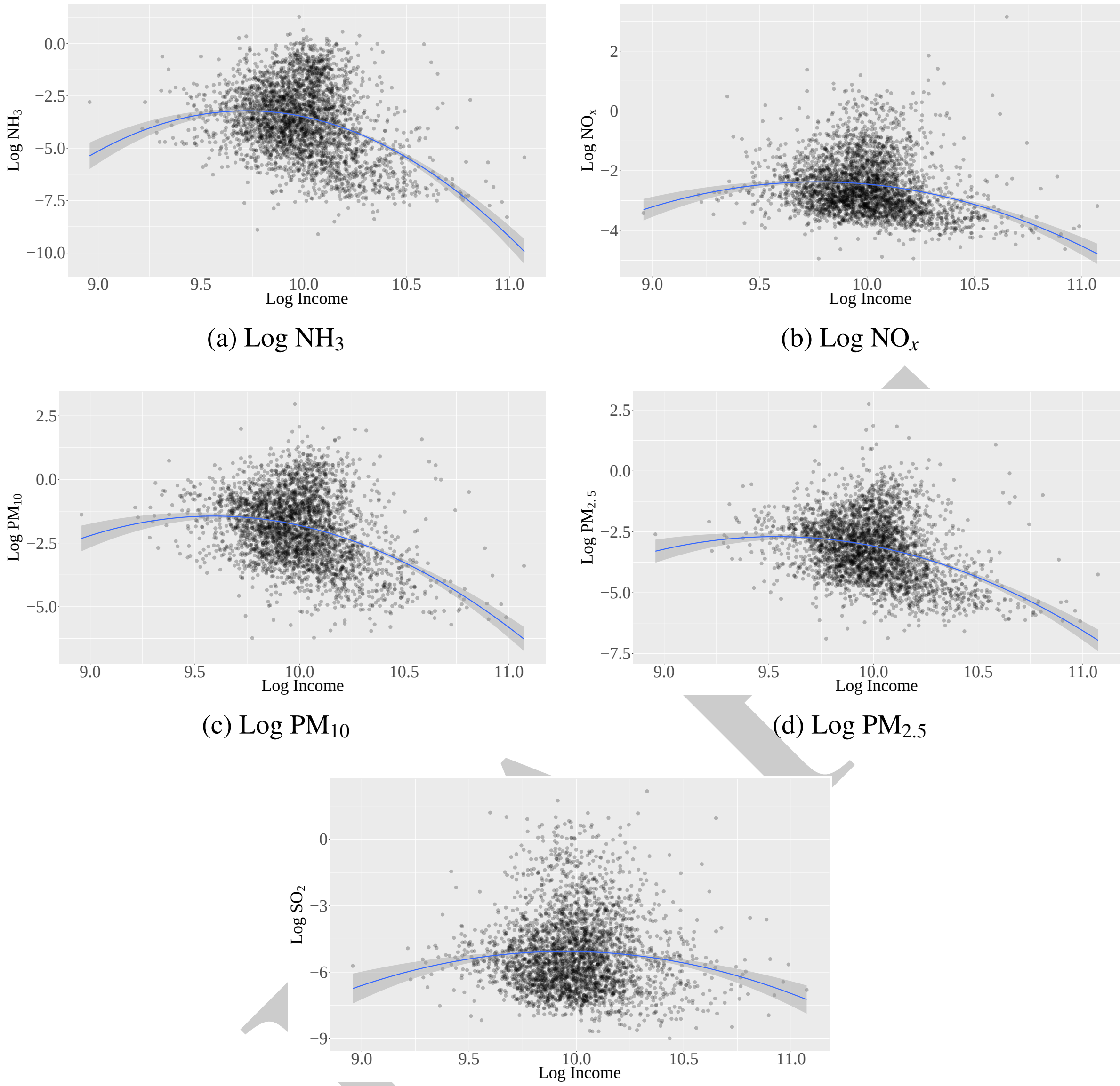
Figure 17 (continued)

(m) LISA Moran Scatterplot for VOC^c



(n) LISA Cluster Map for VOC^c

Figure 17 (continued)



(a) Log NH_3

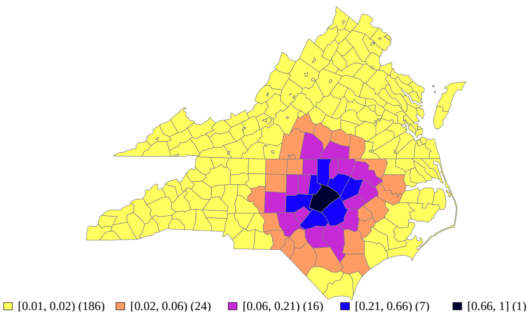
(b) Log NO_x

(c) Log PM_{10}

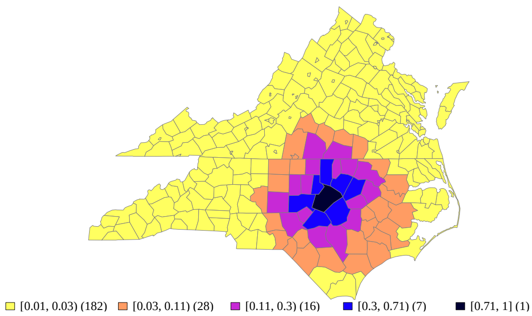
(d) Log $PM_{2.5}$

(e) Log SO_2

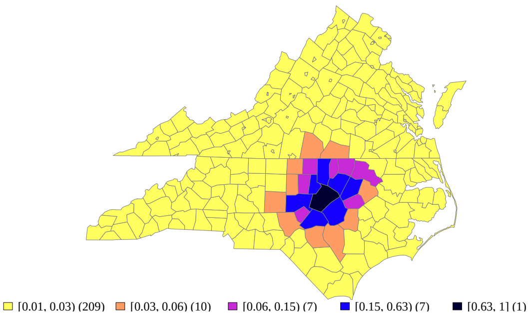
Figure 18: Environmental Kuznets Curves by Dependent Variable



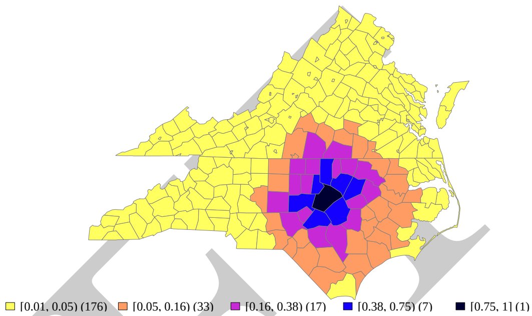
(a) $CO^{c,\circ,q1}$



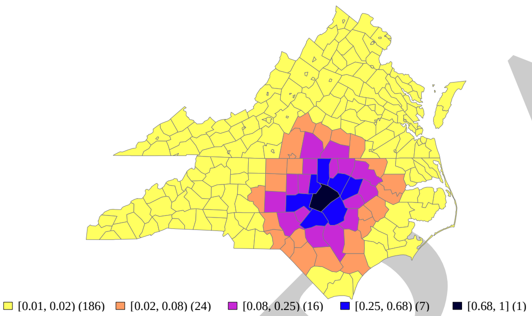
(b) $NH_3{}^{\dagger,b,\bullet,q1}$



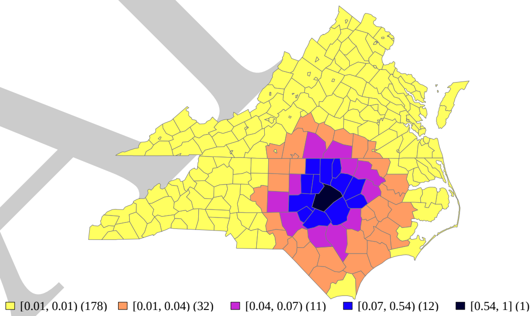
(c) $NO_x{}^{\dagger,b,\circ,q1}$



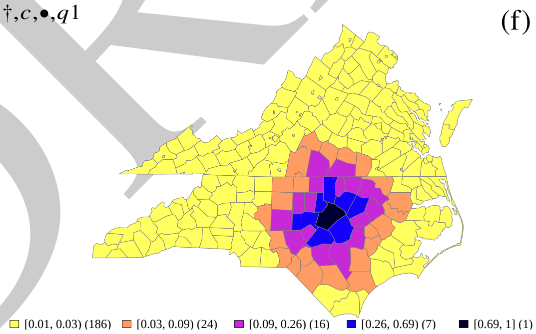
(d) $PM_{10}{}^{\dagger,c,\bullet,q1}$



(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$



(f) $SO_2{}^{\dagger,b,\bullet,q2}$



(g) $VOC^{b,\bullet,q1}$

Figure 19: Spatial Correlation of Dependent Variables for All Counties with Wake County, NC by Base Model Spatial Regression

(a) $CO^{c,\circ,q1}$

(b) $NH_3{}^{\dagger,b,\bullet,q1}$

(c) $NO_x{}^{\dagger,b,\circ,q1}$

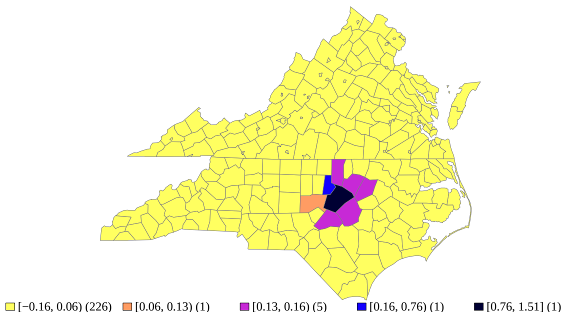
(d) $PM_{10}{}^{\dagger,c,\bullet,q1}$

(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$

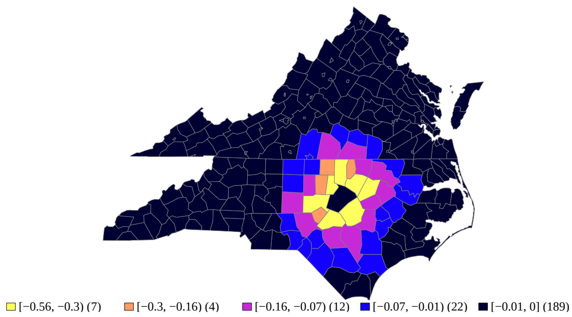
(f) $SO_2{}^{\dagger,b,\bullet,q2}$

(g) $VOC^{b,\bullet,q1}$

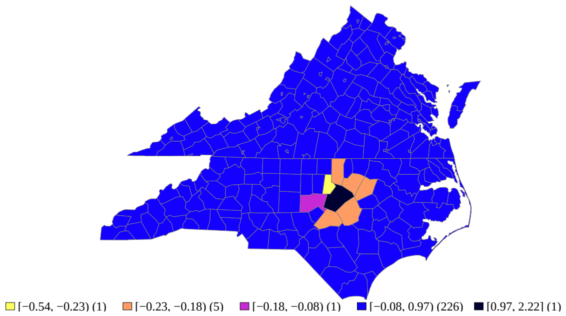
Figure 20: Impact (%) of One Percentage Point Increase in Evangelical Protestants in Wake County, NC by Base Model Spatial Regression



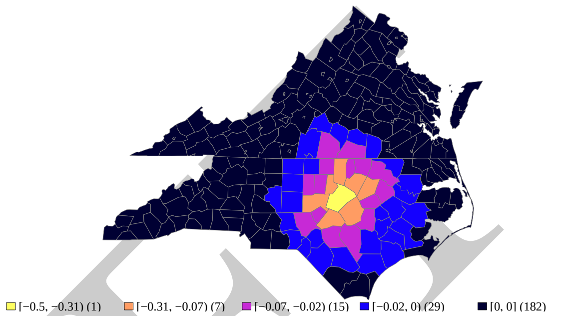
(a) $\text{CO}^{c,\circ,q1}$



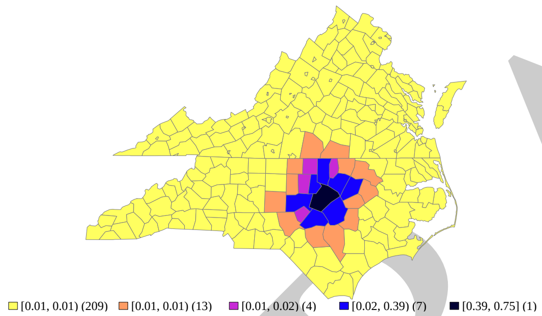
(b) $\text{NH}_3{}^{\dagger,b,\bullet,q1}$



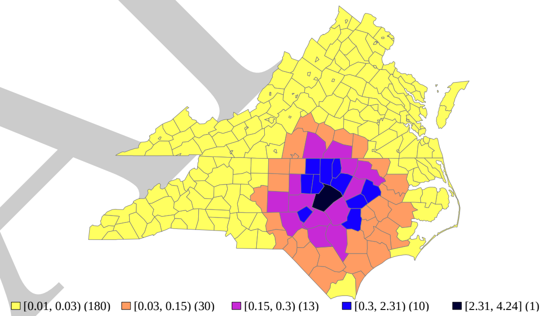
(c) $\text{NO}_x{}^{\dagger,b,\circ,q1}$



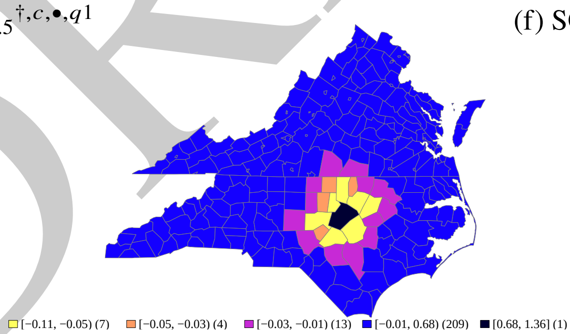
(d) $\text{PM}_{10}{}^{\dagger,c,\bullet,q1}$



(e) $\text{PM}_{2.5}{}^{\dagger,c,\bullet,q1}$

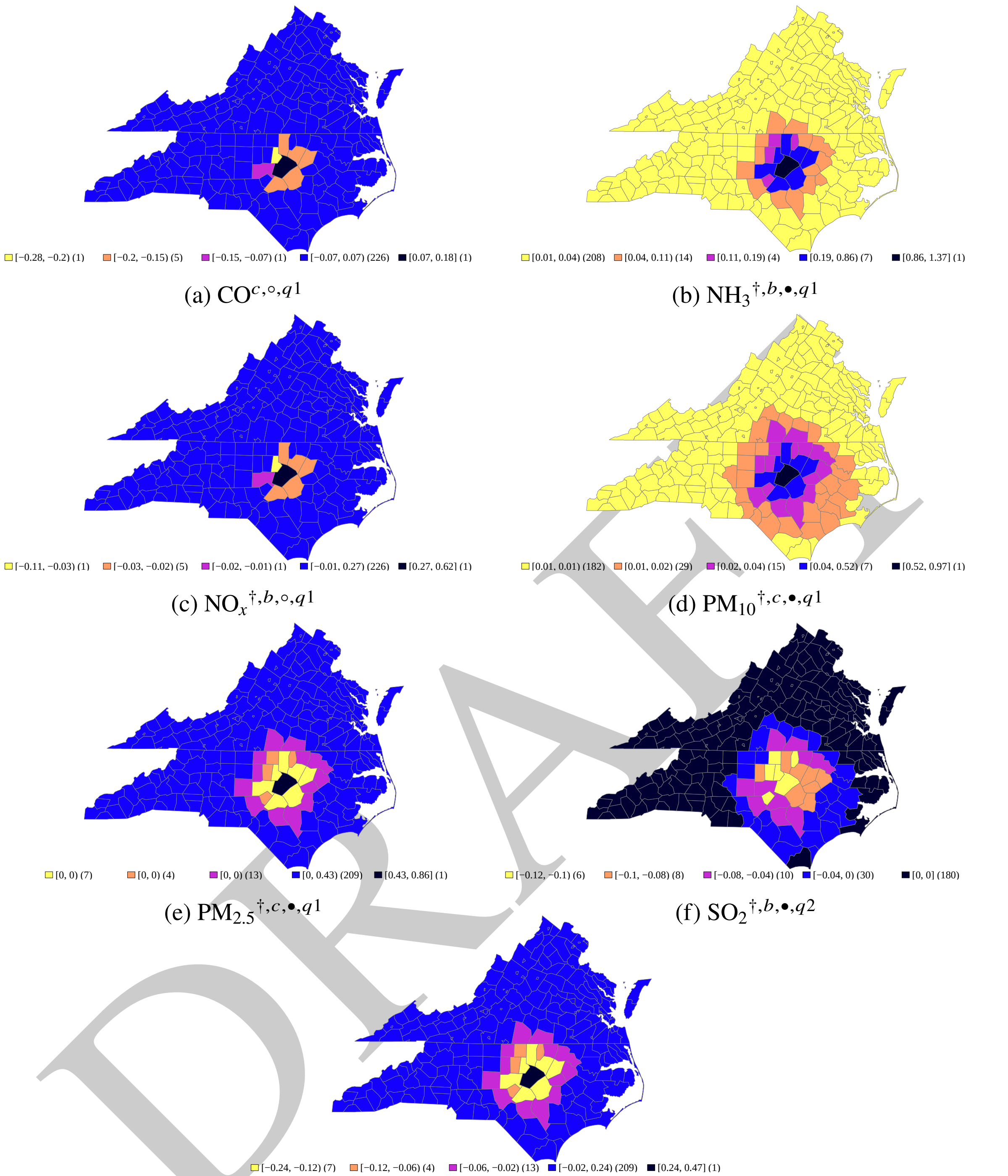


(f) $\text{SO}_2{}^{\dagger,b,\bullet,q2}$



(g) $\text{VOC}^{b,\bullet,q1}$

Figure 21: Impact (%) of One Percentage Point Increase in Black Protestants in Wake County, NC by Base Model Spatial Regression



(a) $CO^{c,\circ,q1}$

(b) ${NH_3}^{\dagger,b,\bullet,q1}$

(c) ${NO_x}^{\dagger,b,\circ,q1}$

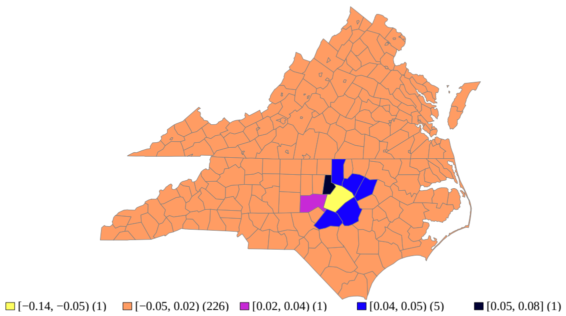
(d) ${PM_{10}}^{\dagger,c,\bullet,q1}$

(e) ${PM_{2.5}}^{\dagger,c,\bullet,q1}$

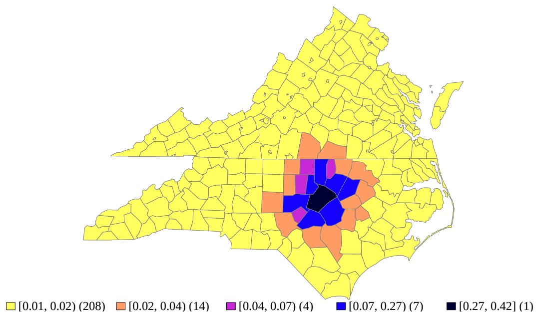
(f) ${SO_2}^{\dagger,b,\bullet,q2}$

(g) $VOC^{b,\bullet,q1}$

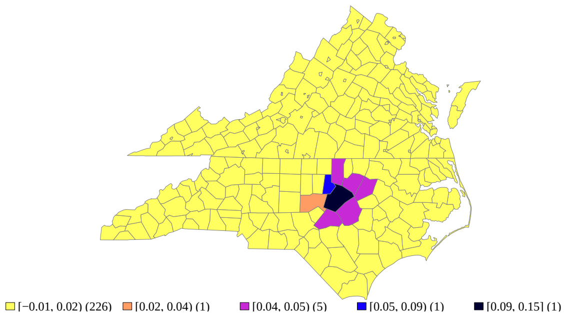
Figure 22: Impact (%) of One Percentage Point Increase in Mainline Protestants in Wake County, NC by Base Model Spatial Regression



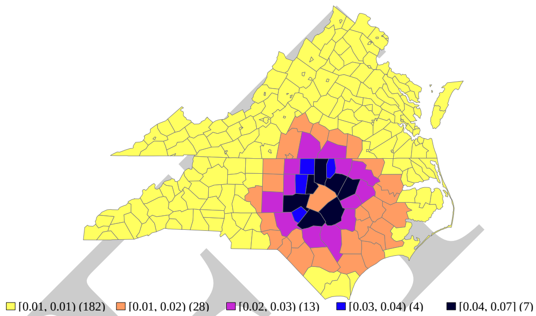
(a) $CO^{c,\circ,q1}$



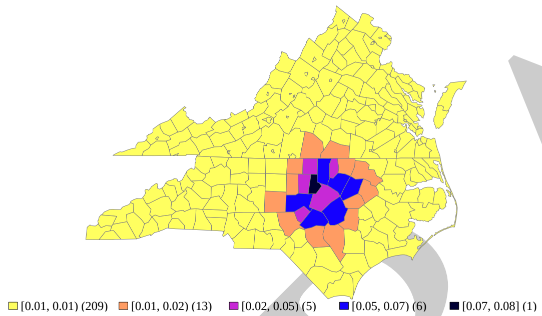
(b) $NH_3{}^{\dagger,b,\bullet,q1}$



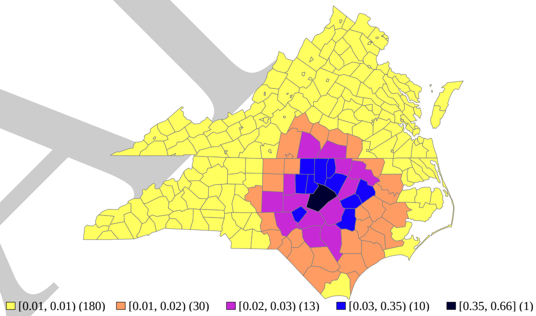
(c) $NO_x{}^{\dagger,b,\circ,q1}$



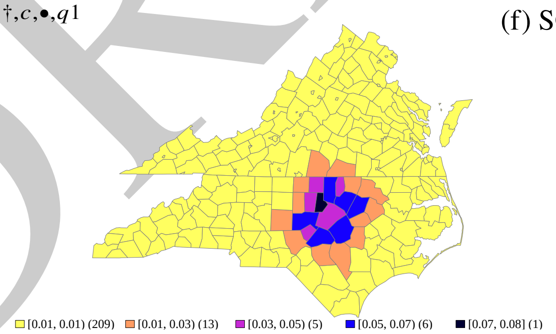
(d) $PM_{10}{}^{\dagger,c,\bullet,q1}$



(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$

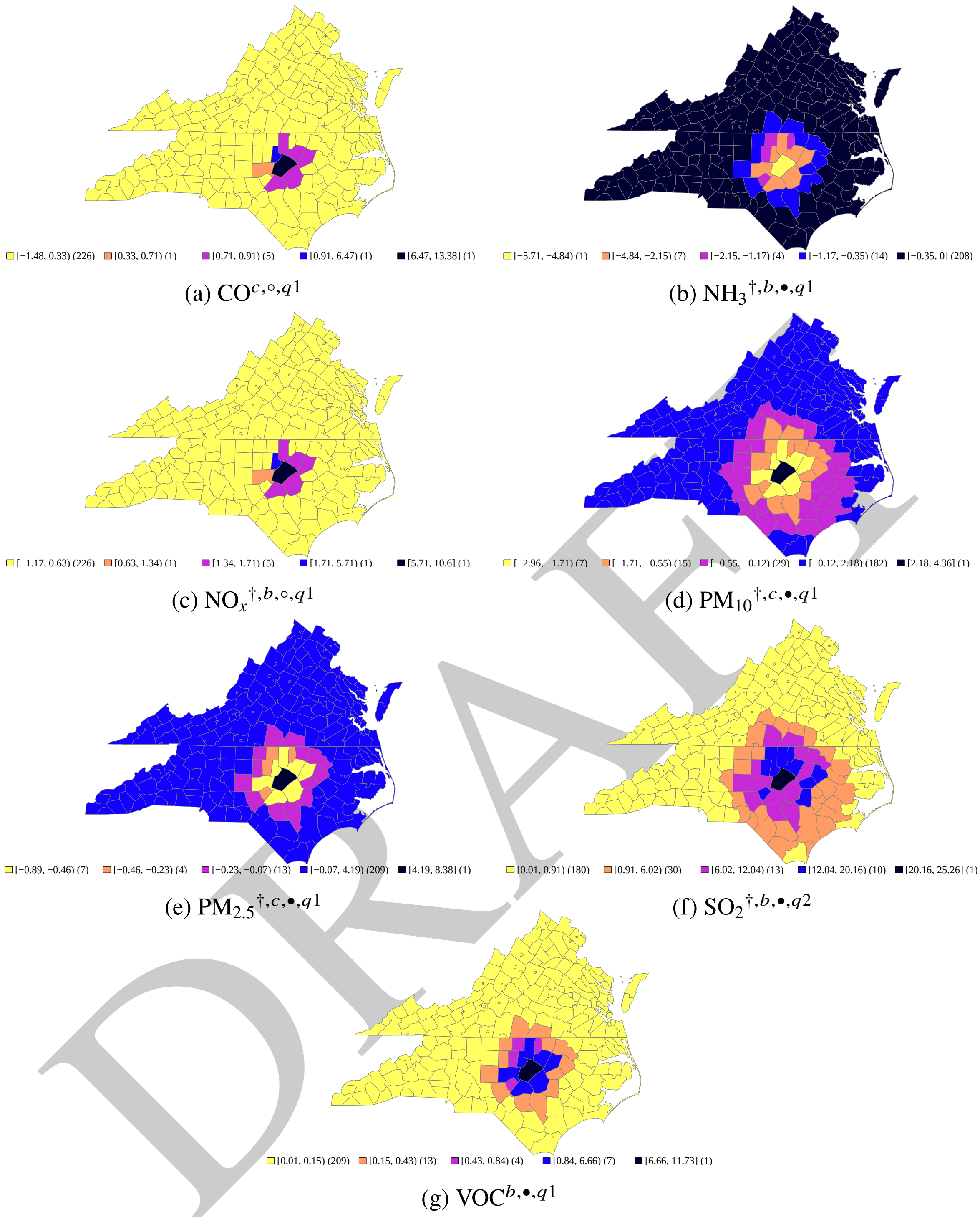


(f) $SO_2{}^{\dagger,b,\bullet,q2}$



(g) $VOC^{b,\bullet,q1}$

Figure 23: Impact (%) of One Percentage Point Increase in Catholics in Wake County, NC by Base Model Spatial Regression



(a) $CO^{c,\circ,q1}$

(b) $NH_3{}^{\dagger,b,\bullet,q1}$

(c) $NO_x{}^{\dagger,b,\circ,q1}$

(d) $PM_{10}{}^{\dagger,c,\bullet,q1}$

(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$

(f) $SO_2{}^{\dagger,b,\bullet,q2}$

(g) $VOC^{b,\bullet,q1}$

Figure 24: Impact (%) of One Percentage Point Increase in Orthodox Christians in Wake County, NC by Base Model Spatial Regression

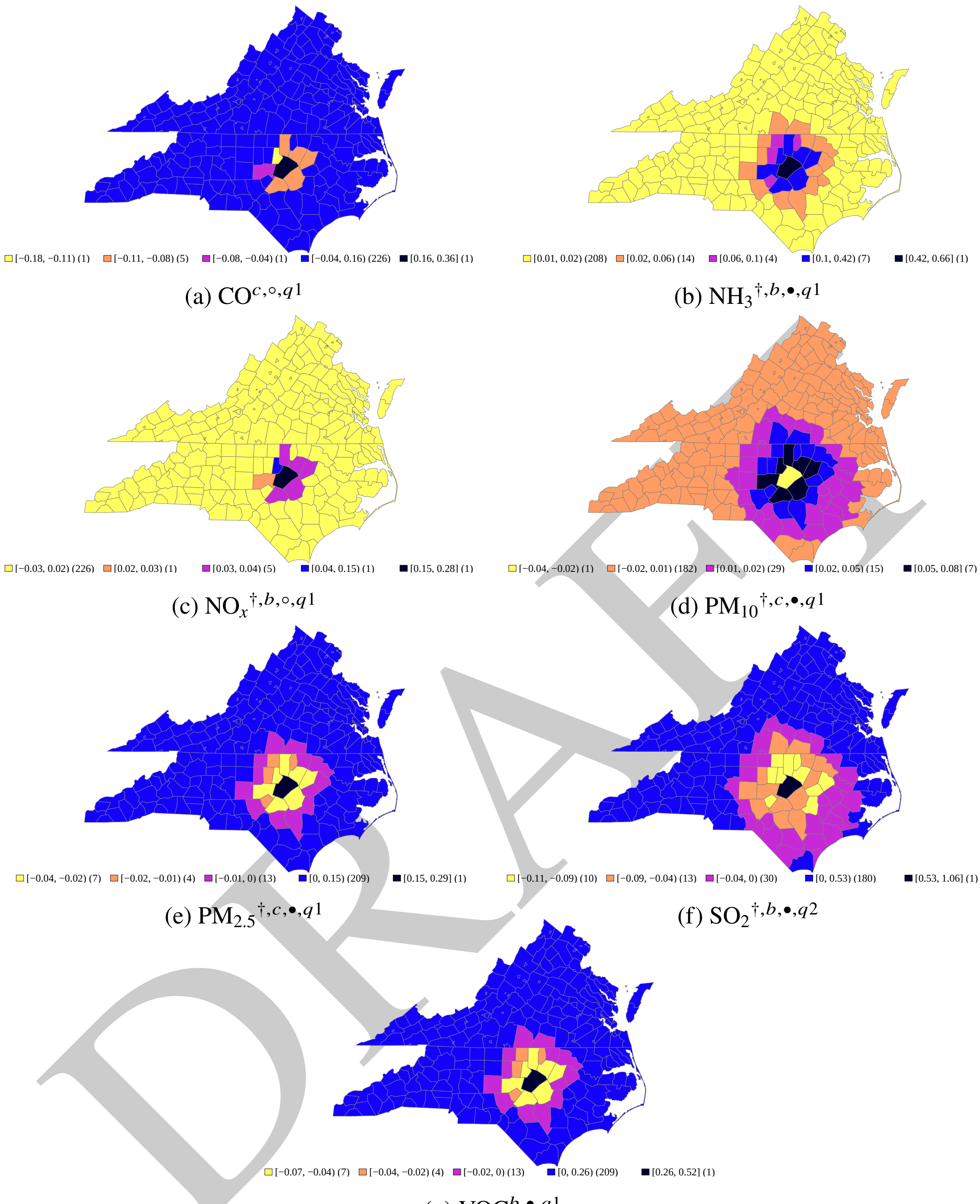
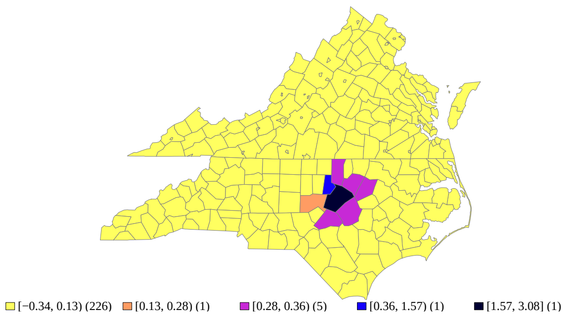
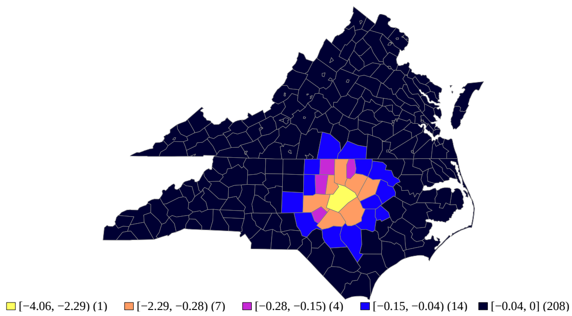


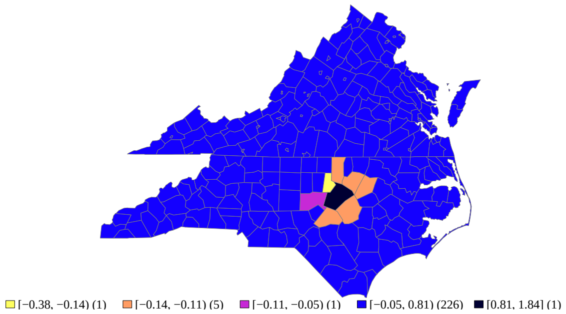
Figure 25: Impact (%) of One Percentage Point Increase in Mormons in Wake County, NC by Base Model Spatial Regression



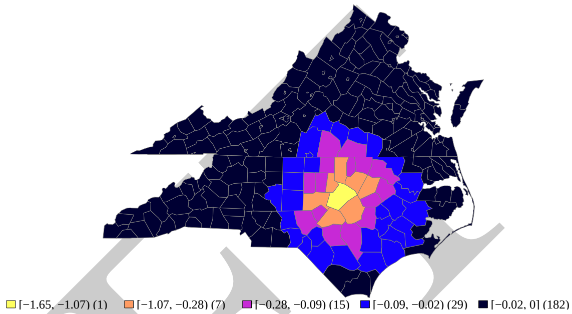
(a) $CO^{c,\circ,q1}$



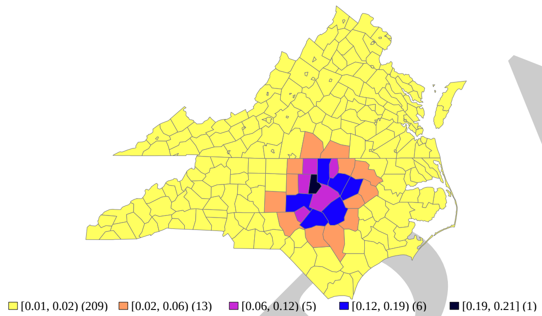
(b) $NH_3{}^{\dagger,b,\bullet,q1}$



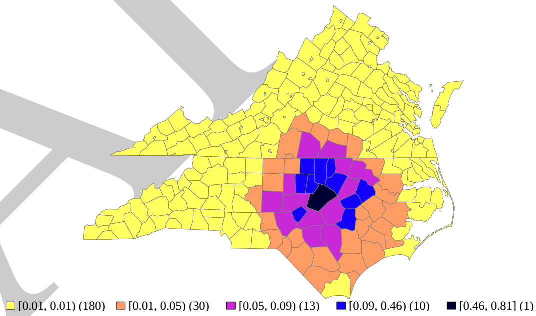
(c) $NO_x{}^{\dagger,b,\circ,q1}$



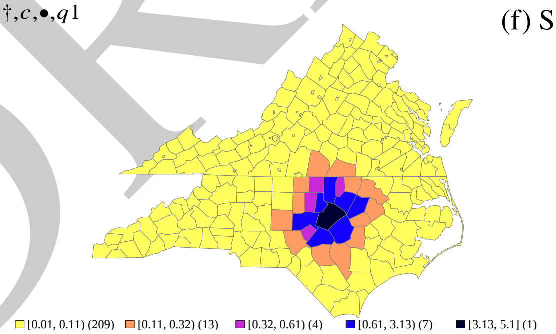
(d) $PM_{10}{}^{\dagger,c,\bullet,q1}$



(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$

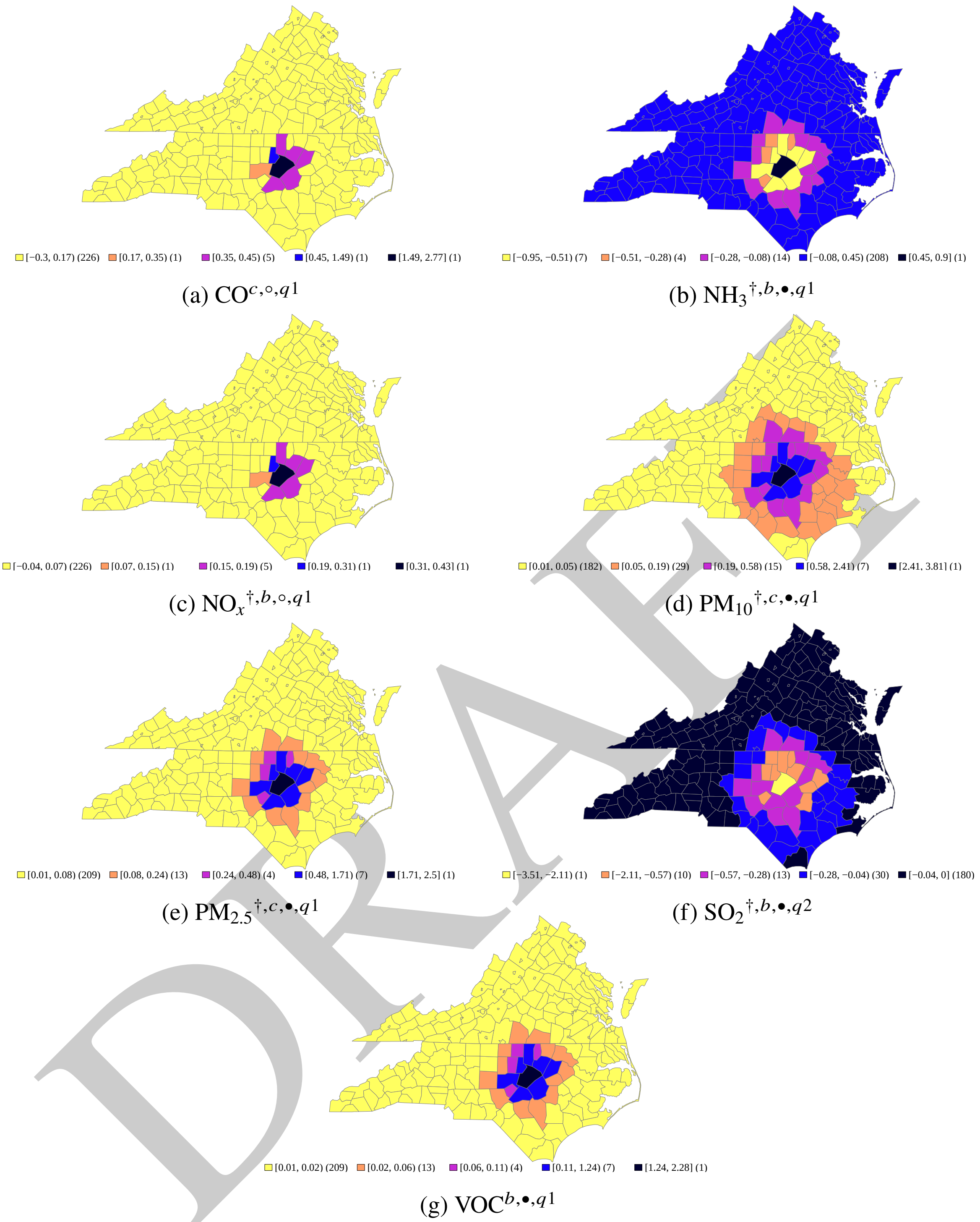


(f) $SO_2{}^{\dagger,b,\bullet,q2}$



(g) $VOC^{b,\bullet,q1}$

Figure 26: Impact (%) of One Percentage Point Increase in Muslims in Wake County, NC by Base Model Spatial Regression



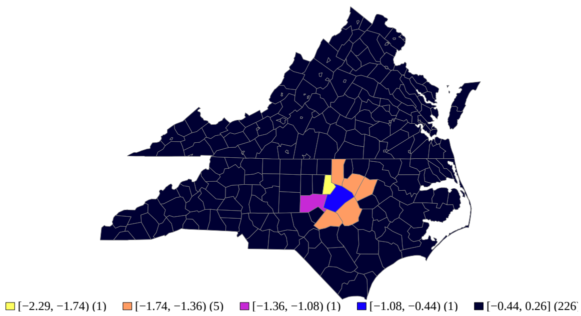
(a) $CO^{c,\circ,q1}$ (b) $NH_3{}^{\dagger,b,\bullet,q1}$

(c) $NO_x{}^{\dagger,b,\circ,q1}$ (d) $PM_{10}{}^{\dagger,c,\bullet,q1}$

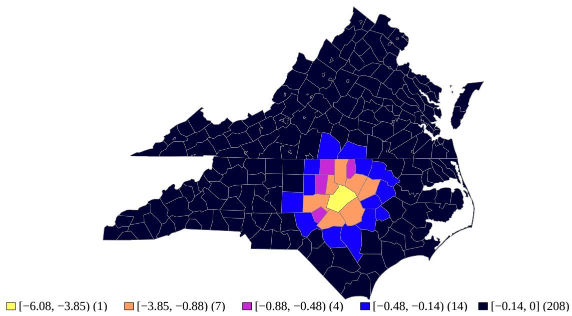
(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$ (f) $SO_2{}^{\dagger,b,\bullet,q2}$

(g) $VOC^{b,\bullet,q1}$

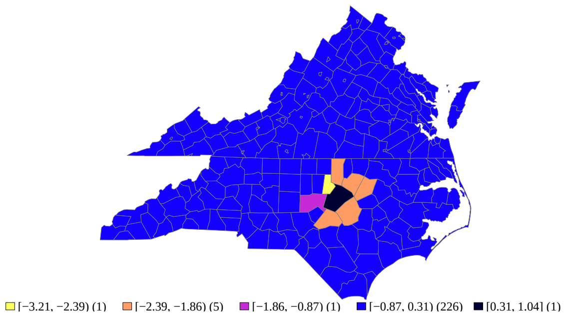
Figure 27: Impact (%) of One Percentage Point Increase in Jews in Wake County, NC by Base Model Spatial Regression



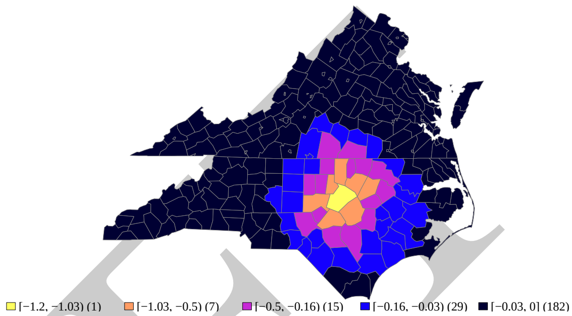
(a) $CO^{c,\circ,q1}$



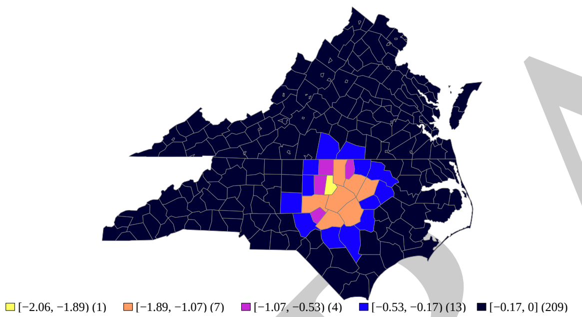
(b) $NH_3{}^{\dagger,b,\bullet,q1}$



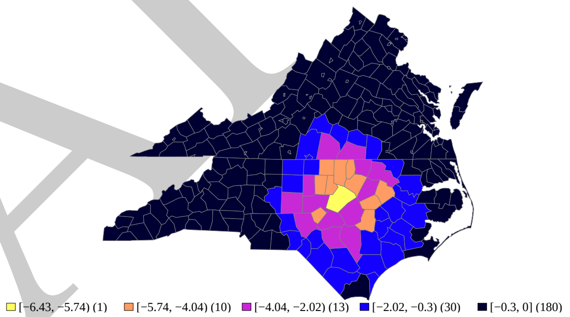
(c) $NO_x{}^{\dagger,b,\circ,q1}$



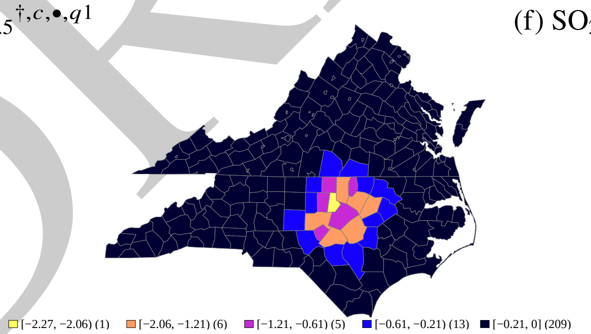
(d) $PM_{10}{}^{\dagger,c,\bullet,q1}$



(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$



(f) $SO_2{}^{\dagger,b,\bullet,q2}$



(g) $VOC^{b,\bullet,q1}$

Figure 28: Impact (%) of One Percentage Point Increase in Hindus in Wake County, NC by Base Model Spatial Regression

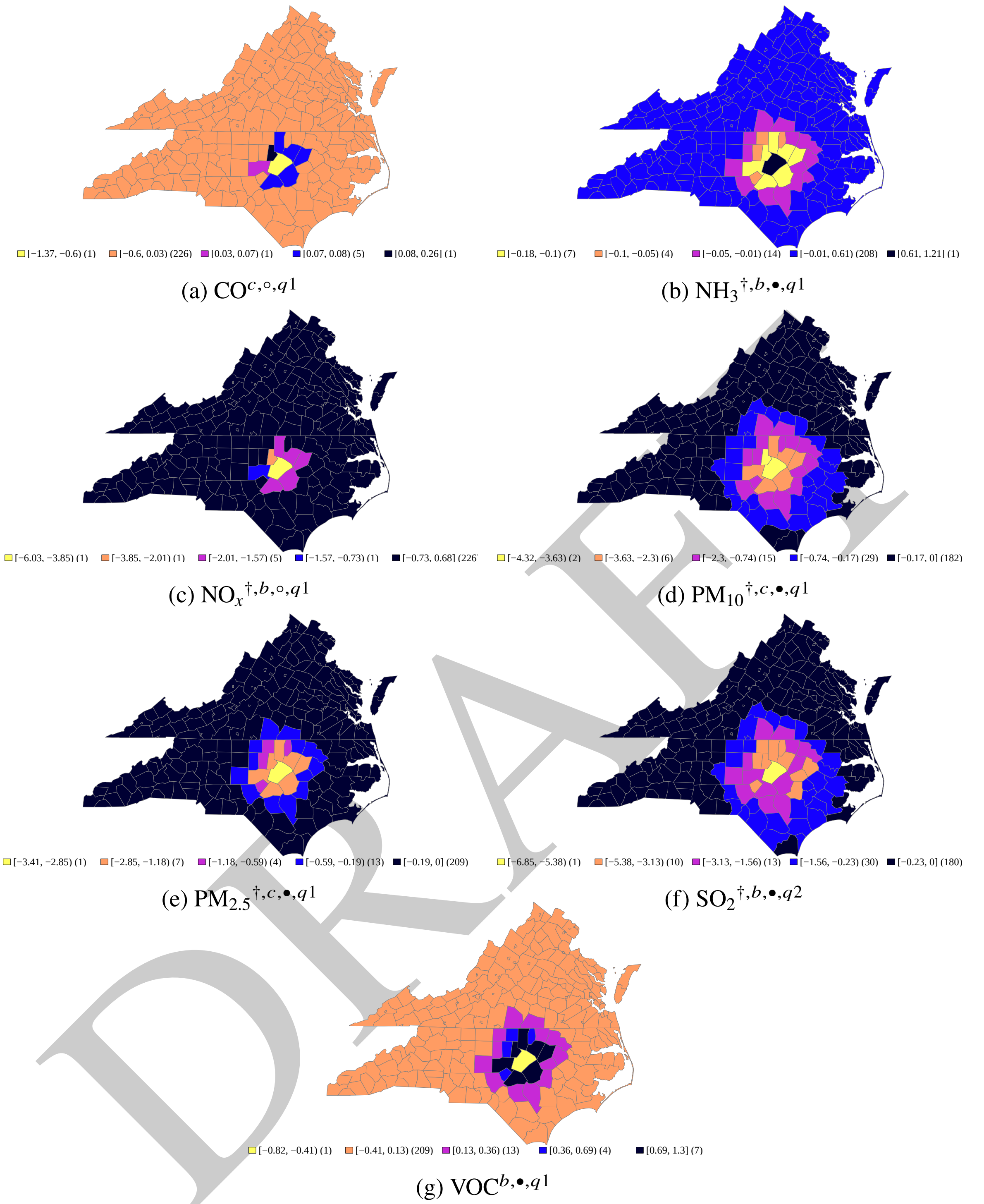
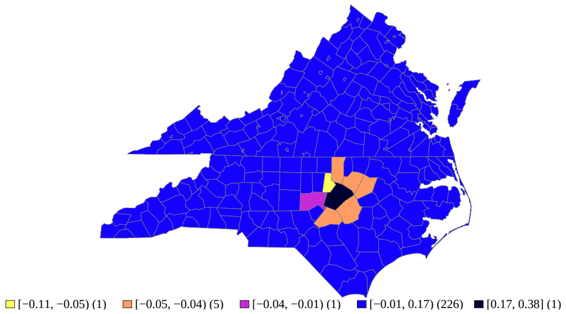
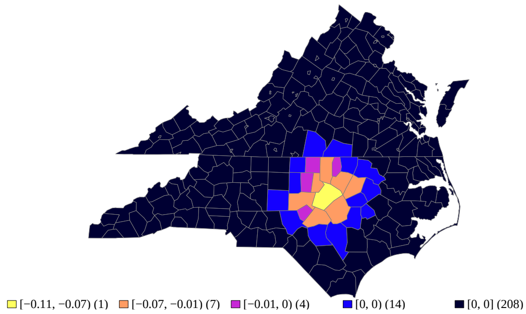


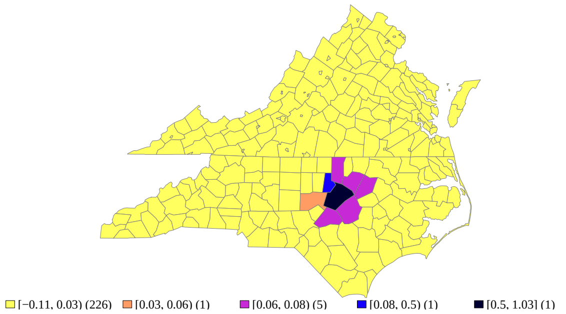
Figure 29: Impact (%) of One Percentage Point Increase in Buddhists in Wake County, NC by Base Model Spatial Regression



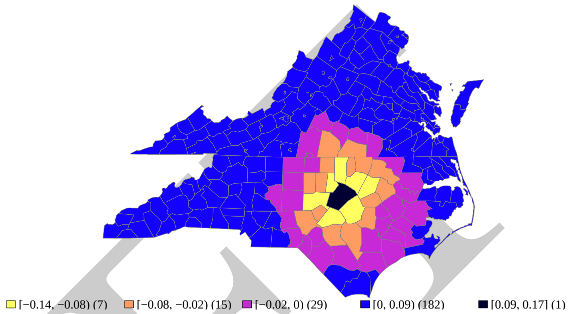
(a) $CO^{c,\circ,q1}$



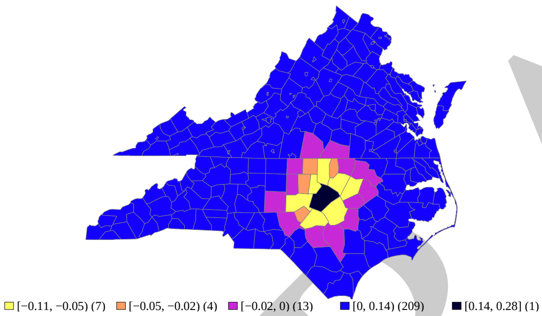
(b) $NH_3{}^{\dagger,b,\bullet,q1}$



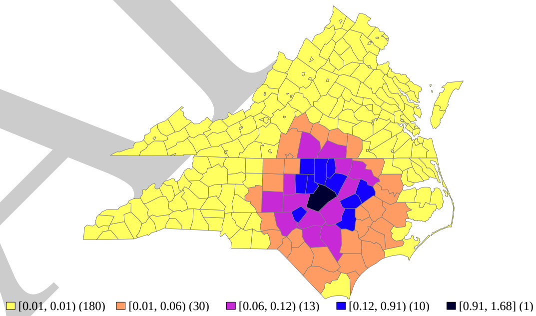
(c) $NO_x{}^{\dagger,b,\circ,q1}$



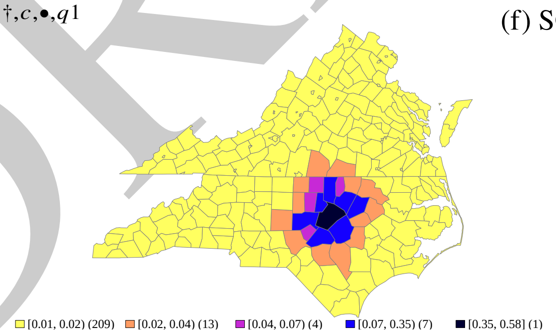
(d) $PM_{10}{}^{\dagger,c,\bullet,q1}$



(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$

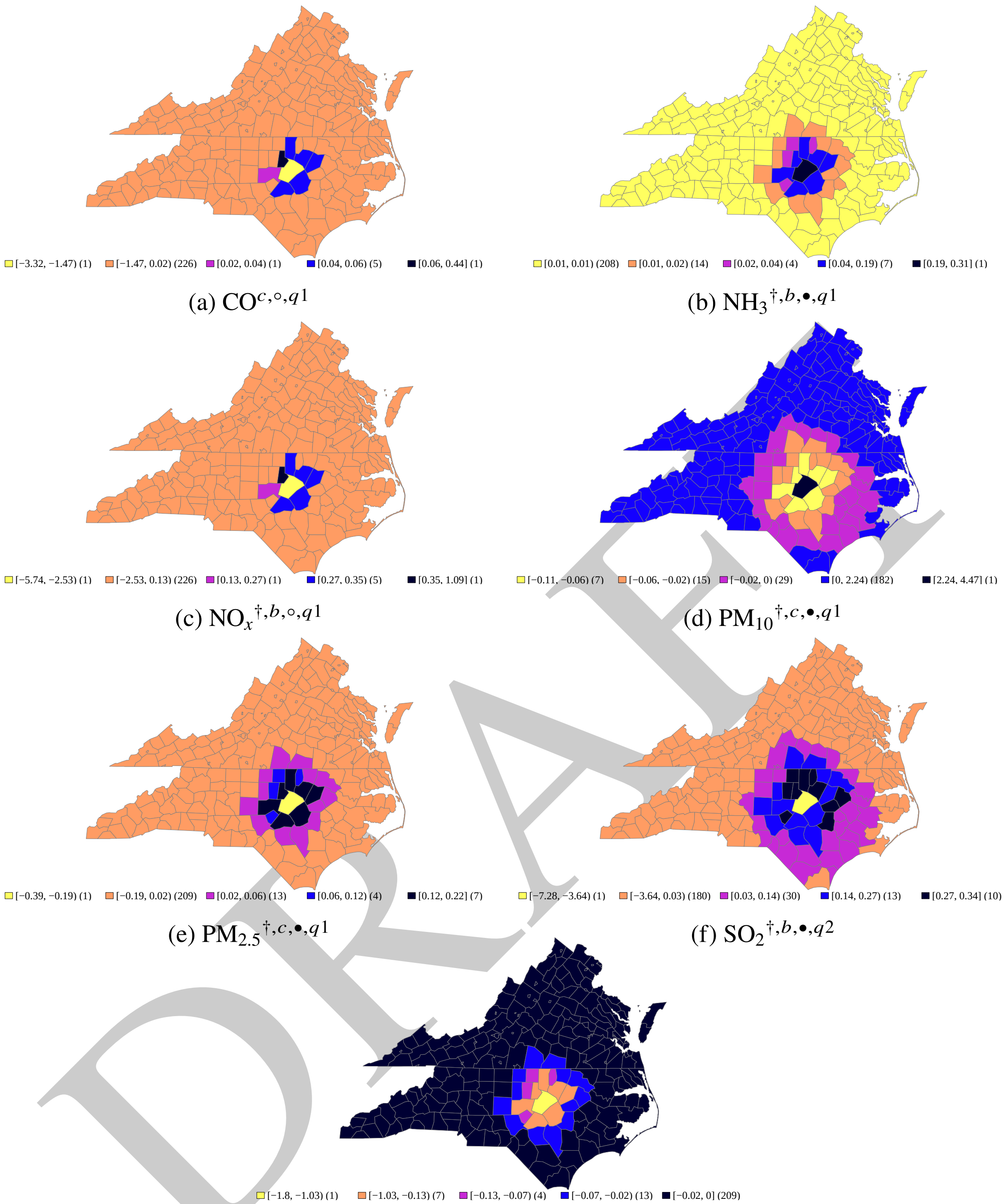


(f) $SO_2{}^{\dagger,b,\bullet,q2}$



(g) $VOC^{b,\bullet,q1}$

Figure 30: Impact (%) of 1% Increase in Income in Wake County, NC by Base Model Spatial Regression



(a) $CO^{c,\circ,q1}$

(b) $NH_3{}^{\dagger,b,\bullet,q1}$

(c) $NO_x{}^{\dagger,b,\circ,q1}$

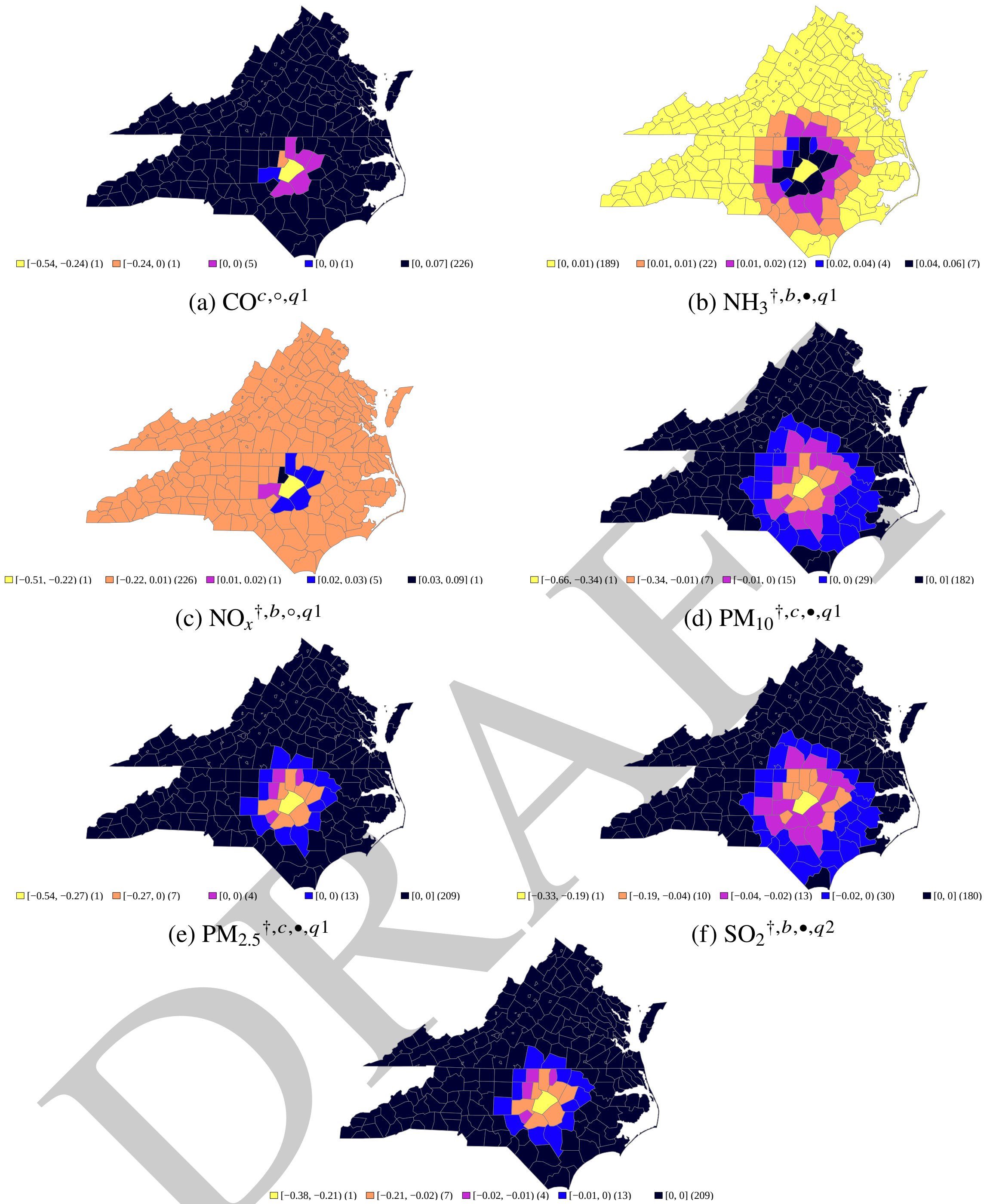
(d) $PM_{10}{}^{\dagger,c,\bullet,q1}$

(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$

(f) $SO_2{}^{\dagger,b,\bullet,q2}$

(g) $VOC^{b,\bullet,q1}$

Figure 31: Impact (%) of 1% Increase in Gas Price in Wake County, NC by Base Model Spatial Regression



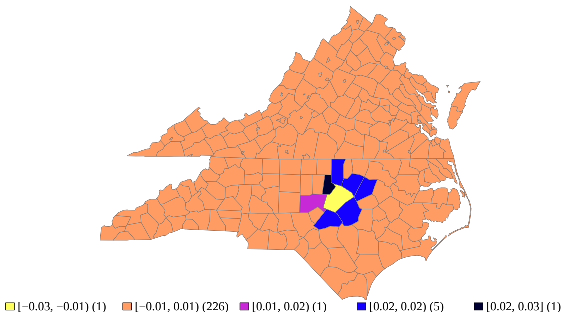
(a) $CO^{c,\circ,q1}$ (b) $NH_3{}^{\dagger,b,\bullet,q1}$

(c) $NO_x{}^{\dagger,b,\circ,q1}$ (d) $PM_{10}{}^{\dagger,c,\bullet,q1}$

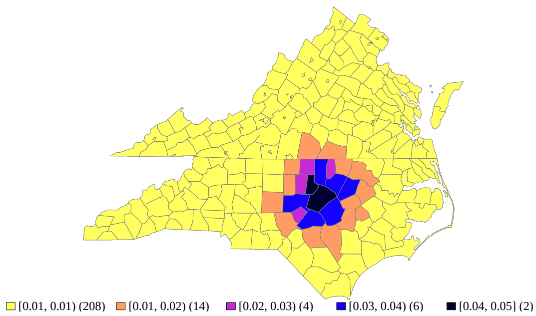
(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$ (f) $SO_2{}^{\dagger,b,\bullet,q2}$

(g) $VOC^{b,\bullet,q1}$

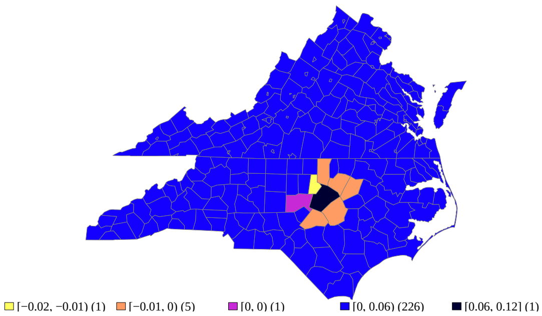
Figure 32: Impact (%) of 1% Increase in Gas Tax/Fee in Wake County, NC by Base Model Spatial Regression



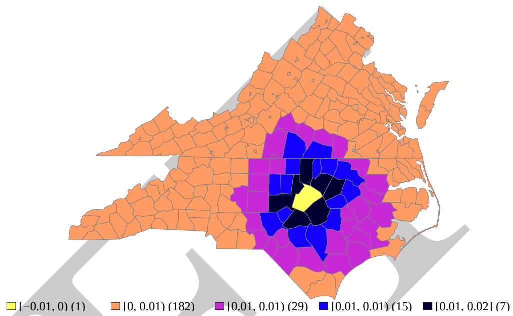
(a) $CO^{c,\circ,q1}$



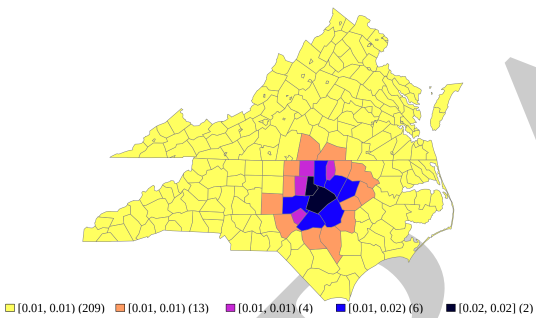
(b) $NH_3{}^{\dagger,b,\bullet,q1}$



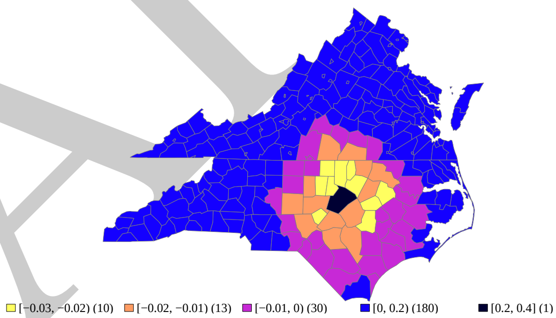
(c) $NO_x{}^{\dagger,b,\circ,q1}$



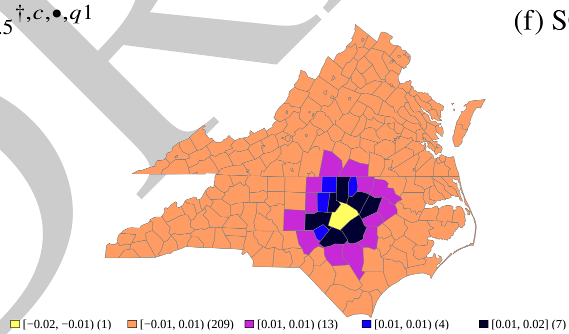
(d) $PM_{10}{}^{\dagger,c,\bullet,q1}$



(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$

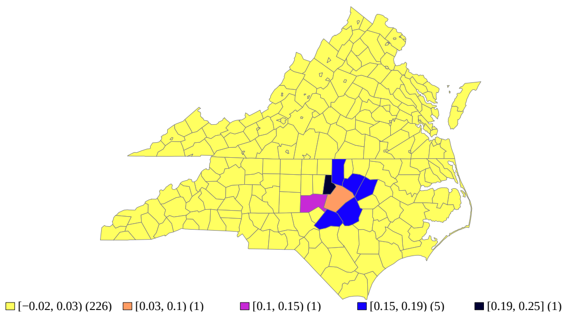


(f) $SO_2{}^{\dagger,b,\bullet,q2}$

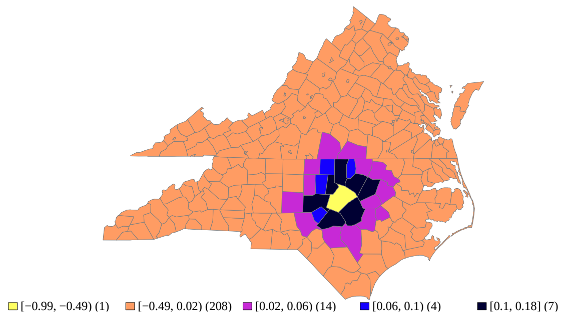


(g) $VOC^{b,\bullet,q1}$

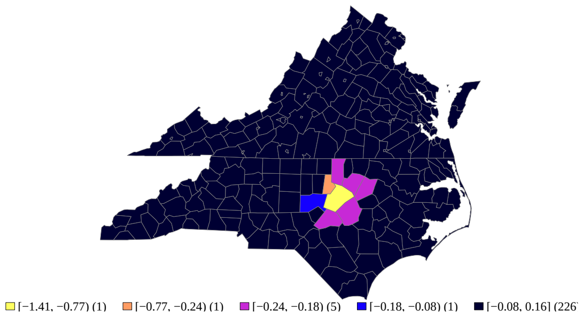
Figure 33: Impact (%) of 1% Increase in Renewable Energy Consumption in Wake County, NC by Base Model Spatial Regression



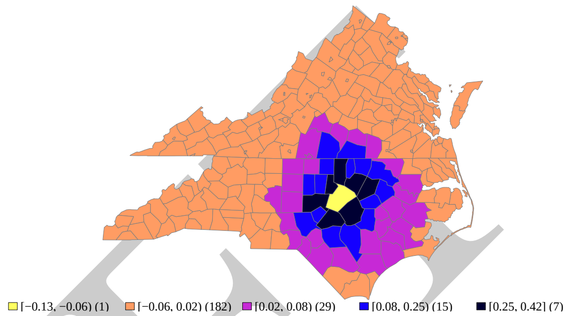
(a) $CO^{c,\circ,q1}$



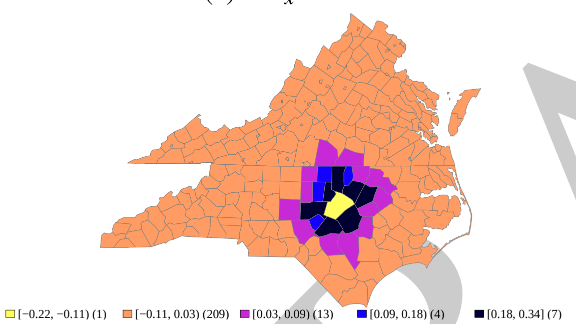
(b) $NH_3{}^{\dagger,b,\bullet,q1}$



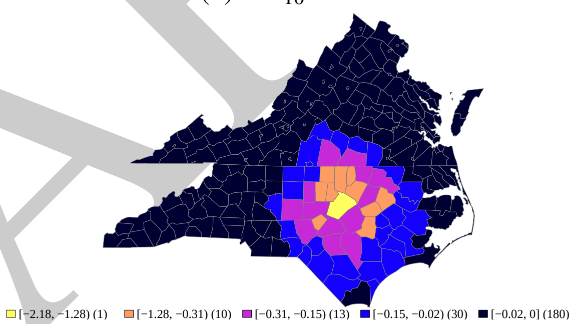
(c) $NO_x{}^{\dagger,b,\circ,q1}$



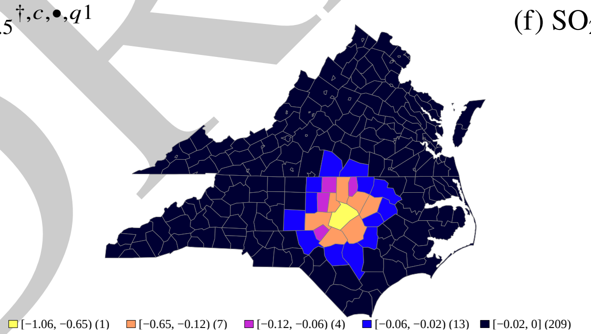
(d) $PM_{10}{}^{\dagger,c,\bullet,q1}$



(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$

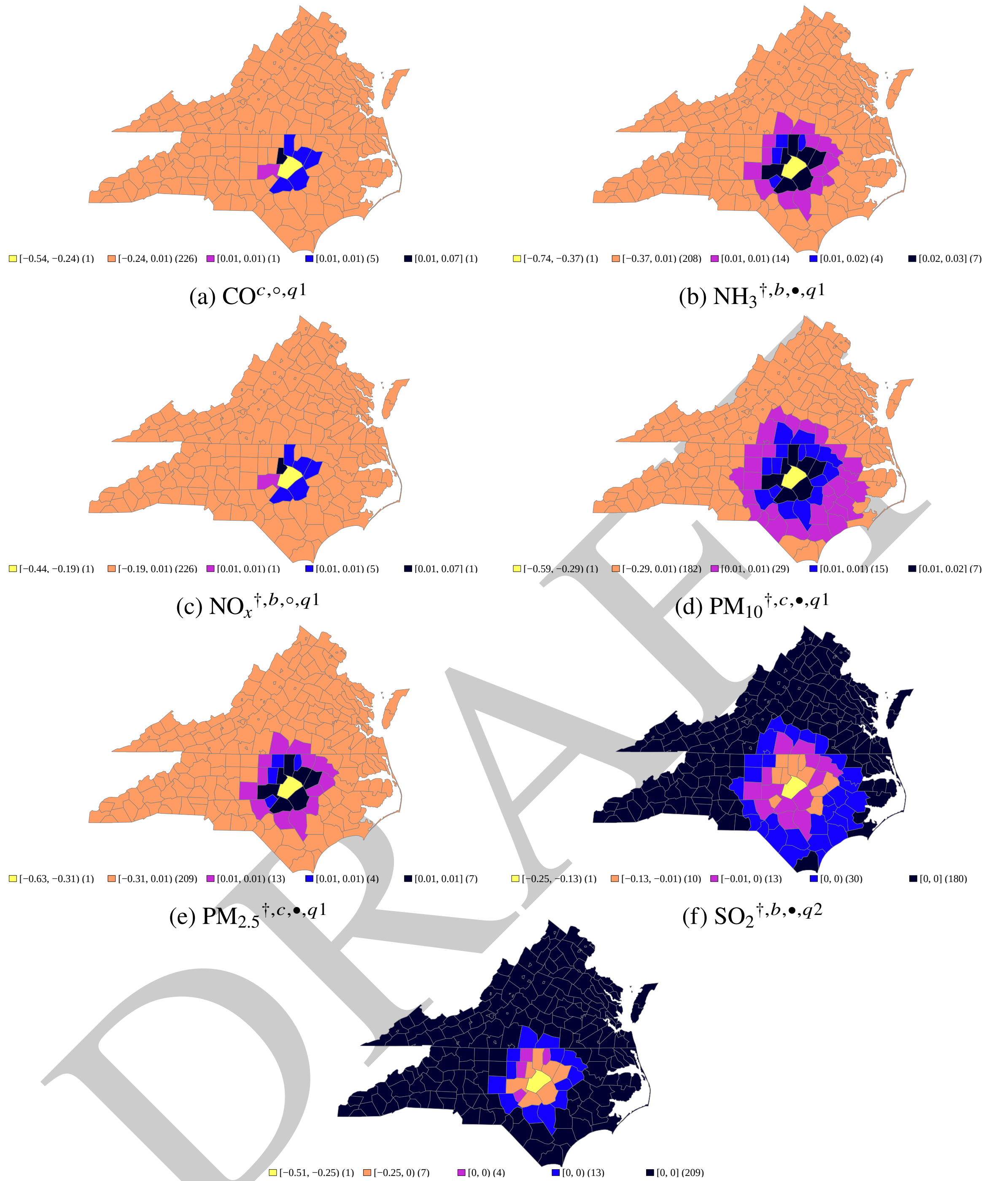


(f) $SO_2{}^{\dagger,b,\bullet,q2}$



(g) $VOC^{b,\bullet,q1}$

Figure 34: Impact (%) of One Percentage Point Increase in Education Variable in Wake County, NC by Base Model Spatial Regression



(a) $CO^{c,\circ,q1}$

(b) $NH_3{}^{\dagger,b,\bullet,q1}$

(c) $NO_x{}^{\dagger,b,\circ,q1}$

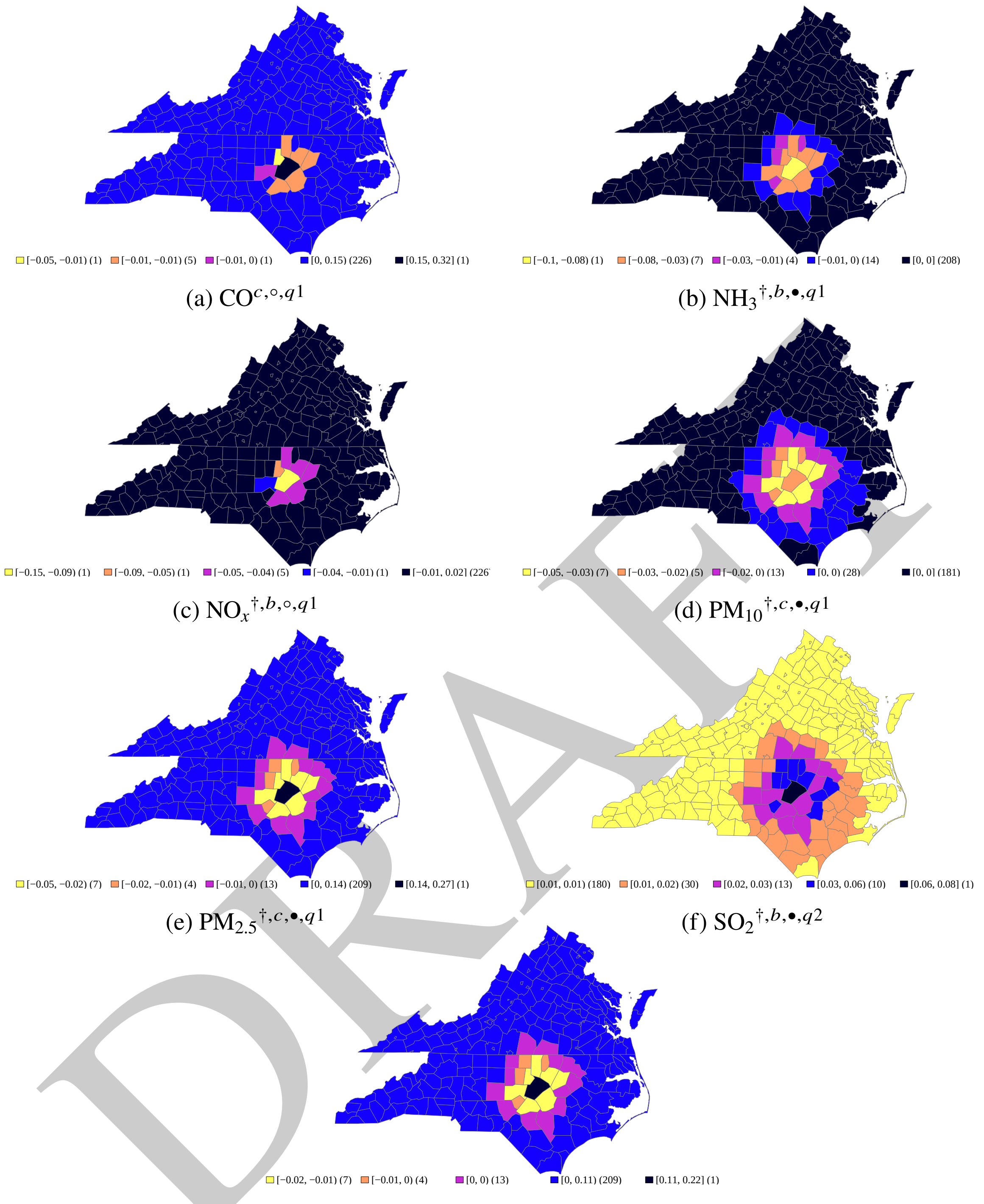
(d) $PM_{10}{}^{\dagger,c,\bullet,q1}$

(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$

(f) $SO_2{}^{\dagger,b,\bullet,q2}$

(g) $VOC^{b,\bullet,q1}$

Figure 35: Impact (%) of 1% Increase in Population Density in Wake County, NC by Base Model Spatial Regression



(a) $CO^{c,\circ,q1}$

(b) $NH_3{}^{\dagger,b,\bullet,q1}$

(c) $NO_x{}^{\dagger,b,\circ,q1}$

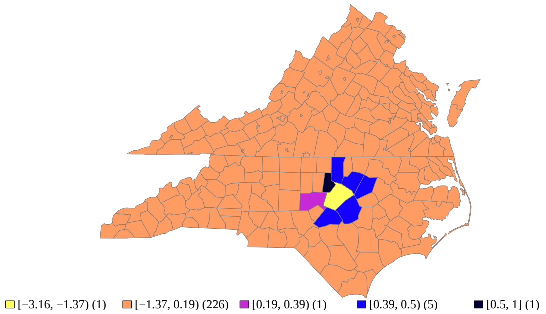
(d) $PM_{10}{}^{\dagger,c,\bullet,q1}$

(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$

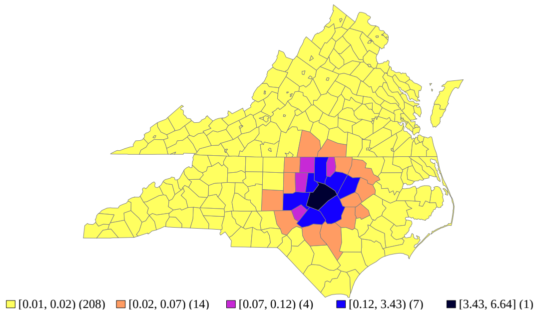
(f) $SO_2{}^{\dagger,b,\bullet,q2}$

(g) $VOC^{b,\bullet,q1}$

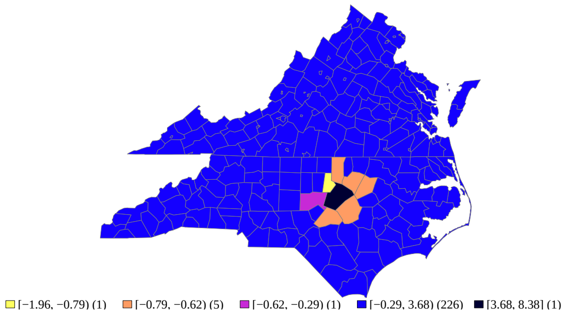
Figure 36: Impact (%) of 1% Increase in Mean Daily Precipitation in Wake County, NC by Base Model Spatial Regression



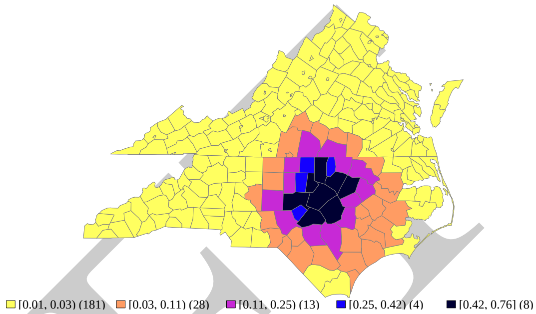
(a) $CO^{c,\circ,q1}$



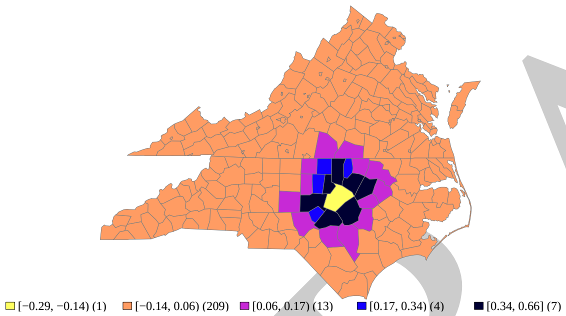
(b) $NH_3{}^{\dagger,b,\bullet,q1}$



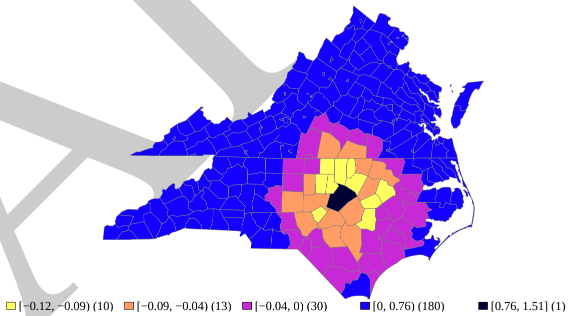
(c) $NO_x{}^{\dagger,b,\circ,q1}$



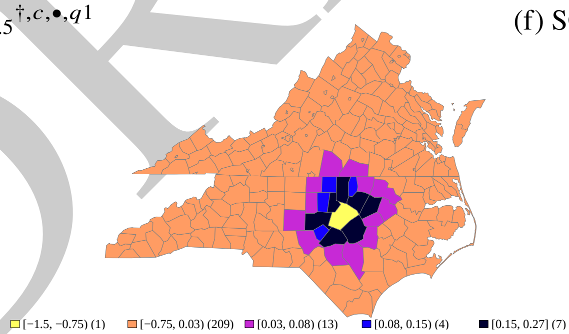
(d) $PM_{10}{}^{\dagger,c,\bullet,q1}$



(e) $PM_{2.5}{}^{\dagger,c,\bullet,q1}$



(f) $SO_2{}^{\dagger,b,\bullet,q2}$



(g) $VOC^{b,\bullet,q1}$

Figure 37: Impact (%) of 1% Increase in Mean Daily Maximum Heat Index in Wake County, NC by Base Model Spatial Regression