5-17-2019

Spatial Analysis of County Level Drug Overdose Deaths and Associated Factors, Over Two Time Periods in the United States

Sunanda Sarkar
Georgia State University

Follow this and additional works at: https://scholarworks.gsu.edu/iph_theses

Recommended Citation
https://scholarworks.gsu.edu/iph_theses/642

This Thesis is brought to you for free and open access by the School of Public Health at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Public Health Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
ABSTRACT

Spatial Analysis of County Level Drug Overdose Deaths and Associated Factors, Over Two Time Periods in the United States

By

SUNANDA SARKAR

April 18, 2019

INTRODUCTION: Recently, drug overdose is being considered as an important public health issue, the magnitude of which is yet to be adequately explored. The United States is experiencing a wide range of drug overdose problems over the past decades, where fatal overdoses have tripled from 1999 to 2016. Geographic approaches to drug overdose death research have emerged in recent years. Studies demonstrated that overdose mortalities are not equally distributed across different geographic areas. Therefore, it is important to consider geographic variations to inform effective prevention and treatment of drug overdoses and prevent premature deaths.

AIM: The aim is to explore spatial distribution of county level drug overdose death rates in the contiguous U.S. over two 5-year time intervals (2007-2011 and 2012-2016); identify and evaluate the extent to which the county level socio-economic and socio-demographic factors are associated with the spatial patterning and explain it.

METHODS: Exploratory spatial cluster analysis was performed to determine whether patterns of observed drug overdose mortality are spatially random or not over two time periods. Both traditional and Empirical Bayes standardization methods were used for spatial autocorrelation test. To determine any change over time, observations in the data are stacked based on time. Time stacked spatial regression analysis was performed to determine the associations between several county level socio-economic and socio-demographic factors and drug overdose death rates in the U.S.

RESULTS: Mean drug overdose death rate increased from early to late time period. Results indicates the presence of significant (at 5% significance level) spatial autocorrelation among the adjacent counties in the drug overdose death rates, and this spatial pattern differs in two time periods. Finally, spatial regression indicates that the effect of different contextual factors are heterogenous over time and across different population.

CONCLUSION: Findings may help inform efforts to prevent, diagnose or treat drug overdoses ahead of time, thus prevent premature deaths by understanding the geographic variations and identifying the areas with growing burdens. Studies focusing on similar associations across different age-groups and insured group may provide better insight.
Spatial Analysis of County Level Drug Overdose Deaths and Associated Factors, Over Two Time Periods in the United States

by

SUNANDA SARKAR

Bachelor of Medicine and Surgery (MBBS), DHAKA MEDICAL COLLEGE, UNIVERSITY OF DHAKA

A Thesis Submitted to the Graduate Faculty of Georgia State University in Partial Fulfillment of the Requirements for the Degree

MASTER OF PUBLIC HEALTH

ATLANTA, GEORGIA 30303
APPROVAL PAGE

SPATIAL ANALYSIS OF COUNTY LEVEL DRUG OVERDOSE DEATHS AND ASSOCIATED FACTORS, OVER TWO TIME PERIODS IN THE UNITED STATES

by

SUNANDA SARKAR

Approved:

___ Lee Rivers Mobley, PhD___
Committee Chair

__ Ruiyan Luo, PhD____
Committee Member

__April 18, 2019_________
Date
Acknowledgments

I would like to start by thanking God, for the wisdom and strength He granted me to finish this process.

I would also like to thank my family, friends, and colleagues for their continued support and encouragement throughout this program and thesis writing.

I am very much grateful to my thesis committee, Dr. Mobley and Dr. Luo, for their guidance throughout this entire process. It has been an honor working with you both. I have gained a tremendous amount of knowledge while working with each of you. This would not have been possible without your help.

Finally, I would like to thank Georgia State University to provide me the opportunity to pursue my study.
Author’s Statement Page

In presenting this thesis as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or to publish this thesis may be granted by the author or, in his/her absence, by the professor under whose direction it was written, or in his/her absence, by the Associate Dean, School of Public Health. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without written permission of the author.

___Sunanda Sarkar______________
Signature of Author
TABLE OF CONTENTS

APPROVAL PAGE ............................................................................................................. II

ACKNOWLEDGMENTS ................................................................................................. III

AUTHOR’S STATEMENT PAGE .................................................................................. IV

LIST OF TABLES .............................................................................................................. VI

LIST OF FIGURES .......................................................................................................... VII

CHAPTER I: INTRODUCTION ......................................................................................... 1

1.1 Overview .................................................................................................................. 1

1.2 Study objectives ........................................................................................................ 4

1.3 Research questions ................................................................................................... 4

CHAPTER II: REVIEW OF THE LITERATURE ............................................................... 5

2.1 Drugs commonly involved in overdose .................................................................. 5

2.2 Overdose related to prescription drugs and insurance coverage ....................... 6

2.3 Overdose related to socioeconomic and socio-demographic determinants ....... 8

2.4 Geographic variations of overdose mortality ....................................................... 9

CHAPTER III: METHODS AND PROCEDURES ......................................................... 11

3.1 Data sources ............................................................................................................ 11

3.2 Conceptual model/ framework for drug overdose mortality .............................. 13

3.3 Data processing and Measures ............................................................................. 16

3.4 Statistical analysis .................................................................................................. 18

CHAPTER IV: RESULTS .............................................................................................. 24

4.1 Descriptive analysis ............................................................................................... 24

4.2 Tests of spatial autocorrelation ............................................................................. 24

4.3 Spatial regression results ....................................................................................... 25

CHAPTER V: DISCUSSION AND CONCLUSION ......................................................... 29

5.1 Spatial clusters for drug overdose deaths ............................................................. 29

5.2 Spatial regression ................................................................................................... 30

5.3 Limitations and strengths ...................................................................................... 33

5.4 Implications and future directions ....................................................................... 34

REFERENCES ................................................................................................................ 35

APPENDIX: TABLES AND FIGURES ......................................................................... 43
List of Tables

Table 1. Variables used for this study, description, metric, sources and years. All the variables are aggregated over two time intervals at county level

Table 2. Descriptive statistics for all county-level variables used in the model

Table 3. Moran’s I statistics for spatial autocorrelation for drug overdose death rates in early & late period (for both traditional method and Empirical Bayes standardization method)

Table 4. Model 1: Regression output from OLS (Ordinary Least Square) model with regression diagnostics

Table 5. Diagnostics for spatial dependence (Lagrange Multiplier tests and Anselin-Kelejian test)

Table 6. Model 2: Regression output from spatial two stage least squares estimation (Lag model)

Table 7. Model 3: Regression output from spatially weighted two stage least squares estimation (Combo/Durbin model)
List of Figures

**Figure 1.** Drug overdose death rates per 100,000 during early (2007-2011) and late (2012-2016) periods using the same cut points

**Figure 2.** Spatial patterns of drug overdose death rates (Traditional method) during early (2007-2011) and late (2012-2016) periods, queen contiguity weight is used

**Figure 3.** Spatial patterns of drug overdose death rates (Empirical Bayes standardization method) during early (2007-2011) and late (2012-2016) periods, queen contiguity weight is used

**Figure 4.** Histogram of drug overdose death rates (early and late)
Chapter I

Introduction

1.1 Overview

Recently, drug overdose is being considered as an important public health issue, the magnitude of which is yet to be adequately explored. However, until now it was considered to be substance abuse, crime under law, or sin per holy books, but not as injury or a public health problem (Paulozzi, 2007; Martins, Sampson, Cerdá, & Galea, 2015). Worldwide, the drug overdose mortality has been increasing considerably. For instance, drug overdose mortality in Oceania (mainly consists of Australia and New Zealand) is about 2.5 times the global average (WDR 2017, n.d.). Similarly, other countries are also experiencing higher rates than the past. Many countries in Europe, including England, Sweden, Norway, Ireland and Estonia, have higher rates for drug mortality than the previous years; Scotland has the highest rate among the countries in Europe (IOAD, n.d.). However, approximately a quarter of worldwide drug overdose deaths happen in the United States— which is the highest among all the countries in the world (WDR 2017, n.d.). While comparing the mortality trends from drug overdose from 2001 to 2015 among the 13 OECD (Organization for Economic Co-operation and Development) member countries, a recent study found the similar result of the United States having the highest mortality rate from drug overdose in 2015 (Chen, Shiels, Thomas, Freedman, & de González, 2018).

Each day, drug overdose accounts for 174 deaths in the United States (Jalal et al., 2018). According to a CDC (Center for Disease Control and Prevention) report published on January 1, 2016, there was a 137% increase in the drug overdose death rate in the United States since 2000 (Rudd, Aleshire, Zibbell, & Gladden, 2016). Another article reported that age-adjusted death rate from drug overdose approximately tripled from 1999 to 2016 (i.e., from 6.1 to 19.8 per 100,000.
population) (Hedegaard, 2017). Jalal et al. (2018) examined all drug overdose mortality in the U.S. from 1979 to 2016 and found the growth to be exponential. This finding is consistent with another report published from the CDC under National Center for Health Statistics (NCHS) (Rossen, Bastian, Warner, Khan, & Chong, 2017). To put this in perspective, this mortality from drug overdose now outnumbers the deaths from road traffic accidents and violence (WDR 2017, n.d.). Thus, the United States is experiencing a wide range of drug overdose problems over the past decades leading to increased economic damage, and this has become an emerging public health issue.

Definition of drug overdose

Drug overdose as defined by the literature is “when someone collapses, has blue skin, has convulsions, has difficulty breathing, loses consciousness, cannot be woken up, has a heart attack or dies while using drugs” (Bohnert, Tracy, & Galea, 2012, p. 3). This definition is also used by other literatures (McGregor, Darke, Ali, & Christie, 1998; Ochoa, Hahn, Seal, & Moss, 2001; Martins, Sampson, Cerdá, & Galea, 2015). This definition implies to any drug causing overdose whether available through prescriptions or illicit, and the majority of fatal overdoses involve opioids. However, prescription opioids are responsible for more overdose deaths (approximately 70%) than any illicit drug (Florence, Luo, Xu, & Zhou, 2016; U.S. DOJ, 2018). Whether prescription or illicit, most of these premature drug overdose deaths are preventable (WDR 2017, n.d.).

Economic burden of drug overdose

The economic burden of all drug overdoses, fatal and non-fatal in the U.S. is huge, and total estimates are yet to be explored. These burdens include costs of healthcare, lost productivity, treatment of addiction, criminal justice involvement, and so on (Abuse, 2019).
Among the recent studies, the largest estimate was published by Florence et al. (2016), who estimated the prescription opioid overdose, abuse and dependence in the U.S. in 2013. The authors found the total estimated cost to be over $78.5 billion for the year 2013. Among the aggregated cost, a little over one third (over $30 billion) was expensed for health care. Again, fatal overdoses accounted for a little over one quarter (around $21 billion) of the total cost (Florence, Luo, Xu, & Zhou, 2016). The authors also found that about 14 percent of the total cost is funded by public health insurance programs and more from other public sources.

**Geographic variations of drug overdose**

Geographic approaches to drug overdose death research have emerged in recent years. Studies demonstrate that drug overdose mortalities are not equally distributed across different geographic areas or population subgroups. It is important to consider geographic variation in order to inform effective prevention, diagnosis and treatment of the condition and to reduce the inequalities (Dwyer-Lindgren et al., 2018). Different studies focus on different aspects like some studies identify state level geographic variations of mortality related to opioid and heroin only (Ruhm, 2017), whereas other studies concentrate on mortalities from drugs, alcohol and even interpersonal violence altogether (Dwyer-Lindgren et al., 2018). However, the geographic patterns of drug overdose mortality rates involving all types of drugs (not including alcohol) focusing on smaller geographic scales like counties are yet to be explored. Detailed evaluation of patterns and associated factors may help understand the problem more and identify approaches that can be applied to prevent premature deaths. This study plans to examine any possible spatial pattern at smaller geographic scales (the county) and associated factors of drug overdose deaths with spatial analysis.
1.2 Study objectives

The objective of this study is to explore the spatial distribution of drug overdose death rates i.e., presence of any possible geospatial clustering at the county level in the contiguous United States over two 5-year time intervals (2007-2011 and 2012-2016), to identify the county level factors associated with this spatial patterning, and to evaluate the extent to which these spatial patterns are explained by county-level factors. To identify the geospatial clusters of drug overdose mortality rates, this study utilizes two methods: The Traditional method of spatial cluster analysis using raw/crude rates and the Empirical Bayes standardization method using smoothing to reduce variance instability caused by small-population areas.

1.3 Research questions

The proposed research questions for this study are as follows:

1. Are the county level drug overdose death rates in the United States spatially correlated among adjacent counties?

2. Do the geospatial patterns differ in the two time periods? If so, how much has the rate increased over time, and where geographically are the greatest increases?

3. Do any socio-economic or socio-demographic factor(s) have significant associations with the observed overdose death rates?
Chapter II

Review of the Literature

This study evaluates the spatial distribution of drug overdose death rates at the county level in the contiguous United States over two 5-year time intervals (2007-2011 and 2012-2016) and presence of any association with county-level socio-economic or socio-demographic factors. Considering the alarming rise in drug overdose deaths stated before, there are several contributing factors that have been found in the literature. For example: insurance coverage, poverty level, employment status, racial background in conjunction with poverty level and urbanization, or declining population in the county may be important predictors. The following literature review provides an overview of the drugs and types of drugs commonly involved in overdose mortality, and how different socioeconomic or sociodemographic factors linked to the epidemic have been described in published literature.

2.1 Drugs commonly involved in overdose

A recent issue of National Vital Statistics Report published from the U.S. centers for disease control and prevention (CDC) in December 2018 identified the drugs that were most commonly involved in drug overdose deaths in the United States during 2011 to 2016. According to this report, drugs that were most commonly involved in the drug overdose deaths fall into three different categories: 1) Opioids: fentanyl, heroin, hydrocodone, methadone, morphine and oxycodone; 2) Benzodiazepines: alprazolam and diazepam; and 3) Stimulants: cocaine and methamphetamine (Hedegaard, Bastian, Trinidad, Spencer, & Warner, 2018). Among the 10 most commonly involved drugs (fentanyl, heroin, hydrocodone, methadone, morphine, oxycodone, alprazolam, diazepam, cocaine, methamphetamine) identified by the report, oxycodone was the highest drug involved in the overdose deaths in 2011, heroin in 2012-
2015, and fentanyl in 2016. On the other hand, cocaine consistently ranked second or third (Hedegaard, Bastian, Trinidad, Spencer, & Warner, 2018). Among all drug overdose deaths, opioids account for approximately 68% of deaths (Scholl, Seth, Kariisa, Wilson, & Baldwin, 2018). This paper studies overdose deaths from all types of drugs and drug classes combined as the underlying cause of death, available from the CDC WONDER website. Details are described in the methods section.

2.2 Overdose related to prescription drugs and insurance coverage

The literature addressing the drug overdose epidemic was primarily focused on death by different types of illicit drugs. However, prescription drugs have proved to have a strong connection to the overdose deaths in the United States over the past several years, and opioids belong in the top three categories of prescribed drugs in the U.S. (Unity Behavioral Health, 2017). Report shows that there has been a significant increase in the use of prescription opioid analgesics among the U.S. adults since 1999 (Frenk, 2015). Prescription opioids are responsible for more than half of the fatal overdoses in the United States currently (U.S. DOJ, 2018). Prescription overdoses usually have an innocent origin as a prescription for a genuine condition. However, lack of information regarding the addictive nature of the drug makes it difficult for the individual to realize ahead of time that he or she may become addicted. Thus, there is a medical component in the causality of this epidemic (Smith, 2017). More use of prescription drugs leads to more morbidity related to overdose. Studies found associations between increased opioid prescriptions and emergency department visits due to opioid overdose (Dasgupta et al., 2006; Wisniewski, Purdy, & Blondell, 2008). This morbidity or ED visits are related to the potency of the drugs i.e., higher for high potency opioids and lower for low potency opioids (Dasgupta et al., 2006). Similar findings demonstrated by another study are that increased numbers of
prescriptions lead to higher sales which in turn is correlated to related overdose deaths (Modarai et al., 2013).

While considering the potential influences of different stakeholders on this epidemic, several have been identified: pharmaceutical companies, distributors, prescriber physicians, health insurers and pharmacies. Drug makers and pharmaceutical companies are given more attention in this regard. There is evidence that marketing of opioid products to the physicians by the pharmaceutical companies is associated with higher rates of seeking treatment for addiction and higher rates of mortality from the prescription opioid overdoses (Smith, 2017; Hadland, Rivera-Aguirre, Marshall, & Cerdá, 2019).

Less attention has been given to the health insurers and pharmacy benefit managers. There seems to be a nexus between the insurers and the pattern of opioid prescribing and related overdose due to addiction. Studies done in different states found a substantial growth in the use of prescription opioid drugs among the population covered by insurance, specifically public insurance. For example, a study in North Carolina showed higher rate of death among the Medicaid beneficiaries from prescription opioid overdose than the general population (Whitmire & Adams, 2010). Another study in Washington state demonstrated that almost half of the opioid overdose deaths occurred among Medicaid enrollees even though a very small percent of the State’s total population were enrolled in Medicaid (CDC, 2009). Changes in the payment patterns for these drugs has also been noticed. According to one study, the financing pattern for opioid pain relievers has shifted substantially from consumers’ out-of-pocket to insurers during the 1999-2012 period. This study also found little change in total expense on opioid drugs, while there was a huge increase in the number of drugs prescribed, suggesting a shift towards the less expensive drugs (Zhou, Florence, & Dowell, 2016).
2.3 Overdose related to socioeconomic and socio-demographic determinants

Although blame for the overdose mortality mostly goes to the supply side (those who make the drugs available), the demand side of this crisis, which is related to people using the drugs, has not been adequately explored. Several studies focused on the overdose related mortality with different socio-economic and socio-demographic determinants. Research conducted in a few U.S. states and among Medicare enrollees showed that areas with higher poverty level have higher concentrations of opioid prescription rates and related mortality, or opioid and heroin overdose-related hospital discharges (Grigoras et al., 2018; Pear et al., 2019). Studies also found that rural areas have higher opioid prescription rates along with overdose rates (Paulozzi & Xi, 2008; Grigoras et al., 2018). As these areas have a greater proportion of Non-Hispanic white population, the mortalities are concentrated among them (Rudd, 2016; Grigoras et al., 2018). A study conducted on racial disparities for the opioid epidemic found an increasing rate of mortality among the Non-Hispanic white population since 1979, while the rate remained stable for the Non-Hispanic black population until 2010. After 2010, this growth was rapid for the both populations (Alexander, Kiang, & Barbieri, 2018). However, this study only looked at Black and White populations in regards of racial disparity on mortality. Frenk (2015), on the other hand, focused on ethnic differences in opioid analgesics use from 2007 to 2012 and showed that the Non-Hispanic population (both white and black) are more likely to use opioid analgesics than the Hispanic population.

Economic factors are other determinants that might have an association with drug overdose mortality. Studies done in two different areas (Luxembourg and New York City) found that unemployment, unstable income, or unequal income distribution are likely to have an impact on fatal overdoses (Galea et al., 2003; Origer, Le Bihan, & Baumann, 2014). A study on opioid poisoning-related hospital discharge done in California found a positive association between
opioid poisoning and lower household income (Cerdá et al., 2017). Brown & Wehby (2017) examined state level drug overdose deaths and associated economic conditions. They found that economic downturns may increase opioid related deaths. A few studies also showed how different economic measures, alone or in groups, are associated with illicit drug use and related mortality. For example, Carpenter, McClellan, & Rees (2017) showed that higher state level unemployment was related to increased use of prescription pain medication and other substance use disorders, because stress of losing a job was associated with higher use of these drugs. This finding is consistent with a qualitative study done in a now deindustrialized area of Pennsylvania which once was a global center of steel production. This study demonstrated how frustration, lack of opportunity, and social isolation due to losing jobs led to the local overdose crisis (McLean, 2016).

**2.4 Geographic variations of overdose mortality**

Few studies have explored overdose mortality by specific geographic regions and time period. A study done in California state demonstrated opioid poisoning to be concentrated in rural areas and suggested a spatial spread (termed as ‘spatial contagion’) from rural and suburban to urban areas. This study also identified a spatial association between income and opioid related hospital discharges (Cerdá et al., 2017). Another study on the geospatial distribution patterns of death from heroin at the county level in the U.S. showed a shift from the West Coast in 2000 to New England and Mid-Atlantic regions, the Great Lakes and the Central Ohio Valley by 2014 (Stewart, Cao, Hsu, Artigiani, & Wish, 2017), which indicates a change in the pattern or evolution over space and time. Another county level spatial study examined all drug poisoning deaths, but the time period was limited to 2007-2009 (Rossen, Khan, & Warner, 2014). This study identified several hotspot and cold-spot clusters for drug poisoning deaths across the U.S.
using a K-nearest neighbor approach. The main hotspots were located in the North Pacific coast, the Southwest, Appalachia, and along the Gulf coast. The main cold-spots were located across the Central Plains and Texas (Rossen, Khan, & Warner, 2014). These studies give evidence regarding those locations where the burdens are higher. Studying this problem over a longer period and comparing it over both time and space will further elucidate the geography of drug overdose deaths and help address the growing burden.
Chapter III

Methods and Procedures

3.1 Data sources

This study focuses on the county-level drug overdose mortality data in the contiguous U.S and the relationship with county-level contextual factors, and comparison over two time intervals using spatial analysis. A good quality spatial data is one which is based on population data or spatially representative data, and not just a random sample from the population which uses the non-spatial sampling method (Mobley, 2013). Good quality spatial data or geographically referenced data is required to conduct spatial analysis for population health research, through which we can evaluate the spatial trends present among the areas in the context of the question asked. Therefore, it helps to identify the areas with greatest concerns or populations at risk. This information in turn is useful for applying public health policy in the areas of concern or where the intervention is needed most.

Data on county level drug overdose mortality in the contiguous U.S. are obtained from CDC WONDER (Wide-ranging Online Data for Epidemiologic Research), an online database developed by Centers for Disease Control and Prevention (CDC) released December 2017. This database includes all mortality data by underlying cause of death. Thus, the data are available in the file “Underlying Cause of Death 1999-2016”. The data file includes national mortality data i.e., raw count of drug overdose deaths per county per year, total population size per county, year code, underlying cause of death (UCD) code, county and state names, and geographic identifiers (county FIPS codes). The mortality data are based on the single underlying cause of death available in the death certificates for the U.S. residents (CDC, n.d.). The underlying cause of death is determined by International classification of disease, 10th revision (ICD-10), defined by
World Health Organization. Based on the ICD-10 classification, drug overdose deaths data over two-time intervals, 2007-2011 and 2012-2016, are downloaded from the CDC website for this study. 5-years period aggregated mortality data per county are collected to maximize the number of the counties with reported mortality rate. As of May 23, 2011, data representing less than 10 counts of death or population per county are suppressed due to CDC privacy policy (CDC, n.d.). Therefore, counties with less than 10 counts of death exhibited blank values which are replaced with zero for this study. The population estimates are taken from the U.S. Census Bureau based on census counts.

The county level contextual factors for predictor variables are compiled from various sources- Small area health insurance estimates (SAHIE) and Small area income and poverty estimates (SAIPE) program of U.S. Census Bureau, Bureau of Labor Statistics, GeoDa center calculated data, and USDA ERS (United States Department of Agriculture, Economic Research service). SAHIE and SAIPE programs are purposed to develop model-based estimates of county and state level health insurance coverage, poverty and income statistics (Bureau, n.d.). SAHIE provides state and county level data on number and percentages of insured and uninsured people by age, sex, race/ethnicity and income level per year. SAHIE data are usually used to assess geographic variations and changes over time in the health insurance coverage in the United States (Bureau, n.d.). County level percent of uninsured people (total population < age 65 years) data are obtained for the two-year periods (2005 and 2014). SAIPE provides annual estimates of income and poverty at school-district, county and state level. The main purpose is to locate poverty-stricken/destitute areas to help allocating the federal funds and local programs (Bureau, n.d.). Data on county level estimated percent of total people in poverty are collected for the income years of 2005 and 2014.
Data on unemployment are obtained from the Local Area Unemployment Statistics (LAUS) program of Bureau of Labor Statistics. This is a federal-state cooperative program providing data related to employment and unemployment for different specified areas (LAUS, n.d.). Annual averages of county level unemployment rate data are available in this database and are obtained for two different years for this study (2005 and 2015). Economic Research Service (ERS) of USDA (United States Department of Agriculture) is complementing USDA’s research mission by broadening the fields of research on various fields like economy and policy relevant themes to highlight economic and social characteristics. Population loss typology code is located under policy relevant types of county typology codes (USDA, n.d.). Population loss typology code data are collected for the two-year periods (2004 and 2015). Data on Poverty rate by race or ethnicity and urbanization are obtained from GeoDa Center. Poverty rate is measured as percent of total population in poverty for each of the three races (Black, White and Hispanic) and for each area (urban and rural) for two different periods (2000 and 2005-2009) by GeoDa center. For earlier period, this rate is calculated from census 2000 tract level data and for later period, this rate is calculated from American Community Survey’s (ACS) aggregated tract level data for 2005-2009 (GeoDa, n.d.).

### 3.2 Conceptual model/ framework for drug overdose mortality

The conceptual model for this research is adapted based on the published literatures on drug overdose deaths in the United States. The variables included in this model are: drug overdose death rates (2007-2011 and 2012-2016), uninsured rate (2005 and 2014), unemployment rate (2005 and 2015), overall poverty rate (2005 and 2014), poverty rate by race/ethnicity and urbanization (2000, 2005-2009 aggregated), and population loss index from the county (2004 and 2015). All these variables are county level variables. The main objective is to
explore the spatial distribution of drug overdose deaths to identify any spatial autocorrelation present among the adjacent counties. In addition, to have a better picture on the distribution of the epidemic of drug overdose deaths across the space and time, the outcome is compared over time based on the predictors. Therefore, all the variables collected are over two time periods as stated above.

The outcome in this model is county level drug overdose death rate. Most of the literatures described the drug overdose death data based on International Classification of Disease (ICD-10), 10\textsuperscript{th} revision of underlying cause of death. According to this death classes, the death codes that fall under drug overdose deaths class are: unintentional drug poisoning/overdose (X40- X44), suicide drug poisoning/overdose (X60- X64), homicide drug poisoning/overdose (X85), and undetermined drug poisoning/overdose (Y10- Y14) (CDC, n.d.). Among these classes, unintentional drug overdose deaths hold the highest rank recently with about 80%-86% of drug overdose deaths followed by suicide (8%-13%) (Hedegaard, 2017). However, to have a detailed idea on the overall picture, all the codes are combined to get all types of drug overdose deaths together for this study. The time period (2007-2011 and 2012-2016) for the outcome variable is chosen because this was the most recent available data in the CDC WONDER website when collected. 5 years interval is enough to over-ride privacy concerns and allow us to obtain the data. Moreover, Affordable Care Act is assumed to play a major role in many aspects of the insurance coverage and associated factors related to prescription medications based on published literature (Smith, 2017).

The exploration of spatial distribution of drug overdose deaths are enriched with additional examination of factors that are assumed to be associated with the clustering. So, data for the predictor variables are also obtained for two time intervals. Percent of total population
uninsured in the county as uninsured rate is included to determine the role insurance coverage in the drug overdose mortality (CDC, 2009; Whitmire & Adams, 2010; Modarai et al., 2013; Burns, 2017). From the perspective of socio-economic and socio-demographic determinants, this study uses unemployment rate, overall poverty rate, poverty rate by race/ethnicity and urbanization variable. Unemployment rate and poverty rate are determinants of economic conditions playing role in the associated factors of drug overdose (Brown & Wehby, 2017; Carpenter, McClellan, & Rees, 2017). However, literature showed mixed results on these determinants; thus, included in the conceptual model for more elaboration. Poverty rate by race/ethnicity and urbanization variables are included to demonstrate the racial disparities based on urbanization (rural versus urban living). Literature showed the role of racial disparities along with living area on drug overdose and related mortality (Paulozzi & Xi, 2008; Alexander, Kiang, & Barbieri, 2018; Grigoras et al., 2018). Population loss index is also included in the conceptual model assuming that social isolation and marginalization might have influence on the drug addiction or overdose (McLean, 2016).

However, the associations with different contextual factors examined in many different published studies ignored any spatial perspective, which is invariably present in socio-economic problems (Koschinsky, n.d.). Study of these factors should consider potential spatial dependence or autocorrelation across the areas, as these are usually shared among people across county boundaries causing spatial correlation. Ignoring spatial autocorrelation may create biased or inconsistent estimates, by violating the assumption for ordinary least square (OLS) regression that the observations under analysis are statistically independent (Mobley et al., 2011; Mobley, n.d.). Thus, examining the factors in consideration of potential spatial autocorrelation among adjacent counties (smaller areal level) and a larger geographic scope including the entire U.S.,
with comparisons over time can provide us with valuable information to help control this epidemic. Description of all the variables with type, metrics, sources and years are summarized in Table 1.

3.3 Data processing and Measures

All the data are downloaded and merged into the Microsoft Excel workbook. The outcome variable is county level drug overdose death rates in the contiguous U.S. over two time intervals, 2007-2011 and 2012-2016, which is continuous. For the outcome variable, 5-years period aggregated death counts per county are downloaded to maximize the number of the counties with reported mortality. Counties with suppressed data that exhibited blank values are replaced with zero. Total population count per county is downloaded for all counties. Crude Mortality rates are calculated from total death count for a county aggregated over the 5-year period and size of the population residing in that county in the MS excel using the formula below:

**Drug overdose death rate**

\[
\text{Drug overdose death rate} = \frac{\text{Number of total drug-related deaths per county}}{\text{population size of that county}} \times 100,000
\]

Drug overdose death rates are calculated for two different periods, 2007-2011 and 2012-2016, and a single rate is obtained for every county in each group. Due to the suppressed data, this file has 1914 counties with positive death rates in 2007-2011, and 2032 counties with positive death rates in 2012-2016. Raw counts for drug overdose death and population counts per county are kept in the data file to be used in the Empirical Bayes standardization method of clustering.

The independent variables are un-insured rate, unemployment rate, overall poverty rate, poverty rate by race or ethnicity and urbanization, and a population loss index. All are county
level variables and in two time intervals. Un-insured rate is a continuous variable. It is measured as percentage of total population under 65 years of age without health insurance and selected for the year 2005 and 2014 from SAHIE program of US Census Bureau. Overall poverty rate is another continuous variable, which is measured as percent of total population in poverty for the income year and selected for the year 2005 and 2014. Annual averages of county level unemployment rate data for the year 2005 and 2015 are downloaded, which is also a continuous variable. The population loss index is coded as dichotomous variable. This variable indicates whether the county has lost population over recent years, where 1= yes and 0=no. The data are selected for the year 2004 and 2015. Poverty rate by race or ethnicity and urbanization is measured as percent of population in poverty for each of the three races (Black, White and Hispanic) and for each area (urban and rural). These variables were calculated by the GeoDa center. Since the data were already aggregated by the GeoDa center, thus are directly collected for this study without further calculation. These are also continuous variables.

The merged data file from MS excel is joined with county geographic shapefile from the U.S. Census Bureau using unique county FIPS codes in the QGIS software. FIPS or the Federal Information Processing Standards are unique numeric codes set for all the counties in the U.S. as identification code (ESRI, n.d.). These FIPS codes are included in the data file as a unique identifier and to be used as geocoding tools to link the data to map. This process in turn helps creating maps in the QGIS and performing regression analysis in the GeoDa software. Since spatial continuity is the basis for spatial regression analysis, Alaska, Hawaii and other spatial islands are excluded from the dataset. Therefore, the final dataset contains 3106 U.S counties.
3.4 Statistical analysis

After data cleaning and measures creation, descriptive analysis of the variables used in the model is performed in the SAS software, version 9.4 (SAS Institute, Cary NC), for both early and late period. A summary of the descriptive statistics is reported in the Table 2.

QGIS is an open-source Geographic Information System (GIS) software which is used to create, edit, analyze, visualize and publish geospatial information (QGIS, n.d.). In addition to the descriptive analysis, QGIS software 3.2 is used to examine and create translational maps to visually demonstrate the trends over time in the two-periods drug overdose death rates (outcome variable) by county (QGIS, 2018). Graduated classification (quantiles as breaks) and sequential color schemes are used for these mapping. The same cut points are used in two maps for both periods to compare geospatial patterns where rates have increased or decreased. Both maps are shown in Figure 1. Mapping is used also to translate the findings of local spatial autocorrelation (LISA) among the adjacent counties for both the traditional method and Empirical Bayes standardization method for drug overdose mortality.

GeoDa is a spatial analytic software, also free and open-source tool, which allows insight in “spatial data analysis by exploring and modeling spatial patterns” (GeoDa, n.d.). Exploratory spatial data analysis to explore presence of any spatial autocorrelation (both global and local) is done in GeoDa (Anselin, Syabri, & Kho, 2006). Spatial autocorrelation indicates how features among neighbors of close proximity are similar to each other as geographic features usually tend to be (ArcGIS, n.d.). Other data processing like building of stacked data set on time for performing time stacked regression analysis is also done in Geoda software. Confirmatory spatial data analysis i.e. model building for spatial regression is performed in GeoDa Space (Anselin & Rey, 2014).
Traditional method and Empirical Bayes standardization method of LISA clustering

Working with variables that are rates or proportions might experience a potential problem in distribution like identification of false outliers while mapping. This occurs due to variance instability which is an intrinsic quality of rates. A reason behind this is varying population densities across the cases (GeoDa, n.d.). Empirical Bayes standardization method is suggested to get rid of spurious outliers and to get robust measures for spatial autocorrelation by using the concept from ‘Bayesian shrinkage estimator’. This method accounts for variance instability caused by small population size in the denominator of the rate variable by transforming the crude rate to have a mean of zero and unit variance (GeoDa, n.d.). Since the outcome variable in this study is rate variable (drug overdose death rate) and population densities across the U.S. counties are not similar, Empirical Bayes standardization approach could alleviate small area estimation problem while analyzing spatial patterns. In this study, both traditional method and Empirical Bayes standardization method are used to identify spatial autocorrelation among the adjacent U.S. counties for drug overdose death rates. Translational maps are created based on the findings from two different methods. The traditional method uses the formula described in the previous section to calculate the raw rate. Calculated raw rate is then used in the LISA clustering analysis. For Empirical Bayes standardization estimation of local clusters, raw counts of drug overdose death in the county are used as the ‘event variable’ and population size in the county is used as ‘base variable’ in the GeoDa software. “LISA for EB rate” option in GeoDa software is selected and the software performs the clustering analysis by transforming the rate variable behind the scene.
Spatial Regression Analysis

Spatial relationship can be examined, explored and modeled by spatial regression analysis. It also helps to explain factors behind the observed spatial patterns (ArcGIS, n.d.). As described in the previous section, unlike ordinary least square (OLS) regression, both the dependent and independent variables under analysis may not be independent across the observations, necessitating use of spatial regression analysis. This spatial autocorrelation can be the result of two spatial stochastic processes- 1) ‘substantive dependence’ where dependence is present in the dependent variable and lag model is required for correct specification; and 2) ‘nuisance dependence’ where dependence is present in the error term, and the error model is required for correct specification (Mobley, n.d.). In the first case, ignoring the spatial autocorrelation among the adjacent counties in the outcome variable can lead to biased effect estimates by overestimating the effect of predictors, which can be corrected using spatial lag regression analysis (Mobley et al., 2011). In the second case, not accounting properly for the spatial error process in the regression can lead to biased standard errors and misleading statistical inference. Therefore, as a first step, the presence of any spatial autocorrelation in the dependent variable (drug overdose death rate) is explored by global and local indices of spatial autocorrelation (Moran’s I and LISA) in GeoDa software. For this study, both traditional method and Empirical Bayes standardization method are used to determine the presence of spatial autocorrelation in the dependent variable. First, to identify the neighbors of a county spatial weights are defined by the queen contiguity where counties that share a common boundary (a common edge or vertex) are considered as neighbors (GeoDa, n.d.). Univariate global Moran’s I provides the information on spatial randomness. The null hypothesis is the death rates are spatially random among the counties. Rejection of the null indicates that spatial autocorrelation...
is present among the adjacent counties. From the analysis in this study, univariate global Moran’s I statistics suggest presence of spatial autocorrelation in the outcome in both early and later period for both traditional and Empirical Bayes standardization methods. Moran’s I statistics for both methods are shown in Table 3.

While univariate global Moran’s I suggests whether there is existence of any spatial clusters or not, it does not specify the exact location of the clusters. Univariate LISA (Local Indicator of Spatial Association) is used to detect the presence of local clustering or association in space. Here, LISA first finds the actual correlation between a county’s value and average of the neighboring counties’ values (also known as spatial lag) for the outcome variable. Then, LISA performs a simulation analysis for correlation of the same variable between a set of randomly selected counties as neighbors other than real spatial neighbors and the same county in question, which is done up to 999 times. This correlation is compared with the actual correlation to see the significance i.e., whether the actual correlation falls in far in the tail of distribution (Anselin, 1995). LISA statistics show four types of geographical clustering- hotspots or high-high (significant spatially autocorrelated clusters of high values in the outcome variable between a county and the neighbors); cool spots or low-low (significant spatially autocorrelated clusters of low values in the outcome variable between a county and the neighbors); low-high (significant spatially autocorrelated clusters where the county has low value but the neighbors have high values in the outcome variable); and high-low (significant spatially autocorrelated clusters where the county has high value but the neighbors have low values in the outcome variable). Figure 2 shows LISA clusters for the traditional method in both periods, and Figure 3 shows LISA clusters for the Empirical Bayes standardization method, also in both periods. Only the statistically significant associations are shown as colors in the map.
To determine whether the observed regression parameter estimates change significantly over time, the observations in the data set are stacked based on time for all the independent and dependent variables where time 0 indicates the earlier period and time 1 indicates late period. In addition, a stacked queen contiguity weight file is created based on time. Spatial regression analysis is conducted using GeoDa Space software on the stacked data file building on two time periods where time is used as a regime. First, the model is estimated using ordinary least square (OLS) using the time as a regime, and specifying separate variance-covariance matrix estimates in each time period, where two separate equations are shown for two periods. The residuals from OLS regression are used as the diagnostics for spatial regression, and a series of Lagrange multiplier tests (diagnostic tests for spatial dependence) are used to determine the correct specification model for spatial autocorrelation. The coefficients/parameter estimates and p values from the OLS regression along with the regression diagnostics are reported in the Table 4.

Second, based on the Lagrange multiplier tests, a two stage least squares regression (lag model) is used for improved model specification. Two stage least square regression analysis allows separate parameter estimates in two different time period along with the spatial lag term (Rho) as a covariate (Anselin & Rey, 2014, p 159). For the heteroskedasticity issues in the OLS model, white correction is used in the lag model to get correct standard errors. Also, to address the observed lack of normality problem, method of moments (MOM) estimation is used instead of maximum likelihood estimation (MLE). Chow tests provide the information on whether any significant change is present over time, both a global change and individual changes in the coefficients associated with specific variables. The null hypothesis is that the coefficient estimates did not change over time. A rejection of the null hypothesis indicates that a significant change over time has occurred. All the diagnostic tests for spatial dependence (Lagrange
Multiplier tests and Anselin-Kelejian test) are shown in Table 5. Coefficients, p values and Chow tests results from the lag model are shown in Table 6.

Lastly, based on the significant Anselin-Kelejian diagnostic test (shown in Table 5) on the residuals from the spatial lag model, a combined error and lag model / Durbin model is estimated (Anselin & Rey, 2014, p 259). Anselin-Kelejian test is a diagnostic test for spatial dependence in the lag model, which provides information on any remaining spatial autocorrelation after the lag model estimation (Anselin & Rey, 2014, p 300). The null hypothesis is that there is no remaining spatial autocorrelation in the model. A rejection of null hypothesis indicates that there is still spatial autocorrelation present, which can be corrected by a Durbin/combo model estimation. This model incorporates both spatial lag coefficient (Rho), the coefficient for lagged dependent variable and error coefficient (Lamda), the coefficient for lagged independent variables in the predictor matrix. Method of moments (MOM) estimation is used here too, which is the default option for combo model. Also, the KP-HET option is used in this model for the heteroskedasticity issue. Coefficients and p values along with Chow test results are reported in the Table 7.
Chapter IV

Results

4.1 Descriptive analysis

A total of 3,106 counties were included in the dataset for the analysis. Alaska, Hawaii and other islands, which are considered as spatial islands, are excluded from the analysis. Variable descriptions are presented in the Table 1, and descriptive statistics are presented in the Table 2 with mean and standard deviation. The mean drug overdose death rate of all counties included in the analysis increased from 9.42 ± 9.92 to 11.91 ± 11.58 (per 100,000) from earlier to later period. McDowell, West Virginia (death rate 88.9), Wyoming, West Virginia (death rate 84.77) and Floyd, Kentucky (death rate 68.5) are the counties with highest drug overdose death rates in the earlier period, while in the later period, the counties with highest rates are again McDowell, West Virginia (death rate 85.61), Rio Arriba, New Mexico (83.64) and Wyoming, West Virginia (83.33). Comparison of rates in these two periods in the Figure 1 shows that the drug overdose rates increased over time, and became more geographically dispersed in the later period.

4.2 Tests of spatial autocorrelation

The Global Moran’s I statistic is statistically significant at \( p < 0.05 \) for drug overdose death rate for the both time periods. The statistic presented in Table 3 indicates the presence of substantial spatial autocorrelation somewhere across the map. The Global Moran’s I for Empirical Bayes standardized spatial clustering method demonstrates the same.

LISA cluster maps shown in Figure 2 (for traditional method) and Figure 3 (for Empirical Bayes standardization method) indicate the presence of statistically significant positive clusters of high rates (hotspots) in red color and low rates (cool-spots) in blue color in drug overdose
death rates during both periods. For the early period, clusters of hot-spots are mostly located in the coastal West region (part of California, Oregon and Washington), part of Southwest region (part of Arizona and Oklahoma, big part of New Mexico), most of the upper and lower part of Southeast region (Tennessee, Kentucky, West Virginia, Florida, part of Georgia, South Carolina and North Carolina), a little part of Northeast region (Pennsylvania), and part of Midwest region (Ohio and Indiana). Cool spots are mostly located in the Midwest and Southwest regions. Hot-spot distribution for Empirical Bayes standardization method in the earlier period demonstrates almost similar result. However, the distribution of cool-spots for Empirical Bayes standardization method is sparser than the traditional method along the Midwest region. For the later period, a shift of hotspot clusters is notable from Southeast to most of the Northeast region. While in the Southeast region there are no significant hotspots for Georgia and Florida for the later period; Kentucky, Tennessee, North Carolina and West Virginia still host significant hotspots. In addition, Oklahoma has more concentrated hotspots than before, and Colorado becomes the new location for hotspots. Cool spots become more significant along the middle part/ central plain area for the later period. LISA maps for Empirical Bayes standardization method for the later period show the similar distribution except for sparser cool-spots distribution along the middle part.

4.3 Spatial regression results

Table 4 shows the results from the first model, Ordinary Least Squares (OLS) estimation using time as a regime for two periods, and here all the predictors were used in both periods. Parameter estimates that are statistically significant at 5% level of significance are shown in bold font. Superscript C indicates the variable coefficients that significantly changed over time based on the Chow test. Global Chow test indicates an overall statistically significant change in effect
estimate over time. Regression diagnostics are also reported in this table. The diagnostics show that multicollinearity condition number is within the normative limits. However, the tests show lack of normality of errors and presence of heteroskedasticity in the model.

Table 5 shows all the diagnostic tests for spatial dependence. The steps for searching correct spatial specification from the diagnostic tests are described here from Anselin & Rey, 2014, p 110. If both the lag and error Lagrange multiplier (LM) tests are not significant, then robust tests are ignored and OLS results are kept without further changes. If both are significant, then the robust tests are examined and the one with a more significant test statistic is assumed the better model specification. Here in the Table 5, both lag and error LM tests are significant for both periods. Therefore, robust tests are examined. Robust LM test for lag is significant, while robust test for error is not. Hence, the lag model is the choice of model specification here.

Table 6 shows the results of two stage least squares estimation (lag model). Results show a large positive lag parameter estimate (0.797 and 0.774 respectively for earlier and later period) which is significant in both periods with a p value of 0.000. This indicates a spillover effect across the county boundaries which means that neighboring counties have impact on the outcomes (death rates) of their neighbors. OLS model estimates are larger (biased upwards by spatial multiplier bias) than lag model estimates and the lag model is needed to address the spatial dependence issues. In Table 6, significant changes of effect estimates over time are indicated with the superscript C based on the Chow test results. The global Chow test indicates that there is a strong significant overall change in effect estimates over time (p-value 0.000) along with a few of the coefficients that show significant changes (unemployment rate, uninsured rate, poverty rate in Black rural and poverty rate in Hispanic rural).
The Anselin-Kelejian test in Table 5 indicates that, after spatial lag modeling there are still spatial autocorrelation effect present in the model. This additional spatial effect is from correlation of spatial error term (Anselin & Rey, 2014, p 259). The null hypothesis for this test is, no spatial effects are present. Rejection of this test indicates that a combined/ Durbin model is the best specification for the final model, with results shown in Table 7. Method of moments (MOM) estimation is the default option for this model. The KP-HET option is used because of strong indication of heteroskedasticity in the OLS model (Anselin & Rey, 2014, p 268).

In Table 7, this Durbin model shows a significant negative error parameter estimate (Lamda), which is -0.581 and -0.524 respectively for earlier and later period with a p value of 0.000. The lag parameter (Rho) remains the same as it was for the lag model in the both periods with high positive values along with statistically significant small p values. Among the parameter estimates- overall poverty rate (p value for earlier period = 0.011, p value for later period = 0.021), unemployment rate (p value for later period = 0.00), un-insured rate (0.00), poverty rate for Black rural (0.005, 0.000), poverty rate for Hispanic rural (0.000), poverty rate for White rural (0.001), poverty rate for White urban (0.000, 0.000) and population loss index (0.000, 0.004) remained unchanged according to their significance across these two models (Lag and Durbin model). However, coefficients and significance for other variables have been changed. For example, poverty rate for Black urban becomes significant for both periods (0.002, 0.000) in the Durbin model instead of only later period in the lag model. Poverty rate for Hispanic urban becomes non-significant for both periods in the Durbin model, while it was significant for only earlier period in the lag model.

Overall poverty rate is significant for both periods (p value 0.011 and 0.021) with positive coefficient indicating that higher drug overdose deaths are associated with increased
poverty. Unemployment rate has positive coefficients in the both periods, but is significant in the later period only (0.00) indicating that a higher unemployment rate is associated with increased deaths from drug overdose. The parameter estimates for the uninsured rate are negative in both periods, but significant in the earlier period (0.00). This indicates that the uninsured rate has a negative association with drug overdose rates meaning that more insurance coverage is associated with higher drug overdose deaths. The poverty rates for Black or White populations in urban areas have positive significant coefficients for both periods indicating that higher poverty in urban areas are associated with increased drug overdose deaths in these two groups of populations. Coefficients for the poverty rates for Hispanic populations in urban areas are not significant. For rural areas, the poverty rate for White populations also has positive coefficient, but significant only in the earlier period (0.001). However, for Black or Hispanic populations in the rural areas have negative association with drug overdose death rates, which is significant for both periods in Black (0.005, 0.000) and for the later period only in Hispanic (0.00) populations. This negative association indicates that increased poverty in rural areas is associated with less drug overdose deaths among these two groups of populations. Population loss index has negative coefficients and significant for both periods (0.00, 0.004). This indicates that counties that experience more population loss have less drug overdose deaths.

In Table 7, significant change of effect estimates over time are indicated with the superscript $^C$ based on the Chow test results. Unemployment rate, uninsured rate, poverty rate in Black rural, poverty rate Hispanic rural and poverty rate in White rural are the variables for which the parameter estimates changed significantly over time. The global Chow test indicates that there is a strong statistically significant overall change in effect estimates over time (p-value 0.000).
Chapter V
Discussion and Conclusion

5.1 Spatial clusters for drug overdose deaths

The United States is facing an epidemic of drug overdoses (Jalal et al., 2018). Using CDC WONDER database of mortality data, this paper focuses on drug overdose deaths, its spatial distribution across the United States at the county level and over time, and its associations with county level contextual factors. The study demonstrates that there was an increase in the drug overdose death rates over the two time periods. However, this increase was not consistently distributed across geography. For example, some counties had high rates in the earlier period, whereas for the later period these were very low. Thus, the distribution varied across the counties and dispersed over time. Death rates were mostly concentrated in the part of coastal areas of West region, part of the Southeast and Northeast regions in the earlier periods, whereas these spread to cover most of the West, Southwest, and Northeast regions and through the Midwest region in the later period.

This study also reveals that there is significant spatial autocorrelation present among the adjacent counties in the drug overdose death rates which answers the first research question. Both traditional method and Empirical Bayes standardization method of drug overdose deaths are used to test for spatial association. Few previous studies mentioned using the Empirical Bayes method for heroin or other drug poisoning death rates due to it being an infrequent event or having small population size to smooth the data (Rossen, Khan, & Warner, 2014; Stewart, Cao, Hsu, Artigiani, & Wish, 2017). In this study, both methods show similar patterns of spatial clustering across the map. However, this spatial cluster pattern differs in two time periods. For example, statistically significant positive clusters of high rates (hotspots) are in the coastal part
of West region, part of Southwest and Northeast regions, and most of Southeast region in the earlier period for both methods. On the other hand, in the later period, these hotspots spread to cover most of West, Southwest and Northeast regions, part of the Midwest region, and lesser part of Southeast region than before, with the Northeast region experiencing the greatest increase, again demonstrated by both methods.

Furthermore, there is a little difference in the significant clusters for low rates (cool-spots) between these two methods. Cool-spots for the Empirical Bayes standardization method are a little sparser than the traditional method. While the primary idea of using the Empirical Bayes standardization method is to get rid of spurious outliers and alleviate small area problems due to varying population densities when considering calculation of rate or proportion variable (GeoDa, n.d.), one could expect a broader pattern for Empirical Bayes standardization method if there was a problem with small denominators. However, this does not seem to be the case in this study as maps from both methods look similar. One reason behind this can be that the data set used in this study does not have small population problem, that means data is not disperse enough to get a significant difference. The Empirical Bayes standardization method can correct for the rates where data are too dispersed. Since this study used 5-years aggregated data from the database to get values for more counties, this could reduce this problem by having less disperse data.

5.2 Spatial regression

Presence of spatial autocorrelation among the adjacent counties in the drug overdose death rates is taken into consideration with the regression modeling. After several misspecification tests are run, a correct model is specified to describe the difference in the outcome of drug overdose deaths as predicted through some socioeconomic and
sociodemographic predictors. The multicollinearity number is within normative limits. However, due to a skewed outcome variable (shown in Figure 4) there are issues with non-normality and heteroskedasticity, which are handled with method of moments (MOM) estimation, and white and KP-HET correction tests for standard errors.

The results suggest that the effect of different contextual factors are heterogenous over time and across different population. The overall poverty rate at the county level is found to have a significant positive association with drug overdose death rates in both periods; that means counties with higher poverty rates have higher drug overdose death rates. This finding is consistent with prior studies which found that high poverty in the county have impact on the positive relationship between higher prescription rates and opioid related deaths (Cerdá et al., 2017; Grigoras et al., 2018).

The results also suggest that the uninsured rate has a negative association with drug overdose deaths. That means counties with greater insurance coverage have higher drug overdose death rates, which may be due to the association of drug overdose with increased prescription pain medication. The study by Grigoras et al. (2018) focused on the prescription behavior of physicians where they found positive associations between higher rates of prescription containing opioids and opioid related deaths. It is noteworthy that pharmaceutical opioids were responsible for 37% of the total 44,000 drug overdose deaths in 2013 (Grigoras et al., 2018). This implies that legal availability of drugs through insurance policies can be a driving factor for drug overdose death. Therefore, more insurance coverage can lead to more exposure to prescription pain medications which might be a source of drug overdose.

In addition, insurance coverage difference resulting from Affordable Care Act (ACA) between these two time intervals can highlight on mechanism behind it. This is commensurate
with studies that demonstrated that Affordable Care Act specifically through Medicaid expansion offers uninsured and vulnerable populations better access to health care, more prescription medications, and less out-of-pocket spending (Mulcahy, Eibner, & Finegold, 2016; Mahendraratnam, Dusetzina, & Farley, 2017). This policy is specifically for people with chronic conditions who need more pain medications. Again, most insurers have a policy of giving easy access or coverage for generic opioid medication as pain reliever which are cheaper but potentially addictive. On the other hand, safer alternatives with less-risky and less-addictive opioids have limited access and are expensive (Analysis, 2017). Thus, it seems that sometimes insurance coverage cannot help in controlling the epidemic situation due to the restricted policies of the insurers. All these mechanisms together are responsible, suggesting that effective implementation of insurance policy can be the primary focus of intervention and policy.

The unemployment rate is found to be significantly positively associated with the drug overdose deaths in the later period indicating that an increase in county level unemployment rate is associated with increased death from drug overdose that is consistent with the findings of Carpenter, McClellan, & Rees (2017) and Brown & Wehby (2017). Losing a job creates frustration, stress, social isolation, and even depression which in turn leads to substance abuse or increased use of pain medication (McLean, 2016).

Another notable association is poverty rate by race/ethnicity and urbanization. Spatial pattern of association between White populations and drug overdose death rates suggest that poverty rate in both urban and rural areas are positively associated with drug overdose deaths. These associations are significant in both periods for urban areas and only in the earlier period for rural areas. That means for White populations, higher poverty rate is associated with greater drug overdose deaths regardless of the locations. Indeed, a previous study documented that areas
with higher opioid prescription rates have higher percentage of White populations and which is associated with higher opioid related mortality (Grigoras et al., 2018). They also found rural areas to have higher prescription rates. However, this paper demonstrates association in both rural and urban areas. Therefore, the underlying mechanism seems to be multifactorial. This is evident when we look at the poverty rate for Black and Hispanic populations in rural areas which is negatively associated with drug overdose deaths; whereas poverty rate in urban areas for the same populations is positively associated. That means, in rural areas this effect is opposite, where Black and Hispanic populations in areas with higher poverty level might have less exposure to the drug overdose. On the other hand, in urban areas poverty in Black populations has significant associations with higher drug overdose deaths, though this association is not significant for poverty in Hispanic populations. Population loss index has negative coefficients and significant for both periods. This means counties experiencing population loss have lower drug overdose deaths.

5.3 Limitations and strengths

One limitation is that the dataset used here for the outcome variable has a good number of suppressed data for privacy policy, which causes the outcome variable to be skewed. This may impact the death rate estimation and model specification. Having more accurate counts instead of suppressed data would be helpful to get a better inference. However, this paper uses method of moments (MOM) to address the issue of non-normally distributed outcomes, which is recommended in this case (Anselin & Rey, 2014). An additional strength is that this study uses population level data which is spatially representative at the county level and uses spatial analysis to address spatial impacts. Another limitation is that this study does not take into consideration of different age groups in the outcome variables. We know that insurance coverage
and prescription patterns vary in different age groups. For example, Grigoras et al. (2018) found higher opioid prescription rates among people with >65 years of age and Medicare populations which can be related to associated mortality. In addition, the relationship of pharmaceutical industries’ role with prescription patterns by the physicians in terms of insurance coverage is not taken into consideration either. On the other hand, this paper uses all drug overdose deaths that includes deaths from all types of drug classes so that it can provide a common idea for all types of drugs; thus, cost-effective policies for areas with greater needs, and easier ways for awareness campaign.

5.4 Implications and future directions

This study informs how drug overdose death rates are changing over time and geographic areas. This shows the areas with growing mortality burdens which can be compared over time. This finding is enriched with associated socioeconomic and sociodemographic factors which may help to address the areas with greater needs. This can be interesting for the researchers who are interested to see the effect of economy and policy on the outcome. Future studies having lesser suppressed data and more contextual factors can provide a better vision. Also, study focusing on different age groups and gender may serve valuable functions. Another focus area can be pharmaceutical industries’ role on pain medication prescription patterns among different insured groups. Further in-depth analysis on these areas might provide with awareness and valuable information for controlling the epidemic situation for policymakers and future researchers.
References


Local Area Unemployment Statistics Home Page. (n.d.). Retrieved February 15, 2019, from 
https://www.bls.gov/lau/

Local Spatial Autocorrelation (1). (n.d.). Retrieved December 9, 2018, from 
https://geodacenter.github.io/workbook/6a_local_auto/lab6a.html#local-moran


Mobley, L. R. (n.d.). “A Primer on Spatial Modeling and Spatial Regression”.


Mobley, L. R. (2013). What are spatial data? When are they sufficient?. *Spatial Demography*, 1(1), 120-130. https://doi.org/10.1007/BF03354890


## Appendix: Tables and Figures

### Table 1

Variables used for this study, description, metric, sources and years. All the variables are aggregated over two time intervals at county level.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Description</th>
<th>Metric</th>
<th>Sources</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome variable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug overdose death rates</td>
<td>Proportion of population died from drug overdose according to their death certificates by county aggregated over 2007-2011 and 2012-2016</td>
<td>Continuous</td>
<td>“Underlying cause of death, 1999-2016” from CDC WONDER website</td>
<td>2007-2011 (Early) &amp; 2012-2016 (Late)</td>
</tr>
<tr>
<td><strong>Predictor variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Un-insured rate</td>
<td>Percent un-insured total population &lt;65 years, early and late.</td>
<td>Continuous</td>
<td>US Census Bureau, SAHIE program</td>
<td>2005 (Early) &amp; 2014 (Late)</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>Proportion of unemployed population, early and late.</td>
<td>Continuous</td>
<td>Bureau of Labor Statistics, LAUS program</td>
<td>2005 (Early) &amp; 2015 (Late)</td>
</tr>
<tr>
<td>Overall poverty rate</td>
<td>Estimated percent of total population in poverty for the income year, early and late</td>
<td>Continuous</td>
<td>US Census Bureau, SAIPE program</td>
<td>2005 (Early) &amp; 2014 (Late)</td>
</tr>
<tr>
<td>Poverty rate for Black rural</td>
<td>Proportions of total population for whom poverty data exists for particular race/ethnicity and urbanicity, early and late</td>
<td>Continuous</td>
<td>Calculated by GeoDa center</td>
<td>2000 (Early) &amp; 2005-2009 aggregated (Late)</td>
</tr>
<tr>
<td>Poverty rate for Black urban</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty rate for Hispanic rural</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty rate for Hispanic urban</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty rate for White rural</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty rate for White urban</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population loss index</td>
<td>The variable indicating whether counties experiencing population loss or not, early and late</td>
<td>Dichotomous</td>
<td>Economic Research Service (ERS) of USDA</td>
<td>2004 (Early) &amp; 2015 (Late)</td>
</tr>
</tbody>
</table>
Table 2

Descriptive statistics for all county-level variables used in the model

<table>
<thead>
<tr>
<th>Variables</th>
<th>Time Interval</th>
<th>Early period (N=3106 counties)</th>
<th>Later period (N= 3106 counties)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Drug overdose death rate</td>
<td>Early period</td>
<td>9.42</td>
<td>9.92</td>
</tr>
<tr>
<td></td>
<td>Later period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Un-insured rate</td>
<td>Early period</td>
<td>17.99</td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td>Later period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>Early period</td>
<td>5.37</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td>Later period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty rate for rural Black</td>
<td>Early period</td>
<td>2.0</td>
<td>4.64</td>
</tr>
<tr>
<td></td>
<td>Later period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty rate for urban Black</td>
<td>Early period</td>
<td>0.62</td>
<td>1.95</td>
</tr>
<tr>
<td></td>
<td>Later period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty rate for rural Hispanic</td>
<td>Early period</td>
<td>1.14</td>
<td>2.98</td>
</tr>
<tr>
<td></td>
<td>Later period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty rate for urban Hispanic</td>
<td>Early period</td>
<td>0.46</td>
<td>1.72</td>
</tr>
<tr>
<td></td>
<td>Later period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty rate for rural White</td>
<td>Early period</td>
<td>7.87</td>
<td>5.26</td>
</tr>
<tr>
<td></td>
<td>Later period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty rate for urban White</td>
<td>Early period</td>
<td>1.26</td>
<td>2.15</td>
</tr>
<tr>
<td></td>
<td>Later period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall poverty rate</td>
<td>Early period</td>
<td>15.34</td>
<td>6.52</td>
</tr>
<tr>
<td></td>
<td>Later period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Counties with population loss</td>
<td>Early period</td>
<td>599 (count)</td>
<td>19.29%</td>
</tr>
<tr>
<td></td>
<td>Later period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>Moran's I</td>
<td>Z-value</td>
<td>P value</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>Traditional method</td>
<td>EB method</td>
<td>Traditional method</td>
</tr>
<tr>
<td>Drug overdose death rate (Early)</td>
<td>0.53</td>
<td>0.52</td>
<td>49.16</td>
</tr>
<tr>
<td>Drug overdose death rate (Late)</td>
<td>0.55</td>
<td>0.54</td>
<td>50.84</td>
</tr>
</tbody>
</table>
Table 4  
**Model 1: Regression output from OLS (Ordinary Least Square) model with regression diagnostics**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Earlier period</th>
<th></th>
<th>Later period</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>P value</td>
<td>Coefficient</td>
<td>P value</td>
</tr>
<tr>
<td><strong>CONSTANT</strong></td>
<td>3.819</td>
<td>0.000</td>
<td>0.674</td>
<td>0.421</td>
</tr>
<tr>
<td>Poverty rate</td>
<td>0.046</td>
<td>0.383</td>
<td>0.085</td>
<td>0.114</td>
</tr>
<tr>
<td>Population Loss</td>
<td>-5.479</td>
<td>0.000</td>
<td>-4.075</td>
<td>0.000</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>0.799</td>
<td>0.000</td>
<td>2.083</td>
<td>0.000</td>
</tr>
<tr>
<td>Un-insured rate</td>
<td>-0.253</td>
<td>0.000</td>
<td>-0.149</td>
<td>0.001</td>
</tr>
<tr>
<td>Poverty rate in Black rural</td>
<td>-0.248</td>
<td>0.000</td>
<td>-0.722</td>
<td>0.000</td>
</tr>
<tr>
<td>Poverty rate in Black urban</td>
<td>0.337</td>
<td>0.001</td>
<td>0.213</td>
<td>0.049</td>
</tr>
<tr>
<td>Poverty rate in Hispanic rural</td>
<td>0.009</td>
<td>0.914</td>
<td>-0.375</td>
<td>0.000</td>
</tr>
<tr>
<td>Poverty rate in Hispanic urban</td>
<td>0.191</td>
<td>0.081</td>
<td>-0.203</td>
<td>0.093</td>
</tr>
<tr>
<td>Poverty rate in White rural</td>
<td>0.611</td>
<td>0.000</td>
<td>0.197</td>
<td>0.000</td>
</tr>
<tr>
<td>Poverty rate in White urban</td>
<td>1.267</td>
<td>0.000</td>
<td>1.035</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Regression Diagnostics**
- Multicollinearity number: 15.847, 15.618
- Normality of errors: 2843.073, 0.000, 1589.124, 0.000
- Heteroskedasticity: 1404.397, 0.000, 1137.679, 0.000, 451.488, 0.000, 444.662, 0.000

**Regimes diagnostics**
- Global Chow Test: 109.827, 0.000

**Note:** Statistically significant effect estimates at significance level of 0.05 are highlighted in the bold font. Superscript C denotes significant change in coefficients over time for the variables at 5% significance level. Global Chow test indicates the overall change of effect over time.
- Test for Normality of errors: Jarque-Bera test
- Tests for Heteroskedasticity: Breusch-Pagan test and Koenker-Bassett test
Table 5
Diagnostics for spatial dependence (Lagrange Multiplier tests and Anselin-Kelejian test)

<table>
<thead>
<tr>
<th>Tests</th>
<th>Earlier Period</th>
<th></th>
<th>Later Period</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>P value</td>
<td>Coefficient</td>
<td>P value</td>
</tr>
<tr>
<td>LM Lag (RS(\rho))</td>
<td>1550.738</td>
<td>0.000</td>
<td>1490.997</td>
<td>0.000</td>
</tr>
<tr>
<td>Robust LM Lag (RS(\rho^*))</td>
<td>205.254</td>
<td>0.000</td>
<td>234.046</td>
<td>0.000</td>
</tr>
<tr>
<td>LM Error (RS(\lambda))</td>
<td>1347.599</td>
<td>0.000</td>
<td>1258.232</td>
<td>0.000</td>
</tr>
<tr>
<td>Robust LM Error (RS(\lambda^*))</td>
<td>2.114</td>
<td>0.146</td>
<td>1.281</td>
<td>0.258</td>
</tr>
<tr>
<td>Anselin-Kelejian test</td>
<td>39.652</td>
<td>0.000</td>
<td>36.166</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Note: LM = Lagrange Multiplier Test. Methodology for steps of diagnostics testing for spatial dependence to determine correct model specification (Anselin & Rey, 2014, p.110) is described in the result chapter.
Table 6
Model 2: Regression output from spatial two stage least squares estimation (Lag model)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Earlier period</th>
<th>Later period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>P value</td>
</tr>
<tr>
<td>CONSTANT&lt;sup&gt;C&lt;/sup&gt;</td>
<td>2.282</td>
<td>0.001</td>
</tr>
<tr>
<td>Poverty rate</td>
<td>0.081</td>
<td>0.038</td>
</tr>
<tr>
<td>Population Loss</td>
<td>-1.404</td>
<td>0.000</td>
</tr>
<tr>
<td>Unemployment rate&lt;sup&gt;C&lt;/sup&gt;</td>
<td>-0.011</td>
<td>0.916</td>
</tr>
<tr>
<td>Un-insured rate&lt;sup&gt;C&lt;/sup&gt;</td>
<td>-0.198</td>
<td>0.000</td>
</tr>
<tr>
<td>Poverty rate in Black rural&lt;sup&gt;C&lt;/sup&gt;</td>
<td>-0.094</td>
<td>0.026</td>
</tr>
<tr>
<td>Poverty rate in Black urban</td>
<td>0.121</td>
<td>0.101</td>
</tr>
<tr>
<td>Poverty rate in Hispanic rural&lt;sup&gt;C&lt;/sup&gt;</td>
<td>0.033</td>
<td>0.591</td>
</tr>
<tr>
<td>Poverty rate in Hispanic urban</td>
<td>0.151</td>
<td>0.045</td>
</tr>
<tr>
<td>Poverty rate in White rural</td>
<td>0.174</td>
<td>0.002</td>
</tr>
<tr>
<td>Poverty rate in White urban</td>
<td>0.658</td>
<td>0.000</td>
</tr>
<tr>
<td>Spatial lag of Death Rate (ρ)</td>
<td>0.797</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Regimes Diagnostics

Global Chow test 38.610 0.000

Note: Statistically significant effect estimates at significance level of 0.05 are highlighted in the bold font. Superscript<sup>C</sup> denotes significant change in coefficients over time for the variables at 5% significance level. Global Chow test indicates the overall change of effect over time.

- White correction test is used to correct the standard errors for heteroskedasticity issue
### Table 7
Model 3: Regression output from spatially weighted two stage least squares estimation
(Combo/ Durbin model)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Earlier period</th>
<th>Later period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>P value</td>
</tr>
<tr>
<td>CONSTANT&lt;sup&gt;C&lt;/sup&gt;</td>
<td>0.966</td>
<td>0.032</td>
</tr>
<tr>
<td>Poverty rate</td>
<td>0.075</td>
<td>0.011</td>
</tr>
<tr>
<td>Population Loss</td>
<td>-1.572</td>
<td>0.000</td>
</tr>
<tr>
<td>Unemployment rate&lt;sup&gt;C&lt;/sup&gt;</td>
<td>0.012</td>
<td>0.877</td>
</tr>
<tr>
<td>Un-insured rate&lt;sup&gt;C&lt;/sup&gt;</td>
<td>-0.127</td>
<td>0.000</td>
</tr>
<tr>
<td>Poverty rate in Black rural&lt;sup&gt;C&lt;/sup&gt;</td>
<td>-0.089</td>
<td>0.005</td>
</tr>
<tr>
<td>Poverty rate in Black urban</td>
<td>0.187</td>
<td>0.002</td>
</tr>
<tr>
<td>Poverty rate in Hispanic rural&lt;sup&gt;C&lt;/sup&gt;</td>
<td>0.016</td>
<td>0.729</td>
</tr>
<tr>
<td>Poverty rate in Hispanic urban</td>
<td>0.112</td>
<td>0.071</td>
</tr>
<tr>
<td>Poverty rate in White rural&lt;sup&gt;C&lt;/sup&gt;</td>
<td>0.158</td>
<td>0.001</td>
</tr>
<tr>
<td>Poverty rate in White urban</td>
<td>0.596</td>
<td>0.000</td>
</tr>
<tr>
<td>Spatial lag of Death Rate (ρ)</td>
<td>0.824</td>
<td>0.000</td>
</tr>
<tr>
<td>Error term (Lamda)</td>
<td>-0.581</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Regimes Diagnostics**

Global Chow test | 40.657 | 0.000

**Note:** Statistically significant effect estimates at significance level of 0.05 are highlighted in the bold font. Superscript <sup>C</sup> denotes significant change in coefficients over time for the variables at 5% significance level. Global Chow test indicates the overall change of effect over time.

- The KP-HET option is used in this model because of heteroskedasticity issues in the OLS model.
Figure 1
Drug overdose death rates per 100,000 during early (2007-2011) and late (2012-2016) periods using the same cut points

Drug Overdose Death Rates 2007-2011 (Early)

Drug Overdose Death Rates 2012-2016 (Late)
Figure 2

Spatial patterns of drug overdose death rates during early (2007-2011) and late (2012-2016) periods, queen contiguity weight is used (Traditional method)

LISA Map of Drug Overdose Death Rates Early (Traditional Method)

LISA Map of Drug Overdose Death Rates Late (Traditional Method)
Figure 3
Spatial patterns of drug overdose death rates during early (2007-2011) and late (2012-2016) periods, queen contiguity weight is used (Empirical Bayes Standardization method)

LISA Map of Drug Overdose Death Rates Early (Empirical Bayes Method)

LISA Map of Drug Overdose Death Rates Late (Empirical Bayes Method)
Figure 4

Histogram of drug overdose death rates (early and late)
## Supplementary Material

### Table

**Model 2: Regression output from spatial two stage least squares estimation (Lag model) – without White correction test**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Earlier period</th>
<th></th>
<th>Later period</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>P value</td>
<td>Coefficient</td>
<td>P value</td>
</tr>
<tr>
<td>CONSTANT(^C)</td>
<td>2.282</td>
<td>0.001</td>
<td>-1.597</td>
<td>0.016</td>
</tr>
<tr>
<td>Poverty rate</td>
<td>0.081</td>
<td>0.048</td>
<td>0.097</td>
<td>0.021</td>
</tr>
<tr>
<td>Population Loss</td>
<td>-1.404</td>
<td>0.000</td>
<td>-0.906</td>
<td>0.029</td>
</tr>
<tr>
<td>Unemployment rate (^C)</td>
<td>-0.011</td>
<td>0.904</td>
<td>0.471</td>
<td>0.000</td>
</tr>
<tr>
<td>Un-insured rate (^C)</td>
<td>-0.198</td>
<td>0.000</td>
<td>-0.033</td>
<td>0.379</td>
</tr>
<tr>
<td>Poverty rate in Black rural (^C)</td>
<td>-0.094</td>
<td>0.045</td>
<td>-0.267</td>
<td>0.000</td>
</tr>
<tr>
<td>Poverty rate in Black urban</td>
<td>0.121</td>
<td>0.121</td>
<td>0.196</td>
<td>0.021</td>
</tr>
<tr>
<td>Poverty rate in Hispanic rural (^C)</td>
<td>0.033</td>
<td>0.594</td>
<td>-0.205</td>
<td>0.000</td>
</tr>
<tr>
<td>Poverty rate in Hispanic urban</td>
<td>0.151</td>
<td>0.077</td>
<td>0.075</td>
<td>0.431</td>
</tr>
<tr>
<td>Poverty rate in White rural (^C)</td>
<td>0.174</td>
<td>0.001</td>
<td>0.042</td>
<td>0.328</td>
</tr>
<tr>
<td>Poverty rate in White urban</td>
<td>0.658</td>
<td>0.000</td>
<td>0.614</td>
<td>0.000</td>
</tr>
<tr>
<td>Spatial lag of Death Rate</td>
<td>0.797</td>
<td>0.000</td>
<td>0.774</td>
<td>0.000</td>
</tr>
</tbody>
</table>

### Regimes Diagnostics

| Global Chow test | 35.552 | 0.000 |

**Note:** Statistically significant effect estimates at significance level of 0.05 are highlighted in the bold font. Superscript \(^C\) denotes significant change in coefficients over time for the variables at 5% significance level. Global Chow test indicates the overall change of effect over time.