

ABSTRACT

Title of Dissertation: **REVEALING VARIATIONS IN ENVIRONMENTAL
INEQUALITY FROM PRODUCTION- AND
CONSUMPTION-BASED PERSPECTIVES**

Guangxiao Hu
Doctor of Philosophy, 2023

Dissertation Directed by: **Professor Laixiang Sun**
Department of Geographical Sciences

Toxic chemicals pose significant threats to ecosystems, climate change, and human health. Unfortunately, pollution inequality is pervasive in the United States, with a disproportionate exposure of racial/ethnic minorities and low-socioeconomic groups to toxic releases. This inequality is especially pronounced in Houston, Texas. Moreover, income inequality has widened over decades, and the distribution of toxic releases has changed over time in the USA.

To better understand the problem of pollution inequality, it is necessary to investigate the embodied toxic release in final demand by states and income groups. Notably, pollution inequality is non-uniform across regions and over time in the USA. The relationship between socioeconomic development and toxic risk necessitates analysis to comprehend the resultant health outcomes at different spatial scales and locations. Further, identifying the spatial heterogeneity of the association between environmental hazards and socioeconomic indicators is critical for addressing environmental inequality. Investigating the spatiotemporal heterogeneity of the impact

of racial disparities and socioeconomic development on toxic risk can reveal disparities between regions and trends in pollution inequality.

This study employs an extended U.S. multi-regional input-output (MRIO) model with toxic chemical release data to analyze the inter-regional transfer of embodied toxic release between states and their unequal distribution between income groups from a consumption-based perspective. Additionally, this study analyzes the spatial non-stationarity in the associations between toxic chemical hazard risk and community characteristics of census block groups in Texas, USA, for 2017 using a multiscale geographically weighted regression (MGWR). Further, this study uses Houston, a city with a history of segregation and discrimination and a diverse racial/ethnic makeup, as an example to analyze the spatiotemporal heterogeneity of the impact of racial disparities and socioeconomic development on toxic risk using geographical and temporal weighted regression (GTWR) models. The study's outcomes are instrumental in determining whether pollution inequality has improved or worsened. Results indicate that non-metallic and metallic products manufacturing sectors are crucial for interregional flows of embodied toxic release from the Great Lake Region to Southeast, Mid-Atlantic, and Northeast regions, and are the most important sectors for most states from the consumption-based perspective. The findings also highlight the significance of identifying the spatial patterns of the association between toxic chemical hazard risks and community characteristics at the census block group level to address environmental inequality.

REVEALING VARIATIONS IN ENVIRONMENTAL INEQUALITY
FROM PRODUCTION- AND CONSUMPTION-BASED PERSPECTIVES

by

Guangxiao Hu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:

Professor Laixiang Sun, Chair/Advisor
Professor Kuishuang Feng, Co-Chair/Co-Advisor
Professor Giovanni Baiocchi
Professor Deb Niemeier
Professor Amy Sapkota

© Copyright by
Guangxiao Hu
2023

Dedication

To my parents:

Hai Hu and Jinghua Chen

Acknowledgments

I would like to express my sincere gratitude to my academic advisor, Professor Laixiang Sun, for his invaluable guidance and unwavering support throughout my Ph.D. journey. His advice, expertise, and encouragement have been instrumental in helping me to navigate the complexities of my research and stay focused on my goals. Whenever I felt overwhelmed or discouraged, Professor Sun was always there to provide me with reassurance and perspective. His words of wisdom and his unwavering faith in my abilities gave me the confidence I needed to push through challenging times. I would also like to express my heartfelt gratitude to my co-advisor, Professor Kuishuang Feng, for providing me with advice not only on the research but also on my future career pathway. His advice and suggestions have helped me to identify my strengths and interests and to develop a plan for achieving my professional goals.

I would also like to thank other committee members equally for their invaluable guidance, encouragement, and support. Professor Giovanni Baiocchi's insistence on strict guidelines helped me to refine my research methods and to present my findings in a clear and concise manner. His attention to detail and thoroughness in reviewing my work challenged me to improve the quality of my research and writing. Professor Deb Niemeier has always been supportive of my research and my career. I have been fortunate to work with Deb on a research project, and her expertise and insights have been invaluable to me. Her commitment to academic rigor and attention to detail has helped me to produce high-quality research that I am proud of. Professor

Amy Sapkota's guidance and mentorship have been invaluable to me, not only during the UMD Global STEWARDS program but also in my dissertation. I am thankful for Amy's leadership and mentorship in the UMD Global STEWARDS program, which has had a profound impact on my academic and professional development. I am particularly grateful for Amy's role as the dean's representative, where she provided crucial feedback and guidance on my dissertation.

I would like to express my sincere gratitude to all of my cohorts at the Human Dimensions of Global Change groups, Cortney Gustafson, Yu Xin, Hannah Younes, Ipsita Kumar, Kaihui Song, Baobao Liu, Yuhao Wang, Xiangjie Chen, and Haolin Shao. Working alongside each of you has been an enriching and transformative experience. Your diverse perspectives, ideas, and insights have broadened my understanding of human geography and challenged me to think critically about the world around us.

To my family, thank you for your unconditional love and unwavering support. Your sacrifices and encouragement have been the driving force behind my academic achievements. I am grateful for your unconditional support and love. To my boyfriend, Wei-Hsiang Wang, thank you for sharing my tears and happiness in life. Your love has kept me motivated, even on the toughest days, and your belief in me has helped me achieve what I thought was impossible. I couldn't have done this without you.

Table of Contents

Preface	ii
Dedication	ii
Acknowledgements	iii
Table of Contents	v
List of Tables	vii
List of Figures	viii
List of Abbreviations	xi
Chapter 1: Introduction	1
1.1 Background	1
1.2 Motivation and research objectives	5
Chapter 2: Tracing toxic chemical releases embodied in US interstate trade and their unequal distribution	12
2.1 Introduction	12
2.2 Model and data	15
2.2.1 Multi-regional input-output model	15
2.2.2 Data	17
2.3 Results	20
2.3.1 Production-based and consumption-based TCR inventories	20
2.3.2 Inter-state transfers of embodied TCR	25
2.3.3 TCR footprint by income groups	29
2.4 Discussion and conclusion	33
Chapter 3: Multiscale Analysis of the Relationship between Toxic Chemical Hazard Risks and Racial/Ethnic and Socioeconomic Groups	39
3.1 Introduction	39
3.1.1 Background	39
3.1.2 Literature Review	40
3.1.3 Research objectives	49
3.2 Materials and Methods	50

3.2.1	Study area	50
3.2.2	Data and Method	51
3.2.3	Global regression model: The Ordinary Least Squares (OLS) model . . .	55
3.2.4	Local regression model: The MGWR model	55
3.3	Results and Discussion	57
3.4	Conclusions and policy implementations	67
Chapter 4:	Is pollution inequality improved or worsen over time? Spatiotemporal heterogeneity of the impact of racial disparities and socioeconomic development on toxic risk	71
4.1	Introduction	71
4.2	Methods and data	79
4.2.1	Study area	79
4.2.2	GTWR model	81
4.2.3	Data	83
4.3	Results and Discussion	88
4.3.1	Analysis of estimation results	88
4.3.2	Spatiotemporal heterogeneity distribution of coefficients	92
4.4	Conclusions and policy recommendations	107
Chapter 5:	Conclusions	111
5.1	Summary of major findings and synthesis of contributions	111
5.2	Limitations and future studies	114
Appendix A:	Appendix for Chapter 3	118
A.1	The OLS model results	118
A.2	Testing regression assumptions	120
A.2.1	Distribution of the dependent variable	120
A.2.2	Linear relationship	120
A.2.3	Normality of residuals	122
A.2.4	Testing the Homoscedasticity Assumption	123
A.2.5	Multicollinearity	124
A.2.6	Moran's I analysis for dependent variable	128
A.2.7	Statistical tests for MGWR results	130
A.3	Parameter estimates for all models	136
Appendix B:	Appendix for Chapter 4	141
B.1	Statistical test	141
B.1.1	t-test for significance	141
B.1.2	Test for multicollinearity	142
B.1.3	Regression diagnostics tests	142
B.2	All results	145
Bibliography		153
Bibliography		153

List of Tables

2.1	Total household income for specified income group.	19
2.2	State abbreviations for US states.	24
2.3	Per capita personal income (dollars) in US in 2017.	32
3.1	Literature review of distributive environmental inequality	45
3.2	Variable code names and description.	54
3.3	Statistical Summary of the variables (N = 15,154 block groups)	56
3.4	The MGWR results for the spatial variability of parameters using adaptive bandwidth	59
3.5	Statistical summary of the local condition numbers of each model specification .	62
4.1	Variable names, units, and detailed description	86
4.2	Statistical Summary of the variables	87
A.1	The OLS model results for the spatial stationarity of parameters	119

List of Figures

1.1	Geographic Distribution of TRI-Reporting facilities in 2019	2
1.2	Toxic Chemical Releases by regions in 2019	3
1.3	Map for 10 EPA Regions.	4
1.4	Total disposal or other releases over time	5
1.5	RSEI Score and corresponding releases over time	6
1.6	Conceptual diagram of proposed research	9
2.1	Production-based and consumption-based TCR and net inflows (red) and outflows (blue) of embodied TCR across states in the US.	21
2.2	Sectoral composition of production-based and consumption-based TCR in US states in 2017.	22
2.3	Embodied TCR in local consumption and domestic export in 2017	25
2.4	Inter-regional embodied TCR flows between the US EPA regions (unit: 1×10^7).	26
2.5	Flows of embodied TCR from top 15 release producing states to the final consumers in the top 15 consuming states.	28
2.6	The distribution of household number and TCR footprint across income groups of all states.	29
2.7	The distribution of TCR footprints in each state (household number and corresponding shares of total TCR).	30
2.8	TCR embodied in household consumption across states and income groups.	31
2.9	Total TCR across income groups in each state to meet household consumption.	37
3.1	Maps of Local R^2 for all MGWR models.	61
3.2	Local coefficients of racial disparity (BLACK and ASIAN) in toxic chemical risk (from Model 4).	63
3.3	Local coefficients of (a) INCOME (from Model 2), (b) VACANT (from Model 3), (c) NHSD (from Model 3), and (d) UNEMP (from Model 5) in toxic chemical risk.	64
3.4	Local coefficients of POP and HISPANIC (from Model 5) in toxic chemical risk.	66
4.1	Redlining map in Houston, Texas	81
4.2	Scheme of geographical and temporal weighted regression (GTWR) model	83
4.3	Spatial distribution of the RSEI score in 2011	84
4.4	Spatial distribution of the RSEI score in 2014	85

4.5	Spatial distribution of the RSEI score in 2017	85
4.6	Temporal variation of the parameter coefficients	89
4.7	Time series change trend of GTWR regression coefficients	91
4.8	The spatiotemporal distribution of regression coefficient estimates of Black_Pc . .	95
4.9	The spatiotemporal distribution of regression coefficient estimates of Asian_Pc . .	96
4.10	The spatiotemporal distribution of regression coefficient estimates of Others_Pc .	97
4.11	The spatiotemporal distribution of regression coefficient estimates of Hispanic_Pc	99
4.12	The spatiotemporal distribution of regression coefficient estimates of Pop_Dens .	101
4.13	The spatiotemporal distribution of regression coefficient estimates of Unemploy_Pc	103
4.14	The spatiotemporal distribution of regression coefficient estimates of Owner_Pc .	105
4.15	The spatiotemporal distribution of regression coefficient estimates of Not_High_Pc	106
A.1	Distribution of dependent variable	120
A.2	Regression diagnostics plots	121
A.3	Partial-residual plots	122
A.4	Histogram of residuals	123
A.5	Lineal model results assuming no multicollinearity	124
A.6	Pair-wise correlation	125
A.7	Partial correlation coefficient matrix	126
A.8	Overall diagnostic checking results	126
A.9	Individual diagnostic checking results	127
A.10	T-test for correlation coefficient results	128
A.11	Monte-Carlo simulation of Moran's I	130
A.12	Monte-Carlo simulation of Moran's I for Model 1 coefficients	131
A.13	Monte-Carlo simulation of Moran's I for Model 2 coefficients	132
A.14	Monte-Carlo simulation of Moran's I for Model 3 coefficients	133
A.15	Monte-Carlo simulation of Moran's I for Model 4 coefficients	134
A.16	Monte-Carlo simulation of Moran's I for Model 5 coefficients	135
A.17	parameter estimates for model 1	136
A.18	parameter estimates for model 2	137
A.19	parameter estimates for model 3	138
A.20	parameter estimates for model 4	139
A.21	parameter estimates for model 5	140
B.1	Regression diagnostics plots	143
B.2	Test of Linearity	144
B.3	The spatiotemporal distribution of regression coefficient estimates of population density 2011-2017	145
B.4	The spatiotemporal distribution of regression coefficient estimates of Black_Pc 2011-2017	146
B.5	The spatiotemporal distribution of regression coefficient estimates of Asian_Pc 2011-2017	147
B.6	The spatiotemporal distribution of regression coefficient estimates of Others_Pc 2011-2017	148

B.7	The spatiotemporal distribution of regression coefficient estimates of Hispanic_Pc 2011-2017	149
B.8	The spatiotemporal distribution of regression coefficient estimates of employment 2011-2017	150
B.9	The spatiotemporal distribution of regression coefficient estimates of housing 2011-2017	151
B.10	The spatiotemporal distribution of regression coefficient estimates of education 2011-2017	152

List of Abbreviations

ACS	American Community Survey
AIC	Akaike information criterion
AICc	corrected Akaike Information Criterion
BIC	Bayesian information criterion
CAA	Clean Air Act
COHGIS	The City of Houston Geographic Information System
CV	cross-validation
CVD	cardiovascular disease
DiD	difference-in-differences
EJ	Environmental Justice
EJScreen	Environmental Justice Screening and Mapping Tool
EPA	Environmental Protection Agency
EPCRA	Emergency Planning and Community Right-to-Know Act
GAM	generalized additive model
GDP	Gross Domestic Product
GEOID	Geographic Identifiers
GIS	Geographic Information System
GTWR	Geographical and Temporal Weighted Regression
GWP	global warming potential
GWR	Geographically Weighted Regression
HAP	hazardous air pollutant
HOLC	Home Owner's Loan Corporation
IMPLAN	the Economic Impact Analysis for Planning
IO	Input-output
MGWR	Multi-scale Geographically Weighted Regression
MRIO	multi-regional input-output

NAICS	North American Industry Classification System
NO_2	Nitrogen Dioxide
O_3	Ozone
OLS	Ordinary Least Squares
P2	Pollution Prevention
PM2.5	particulate matter with aerodynamic diameter $\leq 2.5\mu m$
PM10	particulate matter with aerodynamic diameter $\leq 10\mu m$
RSEI	Risk-Screening Environmental Indicators
RSEI-GM	Risk Screening Environmental Indicators Geographic Microdata
SES	socioeconomic status
STDR	the Spatio-temporal Distance Ratio
TCR	toxic chemical releases
TIGER	Topologically Integrated Geographic Encoding and Referencing
TRI	Toxic Release Inventory
VIF	variance inflation factor
U.S or USA	the United States of America

Chapter 1: Introduction

1.1 Background

Toxic chemicals have impacts not only on climate change [1], but also on the ecosystem and human health [2]. Toxic chemicals, including fluorinated greenhouse gases, methane, ozone and etc., are greenhouse gases with high global warming potential (GWP) [3, 4], which can accelerate the climate change. Ecosystems may be contaminated through the accumulation of toxic chemicals. For example, excessive nutrient loading can be detrimental to aquatic ecosystems [5]. Toxic chemicals are harmful to human health and well-being [6], from acute poisonings and events to chronic effects like cancer and non-cancer potential [2, 7]. For example, Thousands of people died of extremely toxic methyl isocyanate gas escape and their exposure in Bhopal, India, in 1984. In 1985, a serious chemical release occurred at a similar plant in West Virginia [8]. An estimated 14,175 chemical incidents occurred in 2013 [9], causing a huge impact on human health. Thus, it is vital to manage the toxic chemical release and reduce the impact of toxic releases on the environment and human health in the US.

The U.S. Environmental Protection Agency (EPA) defined environmental equality as “the fair treatment and meaningful involvement of all people regardless of race, color, culture, national origin, income, and educational levels with respect to the development, implementation, and enforcement of protective environmental laws, regulations, and policies” [8]. Environmental

injustice or environmental inequality is said to occur if certain racial or socioeconomic groups experience a disproportionate share of the negative environmental consequences of industrial activity [10]. According to Minnesota Pollution Control Agency, only about 6% of communities in Minnesota are near one or more of the facilities that emit pollution at a level above health guidelines. However, 14% of communities of color, which include indigenous peoples, and 9% of low-income communities are located near one or more of these facilities [11].

The current distribution of toxic chemical release is largely unequal across regions within the US [12]. For example, the Toxic Release Inventory (TRI) reporting facilities located concentratedly at the western costal area, northeastern coastal area and the Great Lake's region, while central areas have few TRI reporting facilities (see Figure 1.1). Also, releases differ significantly between EPA regions (see Figure 1.2), where Region 10 Pacific Northwest and Alaska, Region 9 West, Southwest and Hawaii, Region 4 Southeast, Region 5 Midwest and Region 6 South Central account for the majority of total releases. The map of EPA regions can be found in Figure 1.3.

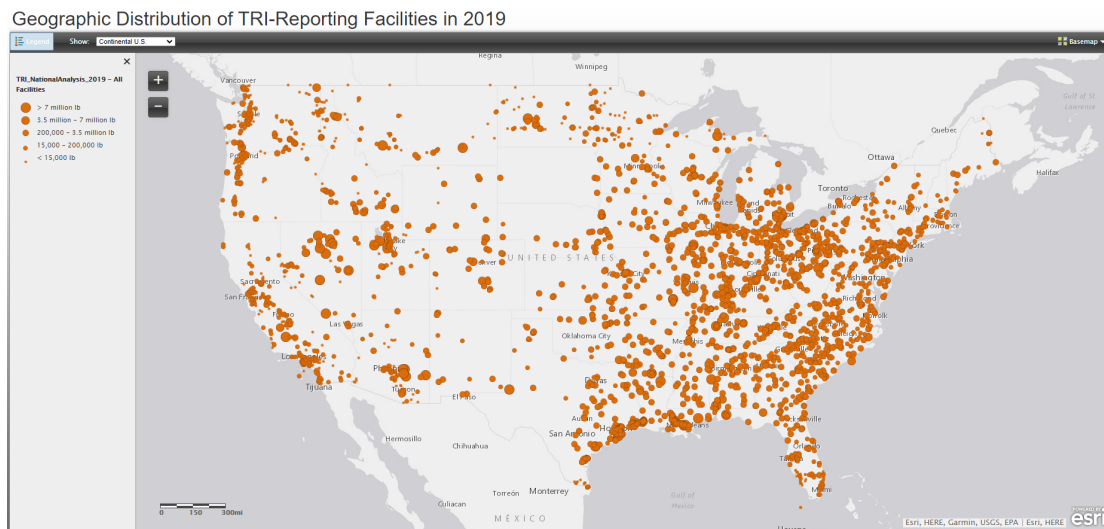


Figure 1.1: Geographic Distribution of TRI-Reporting facilities in 2019

Source:

<https://gispub.epa.gov/trina2019/facilities.html?webmap=90002265bec645dfbf0703c55d03093e>

Releases by Region, 2019

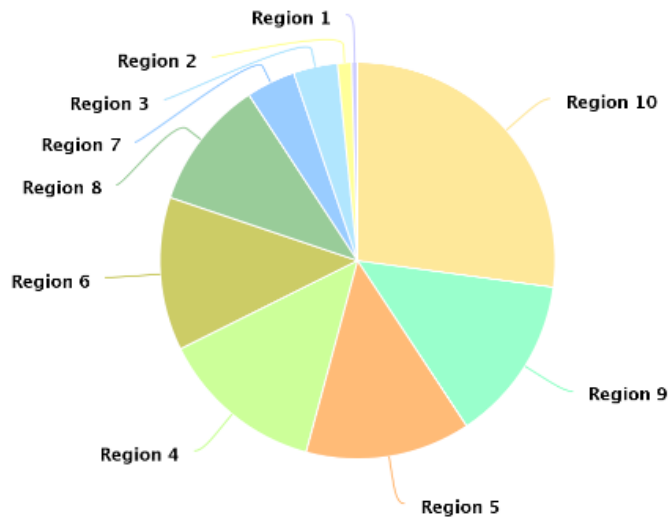


Figure 1.2: Toxic Chemical Releases by regions in 2019

Note: In 2019, Region 1 has 950 facilities, Region 2 has 1040 facilities, Region 3 has 1905 facilities, Region 4 has 4586 facilities, Region 5 has 5330 facilities, Region 6 has 2956 facilities, Region 7 has 1515 facilities, Region 8 has 718 facilities, Region 9 has 1641 facilities, and Region 10 has 742 facilities.

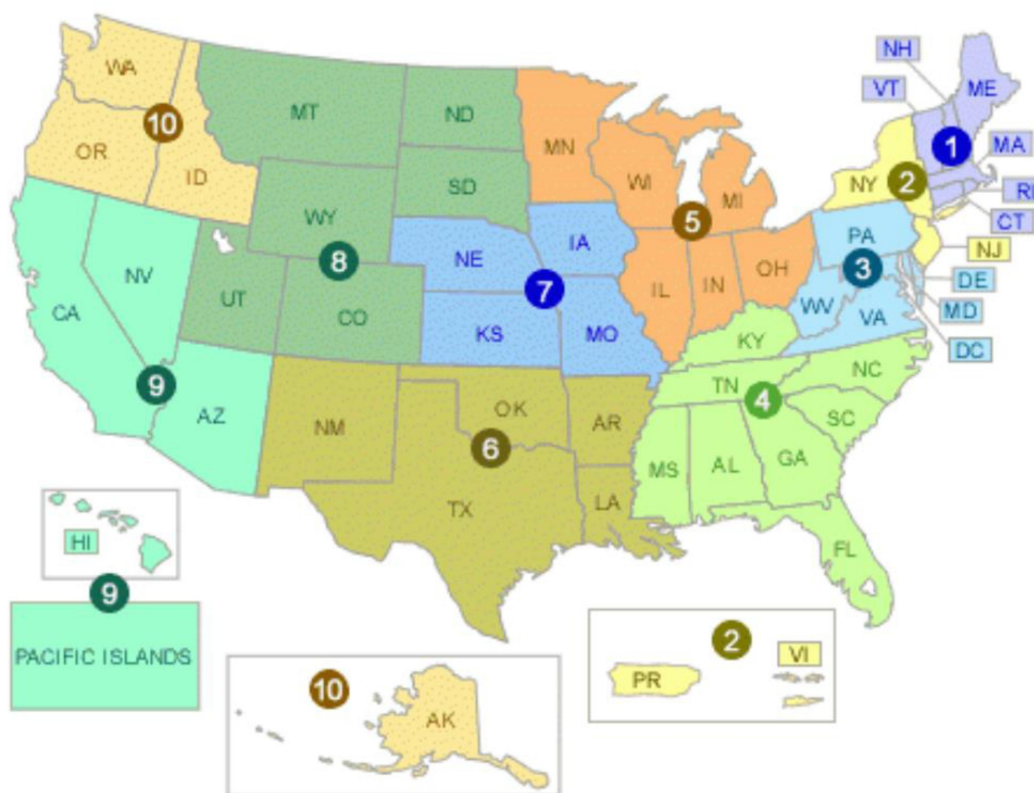


Figure 1.3: Map for 10 EPA Regions.

Note: Source: <https://www.epa.gov/aboutepa/regional-and-geographic-offices>. The EPA regions are distinct administrative units with different bureaucratic cultures, state regulations, and data sources. Region 1 (New England) includes CT, ME, MA, NH, RI, and VT; Region 2 (Northeast) includes NJ, NY, Puerto Rico, and the U.S. Virgin Islands; Region 3 (Mid-Atlantic) includes DE, DC, MD, PA, VA, WV and 7 federally recognized tribes; Region 4 (Southeast) includes AL, FL, GA, KY, MS, NC, SC, and TN; Region 5 (Great Lakes) includes IL, IN, MI, MN, OH, and WI; Region 6 (South Central) includes AR, LA, NM, OK, and TX; Region 7 (Midwest) includes IA, KS, MO, and NE; Region 8 (Mountains and Plains) includes CO, MT, ND, SD, UT, and WY; Region 9 (Pacific Southwest) includes AZ, CA, HI, NV, American Samoa, Commonwealth of the Northern Mariana Islands, Federated States of Micronesia, Guam, Marshall Islands, and Republic of Palau; Region 10 (Pacific Northwest) includes AK, ID, OR, WA and 271 native tribes.

The toxic releases and related toxic risk (Risk-Screening Environmental Indicators (RSEI) score) are changing over time (Figure 1.4 and Figure 1.5). For example, from 2007-2019, total releases of TRI chemicals decreased by 19%, and reduced hazardous air pollutant (HAP) emissions, such as hydrochloric acid, from electric utilities, were the most significant contributor

to the decline. The number of facilities that reported to TRI declined by 9%. The overall RSEI score decreased by 67% from 2007-2019. Of the types of releases modeled by RSEI, air releases, by far, contributed the most to the RSEI Scores.

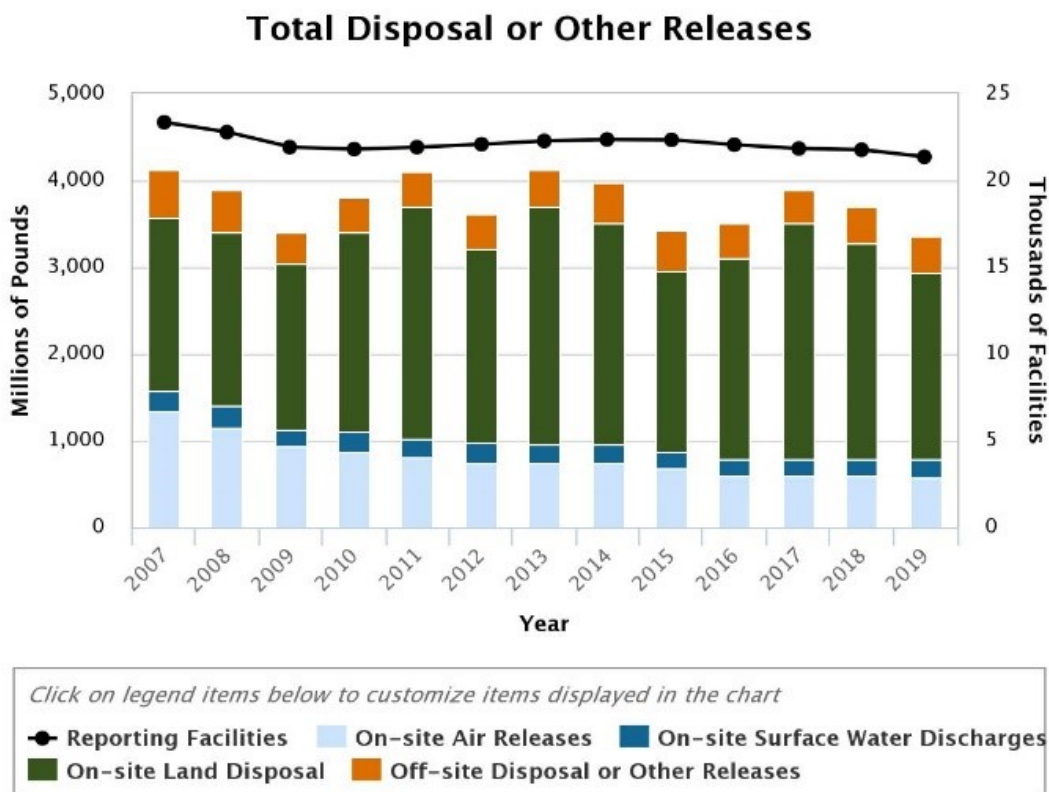


Figure 1.4: Total disposal or other releases over time
Source: <https://edap.epa.gov/public/extensions/EasyRSEI/EasyRSEI.html>

1.2 Motivation and research objectives

Environmental inequality, that is, the disproportionate share of the negative environmental consequences in certain racial or socioeconomic groups, is a long-lasting problem in the US. Most researchers found evidence of racial minorities exposing to more pollution, e.g., [13–15]. low-income groups were also found to be disproportionately exposed to pollution [14, 16, 17]. Since the current distribution of toxic chemical release is largely unequal across regions within

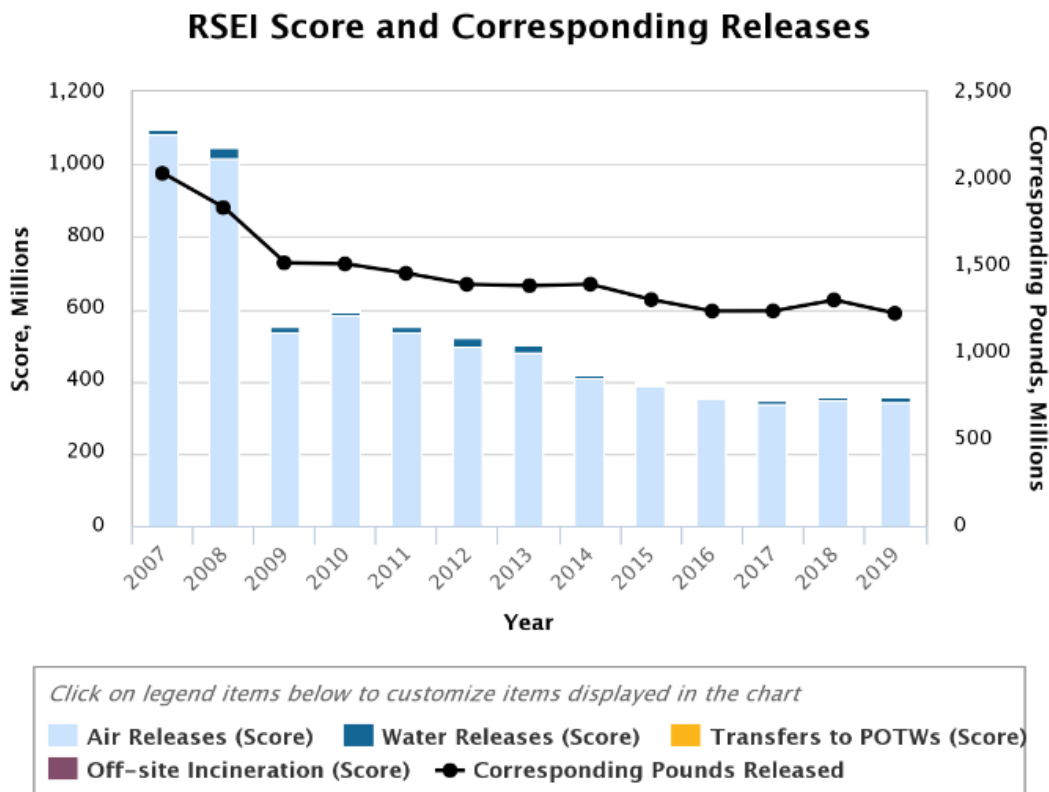


Figure 1.5: RSEI Score and corresponding releases over time
 Source: <https://edap.epa.gov/public/extensions/EasyRSEI/EasyRSEI.html>

the US, much research focused on the regional variations in environmental inequality. For example, Thind et al. [18] estimated exposures to and health impacts of PM_{2.5} (particulate matter with aerodynamic diameter $\leq 2.5\mu\text{m}$) from electricity generation in the US for each US state, by income and by race. The products produced in one region are not only consumed by local consumers, but also traded to meet other regions' consumption. Consumption-based environmental accounting [19–21] can allocate the national toxic release to regions and sectors that consumes the final goods and services, thus demonstrating how the demands of consumer regions cause the environmental problems in producer regions [22,23]. Thus, revealing variations in environmental inequality from both production- and consumption-based perspectives can bring a comprehensive understanding of the inequality problems.

Input-output (IO) model is a commonly used tool in consumption-based environmental accounting. Koh et al. [1] have calculated the toxicological footprint (consumption-based toxic chemical release) of the U.S. between 1998 and 2013. Tessum et al. [24] showed that PM_{2.5} exposure is disproportionately caused by the consumption of goods and services of the non-Hispanic white majority, but disproportionately inhaled by black and Hispanic minorities. These studies, however, only focused on toxic emissions at the national level. Identifying the disparities and inequality between states in the US from a consumption-based perspective is valuable for an effective national pollution mitigation plan according to the characteristic of each state. The multi-regional input-output (MRIO) model can examine environmental impacts embodied in consumption and interregional trade [25–27], thus revealing regional differences in environmental impacts and inequality.

For production-based perspectives, regression models are commonly used models to evaluate the relationship between pollution and demographics. Global regression models implicitly assume spatial stationarity in the relationships between explanatory variables and dependent variable(s), meaning that they assume these relationships do not vary over space. For example, most researchers found evidence of racial minorities and low-income groups exposing to less pollution, e.g., [13–16], etc. In order to capture regional variations, Brunson et al. [28] introduced Geographically Weighted Regression (GWR) as an extension of general regression models and based on kernel-weighted regression. For example, Mennis and Jordan [29] used GWR analysis and showed that the relationships among race, class, employment, urban concentration, and land use with air toxic release density in New Jersey vary significantly over space. However, GWR assumes that the relationships between negative environmental consequences and demographics vary at the same spatial scale, which can hide important local variations by spatial scale in the relationships

among independent and dependent variables in models of environmental equity. MGWR (Multi-scale Geographically Weighted Regression) is an extension of GWR that allows studying the relationships at varying spatial scales and achieves that by using varying bandwidth as opposed to a single, constant bandwidth for the entire study area [30].

Pollution inequality is not uniform across regions and over time in the US. Income inequality has been widening over decades and toxic releases have been changing over time (Figure 1.4 and Figure 1.5). Investigating the spatiotemporal heterogeneity of the impact of socioeconomic development on toxic risk can reveal the differences between regions and the trend of pollution inequality. Some researchers used spatiotemporal models to reveal the relationship between toxic releases and social changes. For example, Aoyagi and Ogunseitan [31] and Park and Kwan [32] used the global model to analyze the spatiotemporal heterogeneity of the impact of socioeconomic development on toxic risk. However, the global model cannot reveal the regional variations in their relationship. Although GWR can capture the geographic context, it cannot capture the temporal one. By including a temporal dimension in GWR, Huang et al. [33] and Fotheringham et al. [34] proposed a temporal extension of GWR: geographical and temporal weighted regression (GTWR), which is a spatiotemporal modeling from a local perspective.

The preceding introduction highlights the uneven distribution of toxic releases across space and time, indicating that pollution inequality in the US is contingent upon the spatial scale and temporal context. The overarching aim of this proposal is to comprehensively comprehend pollution inequality from both production and consumption perspectives and to analyze its variation across space and time. To achieve this objective, this dissertation research puts forth three primary research questions, which are elucidated as follows.

Research question 1: How much toxic chemical releases are triggered by the consumption

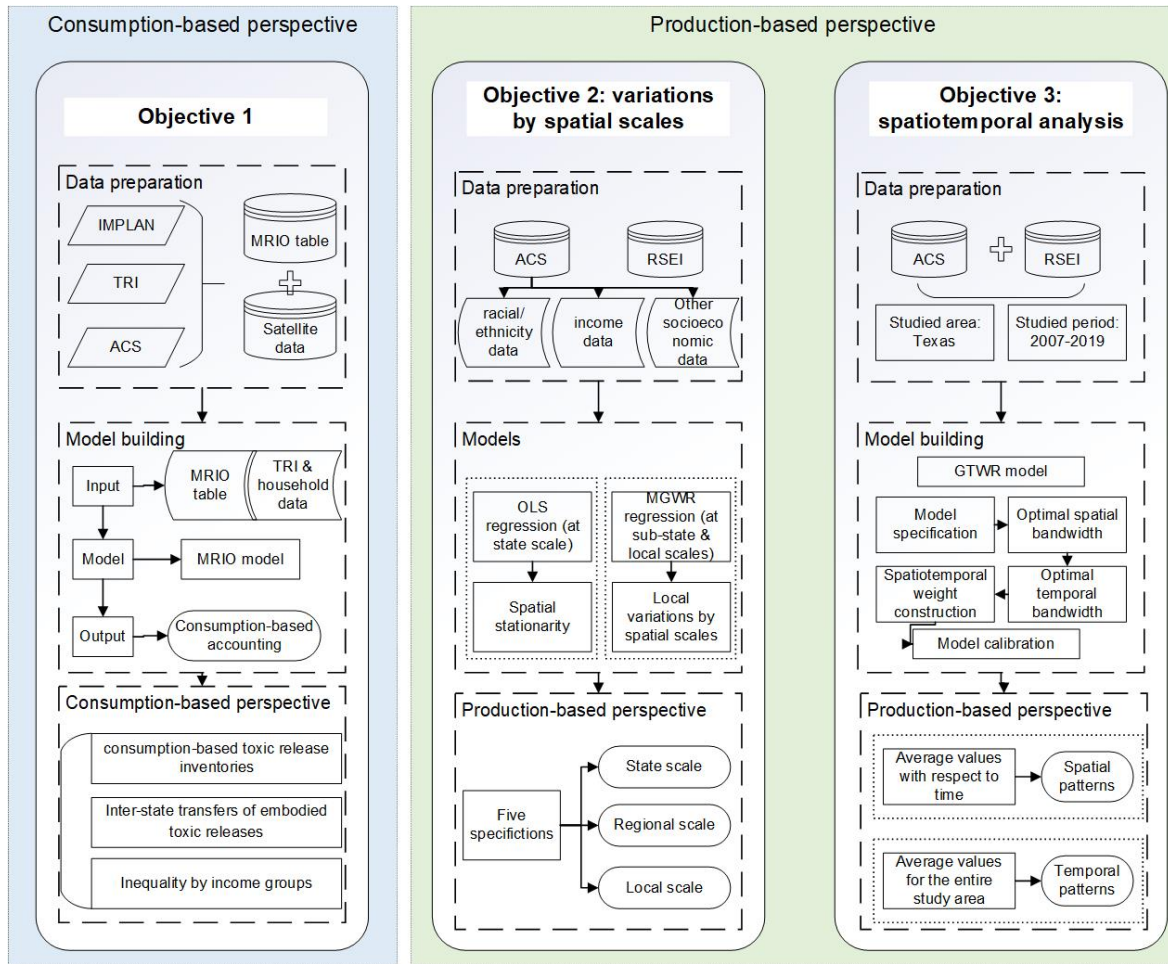


Figure 1.6: Conceptual diagram of proposed research

of final demand and what is the outsourcing of embodied toxic releases between states in the US?

The objective of this paper is to investigate the relationship between toxic releases, consumption, and income inequality across different states in the US. To achieve this goal, this paper will first examine the toxic releases triggered by the final demand of each state, as well as those outsourced to other states. Secondly, this paper will estimate the toxic release inequality resulting from the consumption of different income groups within each state. To accomplish these objectives, this paper plans to utilize a multi-regional input-output (MRIO) model, which will use data from the American MRIO table and toxic chemical release inventory in 2017 to calculate the production and consumption-based toxic release of each state, as well as the inter-regional transfer of embodied

toxic release between states. Additionally, this model will estimate the toxic release footprint by income groups and the inter-state toxic release transfer triggered by different groups by states. By using the MRIO model, this paper will provide quantitative information for understanding how pollution equality of toxic release varies among income groups and regions across the US, and for identifying the transfer of responsibility hidden in products exchanged among states. Ultimately, our findings may inform the development of policies and interventions aimed at reducing toxic release inequality and promoting environmental justice for all communities.

Research question 2: Is there a relationship between socioeconomic indicators and toxic release risk? Local modeling of Texas toxic release

The objective of this paper is to investigate whether racial minorities and individuals with low socioeconomic status have been disproportionately affected by toxic risk and whether the relationship between these factors varies across different regions. To achieve this goal, we plan to employ a Multiscale Geographically Weighted Regression (MGWR) model to assess the impact of race/ethnicity and socioeconomic status (SES) on environmental quality in Texas. Specifically, we will analyze block group level data on industrial pollution risk from the Risk Screening Environmental Indicators database maintained by the Environmental Protection Agency (EPA), along with income and demographic variables obtained from the American Community Survey (ACS). The MGWR model will allow us to examine the relationship between race/socioeconomic status and toxic risk at various spatial scales, from local to regional. By doing so, we aim to provide a comprehensive understanding of the impact of these factors on toxic risk and the extent to which this impact varies across different regions. Ultimately, our findings may inform the development of more effective policies and interventions aimed at reducing pollution inequality and improving environmental quality for all communities.

Research question 3: Is pollution inequality improved or worsen over time? Spatiotemporal heterogeneity of the impact of socioeconomic development on toxic risk.

This paper aims to investigate whether pollution inequality across regions in the United States has improved over time. Specifically, we seek to determine if racial minorities and low socioeconomic entities have been disproportionately affected by toxic risk over time. By examining the spatiotemporal heterogeneity of the impact of socioeconomic development on toxic risk, we can identify differences between regions, enabling local governments to develop region-specific regulations. Moreover, we can identify the trend of pollution inequality and its key determinants over time, which is crucial for developing effective policy interventions. To our knowledge, no previous studies have explored the relationship between socioeconomic indicators and toxic risk in the US, incorporating both a local model and a temporal dimension. Our study focuses on Texas and employs a geographical and temporal weighted regression (GTWR) model to analyze the spatiotemporal heterogeneity of the impact of socioeconomic development on toxic risk in the state.

Chapter 2: Tracing toxic chemical releases embodied in US interstate trade and their unequal distribution

2.1 Introduction

The current distribution of toxic chemical releases (TCR) is largely unequal across regions within the US [12]. According to the Toxic Release Inventory (TRI), toxic release facilities are concentrated in the western coast, southeastern coast, and the Great Lake region, while midwestern regions have much fewer TRI facilities (see Figure. 1.1). Also, the total TCR differs significantly between regions in the US (see Figure. 1.2). Although Pacific Northwest and Pacific Southwest have relatively smaller numbers of facilities, the region contributes a large number of total releases. Many existing studies regarding TCR in the US focused on TCR from the production perspective, analyzing the inequality across regions and the health impacts across demographic groups. For example, Ard [15] explored air pollution exposure across regions and demographic groups and found that the variation of the exposure has grown substantially in the United States over 1995–2004 using the United States Environmental Protection Agency (EPA)'s Risk Screening Environmental Indicators Geographic Microdata (RSEI-GM) and multilevel regression models. The results showed that African Americans are more exposed to air toxins than Whites even though the amount is decreasing for everyone. Using multilevel models, Ard and Smiley

[35] investigated the relationship between racialized poverty segregation and hazardous industrial sites, and the results showed that counties with higher economic and racial segregation levels have a significantly higher density of facilities, which are disproportionately clustered at communities with higher percentages of poor African American and Hispanic people.

However, the products produced in one region are not only consumed by local consumers, but also traded to meet other regions' consumption. Therefore, consumption in one place may cause pollution in other places via interregional trade of goods and services [36]. The embodied TCR in regional trades may help us identify the transfer of responsibility hidden in products exchanged among regions. Consumption-based environmental accounting [19–21,37] can allocate the production-based TCR to regions that consume the final goods and services, thus demonstrating how the demands in consumer regions cause the environmental problems in producer regions [22,23]. Koh [38] calculated the toxicological footprint (consumption-based TCR) of the U.S. between 1998 and 2013. Fujii [39] also calculated the consumption-based toxicity of chemical emissions from the US industrial sector over the 1998–2009 period. These studies, however, only focused on toxic emissions at the national level. Through tracing TCR embodied in trade among U.S. states, consumption-based accounting is important for identifying the ultimate drivers of the releases, thus, opening up opportunities for cooperative solutions among states [40].

Aside from regional disparities, the consumption of different income and racial groups has contributed to TCR exposure and other criteria air pollution exposure disproportionately in the U.S. [24]. For example, Wang and Feliberty [41] showed that minority and low-income groups are more likely to be located within buffers around the toxic release inventory facilities. Clark et al. [42] also showed that low-income nonwhite young children and elderly people are disproportionately exposed to residential outdoor NO_2 . Tessum et al. [24] showed that $PM_{2.5}$

exposure is disproportionately caused by the consumption of goods and services of the non-Hispanic white majority, but disproportionately inhaled by black and Hispanic minorities. Thind et al. [18] found that average exposures to PM 2.5 are the highest for blacks, and are higher for lower income than for higher income. Using EPA's TRI data, Charette et al. [43] assessed residential clusters associated with pollutant releases, and results showed that the highest amount of releases was concentrated at clusters with a high percentage of the working class. Income inequality is extremely high in the U.S., for example, the mean household income in the top 20% was 10.3 times greater than that in the bottom 20% in 1975 while the gap was enlarged to 16.6 times in 2019 [44,45]. Income groups differ significantly in terms of consumption patterns, thus contributing to the environmental impacts very differently [46]. It is essential to understand consumption-based TCR or TCR footprint, which accounts for all toxic emissions (based on the TRI) along entire domestic supply chains [1], of different income groups so that policy decision-makers can be informed by demand-side toxic reduction policy recommendations with lower socioeconomic costs [47].

In summary, previous research mainly focused on TCR in the US from the production perspective, analyzing pollution exposure across regions and demographic groups [15, 35]. A few researchers calculated consumption-based toxicity of chemical emissions in the US [38,39], however, these studies were conducted at the national level. There has been a lack of research attention to the flows of TCR embodied in interstate trade, which is termed consumption-based toxic chemical footprints. Consequently, there has been an absence of research analyzing the disproportionate exposure of consumption-based TCR faced by different income groups across states in the US. In this study, we fill the above gaps by first assessing regional consumption-based accounting of pollution inequalities and outsources of embodied TCR between states in the US;

and then estimating the TCR footprint of different income groups at the state level. For this purpose, we developed an environmentally extended multi-regional input–output (MRIO) model for the US using the MRIO table from IMPLAN (the Economic Impact Analysis for Planning) and the toxic chemical release inventory data in 2017 from US EPA, to calculate the production- and consumption-based TCR of each state, the inter-regional transfer of embodied TCR between states, as well as the TCR footprint across income groups. The MRIO model is a commonly used method to examine environmental impacts embodied in consumption and interregional trade [25–27, 48–50]. By using the MRIO model, this paper provides quantitative information on how pollution inequality of toxic release varies across regions and income groups in the US. Such information can facilitate the formation of shared responsibilities between consumers and producers in the efforts to mitigate the adverse impact of the toxic release on the environment and human health.

2.2 Model and data

2.2.1 Multi-regional input-output model

In a multi-regional input–output model, the product flows within a region and the product inflows and outflows among regions are taken into consideration. Suppose there are n regions and there are m producing sectors in each region. Z_{ij}^{rs} is the monetary flow from sector i in region r to sector j in region s . The elements of X are denoted as x_i^r , representing the total output of sector i in the region r . The elements of the direct consumption coefficient matrix A are denoted as a_{ij}^{rs} , which is given by $a_{ij}^{rs} = z_{ij}^{rs}/x_j^s$, representing the amount of commodity i in the region r directly required to produce per unit of sector j in region s . Y consists of intraregional (y^{rr} and

y^{ss}) and interregional final demand y^{rs} and y^{sr}). The element y_i^{rs} represents the final demand for commodity i in the region r from region s .

The basic equation of the multi-region input-output model (MRIO) is

$$x = Ax + Y \quad (2.1)$$

where,

$$A = \begin{bmatrix} A^{1,1} & A^{1,2} & \dots & A^{1,n} \\ A^{2,1} & A^{2,2} & \dots & A^{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A^{n,1} & A^{n,2} & \dots & A^{n,n} \end{bmatrix} \quad X = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix} \quad Y = \begin{bmatrix} y^{1,1} & y^{1,2} & \dots & y^{1,n} \\ y^{2,1} & y^{2,2} & \dots & y^{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ y^{n,1} & y^{n,2} & \dots & y^{n,n} \end{bmatrix}$$

To get x , the equation 2.1 can be rewritten as:

$$x = (I - A)^{-1}Y \quad (2.2)$$

where, $L = (I - A)^{-1}$ is the Leontief inverse matrix; I is the identity matrix.

To calculate the interregional TCR flows, a matrix t_c is needed to represent the direct TCR coefficient of different sectors in different regions. The interregional TCR flows can be calculated as follow:

$$TF = t_c(I - A)^{-1}Y \quad (2.3)$$

where TF is the TCR embodied in commodities flowing between regions; t_c is a matrix

of sectoral direct TCR coefficient (total toxic chemical releases of sector j divided by the total output of sector j in each state).

Production-based TCR of region r can be estimated as [21]

$$TCR_P^r = \sum_{s=1}^n \sum_{i=1}^m TF_i^{rs} \quad (2.4)$$

Consumption-based TCR of region s can be estimated as [21]

$$TCR_C^r = \sum_{r=1}^n \sum_{i=1}^m TF_i^{rs} \quad (2.5)$$

To get the TCR related to each income group's final consumption by state level in the US [51,52], we can get:

$$T_{incomegroup} = t_c(I - A)^{-1}HY_{incomegroup} \quad (2.6)$$

where $T_{incomegroup}$ is the embodied TCR associated with each income group's final consumption (TCR footprint) in each state. $HY_{incomegroup}$ is the household final demand for each income group in each state.

2.2.2 Data

The American multi-regional input-output (MRIO) tables were compiled using single-regional input-output tables for each of the 51 states and the trade flow data between states. The trade flow data between states were collected from the county-level trade data and aggregated at the state level. The MRIO data needed were extracted from the Economic Impact Analysis for

Planning (IMPLAN) database [53]. There are 546 sectors in IMPLAN, representing all industries in the US classified by North American Industry Classification System (NAICS) codes [53].

The toxic release data was collected from Toxics Release Inventory (TRI) Program conducted by the United States Environmental Protection Agency (EPA). TRI is a compilation of data submitted by certain facilities subject to the reporting requirements of Section 313 of the federal Emergency Planning and Community Right-to-Know Act (EPCRA) [54]. All facilities are required to report to TRI if they meet the reporting criteria: (1) are subject to EPCRA 313(b)2 determination, or (2) are in a covered industry sector and exceed the employee threshold [55]. Those facilities are mainly involved in hazardous waste treatment, manufacturing, electric power generation, and metal mining. The government passed EPCRA in 1986 in response to the disastrous gas leak in Bhopal, India, in 1984, which raised the concern about environmental and safety hazards of toxic releases. Toxics Release Inventory compiled the information of certain toxic chemicals that may pose a threat to human health and the environment. The information includes how much of each chemical is released to the environment (air or water) and managed through recycling, energy recovery, and treatment of each U.S. facility. The current TRI toxic chemical list contains 775 individually listed chemicals and 33 chemical categories, which may cause cancer or other chronic human health effects, significant adverse acute human health effects, or significant adverse environmental effects [56].

EPA's Risk-Screening Environmental Indicators (RSEI) is a screening-level model incorporating TRI data with measures of human exposure and toxicity, which assesses the potential impact of industrial chemical releases from pounds-based, hazard-based, and risk-related perspectives [56]. The RSEI model uses reported data and information on facilities' on-site chemical releases and chemical transfers from these facilities to off-site facilities to model RSEI risk-related results.

Table 2.1: Total household income for specified income group.

Specification	Explanation
Households LT15k	Those earning less than \$15,000
Households 15–30 K	Those earning between \$15,000 and \$30,000
Households 30–40 K	Those earning between \$30,000 and \$40,000
Households 40–50 K	Those earning between \$40,000 and \$50,000
Households 50–70 K	Those earning between \$50,000 and \$70,000
Households 70–100 K	Those earning between \$70,000 and \$100,000
Households 100–150 K	Those earning between \$100,000 and \$150,000
Households 150–200 K	Those earning between \$150,000 and \$200,000
Households 200 K+	Those earning greater than \$200,000

Source: Candi Clouse (2022) [58].

The RSEI model can provide comparable values for each kind of toxic release using a continuous system of numerical weights developed by the EPA. In this paper, we used “TRI Pounds”, which reflect the number of pounds released or transferred that are reported to TRI and are available for all releases and transfers. In order to make the sectors of TRI/RSEI data and IMPLAN data consistent, this paper matched the sectors between these two datasets using Primary-six-digit North American Standard Industry Classification System (NAICS) code.

Household number data of each state by income level was collected from American Community Survey (ACS) by the United States Census Bureau (“2017 Data Profiles — American Community Survey — US Census Bureau,”) [57]. The income groupings in ACS are different from the IMPLAN, thus matching groupings is needed between these two datasets. We assumed that the number of households within an income group is equally distributed, then re-aggregated the number of households in ACS according to the income grouping in the IMPLAN dataset. The household income for the specified income group is listed in Table 2.1.

2.3 Results

2.3.1 Production-based and consumption-based TCR inventories

Figure 2.1 shows the production-based and consumption-based TCR and net flows of embodied TCR in each US state in 2017 (The abbreviations of states can be found in Table 2.2). The TCR from industrial production is unequally distributed across regions in the US with a high concentration in a few south central and Great Lakes Region states. The TCR from the top 10 states accounts for about 53% of the national total release. Texas is ranked first with a total amount of 54.34 million pounds TCR, which is 56.7% higher than the TCR in Indiana, the second largest TCR state, and 70% higher than that in Ohio (the third largest TCR state). Michigan also produce a large amount of TCR and accounted for 18.15% of the national total TCR in 2007. In sectoral detail, Figure 2.2 presents a comparison of sectoral TCR from both the production and consumption perspectives for all states. From the production perspective, every state has its own dominant industries. For example, non-metallic and metallic products manufacturing sectors are the most important sources of the TCR for most states. Texas, as the highest production-based TCR state, has more than 80% TCR from those two sectors. Mississippi, Louisiana, and Arizona also have more than 80% TCR from those non-metallic and metallic products manufacturing sectors. In addition, mining and supportive activities are the largest contributing sector to the TCR in Nevada and Alaska, and utilities are the largest TCR sector in North Dakota.

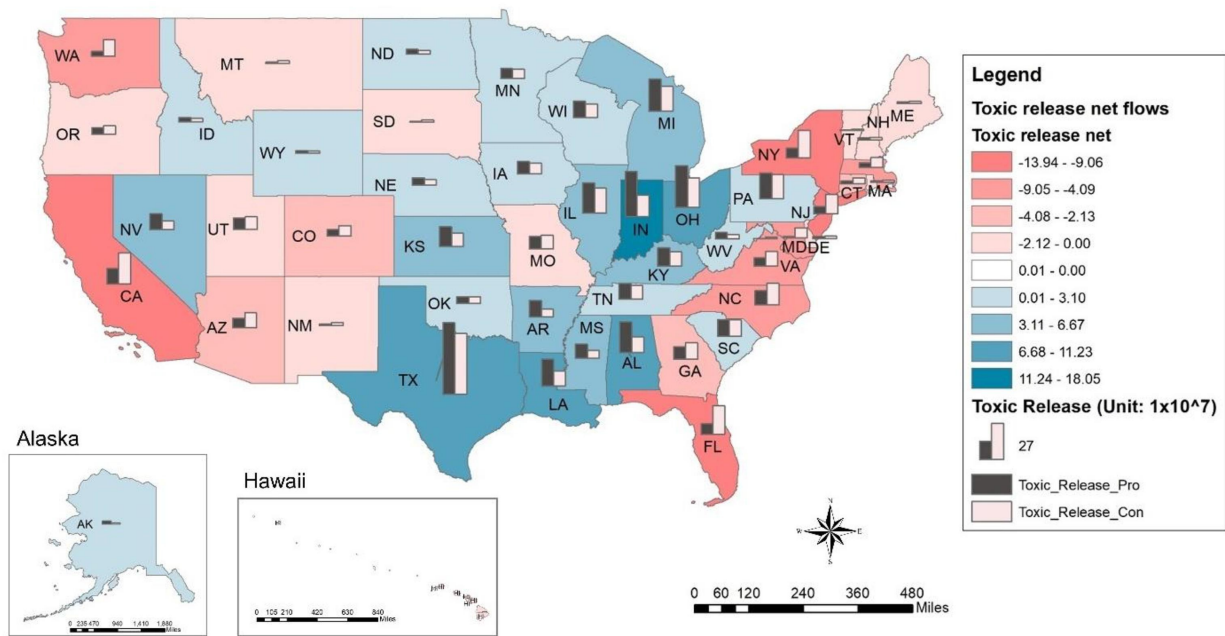


Figure 2.1: Production-based and consumption-based TCR and net inflows (red) and outflows (blue) of embodied TCR across states in the US.

USA states shapefile:

<https://www.arcgis.com/home/item.html?id=1a6cae723af14f9cae228b133aebc620>

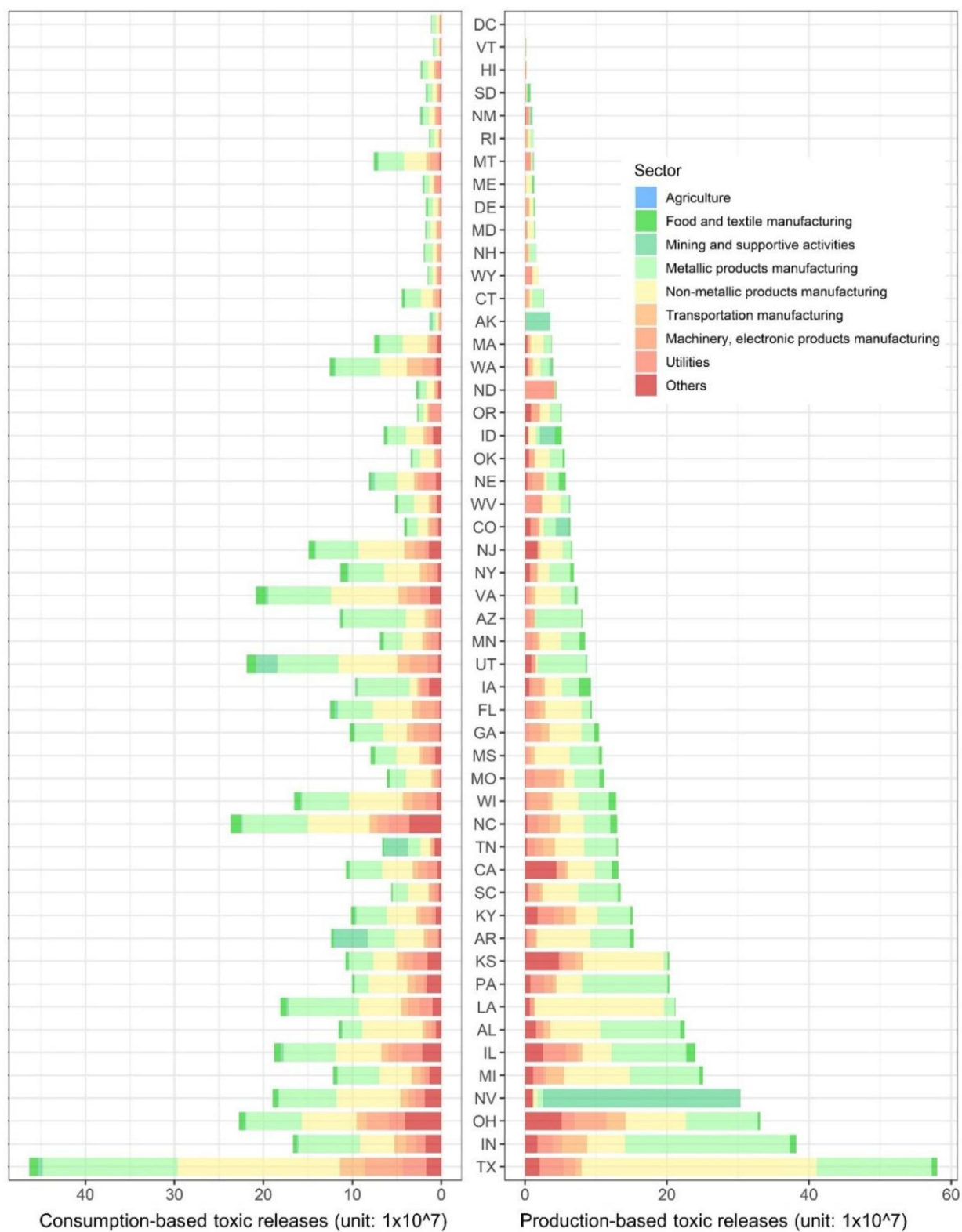


Figure 2.2: Sectoral composition of production-based and consumption-based TCR in US states in 2017.

Compared with the production-based TCR, the consumption-based TCR are less unequally distributed across regions in the US. Texas still ranks the top state for consumption-based TCR with an amount of about 46.31 million pounds. Some west and east coastal states, such as Washington, California, and North Carolina, and some Great Lakes Region states, including New York and Michigan, have a similar number of consumption-based TCR. While most central states, such as Montana and Wyoming, produce and consume relatively small amounts of TCR. In sectoral detail, the non-metallic and metallic products manufacturing sector are the most important sectors, accounting for more than 50% of the consumption-based TCR in most states.

From Figure 2.1 we can see that net importers of embodied TCR are mainly focused on the east and west coastal areas, while net exporters are focused on central areas. For example, Indiana, Texas, Louisiana, and Alabama are among the largest net exporters of embodied TCR. Those states are also poor states with low per capita personal incomes (Table 2.3), while the coastal states, such as Washington, California, New York, New Jersey, and Florida, are the main net importers of embodied TCR with high personal incomes.

Figure 2.3 showed the embodied TCR in local consumption and domestic trade in 2017. In this figure, we can see that a large amount of embodied TCR flowed in and out of the states via inter-state trade, while the embodied TCR in local demand accounted for a very small proportion of total releases in every state. For example, the local consumption of Texas cause only around 10 million pounds of TCR that were produced from its domestic firms, while the other 20 million pounds of TCR were imported from other states. In the meantime, a similar amount of TCR in Texas was caused by other states to meet the final demand in those states.

Table 2.2: State abbreviations for US states.

US state	Abbreviation	US state	Abbreviation	US state	Abbreviation
Alabama	AL	Kentucky	KY	North Dakota	ND
Alaska	AK	Louisiana	LA	Ohio	OH
Arizona	AZ	Maine	ME	Oklahoma	OK
Arkansas	AR	Maryland	MD	Oregon	OR
California	CA	Massachusetts	MA	Pennsylvania	PA
Colorado	CO	Michigan	MI	Rhode Island	RI
Connecticut	CT	Minnesota	MN	South Carolina	SC
Delaware	DE	Mississippi	MS	South Dakota	SD
District of Columbia	DC	Missouri	MO	Tennessee	TN
Florida	FL	Montana	MT	Texas	TX
Georgia	GA	Nebraska	NE	Utah	UT
Hawaii	HI	Nevada	NV	Vermont	VT
Idaho	ID	New Hampshire	NH	Virginia	VA
Illinois	IL	New Jersey	NJ	Washington	WA
Indiana	IN	New Mexico	NM	West Virginia	WV
Iowa	IA	New York	NY	Wisconsin	WI
Kansas	KS	North Carolina	NC	Wyoming	WY

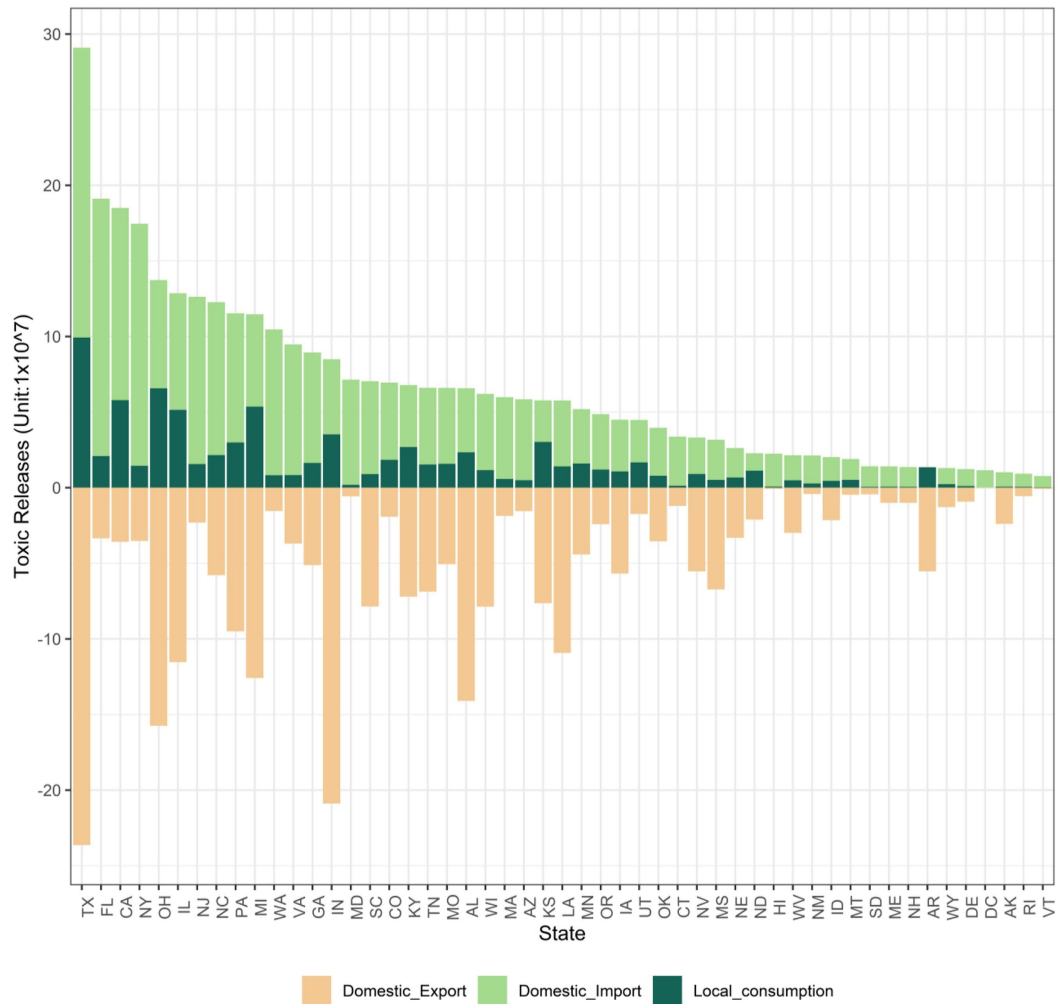


Figure 2.3: Embodied TCR in local consumption and domestic export in 2017

2.3.2 Inter-state transfers of embodied TCR

Figure 2.4 showed the inter-state transfers of embodied TCR in the US EPA regions (see Figure 1.3). The results showed that the main outsourcing of embodied TCR is from release-intensive central regions to rich coastal regions. For example, the largest interregional transfer of TCR is from Great Lake Region to the Southeast (16.2 million pounds). Although Great Lake Region (including Illinois, Indiana, Michigan, Minnesota, Ohio, and Wisconsin) only has 16% of the U.S. population, it owns 25% of all the TRI facilities, most of which are concentrated

in toxic release-intensive sectors of the fabricated metals (i.e., manufacture of metal products) and chemical manufacturing sectors [59]. Except for the Southeast, Great Lake Region is also outsourced to neighboring rich regions like Mid-Atlantic (9.52 million pounds) and Northeast (8.19 million pounds). On the contrary, final demand consumption in the Southeast region (including Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee, and 6 tribes), which accounts for 1/5 of the U.S. population, is mainly being supported by production and associated releases occurring in neighboring regions, including Great Lakes (16.2 million pounds) and South Central (9.89 million pounds).

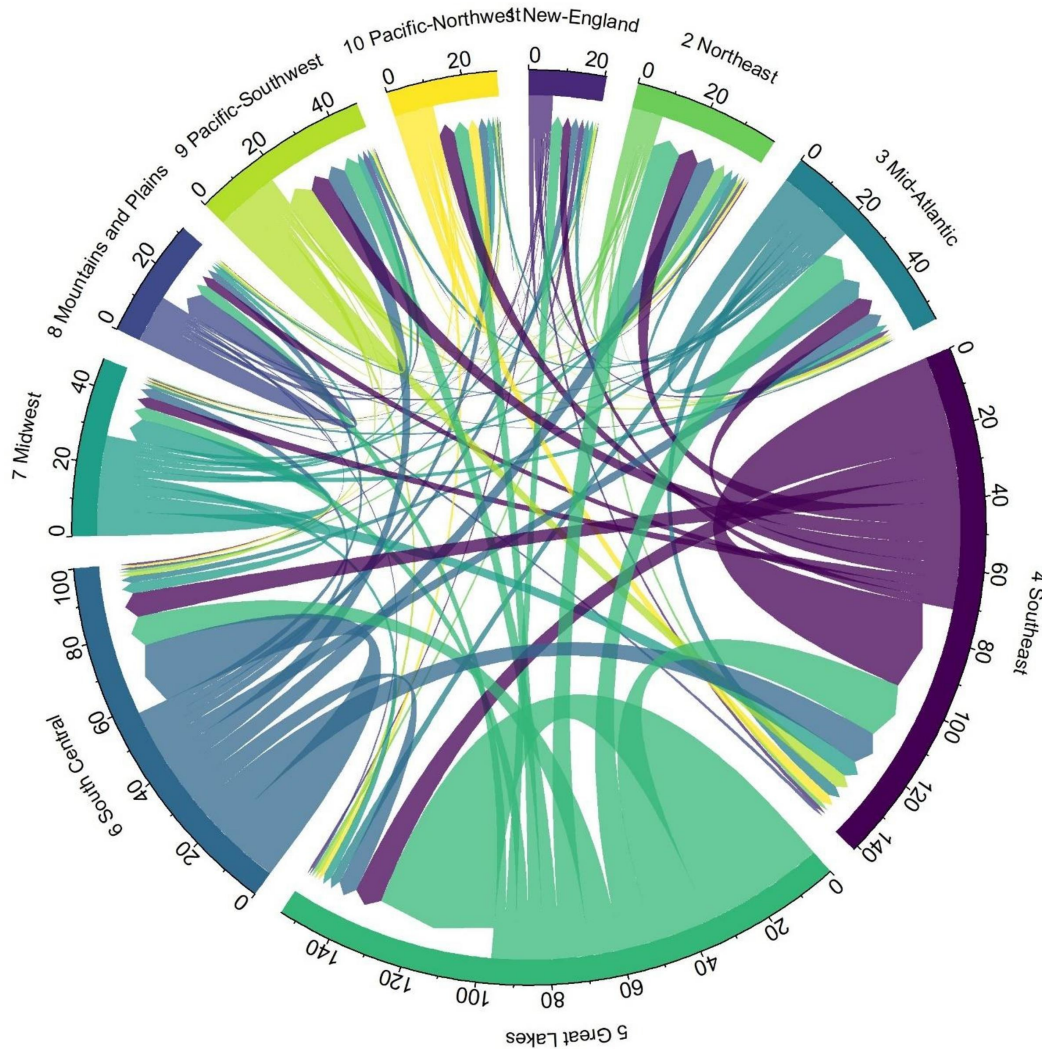


Figure 2.4: Inter-regional embodied TCR flows between the US EPA regions (unit: 1×10^7).

In Figure 2.5, we selected the top 15 states in production, which produced the greatest TCR (first column), and the top 15 states in consumption, whose consumption triggered the largest amount of TCR (third column) to show the main embodied TCR flows between producers and final consumers. The top 15 toxic release-producing states, including Texas, Illinois, Ohio, Michigan and etc., account for 63.15% of total TCR, and it is intuitive that they are net exporters of the embodied TCR (see Figure 2.1). The top 15 toxic release “consuming” states can be divided into two types of groups: one group includes the top producing states, for example, Texas, Ohio, and Illinois; another group is the top net importers of embodied TCR, for example, California, Florida, New York, and Washington. The sectoral embodied TCR flows show different patterns between these two groups. In the first group, Texas is the largest state in TCR production, and non-metallic products manufacturing accounts for the largest percentage of the production (53.25%), followed by metallic products manufacturing (27.90%). Meanwhile, a great majority of TCR from non-metallic and metallic products manufacturing is consumed by Texas. Similar patterns showed in other states in this group. For another group, non-metallic and metallic products manufacturing are also major consumption sectors in California, Florida, New York, and Washington, accounting for around 70% of the total consumption in each state. Although these states are not major TCR production states, they import large amounts of TCR from producing states with high TCR intensity.

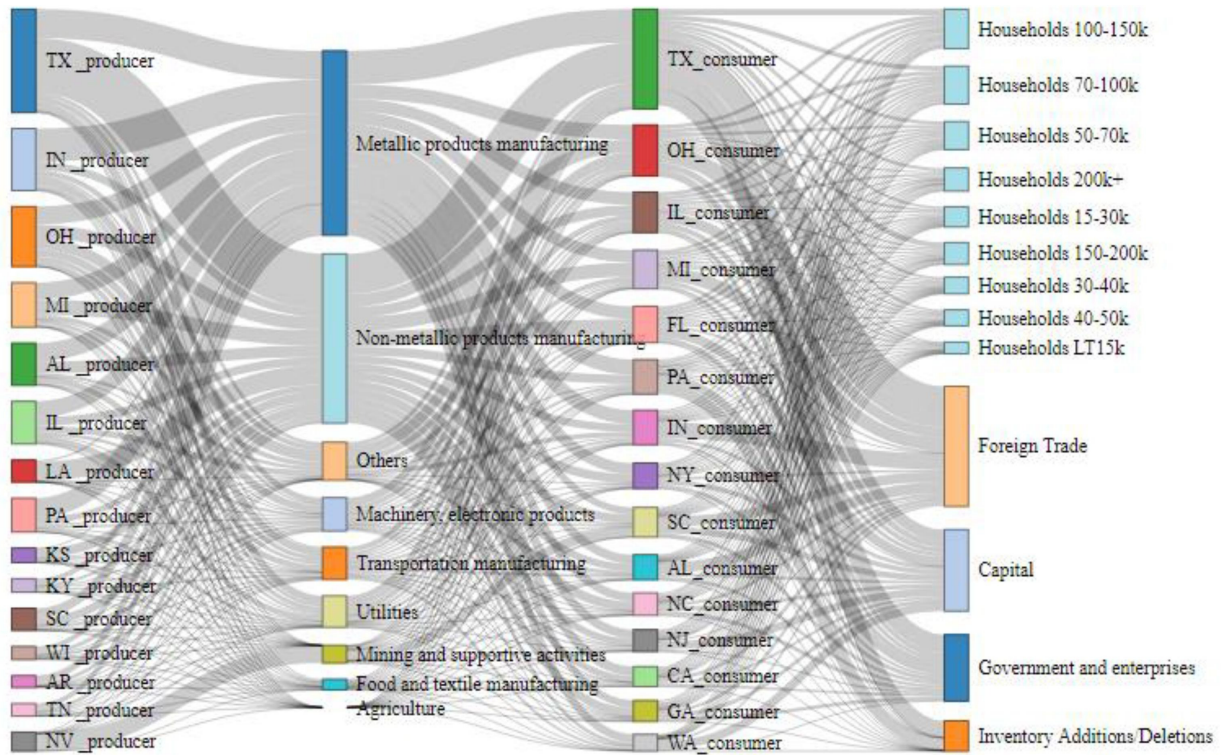


Figure 2.5: Flows of embodied TCR from top 15 release producing states to the final consumers in the top 15 consuming states.

From the final demand side, the largest portion of embodied TCR is triggered by households' final demand of all income groups (41.88%) for the top 15 consuming states, followed by foreign trade (23.39%) and capital formation (15.88%). In detail, Illinois's households with income 100–150k triggered the largest amount of TCR, around 1.593 million pounds, followed by Ohio's households with income 70–100k (1.574 million), 100–150k (1.567million), and Florida's households with income 70–100k (1.412 million). Foreign trade is also a major part that causes the embodied TCR, and Texas's export accounts for the largest proportion of embodied TCR (20%), followed by Ohio (13.99%) and Indiana (13.27%).

2.3.3 TCR footprint by income groups

Since households' final demand contributed to about half of the total TCR, in this section, we provide a detailed analysis of household consumption and associated TCR across income groups. Figure 2.7 showed the unequal distribution of TCR footprint by income groups in each state. In general, high-income groups, although with small populations, account for a high percentage of the total TCR footprint; while low-income groups with large populations account for a low percentage of the total TCR footprint. For example, the consumption of the top 35% of rich households (groups of larger than 100k) contributes to more than 50% of total TCR triggered by the final demand of all households in the U.S., where a similar unequal distribution occurs in most states (Figure 2.7). The main reason is that the TCR footprint per household in high-income groups is larger than that in low-income groups (Figure 2.6).

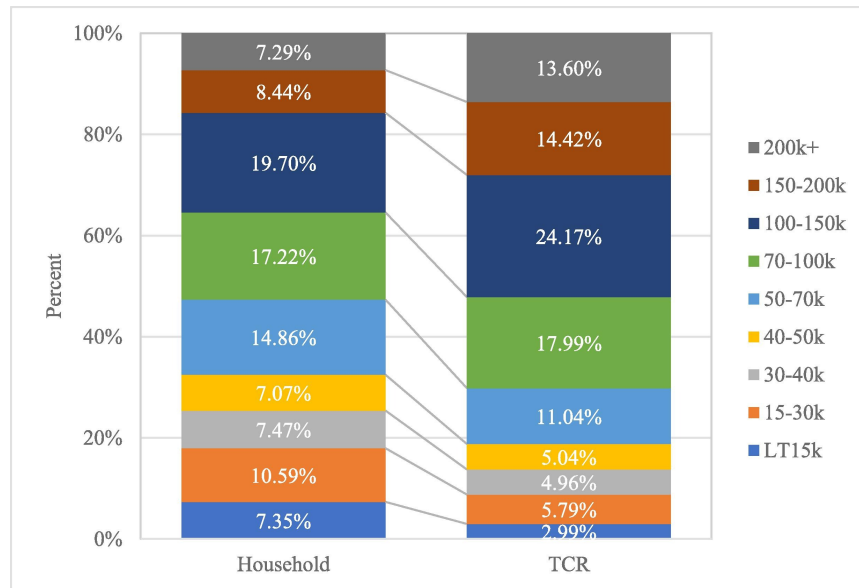


Figure 2.6: The distribution of household number and TCR footprint across income groups of all states.

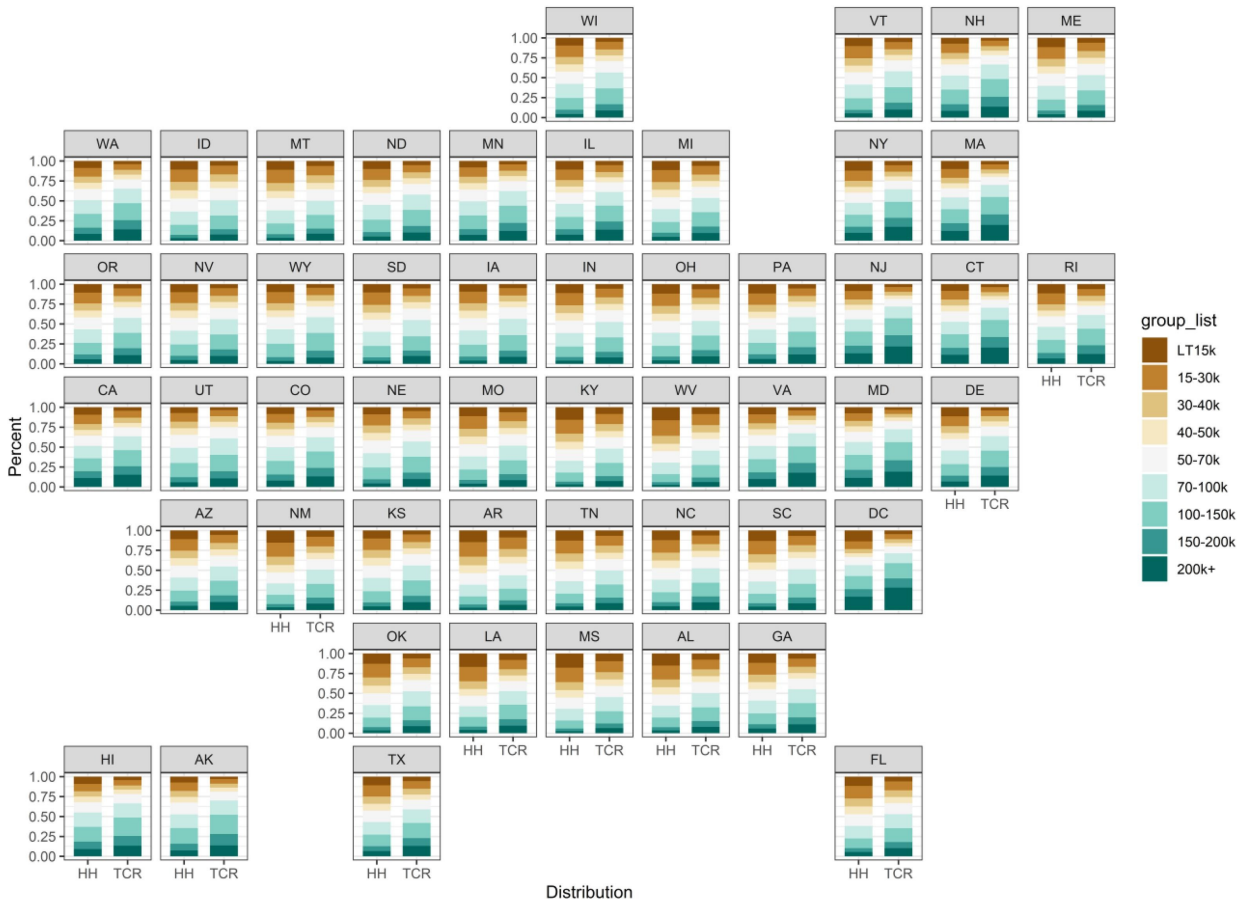


Figure 2.7: The distribution of TCR footprints in each state (household number and corresponding shares of total TCR).

Note: HH means household number.

We also found that the distributions of population and TCR footprint between rich and poor states show different patterns. On the one hand, poor states with low per capita income (Table 2.3) have a larger percentage population of low-income groups (groups of less than 30k) and thus have a higher percent of corresponding TCR footprint than the national average level. For example, the population of low-income groups in poor states Kentucky, Mississippi, and West Virginia account for more than 30% of the total population, while the national percentage is only 17.95%. The corresponding percent of TCR footprint in the above three states is more than 20%, compared to the national average of 8.78%. In contrast, the population of the high-income group (groups

larger than 100k) is smaller than the national average level, and so is the corresponding TCR footprint. On the other hand, it is intuitive that the rich states have a larger percentage of high-income groups than the national average (35.44%), for example, high-income groups account for 42.57%, 40.17% and 40.33% of the total population in District of Columbia, Maryland, and New Jersey, respectively. Although the per household TCR footprints in rich states are smaller than in poor states like Mississippi (Figure 2.8), the high number of populations in rich states still make a high percentage of TCR footprint. The consumption of high-income groups in the District of Columbia, Maryland, and New Jersey are contributing 58.82%, 56.21%, and 57.05% of the total TCR footprint, respectively (compared to the national mean of 52.20%).

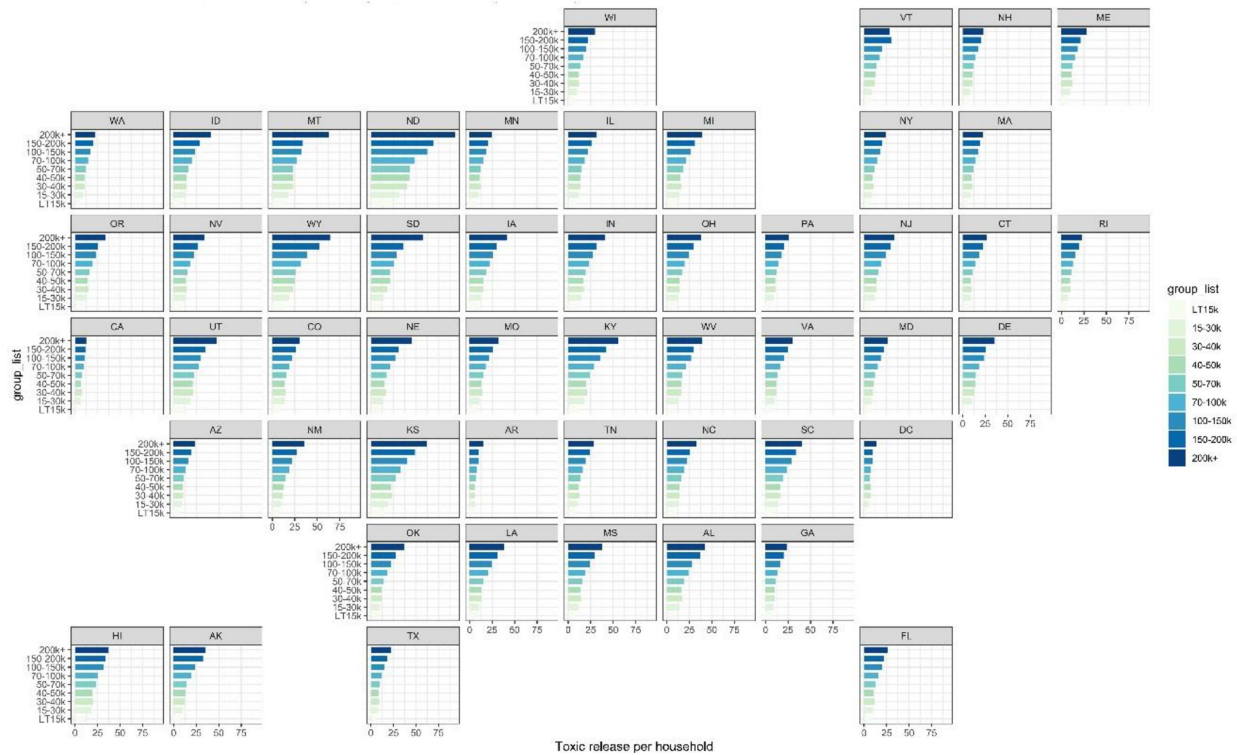


Figure 2.8: TCR embodied in household consumption across states and income groups.

Table 2.3: Per capita personal income (dollars) in US in 2017.

State	Per capita personal income	State	Per capita personal income	State	Per capita personal income
District of Columbia	78,974	North Dakota	52,610	Missouri	45,307
Connecticut	71,699	Rhode Island	52,600	Indiana	45,217
Massachusetts	68,405	Vermont	51,632	Tennessee	45,193
New Jersey	64,964	Delaware	50,738	Georgia	44,865
New York	62,913	Nebraska	50,617	North Carolina	44,376
Maryland	60,714	South Dakota	49,738	Utah	44,142
California	60,581	Wisconsin	49,239	Louisiana	43,903
New Hampshire	58,689	Kansas	48,846	Oklahoma	43,769
Washington	58,400	Oregon	48,719	Arizona	42,566
Alaska	57,295	Florida	48,473	Idaho	42,218
Wyoming	56,421	Texas	48,402	South Carolina	42,178
Virginia	55,582	Iowa	47,629	Arkansas	41,622
Colorado	55,550	Nevada	47,615	Alabama	41,000
Minnesota	54,930	Ohio	46,804	Kentucky	40,874
Illinois	54,247	Maine	46,525	New Mexico	39,727
Hawaii	53,382	Montana	46,138	West Virginia	38,891
Pennsylvania	53,277	Michigan	45,931	Mississippi	36,510

2.4 Discussion and conclusion

From the previous analysis, we gain some new insights regarding the regional distributional patterns of production-based and consumption-based TCR, the outsourcing of embodied TCR, and the unequal distribution by regions and demographic groups. This analysis is useful for the distribution of pollution reduction responsibility between different entities, such as states and demographic groups.

The production of TCR is unequally distributed across regions in the US, mainly concentrating on certain central and southern states and the Great Lakes Region, such as Texas, Indiana, Ohio, etc. The disparities of production-related TCR in each region can be partly explained by the types of main sectors located in each region. The production activities of certain types of products are geographically concentrated and generate large quantities of TRI chemicals. For example, almost all the mining and supportive activities are concentrated in Nevada and Alaska. Non-metallic product manufacturing is mainly focused in Texas, Louisiana, Kansas, and Michigan (Figure 2.2). These states produce products to not only meet their own final demand but also satisfy the final demand in other states. In order to enjoy the benefit from economies of scale, industrial concentration occurs in the US, thus causing the concentration of TCR. Under the production-based principle, the producer is responsible for the TCR from the production of goods and services [20]. In order to reduce the TCR, the states with high releases should implement effective measures. For example, in 2019, 5% of facilities in Arkansas, Louisiana, New Mexico, Oklahoma, and Texas have reported implementing new source reduction activities [60]: a motor vehicle parts manufacturer improved the effectiveness of the producing system by updating the zinc rinse system and reducing zinc waste. Similar to EPA's "good neighbor" plan, which

requires more than 20 states, including Texas, Louisiana and others, to cut ozone pollution from power plants and industrial sources that contribute to ozone pollution in neighboring states [60]. Under the production-based principle, the government could propose plans that control toxic emissions from large TCR producers. However, there are limits to the production-based principle, which cannot distinguish the consumption between export and domestic, thus exacerbating the responsibility of the producer [20].

The consumption-based TCR is less unequally distributed across regions in the US (Figure 2.1). From the previous analysis, we found that the production activities of particular goods and services are geographically concentrated. Thus, in order to fulfill the final demand, states have to import products they don't produce enough from other states, thus triggering the embodied TCR of other states. For example, the main economic sectors in California and New York are tertiary industries such as finance, insurance, real estate, rental, and leasing, professional and business services, and information [61], thus, CA and NY need to import necessary goods like metallic and non-metallic products to fulfill their intrastate demand (Figure 2.1). In the US, local consumption accounted for a small percentage of total consumption-based TCR, while inter-regional transfer between states accounts for a large proportion (Figure 2.3). Therefore, policies under the consumption-based principle, which take inter-regional trade into account, would be more appropriate. Under the consumption-based principle, the consumer is responsible for the TCR from the production of goods and services [20, 21], even if they are imported from other states. The net importers of embodied TCR are mainly the coastal states (Figure 2.1), which means they should shoulder more TCR reduction responsibility from the consumption-based principle than the production-based principle. For example, Texas, Florida, Illinois, New York, New Jersey, Pennsylvania, California, etc. are main consumers (Figure 2.5) of TCR, and their

demand triggered other states' production, such as Texas, Indiana, Ohio, Michigan, Alabama, Illinois, etc. Those consumer states have generated the demand for the exported products and thus caused the production for export to take place, therefore, they should shoulder the responsibility for the TCR reduction. However, consumer responsibility requires political decision-making in other regions. The United States is a federal republic where the federal government has certain control over states' policies, but state governments also have significant authority in their respective states [62]. It is hard to make certain consumer states shoulder TCR responsibility without a unanimous agreement between states. Some researchers brought up new insights of shared responsibility – the responsibility for impacts can be shared between producer and consumer – to make a fair distribution of TCR reduction responsibility [19,21].

Our results showed that the largest portion of embodied TCR is triggered by households' final demand of all income groups and for example, Illinois's households with income 100–150k triggered the largest share of TCR in comparison to other income groups in the state (Figure 2.7). It is interesting to dig into the unequal distribution of TCR footprint between different income groups in each state. The consumption of the top 35% of households (groups of 100k) contributes to more than 50% of total TCR in the U.S. The TCR footprint per household in high-income groups is larger than that in low-income groups (Figure 2.8), which is very intuitive since rich-income groups consume more goods and products than poor-income groups and thus trigger more TCR during the production process of goods and services. Rich states, such as New York, the District of Columbia, and California, have a larger percentage of high-income groups than the national average, thus having a higher TCR footprint than the national one (Figure 2.7). The income inequalities in those states had gaps wider than the national gap [63], with rich people spending more money on housing, transportation and etc., which triggered a

disproportionally larger share of the TCR footprint in the total. Given the high inequality, in order to solve this problem, the government needs to propose policies that increase the bargaining power of low-income groups and promote political participation by all citizens in economic and environmental issues. More interestingly, although rich states have high total household TCR footprints (Figure 2.9), the per household TCR footprints are pretty low (Figure 2.8). Poor states, in contrast, have high per-household TCR footprints (Figure 2.8) despite the low total household TCR footprints. The regional disparities may be partly explained by the household consumption patterns in different states. For poor states, such as Mississippi, Alabama, and Kentucky, the population densities are relatively low [64]. People in these states usually have larger houses, more vehicles, and more other resources per capita, thus, the large number of final demands from these states' households triggers more TCR. Therefore, in order to reduce household TCR, it is necessary to promote green consumption patterns in all households, especially in high-income groups in poor states, since they have a larger TCR reduction potential.

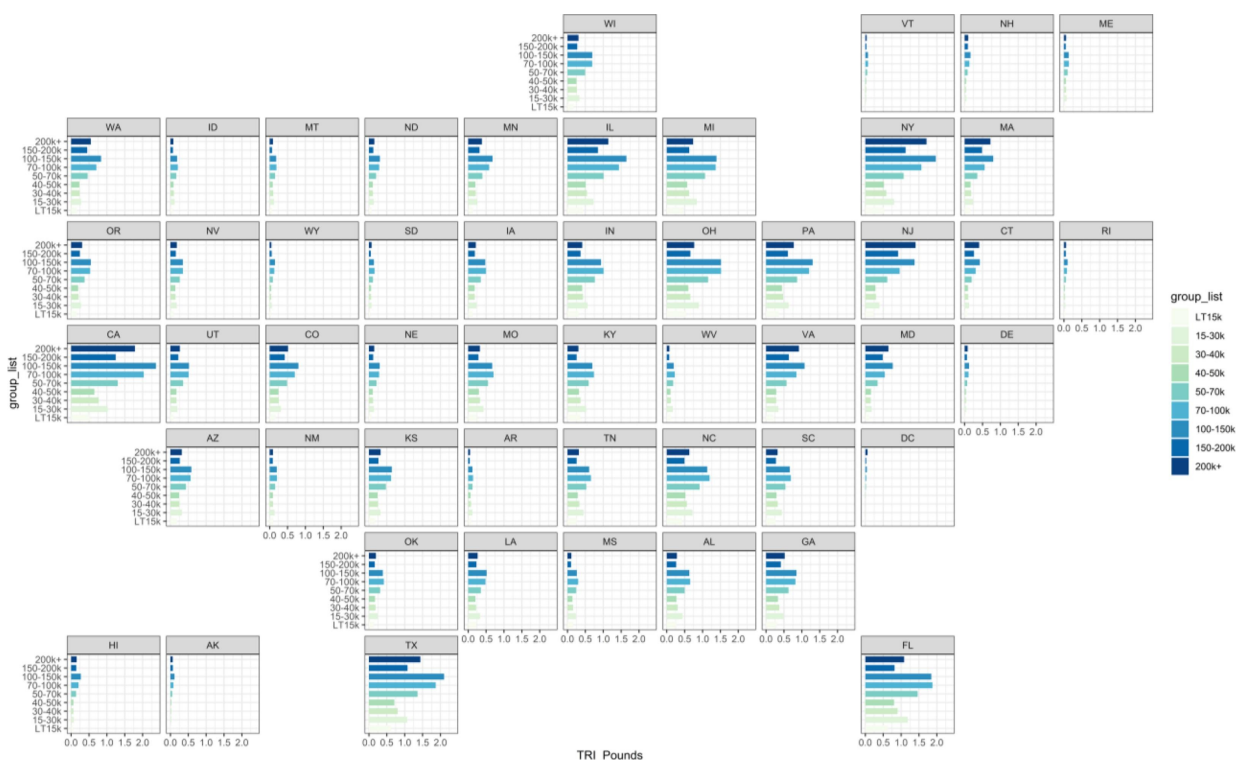


Figure 2.9: Total TCR across income groups in each state to meet household consumption.

In conclusion, this paper filled two important gaps in the literature: The lack of research attention to the interstate flows of consumption-based toxic chemical footprints and the absence of analyzing the disproportionate exposure of consumption-based TCR exposure by income groups across states in the US. This paper provided qualitative information about spatial disparities of TCR in terms of pollution transfer triggered by final consumption across states and income groups. This new set of information can facilitate the government and other stakeholders to propose mitigation policies and actions from both production-based and consumption-based perspectives. Although the input-output model can provide information about consumption-based perspectives, it also has some drawbacks that are important to consider. For example, the input-output model assumes a fixed relationship between inputs and outputs, without accounting for changes over time, which limits its ability to capture dynamic changes. The model also assumes a linear

relationship between inputs and outputs. Moreover, the model treats sectors as homogeneous entities, assuming that all firms within a sector produce identical goods or services [65].

Chapter 3: Multiscale Analysis of the Relationship between Toxic Chemical Hazard Risks and Racial/Ethnic and Socioeconomic Groups

3.1 Introduction

3.1.1 Background

Technological hazards are an inevitable product of technological innovation and human development [66]. The release of acutely toxic chemicals into the air, water, and land is one type of technological hazard, that exerts an impact on climate change [1, 3, 4], ecosystem [67], and human health [2, 6, 7]. The United States of America (USA) Environmental Protection Agency (EPA) defined environmental justice as “the fair treatment and meaningful involvement of all people regardless of race, color, culture, national origin, income, and educational levels with respect to the development, implementation, and enforcement of protective environmental laws, regulations, and policies” [8].

Distributive environmental injustice or environmental inequality may occur if certain racial/ethnic or socioeconomic groups experience a disproportionate share of the negative environmental consequences of industrial activity [10]. There are many screening tools identifying the areas where distributive environmental injustice exists, for example, EPA’s EJScreen (Environmental Justice Screening and Mapping Tool) provides environmental justice indexes to show areas where vulnerable populations

may be disproportionately impacted by pollution. The Council on Environmental Quality proposed Climate and the Economic Justice Screening Tool to identify which communities are identified as disadvantaged based on environmental and socioeconomic indicators. The research on distributive environmental inequality is essential for hazard risk management because it identifies social vulnerability underlying the inequality and thus provides targeting information for risk assessment and mitigation [68].

3.1.2 Literature Review

3.1.2.1 Hazards risk management and social vulnerability

According to the Federal Emergency Management Agency (FEMA), hazard risk management is an ongoing process that involves a set of defined and iterative component parts to exercise control over risks associated with hazards facing a community [69]. The process of hazard risk management involves several stages, including hazard identification, risk assessment, risk analysis, and risk treatment [66]. The identification of risks is a fundamental stage in risk management, which involves the identification of potential sources of risk and the resources that may be exposed to such risks [70]. Hazard risk assessment, on the other hand, investigates the potential threats posed by identified hazards, including their likelihood and consequence [66]. Hazard risk analysis, when combined with vulnerability analysis, can aid in determining the factors that cause risks, prioritizing the ranking of risks, and understanding the rationale behind such rankings. In analyzing hazard risks, social vulnerability must be considered to dedicate appropriate resources and time to measures that offer the greatest benefits during risk treatment. In the context of hazard, disaster, and emergency management research, social vulnerability refers

to the socioeconomic and demographic factors that affect community resilience [71]. Social vulnerability is a byproduct of social inequalities that are influenced by various factors such as income, age, race, culture, and power relationships [71, 72]. Socioeconomic status, including income, education, and employment status, is a significant factor contributing to social vulnerability. For instance, Fothergill and Peek found that the poor are more likely to perceive hazards and less prepared for them [73]. Additionally, Masozera et al. demonstrated that low-income populations were more vulnerable to natural disasters, such as Hurricane Katrina, than their wealthier counterparts [74]. Similar findings have been observed for other vulnerable groups, including children, elders [71, 75], racial and ethnic minorities, and people living in low-quality housing [72, 76, 77]. Racial and ethnic minorities are vulnerable to hazards due to cultural or language barriers that may exclude them from communications and action [74].

3.1.2.2 Distributive environmental justice in the U.S.

The distributive environmental justice literature typically addresses the following 3 W's questions: which vulnerable communities are exposed to environmental hazards disproportionately, where those communities are located, and what are the costs and benefits of risk reduction and management measures [68, 78, 79]. In terms of technological hazards, racial-ethnic and socioeconomic disparities in hazard exposure in the United States are well documented (see Table 3.1). Some researchers investigated whether exposure to hazardous waste differ by race/ethnicity, age and other socioeconomic status using statistical analyses [42, 43, 80]. Regression models are commonly used to evaluate the relationship between environmental hazards and vulnerable groups, including both global and local models. In a global regression model, all data of the

research are used simultaneously to fit a single model and this type of research implicitly assumes a spatially stationary relationship between explanatory variables and dependent variable(s). For example, Wilson et al. [81] and Naidu et al. [82] used standard econometric regression models to assess spatial disparities in the distribution of TRI (Toxics Release Inventory) facilities and exposure to air pollution by race/ethnicity and socioeconomic status. Their results showed that race was significantly correlated to the distribution of environmental burdens. In fact, most global-regression-based research found evidence of racial minorities being disproportionately exposed to pollution [13, 14, 83–89]. In terms of income relevance, this stream of literature found that low-income groups are often disproportionately exposed to pollution [14, 16, 17, 31, 85, 90]. Furthermore, other socioeconomic indicators, for example, home ownership, may also have an effect on the disproportionate distribution of pollution across various race/ethnic and socioeconomic groups [13, 31].

3.1.2.3 Spatial heterogeneity in environmental justice

Due to the unequal distribution of environmental exposure and its correlation with racial-ethnic and socioeconomic indicators across various study areas, researchers have employed spatial analysis techniques to investigate whether environmental inequality remains constant over space [91]. Tools such as EJScreen have revealed significant heterogeneity between areas where vulnerable populations are disproportionately impacted by pollution. It is crucial to understand the spatial heterogeneity of the relationship between environmental hazards and socioeconomic indicators at various spatial scales and locations to comprehend the resultant health outcomes. Accounting for the spatial distribution of environmental inequality enables policymakers to develop interventions

at the local level that target vulnerable groups with the goal of building a health-promoting and equitable environment [92]. Additionally, a comprehensive assessment provides multidimensional perspectives in environmental justice debates [93] and facilitates the mobilization of local stakeholders [92]. To address the spatial stationarity assumption and allow for spatially varying parameters, Brunsdon et al. introduced the Geographically Weighted Regression (GWR) model, an extension of general regression models based on kernel-weighted regression, which enables researchers to capture spatial dependence and spatial heterogeneity [28]. Several researchers have demonstrated the effectiveness of GWR in understanding the spatial variations of relationships between environmental conditions, social characteristics, and health inequalities [29, 91, 94–97]. For instance, Mennis and Jordan employed GWR analysis and demonstrated that the relationships among race, class, employment, urban concentration, and land use with air toxic release density in New Jersey vary significantly over space [29]. In some of these areas, the minority–air-toxic release association is mediated by high poverty rates, in other areas, by the presence of industrial, commercial, and transportation land uses. Lersch and Hart, as well as Chamberlain, analyzed the relationships between the spatial distribution of industrial facilities and community characteristics in Hillsborough County, Florida, and Houston-Sugarland-Baytown Metropolitan Statistical Area, respectively [94, 97]. Gilbert and Chakraborty also utilized GWR to examine potential health risks associated with exposure to hazardous air pollutants related to race/ethnicity and socioeconomic status. The results showed that in nearly all (97%) of the census tracts examined, a higher percentage of Black residents was found to be significantly associated with an increased risk of cancer. Conversely, a majority of the tracts (65%) in the state showed that a higher percentage of Hispanic residents was associated with elevated cancer risk [91]. Jia et al. explored the relationship between racial composition and cancer risks from air toxic exposure [96]. The results showed that airborne toxic

pollutants are not equally distributed in areas with a high concentration of African Americans. This unequal exposure to pollutants is anticipated to result in an increased risk of cancer and other negative health outcomes among minority groups.

Table 3.1: Literature review of distributive environmental inequality

Authors	Study area	Spatial resolution	Study period	Models	Pollution studied	Indicators
Other models						
Bell and Ebisu (2012) [80]	USA	215 census tracts	2000-2006	population-level exposures for different groups	PM2.5 components	race/ethnicity, age, and socioeconomic status
Clark et al. (2014) [42]	USA	207,492 Census Block Group	2006	population-weighted mean NO2 concentrations	NO2	race-ethnicity, poverty status, household income, education status, and age
Charette et al. (2021) [43]	Upstate New York	2,123 block group	2000	Hierarchical clustering analysis	chemical releases	education, wealth, income, race/ethnicity, employment, housing, and land use
Hwang and Kannegalla (2021) [98]	Texas	county and census tract levels	2015	GIS-based spatial analysis	TRI	race, ethnicity, educational attainment, income level
Global models						
Wilson et al. (2012) [81]	Charleston, SC	census tract	2008 TRI data and 2000 census data	linear regression models	chemical releases	race, ethnicity and socioeconomic
Naidu et al. (2012) [82]	Ohio	88 counties	1996-2009	econometric model	RSEI risk score	race, class, home price, wealth

Authors	Study area	Spatial resolution	Study period	Models	Pollution studied	Indicators
Chakraborty et al. (2014) [13]	Greater Houston Metropolitan Statistical Area, Texas	1,066 census tracts	2007-2011	spatial lag models	Pollution Risk	race/ethnicity, socioeconomic status, Gini index, population density
Zou et al. (2014) [14]	USA	52,613 urban census tracts	1999	logistic regression modeling and global spatial autocorrelation	benzene pollution	age, racial, educational attainment and income differences
Zwinkl and Moser (2015) [16]	USA	74,000 census tracts	2010	OLS, spatial lag model, spatial error models	pollution exposure score	race, class, income inequality
Zwinkl et al. (2014) [17]	USA	143,716 census block-group in urbanized areas	2007 RSEI, 2000 demographic variables	econometric model with special fixed effects	industrial air toxic exposure	income, race and ethnicity
Ard (2015) [15]	USA	216,159 census block groups	1995-2004	three-level multilevel model	air toxins	race
Aoyagi and Ogunseitan (2015) [31]	Los Angeles urban area	1,942 census tracts	1989-2002	OLS, spatial lag model	TRI emissions and toxicity risks	socio-demographic factors (% Hispanic, median household income, % owner-occupied housing)
Kravitz-Wirtz et al. (2016) [86]	USA	census blocks	1990-2009	multilevel model	NO ₂ , PM _{2.5} and PM ₁₀	race, sociodemographic and economic characteristics

Authors	Study area	Spatial resolution	Study period	Models	Pollution studied	Indicators
Pope et al. (2016) [88]	Phoenix metropolitan region	census groups for O3, 2,171 for PM10)	2008–2010	hierarchical regression method	O3, PM10	class, age, race and ethnicity
Johnson et al. (2016) [85]	Atlanta Metropolitan Statistical Area	657 census tracts	2006–2011 TRI data and 2000 census data	linear regression model	TRI facilities, air releases	race/ethnicity, education, housing poverty,
Post et al. (2018) [89]	Lubbock County, Texas	204 census block groups	2012	OLS, spatial lag model	hazardous waste	Median income, % minority
Li et al. (2019) [87]	Texas	census tract	2000–2010	multinomial model, OLS	air pollution	Racial, ethnic, and income
Collins and Grineski (2019) [84]	metropolitan Salt Lake City, Utah	219 census tracts	2010	multivariate generalized estimating equation	air pollution	Mormon prevalence, racial and ethnic composition, and socioeconomic status
Local models						
Mennis and Jordan (2008) [29]	New Jersey	1,933 census tracts	2000	GWR	air toxic release	race, class, employment, urban concentration, and land use
Gilbert and Chakraborty (2011) [91]	Florida	3,151 census tracts	1999	OLS and GWR	hazardous air pollutants	race/ethnicity and socioeconomic status

Authors	Study area	Spatial resolution	Study period	Models	Pollution studied	Indicators
Jia et al. (2014) [96]	Memphis/Shelby County, Tennessee	census tract level	2005	GWR	cancer risks from air toxics	racial composition (% African American, population density)
Lersch and Hart (2014) [97]	Hillsborough County, Florida	318 census tracts	2008	GWR	lead exposure and crime	% black, % Hispanic, % with a bachelor's degree
Tu et al. (2016) [99]	Georgia	census tract level	2001	GWR	O ₃ and PM _{2.5}	age, parity, and six other socioeconomic, behavioral, and land use factors
Bakhtsiyarava and Nawrotzki (2017) [100]	USA	2,054 Public Use Microdata Area	2005-2011	GWR	toxin exposure	Males, employed in manufacturing, employed in mining, employed in construction, household income, retiree dep. Ratio, area urban, average precipitation, average max temperature
Chamberlain (2017) [94]	Houston-Sugarland-Baytown Metropolitan Statistical Area	1,070 census tracts	2015	OLS and GWR	toxic releases	Population, non-white, poverty, nativity, manufacturing, home value, entropy

3.1.3 Research objectives

However, the GWR operates on a “single bandwidth” for all the variables and thus it overlooks the inherent spatial variation in variables that explain negative environmental outcomes associated with environmental injustice within and across geographic units. This limitation can be addressed by Multi-scale Geographically Weighted Regression (MGWR), which allows each variable to be calibrated at flexible processing scales. In more detail, the MGWR specifies the processing bandwidth unique to each variable by considering the underlying spatial unit [30]. The uniqueness is ensured by the MGWR’s adaptive Bi-square kernel utility, which removes the effect of observations that fall outside the neighborhood-specified bandwidth and uses the corrected Akaike information criterion (AICc) to select the optimal bandwidth for each explanatory variable. Capturing the spatial heterogeneity of environmental inequality by multiple spatial scales can provide perspectives on policy recommendations for specific geographic units. Besides, eliminating the restriction that all relationships hold at the same spatial scale can minimize over-fitting, reduce bias in the parameter estimates, and mitigate the collinearity due to similar functional transformations [30, 101–103]. AICc criterion evaluates a statistical model by considering two important factors: (1) how well the model fits the data (i.e., the goodness of fit), and (2) how complex the model is (i.e., the number of parameters used in the model). By taking both factors into account, the AICc helps to prevent both overfitting and underfitting [104–106].

To the best of our knowledge, there has been a lack of studies addressing the important question of whether the relationship between negative environmental outcomes and racial, ethnic, or socioeconomic indicators would stay unchanged across multiple spatial scales in the USA. Taking Texas as an illustrative example, this study addressed this gap by investigating the spatial

non-stationarity in the associations between toxic chemical hazard risk and community characteristics such as race/ethnicity and socioeconomic status at the census block group level for 2017 using the MGWR. Data on industrial pollution risks are from EPA's Risk Screening Environmental Indicators. Data on income and demographic variables are from the American Community Survey (ACS). In addition to enriching the existing literature on distributive environmental justice and hazard risk management, findings from this paper will provide geospatially explicit information about the relationship between race/socioeconomic status and toxic risk, which can serve as guidance for intervening in the processes that drive socio-spatial disparity in toxic chemical hazard risks in Texas. The analysis approach of this paper can also be applied to other regions that face similar environmental injustice problems and need more robust hazard risk management.

3.2 Materials and Methods

3.2.1 Study area

Texas is the second-largest state in the USA in terms of area and population. It covers approximately 695,662 km² of land area with 5,265 census tracts and 15,811 census block groups according to Guide to 2010 State and Local Census Geography. According to the EPA's 2019 TRI Factsheet, Texas accounts for about 8.4 percent of the total number of toxics release inventory (TRI) facilities and about 13% of the total production-related waste managed in the USA [55]. For example, hundreds of petrochemical manufacturing and distribution facilities are located along the Gulf of Mexico and major metropolitan areas such as Dallas-Fort Worth-Arlington and San Antonio New Braunfels, leading to considerable risks of hazardous material releases in the operation or transportation processes, which threaten public safety and human health and damage

the natural and manmade environment.

In terms of spatial distribution of population, racial segregation has existed in Texas since the 1820s, and was developed as a method of group control after the Civil War [107]. Institutionalized segregation became legalized in places with large number of Black population and then was extended by the late 19th century [107]. In the 1930s, the Federal Home Owner's Loan Corporation (HOLC) drew maps for nearly 250 cities to document the relative riskiness of lending across neighborhoods, drawing the lowest-rated neighborhoods with a vast majority of African American residents in red (the so-called practice of "redlining") [108, 109]. Cities in Texas, including Amarillo, Austin, Beaumont, Dallas, El Paso, Fort Worth, Galveston, Houston, Port Arthur, San Antonio, and Waco were included in this program. The economic and racial segregation created by "redlining" persists in many cities [110]. (Mitchell and Franco, 2018). Until recently, racial inequality still exists in Texas. In 2018, about one in five Black (20%) and Hispanic (21%) Texans were living in poverty, as were 12% of Asian-American and 9% of White Texans. The white unemployment rate in Texas was 3.2% while being 5.3% for Black residents [111].

The above two sets of significant features of Texas make it an ideal study area for investigating the geospatial patterns of the association between toxic chemical hazard risks and community characteristics at the census block group level, for which the MGWR approach is indispensable.

3.2.2 Data and Method

The toxic chemical risk data can be downloaded directly from EPA's Risk-Screening Environmental Indicators (RSEI) model. The RSEI is a screening-level model which incorporates Toxics Release Inventory (TRI) data with measures of human exposure and toxicity and assesses the potential

impact of industrial chemical releases from pounds-based, hazard-based, and risk-related perspectives [112]. “TRI Pounds” (pounds-based) reflect only the number of pounds released or transferred that are reported to TRI and are available for all releases and transfers. Hazard-based results (“RSEI Hazard”, or “toxicity-weighted pounds”) are calculated by multiplying the pounds released by the appropriate chemical-specific toxicity weight (the toxicity weight also depends on the exposure pathway), is a result that accounts for the size of the release and the chemical’s toxicity. An RSEI Risk-Screening Score is calculated as toxicity weight multiplied by the exposed population multiplied by the estimated dose. This research used the RSEI score in 2017 as a dependent variable because it captures the toxic chemical releases, the chemical’s toxicity, and human exposure. Because the RSEI score is skewed to the left – many block groups have little or no industrial air pollution – we add 1 (so as to include observations with zero values) and take its logarithm.

According to Galobardes [113], socioeconomic status includes the following dimensions: education, housing, income, and occupation-based measures. Thus, the socioeconomic variables we selected include median household income (INCOME), the square of median household income (INCOME2), the percentage of owner-occupied housing (OWNER), the percentage of vacant housing (VACANT), the percentage of adults without high school diploma (NHSD), the percentage of unemployment (UNEMP) and Gini index (GINI). For racial-ethnic variables, we selected the percentage of the Black population (BLACK), the percentage of the Asian population (ASIAN), the percentage of other races (OTHERS), and the percentage of Hispanic or Latino (HISPANIC). The variable code names and descriptions are listed in Table 3.2

In regression models that examine the relationship between income variables and environmental quality, two types of endogeneity problems may arise. The first type is the simultaneity bias,

which arises due to the reverse causality between Gross Domestic Product (GDP) and environmental degradation [114]. When economic activity increases, it can lead to an increase in pollution levels. Simultaneously, pollution can cause harm to people's health, which in turn can reduce GDP. Furthermore, GDP and pollution may be jointly produced in the production process, leading to a situation where GDP and pollution are simultaneously determined. The second type of endogeneity problem is due to omitted variable bias. This occurs when a relevant variable is left out of the regression model, and this variable is correlated with both the dependent and independent variables. In the context of environmental quality and income, it is possible for other factors, such as government policies, cultural factors, and geographical factors, to impact both of these variables. The exclusion of these variables from the model may result in leading to biased estimates. To address this concern, this study employs a spatial econometric method to investigate the influence of racial/ethnic factors and socioeconomic status on toxic risk. This approach allows for an examination of the spatial spillover effect between regions, and to some extent, mitigates endogeneity issues that could arise due to unobserved variables [115].

Table 3.2: Variable code names and description.

Variable code name	Description
Log_RSEI_score	Natural logarithm of (1+ RSEI score)
Intercept	Intercept
POP	Population density per km2
BLACK	The percentage of Black or African American population
ASIAN	The percentage of Asian population
OTHERS	The percentage of other races
HISPANIC	The percentage of Hispanic or Latino
INCOME	Median household income in the past 12 months (in 2019 inflation-adjusted dollars)
INCOME2	The square of median household income
OWNER	The percentage of owner-occupied housing
VACANT	The percentage of vacant housing
NHSD	The percentage of adults without high school diploma (for population 25 years and over)
UNEMP	The percentage of unemployment, which is measured by the percent of the labor force that are unemployed.
GINI	Gini index, a summary measure of income inequality, ranging from 0 (perfect equality) to 1 (perfect inequality)

All the above demographic data were collected from American Community Survey (2019 ACS 5-year Estimates) [116]. For data cleaning, we removed 46 block groups with 0 population, 604 block groups without median house income data, 1 block group without education data, 1 block group without unemployment rate, and 4 block groups without Gini index. After cleaning, there remain 15,154 block groups and the statistical summary of the variables can be found in Table 3.3

3.2.3 Global regression model: The Ordinary Least Squares (OLS) model

The equation for the OLS model can be written as:

$$y_i = \beta_0 + \beta x_i + \varepsilon_i \quad (3.1)$$

where y_i is the dependent variable log_RSEI_score in block group i , β_0 is the intercept; x_i is the vector of selected explanatory variables, β is the vector of regression coefficients, ε_i is a random error term.

3.2.4 Local regression model: The MGWR model

The equation for the MGWR model can be written as:

$$y_i = \sum_j \beta_{ij}(u_i, v_i) X_{ij} + \varepsilon_i, i = 1, 2, \dots, N \quad (3.2)$$

where y_i is the dependent variable log_RSEI_score in block group i , $\beta_{ij}(u_i, v_i) X_{ij}$ refers to the differential bandwidth at the feature space; X_{ij} is the value of the j - th regression

Table 3.3: Statistical Summary of the variables (N = 15,154 block groups)

	min	max	range	median	mean	SE.mean	std.dev	coef.var
Dependent variable								
Log_RSEI_score	0.00	17.20	17.20	0.00	0.24	0.01	1.38	5.75
Independent variables								
<i>Race/Ethnicity</i>								
BLACK	0.00	100.00	100.00	4.40	11.54	0.14	17.20	1.49
ASIAN	0.00	87.51	87.51	0.32	3.73	0.06	7.69	2.06
OTHERS	0.00	90.17	90.17	5.93	9.07	0.08	10.11	1.12
HISPANIC	0.00	100.00	100.00	29.98	38.79	0.24	30.09	0.78
<i>Income</i>								
INCOME	4484	250000	245516	56563	65600	302	37186	0.57
GINI	0.10	0.84	0.73	0.39	0.39	0.00	0.08	0.21
<i>Housing</i>								
OWNER	0.00	100.00	100.00	69.14	63.43	0.21	26.41	0.42
VACANT	0.00	92.03	92.03	9.15	11.52	0.09	10.91	0.95
<i>Education</i>								
NHSD	0.00	86.51	86.51	13.63	18.03	0.13	15.63	0.87
<i>Occupation</i>								
UNEMP	0.00	58.74	58.74	3.99	5.26	0.04	5.28	1.00
<i>Other variables</i>								
POP	0.06	27721.04	27720.99	1168.11	1558.19	14.76	1816.63	1.17

SE.mean means standard error of the mean; std.dev means standard deviation; coef.var means the coefficient of variation.

parameter; ε_i refers to the regression error.

To get the bandwidth that represents the optimum number of nearest neighbors in the GWR, the adaptive kernel is used to find the optimum bandwidth automatically using the chosen statistical criteria [101,102]. By allowing multiple bandwidths in the MGWR, the model accounts for an optimal number of neighbors for each parameter estimate, thus allowing better predictions for the response variables, theoretically. The MGWR recasts the GWR as a generalized additive model (GAM) and uses a back fitting algorithm, an iterative process of calibrating a series of GWR models based on the models' partial residuals until the MGWR model converges to a solution [101]. The calculation is based on iterations of optimum bandwidth for each parameter estimate.

There are five specifications for five models according to different combinations of explanatory variables. Specification 1 (Model 1) only includes population density, race and ethnicity; Specification 2 (Model 2) includes population density, race, ethnicity, income and Gini; Specification 3 (Model 3) includes population density, race, ethnicity and socioeconomic indicators except for income; Specification 4 (Model 4) includes all indicators (exclude NHSD after collinearity test, see Appendix A); Specification 5 (Model 5) includes all indicators (exclude NHSD after collinearity test, see Appendix A) and income square.

3.3 Results and Discussion

Results from all MGWR models are firstly summarized in order to provide the model fit, significance, and bandwidth of all models in one table (Table 3.4). Results from OLS models are also provided in table A.1, which include estimated parameters, model fit, and significance,

as a comparison to spatial nonstationary results of MGWR models. More detailed statistical summaries and maps of parameter estimates for all MGWR models are presented in Appendix. For model performance, Local models have higher R^2 values, lower Akaike information criterion (AIC), and corrected Akaike information criterion (AICc), than global models. For example, the R^2 in local models are more than 0.2 in all five specifications, while being less than 0.02 in global models. These sharp contrasts indicate very limited explanatory power of the OLS models. The values of AIC in local models are around 50000 for all specifications, while the values in global models are more than 52600. It means that the local models outperform global models. Except for poor model fit for the OLS model, Moran's I analysis and Monto Carlo simulation (Appendix) showed that the RSEI score values are strongly clustered, which means that a global model may cause bias in results due to the omission of spatial characteristics. The high spatial concentration of parameter estimates shown in Figure 3.2 - 3.4 also highlighted the preference for local models.

Table 3.4: The MGWR results for the spatial variability of parameters using adaptive bandwidth

	Model 1			Model 2			Model 3			Model 4			Model 5		
	Parameter mean	STD	Bandwidth	Parameter mean	STD	Bandwidth	Parameter mean	STD	Bandwidth	Parameter mean	STD	Bandwidth	Parameter mean	STD	Bandwidth
Intercept	-0.034**	0.033	7452	-0.044*	0.239	43	-0.049*	0.241	43	-0.040**	0.038	7452	-0.052*	0.241	43
BLACK	0.039***	0.001	15153	0.048***	0.001	15153	0.036***	0.001	15151	0.031***	0.001	15151	0.038***	0.000	15151
ASIAN	0.030***	0.001	15153	0.035***	0.001	15153	0.032***	0.001	15151	0.028***	0.001	15151	0.033***	0.001	15151
OTHERS	0.007	0.009	12793	0.005	0.002	15153	0.003	0.002	15151	0.005	0.009	12793	0.004	0.003	14632
HISPANIC	0.096*	0.101	1284	0.067*	0.089	1546	0.054*	0.079	1572	0.093*	0.103	1368	0.069*	0.081	1799
POP	-0.256*	0.459	43	-0.186*	0.158	1043	-0.192*	0.159	1067	-0.266*	0.459	43	-0.193*	0.160	1043
INCOME				-0.010**	0.031	9727				0.006	0.020	11082	0.004	0.035	9890
INCOME2													0.005	0.001	15151
GINI				0.008	0.016	11665				-0.004	0.012	11660	0.004	0.015	11660
OWNER							-0.030	0.001	15151	-0.028	0.015	8461	-0.033	0.002	15151
VACANT							0.006**	0.043	2283	0.001**	0.030	3160	0.006**	0.045	2269
NHSD							0.013**	0.028	4062						
UNEMP							0.006	0.006	12053	0.007	0.006	12153	0.006**	0.009	10064

n	15154	15154	15154	15154
R2	0.218	0.202	0.205	0.22
AIC	41424.111	41836.562	41804.017	41421.11
AICc	41586.919	42016.237	41990.762	41993.267
BIC	49585.467	50394.597	50521.595	49749.02

Note: "****" means that variables are significant at the global (state) scale; "***" means that variables are significant at the sub-state regional scale; "**" means that variables are significant at the local scale. Significance in this table means less than 5%.

Among the MGWR models, Model 4 has the highest R^2 values and lowest AIC, indicating that it is the most parsimonious model. Local R^2 values in Model 4, also in other models, showed local variations across Texas. Western, southeastern and central Texas have higher R^2 than other parts of the state (see Figure 3.1), indicating decent prediction power of the models in these areas. On the contrary, the local R^2 values were low in northern Texas, indicating poor prediction power of the models across these areas. Calibrating a MGWR model produces a vector of optimal bandwidths (“Bandwidth” in Table 3.4, indicating the total number of nearest neighbors) that describe the spatial scale at which each process in the model varies [117]. According to the optimal bandwidths of each variable, the relationships between dependent and independent variables for all the models can be explained at the global (for this paper, the state scale), regional (sub-state), and local scales. A global scale means that the bandwidth is large enough to include most block groups in the state when process the relationships between the response variable and explanation covariates for each location, and thus the corresponding results are statistically consistent across the entire study area. In all models, the parameter estimates for BLACK and ASIAN are significant at the global scale with bandwidths indicating that most block groups are included in each local subset. In model 2, the parameter estimates for INCOME are significant at the regional scale with bandwidths implying the relevance of several thousand nearest neighboring block groups. Similarly, the parameter estimates for VACANT in models 3, 4 and 5, for NHSD in model 3, as well as for UNEMP in model 5, are also significant at the regional scale. In all models, the parameter estimates for HISPANIC and POP are significant at the local scale, meaning their relationships with the dependent variable vary locally and the spatial heterogeneity is large. The parameter estimates for OTHERS, INCOME2, and GINI are not significant in any models.

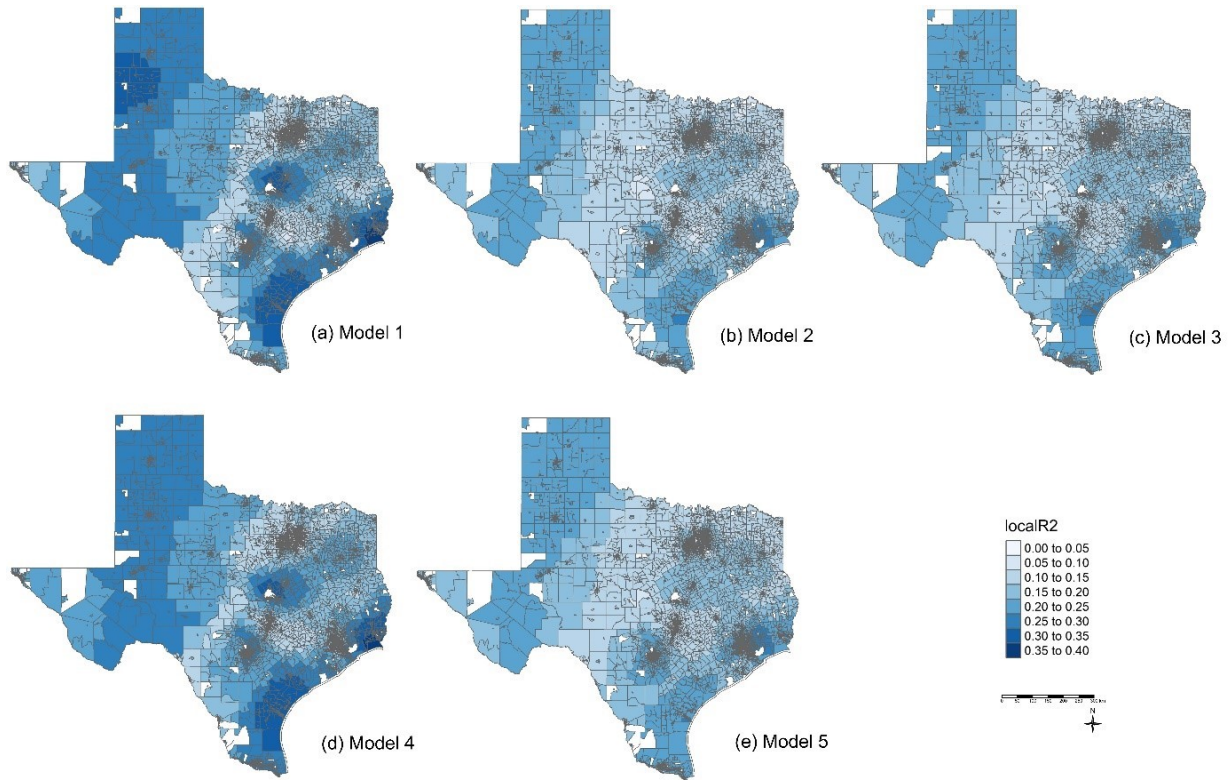


Figure 3.1: Maps of Local R^2 for all MGWR models.
Note: Blocks without color have no data.

In order to assess the presence of local collinearity in spatial regression models, it is necessary to analyze the local condition numbers for each model, as presented in Table 3.5. The local condition number is a metric that measures the degree of collinearity or multicollinearity among the independent variables in a spatial regression model. Specifically, it is computed as the ratio of the largest eigenvalue to the smallest eigenvalue of the design matrix for each observation. A high local condition number suggests that the independent variables are highly correlated in a specific location, which can result in unstable estimates of the regression coefficients. A commonly used rule of thumb for the local condition number is that values greater than 30 indicate that the collinearity between independent variables is becoming problematic and may result in unstable coefficients [117]. From Table 3.5, the summary shows that the local condition

numbers do not exceed the rule-of-thumb 30, meaning there is no local multicollinearity.

Table 3.5: Statistical summary of the local condition numbers of each model specification

	Model 1	Model 2	Model 3	Model 4	Model 5
Min.	1.264	1.457	1.813	1.87	2.272
1st Qu.	1.543	1.738	2.426	2.151	3.951
Median	1.709	1.899	2.589	2.314	6.224
Mean	2.075	2.609	2.788	2.474	5.719
3rd Qu.	2.239	2.201	2.97	2.55	7.003
Max.	4.929	22.206	5.201	4.915	8.906

Although Table 3.4 provides the average parameters for each explanatory variable, regional variations of relationships can be hidden for parameter estimates at the regional and local scales. Further analysis needs to be done to determine the relationships between racial/ethnic, socio-economic indicators, and toxic risk, and how the relationships vary over space, as presented in Figure 3.2 - 3.4. In Figure 3.2 - 3.4, we used sequential color scales to represent parameter estimate values, with warm colors (mainly in red) representing significantly positive, cool colors (mainly in blue) representing significantly negative, and blank representing insignificant parameters or no data.

Figure 3.2 shows that racial minorities, including BLACK and ASIAN, have significant positive associations with toxic risk at the global (state) scale. These positive associations suggest that densely concentrated racial minority groups are more likely to be exposed to toxic risk in Texas. The coefficients of the BLACK under Model 4 range from 0.028 to 0.036 with a mean value of 0.031. It means that holding all other variables constant, with every one percentage-point increase in the share of the Black population within the block group, the toxic risk would increase by approximately 3.15%. For ASIAN, holding all other variables constant, with every percentage-point increase in the share of Asian, the toxic risk will increase by 2.84% (since the

parameter mean is 0.028 in Model 4).

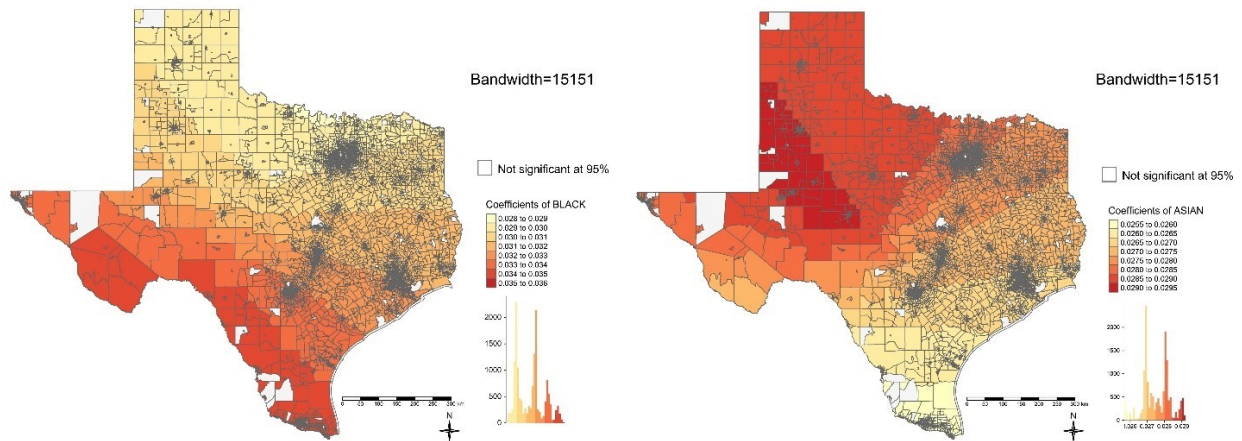


Figure 3.2: Local coefficients of racial disparity (BLACK and ASIAN) in toxic chemical risk (from Model 4).

Note: Blocks without color have insignificant coefficients or no data.

Figure 3.2 reveals an important result that even though we used local modeling techniques to analyze the spatially varying relationships between variables, their relationships are consistent over the entire State of Texas. This result resonates with the global model result (see Table A.1) and those in environmental inequality literature using global models [81, 82, 87, 89], indicating that the minority racial groups are more likely to experience the negative environmental outcomes. Three complementary explanations have been put forward to analyze these observed environmental inequalities: rational choice, sociopolitical, and racial discrimination [118]. Areas with predominantly minority communities often have cheaper land values, which attract industry actors to establish polluting facilities because it would be an economically rational choice for them. Sociopolitical reasons take minority communities as the “path of least resistance” because they have fewer resources and political clout to oppose the siting of unwanted facilities, resulting in industries located in those communities [119]. Racial discrimination indicates that minority communities

are intentionally targeted for environmental hazards because of beliefs in racial superiority and/or an intention to protect whites from environmental harm [118, 120].

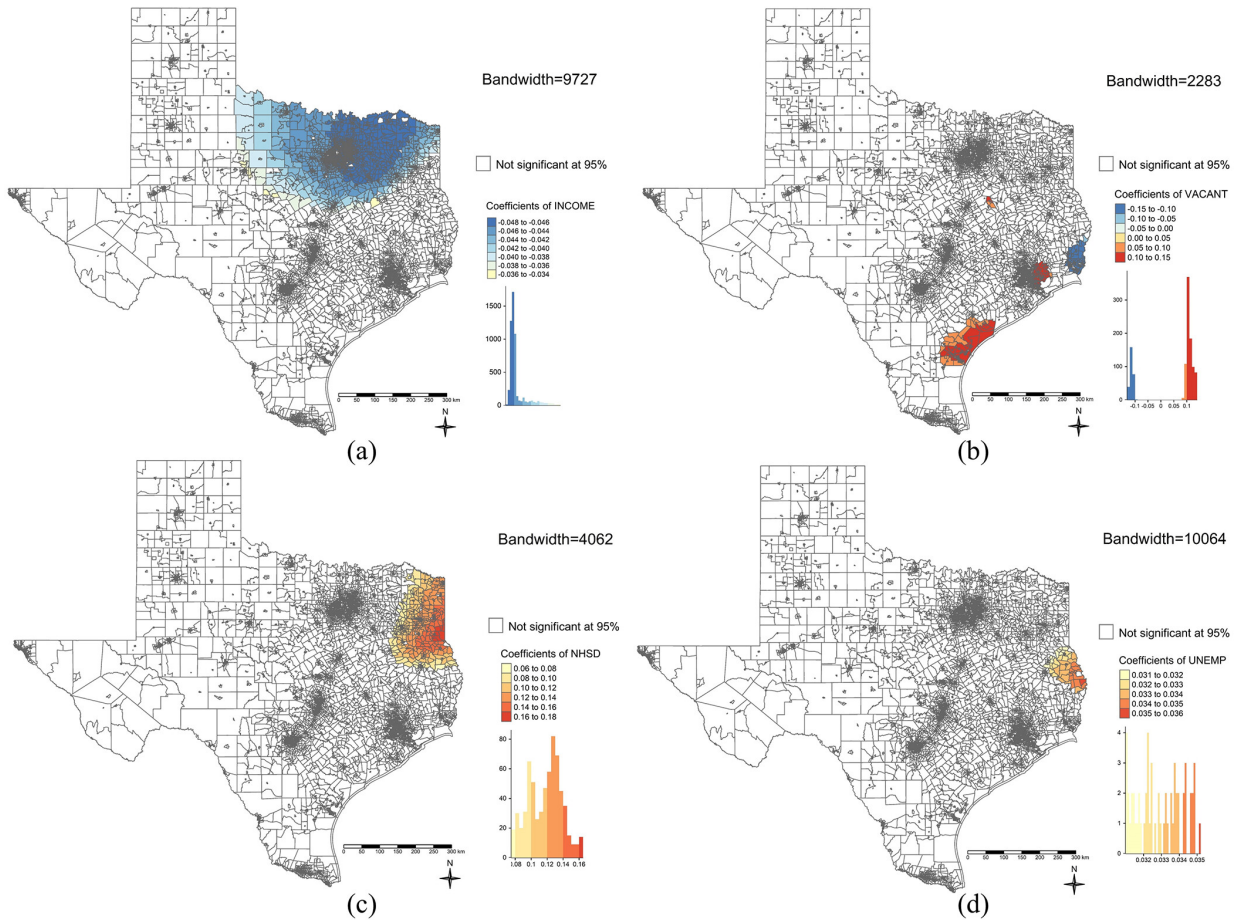


Figure 3.3: Local coefficients of (a) INCOME (from Model 2), (b) VACANT (from Model 3), (c) NHSD (from Model 3), and (d) UNEMP (from Model 5) in toxic chemical risk.

Note: Blocks without color have insignificant coefficients or no data.

Unlike racial minorities, Figure 3.3 presents that the statistical associations between toxic risk and socio-economic variables are not uniform across Texas, showing regional patterns of spatial heterogeneity. From Figure 3.3 it can be seen that INCOME has a significant negative association with toxic risk around Dallas-Fort Worth-Arlington Metropolitan Statistical Area. It means that people with low household income are more geospatially concentrated in these areas with high exposure to toxic chemical risk. In more detail, low-income communities living in

block groups around the urbanized area in Dallas [121] experience higher toxic risk than their counterparts in rural areas nearby (Figure 3.3). Low-income neighborhoods and communities, similar to racial minorities, are also disproportionately targeted as pollution facilities sites by industries following the path of least resistance. Although previous studies like Chakraborty et al. [13] and Zou et al. [14] using global models showed the same relationships, our result recommends the need for regional-level research in environmental inequality to determine the zooming-in locations with the disproportionate distribution of toxic risk.

Other socio-economic indicators tend to have positive relationships with toxic risk. For example, there are positive relationships between NHSD or UNEMP and toxic risk in the eastern area of Texas. Block groups with lower education attainment (in our study, high percentages of adults without high school diplomas) are more likely to experience toxic risk, especially in areas around the northeastern area of Texas. High unemployment rates in block groups are also associated with a high level of toxic risk. VACANT has a regional pattern of spatial heterogeneity. There is a negative relationship between VACANT and toxic risk around the city of Beaumont, while a positive relationship is around the city of Corpus Christi and Houston. This result demonstrates how particular explanatory variables can contribute diversely to toxic risk in different parts of Texas, thus emphasizing the limitation of the global model and the preference for using local models like MGWR to take local processes into consideration.

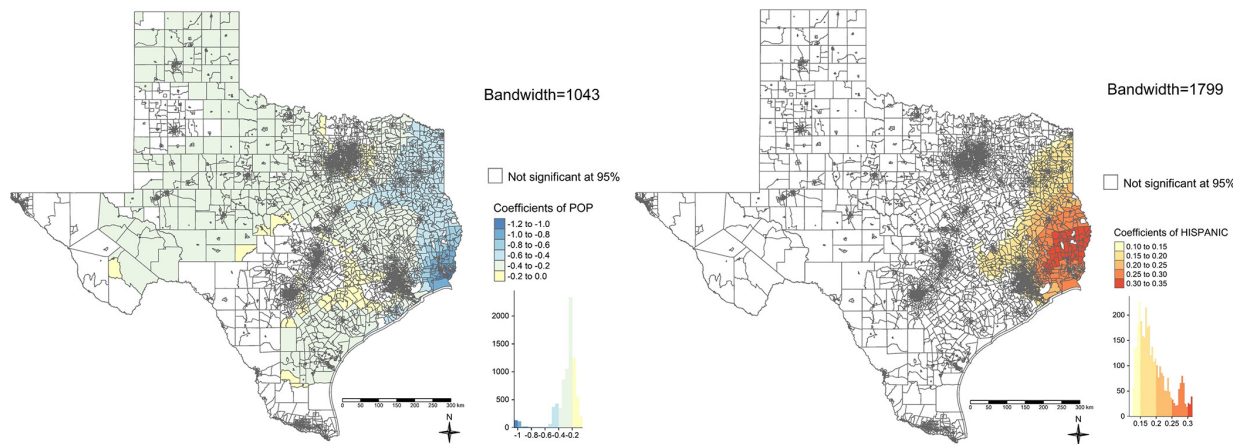


Figure 3.4: Local coefficients of POP and HISPANIC (from Model 5) in toxic chemical risk.
 Note: Blocks without color have insignificant coefficients or no data.

Figure 3.4 indicates that the parameter estimates surfaces for population density (POP) tend to exhibit local patterns of spatial heterogeneity. Population density has a negative relationship with toxic risk across northern and eastern Texas; while for western, central, and southern Texas, the relationship is not significant. The absolute values of parameter estimates in eastern Texas are larger (up to -1.2 to -1) than in other regions like northern Texas (less than -0.4). Population density is a commonly used control variable in environmental justice analysis [13, 91]. Previous research about the effects of population density on negative environmental outcomes, for example, air and water pollutants, toxic materials, and greenhouse gases, have generated contradictory results. For example, Borck and Schrauth [122], Naidu et al. [82] and Chakraborty et al. [13] showed that higher population density tends to associate with more air pollution, environmental risk, and health risk. However, Li et al. [87] found that excess emissions are negatively correlated with population density. Gilbert and Chakraborty [91] indicated a significantly positive influence of population density on cancer risk in Tampa, and an opposite effect in Jacksonville, meaning that the relationships are different in different areas. The unequal distribution of the relationship

across Texas in our study, together with the contradictory results from previous studies, emphasizes the need to conduct detailed spatial analysis in determining the relationship between population density and toxic risk across geographical locations.

According to the United States Census Bureau, Hispanic or Latino are officially the nation's largest minority group, accounting for 18.4% of the USA's total population in 2019, whereas this share is up to 39.3% in Texas. Unlike racial minorities like Black and Asian people, which experience consistent relationships with the toxic risk over the entire State of Texas, the relationships between the ethnic minorities Hispanic or Latino and the toxic risk are not stationary across Texas. Instead, the pattern of the relationship varies locally and is only significant in certain areas, such as eastern Texas. The percentage of Hispanic or Latino has a positive relationship with toxic risk in eastern Texas, with the parameter ranging from 0.1 to 0.35. The estimates around the major town of Beaumont are larger than other surrounding areas. This may be partly because Hispanics are often segregated or constrained to rent homes in neighborhoods that are disproportionately exposed to multiple types of environmental externalities and pollution [13, 123].

3.4 Conclusions and policy implementations

To summarize, this study analyzed the spatial non-stationarity in the associations between toxic chemical risk and community characteristics including socioeconomic status, race/ethnicity, and population density across census block groups in Texas, USA, in 2017. It employed the Multi-scale Geographically Weighted Regression (MGWR) technique to investigate whether the direction and extent of the associations vary spatially at different geographical scales in the state. The results showed that vulnerable racial minorities, measured as the percentages of Black or

Asian population in a census block group, have significant positive relationships with toxic risk and the relationships keep stationary across the state, which means that racial minorities are exposed to high toxic chemical risk wherever they locate in Texas. Unlike racial minorities, the statistical associations between toxic risk and socio-economic variables are not stationary across the state, showing regional patterns of spatial variation. Income has a significant negative association with toxic risk around Dallas-Fort Worth-Arlington Metropolitan Statistical Area, meaning that low-income communities are more geospatially concentrated in these areas with high exposure to toxic chemical risk. The share of the population without high school diplomas and the unemployment rate both have positive relationships with toxic risk in the eastern area of Texas. This means that people with low socioeconomic status are more geospatially concentrated in eastern areas with high exposure to toxic chemical risk. There is a negative relationship between the percentage of vacant housing and toxic risk around the city of Beaumont, while this relationship is positive around the city of Corpus Christi and Houston. The relationship between toxic risk and the population share of ethnic minority Hispanic/Latino varies locally and is only significantly positive in eastern areas of Texas. Population density has a locally varying relationship with toxic risk and the relationship is significantly negative in northern and eastern Texas.

The findings of this study may provide evidence that supports the environmental justice movement [91]. The movement was started by people of color to address the inequity in the distribution of environmental protection, including exposure to toxic chemical hazards by various communities. Vulnerable minorities, especially African Americans, Asians, and Hispanics/Latinos were disproportionately affected by toxic chemical hazards in Texas, the riskiest state in the USA according to the rank of the RSEI risk-screening score in 2019. Our findings further highlight the

importance of identifying the spatial patterns for addressing environmental inequality in Texas. This more spatially explicit analysis might be extended to other states with high-level of toxic chemical risk.

The results of this study may provide useful information for identifying hotspots with disproportionate toxic risk and thus facilitating better-targeted hazard risk management. In other words, with the detailed zooming-in information on the distribution of disproportionate toxic risk across geographical locations and diverse communities, the policymakers such as the Texas Environmental Agency and Texas Department of State Health Services may better identify the combinations of the most vulnerable locations and communities for targeted monitoring and mitigation, thus more effectively reducing the disproportionate impacts of toxic chemical hazards on most vulnerable communities and improving environmental health and well-being of all residents. For example, our finding indicates that racial minorities are exposed to highly toxic chemical risk wherever they locate in Texas, thus, it might be more useful to implement a state-wide policy in order to manage racial injustice regarding toxic exposure. On the other hand, a more local-level policy would be more useful when it comes to managing the toxic exposure of income inequality. Although the EJScreen tool can also provide highlighted areas with higher environmental injustice, it lacks information on the statistical relationships between toxic chemical hazards and racial/ethnic and socioeconomic indicators. This paper provided information on how the direction and extent of the associations between hazards and socioeconomic indicators vary spatially at different geographical scales, which may provide policymakers with evidence to tackle real-world situations at various scales, ranging from census-block neighborhoods to the state level.

Furthermore, the geographically specific findings of this study may imply that a broader constellation of actors at the county, town, and village levels needs to be involved in environmental

justice planning or engagement, in addition to a few agencies at the State level or in big cities. The detailed presentation and reasoning of the spatial patterns for environmental inequality provided science-based information and academic support that may empower the community stakeholders in their push to achieve environmental justice at the local level. This spatial-explicit information might help emergency managers and urban planners to identify which part of their community is more vulnerable to technological hazards, and further help them to better prepare for, respond to, and recover from actual damage and casualties resulting from the hazardous substance events.

The MGWR model exhibits several limitations that require further investigation. Notably, the lack of interpretability represents a primary challenge for this model owing to the different metrics of bandwidth utilized [102]. Specifically, the inclusion of certain Minkowski metrics, such as the timelike dimension, poses difficulties in the intuitive interpretation of estimated bandwidths regarding spatial scale using standard concepts like distance, nearest neighbors, or travel time. MGWR performed multiple regression models at different spatial scales, which can be computationally intensive and time-consuming, especially when dealing with large datasets. This limitation presents a notable impediment when researchers require multiple iterations of model simulations or assessments of parameter sensitivity [124]. Despite the aforementioned drawbacks, the MGWR model integrates various spatial scales, encompassing global, regional, and local scales, within a single computational framework. As such, it remains a valuable tool for elucidating spatial dynamics relevant to the practice of environmental movement.

Chapter 4: Is pollution inequality improved or worsen over time? Spatiotemporal heterogeneity of the impact of racial disparities and socioeconomic development on toxic risk

4.1 Introduction

“Empirical research is now catching up with the reality: that America is segregated and so is pollution” [125]. Racial segregation in the United States is a pervasive and persistent issue that has shaped American society since its founding [126, 127]. It refers to the systematic separation and exclusion of individuals from different racial and ethnic backgrounds in various aspects of social life, including housing, education, employment, and public facilities. This discriminatory practice was institutionalized through a series of laws, policies, and customs that were designed to reinforce racial hierarchies and maintain white supremacy. The origins of racial segregation in the United States can be traced back to the colonial period and the dispossession of Native Americans and the enslavement of Africans [127]. The rise of Jim Crow laws in the late 19th and early 20th centuries formalized segregationist practices, including the establishment of separate but unequal public schools, transportation, and other public accommodations [128]. Redlining is another significant aspect of racial segregation in the United States that played a role in exacerbating racial disparities in housing and wealth. In the 1930s, the Home Owners’

Loan Corporation (HOLC), a federal agency, evaluated neighborhoods in approximately 250 American cities using a four-grade scale, ranging from "Best" to "Hazardous" (grades A to D), which were then depicted on a series of maps. Neighborhoods inhabited by African American, Asian American, Latinx, or immigrant populations were significantly more likely to receive unfavorable grades of D or C [129]. The discriminatory practices of redlining have had such far-reaching effects that their impact continues to persist to this day. Numerous studies have demonstrated that specific racial/ethnic or socioeconomic groups encounter an inequitable burden of the negative environmental consequences of industrial activity [14, 24, 82, 83, 85, 98]. For example, ethnic minorities, such as African American, Asian American, and Hispanic groups, encounter a disproportionate exposure to pollution [13, 14, 84]. Also, people of low socioeconomic status, such as low-income groups and low-educated groups, are often disproportionately exposed to pollution [17, 31, 85, 90].

However, much of the research has focused on cross-sectional data and lacked temporal analysis. In recent decades, there have been substantial reductions in various pollutants, such as Nitrogen Dioxide, PM_{2.5}, PM₁₀ (particulate matter with aerodynamic diameter $\leq 10\mu\text{m}$), Carbon Monoxide, Lead, and others, as reported by the United States Environmental Protection Agency [130]. As depicted in Figure 1.5, the RSEI score and the corresponding pounds released exhibited an overall declining temporal trend from 2007 to 2019. The total RSEI score from 2007 to 2019 demonstrated a decline of approximately 65%, mostly attributed to air releases. Income inequality, which has been a long-lasting problem in the US, has been widening over decades. As reported by the Center on Budget and Policy Priorities, income inequality in Texas has shown a consistent deterioration trend from the 1970s to the mid-2000s. Notably, from the late 1990s to the mid-2000s, the average income of the poorest households declined by 10%,

while the average income of the wealthiest households increased by 10.7% [131]. Given the evolving patterns of pollutants and socioeconomic segregation, an intriguing question would be to investigate the temporal variations in environmental inequalities.

Many researchers have conducted studies on the persistence of environmental inequality over the long term across various countries. For example, using multivariable linear regression, Kirby-McGregor et al. [132] examined whether the reduction in ambient PM_{2.5} concentrations from 2001 to 2016 in Canada was equitable across communities with distinct sociodemographic characteristics, accounting for geographical boundary modifications. Their results indicated that communities with lower education and income experienced smaller absolute decreases in PM_{2.5}. Rüttenauer and Best (2021) [133] investigated the sociodemographic changes before and after the establishment of new highly polluting industrial facilities in Germany from 2008 to 2017, revealing that the average income within the area decreased after the establishment of new facilities. Similarly, Padilla et al. (2014) [134] analyzed the spatial and temporal relationship between ambient air nitrogen dioxide concentrations and socioeconomic and demographic factors in four French metropolitan areas during 2002-2005 and 2006-2009. The study found that areas characterized by higher socioeconomic status were associated with elevated mean levels of NO_2 in Paris, whereas in Lille and Marseille, areas with greater levels of deprivation exhibited the highest levels of exposure to NO_2 . Ma et al. (2019) [135] explored socio-spatial and temporal inequalities in air pollution in Beijing, China from 2000 to 2010. Their study demonstrated that environmental injustice increased over time for migrant populations and the elderly, who experienced a disproportionate and increasing burden of deteriorating air quality, and highlighted the spatial and temporal variations in environmental inequality.

Within the context of the United States, numerous researchers employ statistical analyses

to measure exposure disparities among various racial, ethnic, and income groups. One common approach involves estimating population-weighted means for different groups over time. For example, Clark et al. [136] utilized this method to calculate outdoor NO_2 exposure levels across race-ethnicity, socioeconomic characteristics, and geographic location for the contiguous United States in 2000 and 2010. Their findings indicated that absolute NO_2 exposure disparities by race-ethnicity decreased from 2000 to 2010. However, relative NO_2 exposure disparities persisted, as nonwhite individuals exhibited higher concentrations of NO_2 than their white counterparts in 2010. In a study conducted by Rosofsky et al. [137], longitudinal air pollution inequality was quantified using population-weighted PM_{2.5} and NO_2 concentrations in Massachusetts between 2003 and 2010. The results indicated that relative air pollution inequalities exhibited a slight increase over time. Despite this, annual-average population-weighted PM_{2.5} and NO_2 concentrations remained highest for urban non-Hispanic black populations and urban Hispanic populations, respectively. Liu et al. [138] measured exposure disparities among racial/ethnic groups and by income for multiple spatial units and years regarding air pollution. The findings indicated that although absolute and relative exposure disparities between racial/ethnic groups decreased over time, minority groups still exhibited the highest national average exposure levels. Jbaily et al. [139] calculated the population-weighted PM_{2.5} concentration in racial/ethnic and income groups at the US zip code level between 2000 to 2016. The findings revealed that regions with higher-than-average white and Native American populations consistently experienced average PM_{2.5} levels that were lower than areas with higher-than-average Black, Asian, and Hispanic or Latino populations, which contradicts previous studies. Furthermore, areas with low-income populations exhibited higher average PM_{2.5} levels than areas with high-income groups for the years 2004-2016. Previous studies have utilized estimated population-weighted mean concentration

to demonstrate the longitudinal trend of pollution inequality. This approach has allowed for the identification of disparities among racial, ethnic, and income groups across different regions and years. However, relying on estimated data introduces a degree of uncertainty and potential inaccuracies. Moreover, these studies have neglected to examine the correlations between racial/ethnic and socioeconomic characteristics and pollution, which is crucial for understanding the underlying mechanisms driving these disparities.

Regression models are very useful tools to investigate the associations between various socioeconomic indicators and pollution levels. By using regression models, researchers can analyze the impact of socioeconomic factors on pollution, and identify potential areas for targeted interventions to reduce pollution and related health disparities. For example, Ard [83] conducted a nationwide investigation using spatially specified multilevel models to examine the relationship between industrial toxins and race and social class. The findings demonstrated that between 1995 and 2004, exposure to air toxins declined for all groups. However, African-Americans continued to experience higher levels of exposure compared to Whites and Hispanics. Kravitz-Wirtz et al. [86] employed multilevel repeated measures models to investigate the association between race/ethnicity and exposure to air pollutants over time. Their findings were consistent with previous research, showing that although air pollution levels declined nationally over time, racial and ethnic disparities in exposure to NO_2 , PM2.5, and PM10 persisted. Specifically, their results indicated that, on average, Blacks and Latinos were more likely to be exposed to higher levels of these pollutants compared to Whites. Li et al. [87] conducted a study in Texas to investigate the correlations between excess emissions and demographic indicators from 2000 to 2012 using a logit and a conditional OLS modeling approach. Their findings suggest that the geographical distribution of excess emissions is not random with respect to neighborhoods’

racial, ethnic, and socioeconomic attributes. Specifically, the study found a positive association between excess emissions and the percentage of the Black population. Colmer et al. [140] utilized a standard OLS model to investigate the pollution disparities in the United States by combining PM_{2.5} concentration data from 1981 to 2016 with demographic data from 65000 US census tracts. The study revealed that absolute disparities in pollution exposure have declined over time, but relative disparities persist. Bravo et al. [141] utilized spatial and temporal data from 2002-2016 to assess the variability and trends in PM_{2.5} and O_3 concentrations across North Carolina, USA, and their associations with community characteristics. The findings revealed that while pollutant concentrations declined over time, exposure disparities increased for communities that were racially and educationally isolated. Additionally, tracts with higher urbanicity and lower educational isolation had higher PM_{2.5} levels. Ard and Smiley [35] employed multilevel models to examine the association between industrial facilities and economic conditions, as well as poverty segregation during the period of 1980 to 2009. Their findings indicated that counties characterized by higher levels of economic and racial segregation had a significantly higher density of industrial facilities. Collins et al. [142] used a multivariable generalized estimating equation to quantify PM_{2.5} exposure disparities of racial minorities from 2012 to 2016. The results confirmed associations between increased minority composition and greater exposure to long-term and short-term PM_{2.5} measures. Kodros et al. (2022) [143] conducted a study to investigate the relationship between racial residential segregation and trace metals in PM_{2.5} from 2010 to 2019. The findings indicate that individuals residing in racially segregated communities are exposed to a form of fine particulate matter containing over three times higher mass proportions of known toxic and carcinogenic metals. Ma et al. (2022) assessed PM_{2.5}-attributable CVD (cardiovascular disease) deaths in different racial/ethnic groups from 2001 to 2016. The findings

revealed that Non-Hispanic Black individuals had the highest burden of PM_{2.5}-attributable CVD mortality. Although the gap between racial/ethnic minorities and non-Hispanic White people narrowed over time, the disparity remains substantial. Currie et al. (2023) employed quantile regression to decompose the primary drivers of convergence in black-white pollution exposure differences related to PM_{2.5} from 2000 to 2015. The results demonstrated that areas with more significant Black populations experienced more substantial declines in PM_{2.5}, contributing to the reduction in the exposure disparity between Black and White individuals.

Those researchers, however, used the global model to analyze the spatiotemporal heterogeneity of the impact of socioeconomic development on toxic risk. Specifically, they conducted regressions between socioeconomic characteristics and pollution at each geographic unit, while assuming spatial stationarity in the relationships between explanatory and dependent variables. According to Tobler's first law of geography: everything is related to everything else. But near things are more related than distant things [144]. Thus, incorporating the concept of spatial dependence is essential when investigating the spatial heterogeneity of environmental inequality [144]. Local models, such as the Geographical Weighted Regression (GWR) model, enable the estimation of parameters specific to each location, thereby incorporating the geographic context. While the GWR model captures the geographic context, it does not incorporate the temporal dimension. Temporal data is a valuable resource for investigating the dynamics of the underlying spatial processes and for forecasting relevant variables. This is particularly relevant in research domains such as the diffusion of contagious diseases, the spread of air or water pollution, and the expansion of urban sprawl [33]. To address this limitation, Huang et al. [33] and Fotheringham et al. [34] proposed a temporal extension of GWR, namely, the geographical and temporal weighted regression (GTWR) model, which allows for spatiotemporal modeling from a local perspective.

The GTWR model incorporated spatial and temporal weighting functions that account for the heterogeneity of data across space and time. These weighting functions allow the model to estimate coefficients that vary across different spatial and temporal contexts, providing a more nuanced understanding of the relationships between variables.

Moreover, previous research mainly focused on air pollution, such as PM_{2.5}, PM₁₀, NO_2 , O_3 , etc. Increasing concerns regarding the adverse effects of toxic chemicals on the environment and human health have garnered considerable attention, particularly in the aftermath of the 2023 Ohio train derailment. This incident resulted in the release of toxic chemicals, including vinyl chloride and butyl acrylate, into the air, soil, and water, thereby highlighting the potential risks associated with toxic chemical releases. The release of toxic chemicals presents a significant environmental concern, with the potential to cause widespread and far-reaching implications for both human health and ecological systems [2, 6, 7]. The sources of these chemicals are diverse, encompassing industrial processes, transportation, and agricultural practices, among other activities. When these chemicals are released into the environment, they can contaminate multiple domains, including the air, water, soil, and food, potentially leading to a range of adverse human health outcomes, such as respiratory complications, cancer, and neurotoxicity [2]. Furthermore, these toxic releases can disrupt the ecological balance, adversely affecting biodiversity and ecosystem services [67].

To the best of my knowledge, no study has yet examined the correlation between socioeconomic indicators and toxic risk in the United States utilizing a local model that incorporates a temporal dimension. The exploration of spatiotemporal heterogeneity in the influence of socioeconomic development on toxic risk may uncover variations across regions, enabling local authorities to tailor regulations accordingly. Furthermore, investigating the trends in pollution inequality may

assist in identifying the primary impact factors over time and therefore facilitate the development of effective solutions to mitigate this issue. In this paper, we aim to answer if pollution inequality across regions improved over time, that is if racial minorities and low socioeconomic entities suffered more from the toxic risk as time goes by. Taking Houston, Texas, a racially and ethnically diverse city with a history of segregation and discrimination, as an example, the objective of this paper, therefore, is to analyze the spatiotemporal heterogeneity of the impact of racial disparities and socioeconomic development on toxic risk in Houston using geographical and temporal weighted regression (GTWR) model to answer the question of whether pollution inequality improved or worsen. We then compare the differences between regions with “redlining” and no “redlining” areas, as well as the differences between redlining grades (A, B, C, and D).

4.2 Methods and data

4.2.1 Study area

The prevalence of pollution inequality persists throughout the United States, particularly in Houston, Texas. Houston is the fourth most populous city in the nation, with an estimated population of 2,304,580 in 2020 (trailing only New York, Los Angeles, and Chicago), and is the largest in the southern U.S. and Texas [116]. Comprising a land area of 640.4 square miles ($1,659 \text{ km}^2$) [116], more than twice of Singapore (280 square miles), Houston is the ninth-most expansive city in the United States (including consolidated city-counties).

Houston is characterized by its expansive and diverse population, as well as a legacy of segregation and discrimination. Despite attempts at emancipation in 1863 [145] and desegregation efforts in 1964 [146], redlining (see Figure 4.1) and other systemic factors have resulted in

persistent disparities in the city. In this figure, residential areas designated as green and blue were considered "Best" and "Still Desirable" and were predominantly occupied by white residents. Conversely, yellow neighborhoods were classified as "Definitely Declining", while red areas were deemed "Hazardous", and predominantly inhabited by Black populations. The practice of redlining legally prevented Black families from accessing fair housing mortgages for decades. Houston, Texas exhibits pronounced racial and ethnic disparities in exposure to pollution, with racial and ethnic minority communities experiencing higher levels than non-minority populations. For example, as of 2019, the Harrisburg/Manchester neighborhood hosts 21 industrial and toxic waste facilities, including generators of hazardous waste and waste treatment and disposal facilities, situated within a three-mile radius. Notably, the Hispanic population comprises 90% of the residents in the Harrisburg/Manchester area, exacerbating concerns of environmental injustice.

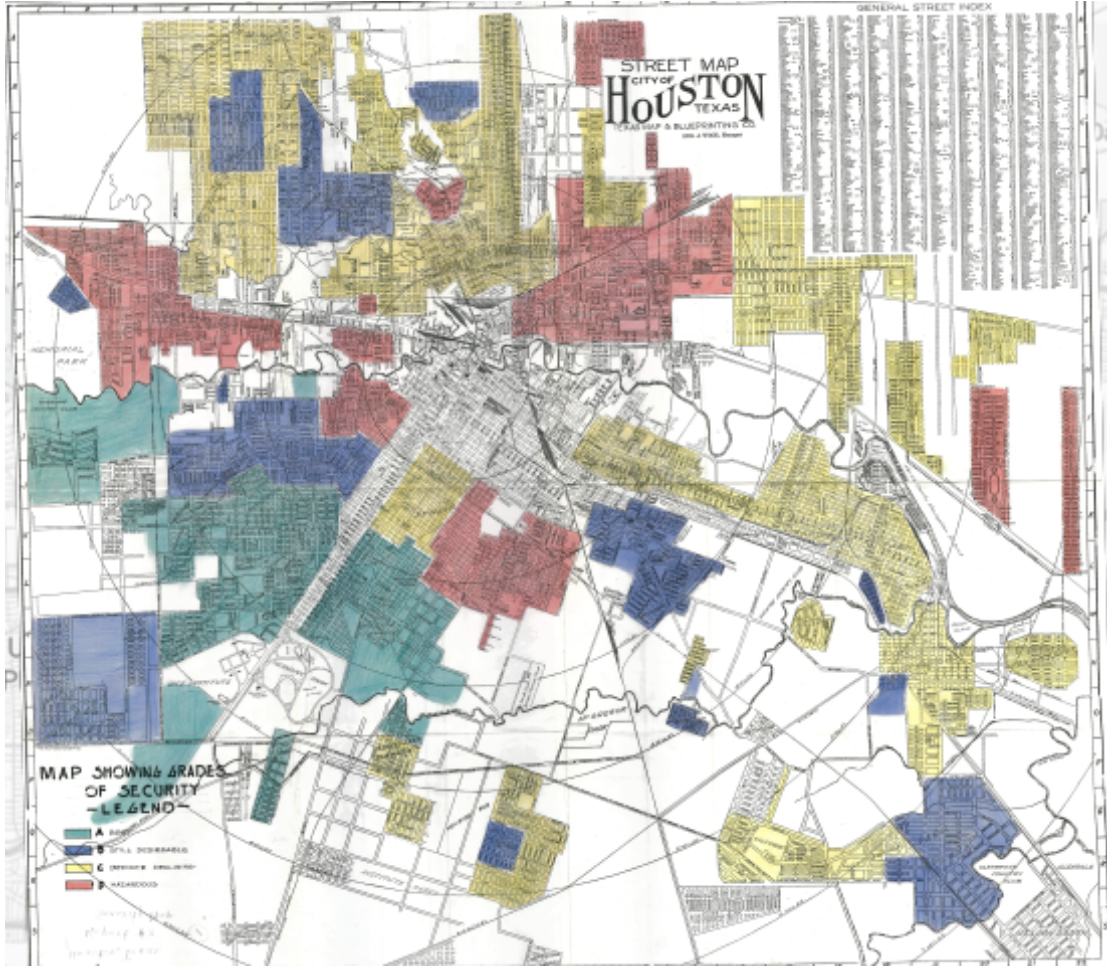


Figure 4.1: Redlining map in Houston, Texas

Source: <https://dsl.richmond.edu/panorama/redlining/loc=12/29.746/-95.444city=houston-tx>

4.2.2 GTWR model

The general route to estimate local parameters in the Geographical and temporal weighted regression (GTWR) model can be found in Figure 4.2. First, a traditional GWR model can be specified using the following equation:

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \quad (4.1)$$

where i is the index of a spatial point, y_i is the dependent variable in local i , x_{ik} is the

$k - th$ independent variable in local i , and ε_i is the error term. $\beta_k(u_i, v_i)$ is the estimated $k - th$ parameter in local i , which can vary across space to measure spatial nonstationary relationships.

If using the matrix, the estimated parameters can be denoted as:

$$\hat{\beta}(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) y \quad (4.2)$$

where W is a diagonal matrix representing the geographical weights of each observation of the $i - th$ point.

After specifying a GWR model using the input data in equation 4.1, an optimal spatial bandwidth needs to be specified for each time period based on goodness-of-fit criteria such as cross-validation (CV) or Akaike information criterion (AIC). The determination of the optimal temporal bandwidth is based on CV or AIC, following the selection of the optimal spatial bandwidth. Once both spatial and temporal bandwidths are determined, the next step is to construct a spatiotemporal weight matrix. Detailed information about the structure of the spatiotemporal weight matrix and bandwidth optimization can be found in [33] and [34].

The GTWR model can be expressed as:

$$Y_i = \beta_0(u_i, v_i, t_i) + \sum_k \beta_k(u_i, v_i, t_i) X_{ik} + \varepsilon_i \quad (4.3)$$

Similarly, the estimation of $\beta_k(u_i, v_i, t_i)$ can be expressed as:

$$\hat{\beta}(u_i, v_i, t_i) = (X^T W(u_i, v_i, t_i) X)^{-1} X^T W(u_i, v_i, t_i) Y \quad (4.4)$$

where $W(u_i, v_i, t_i) = \text{diag}(\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in})$ and n is the number of observations. $\alpha(1 \leq$

$j \leq j$) are space-time distance functions of (u, v, t) corresponding to the weights when calibrating a weighted regression adjacent to observation point i .

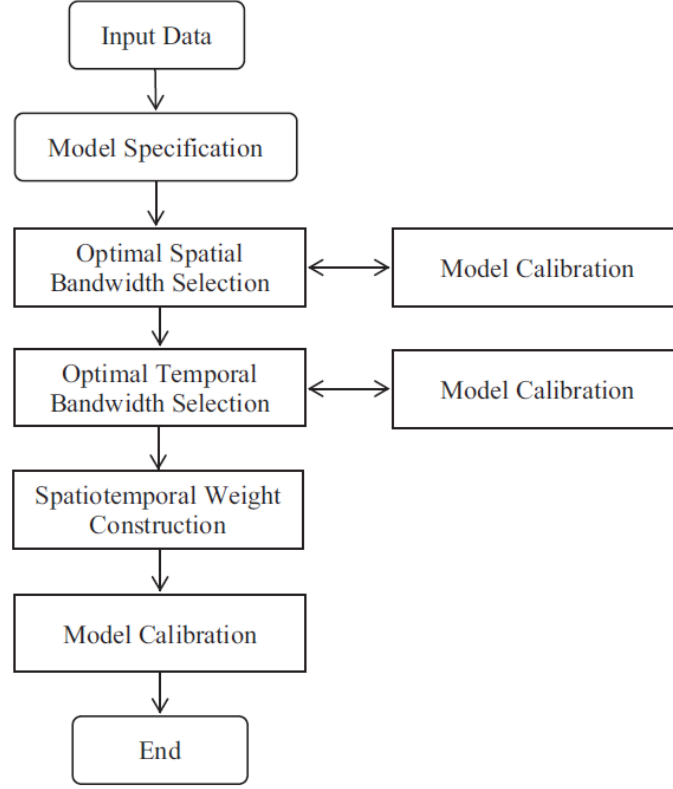


Figure 4.2: Scheme of geographical and temporal weighted regression (GTWR) model
Source: Fotheringham et al. (2015) [34]

4.2.3 Data

The independent variable is the toxic risk, which can be calculated using Risk-Screening Environmental Indicators (RSEI) Model. The RSEI scores for the years 2011-2017 are obtained from the U.S. Environmental Protection Agency website [147]. The dependent variables encompass population, race, and ethnicity data, including population density, percent of Black, percent of Asian, percent of other races, and percent of Hispanic/Latino, as well as economic data, such as median income. Other socioeconomic data, including percent of owner-occupied housing,

percent of vacant housing, percentage of houses with mortgages, percentage of adults without a high school diploma, and percent of unemployment, are also considered, see Table 4.1. The aforementioned data are obtained from the American Community Survey 5-Year Data from 2009-2013 to 2015-2019 [148]. It is important to note that the race/ethnicity data and socioeconomic data are available at the block group level, while the RSEI score data are available at the facility level. The statistical summary of all the variables can be shown in Table 4.2.

The figures from 4.3 to 4.5 portray the spatial distributions of the Risk-Screening Environmental Indicators (RSEI) score in Houston, Texas, for the years 2011, 2014, and 2017. The figure indicates that the RSEI scores are primarily concentrated in the eastern and western regions of Houston, as well as along the Buffalo Bayou River in 2011. The distribution pattern of the RSEI scores in 2014 indicates a decrease in the eastern and northeastern regions. Between 2014 and 2017, the RSEI scores in the western region declined while those in the northern region increased.

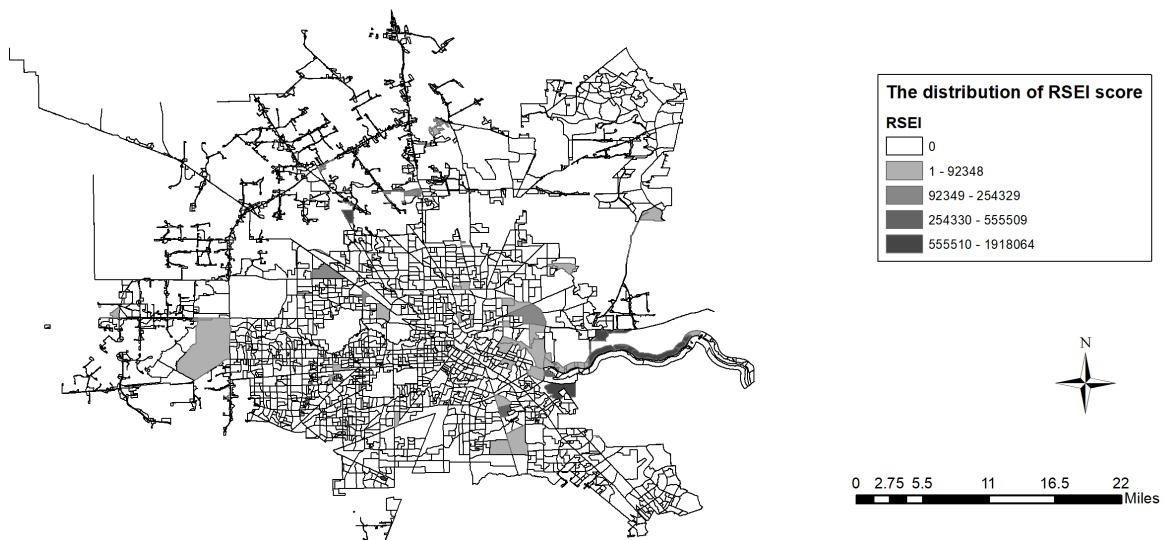


Figure 4.3: Spatial distribution of the RSEI score in 2011



Figure 4.4: Spatial distribution of the RSEI score in 2014

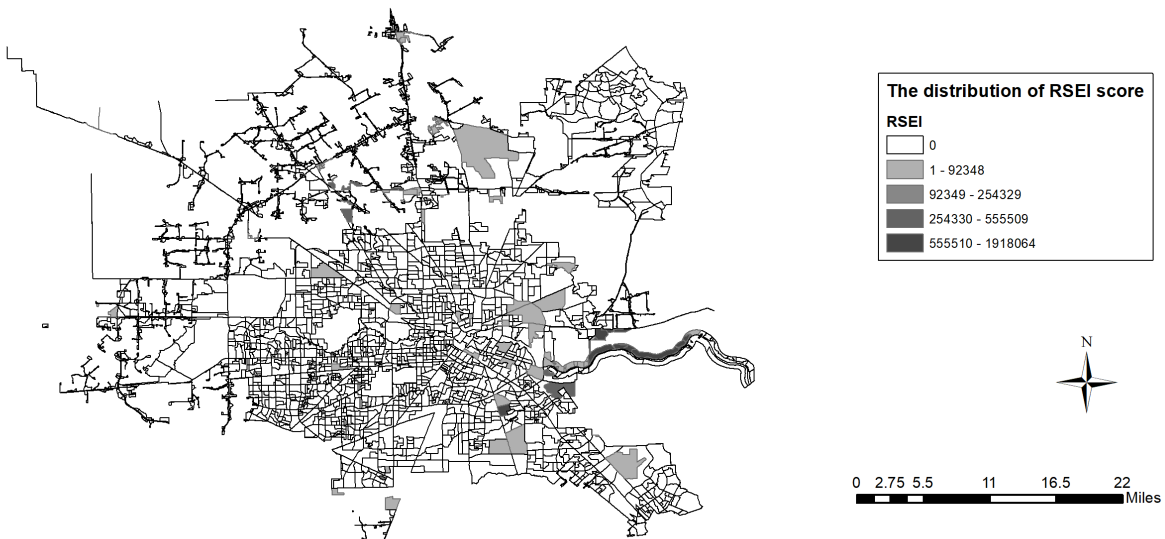


Figure 4.5: Spatial distribution of the RSEI score in 2017

To reconcile the RSEI score data acquired from EPA with the race/ethnicity and socioeconomic data obtained from ACS, we executed a series of steps. Initially, we extracted the 12-digit GEOID (Geographic Identifiers) from the ACS table to represent each census block group. As the RSEI score data were obtained at the level of TRI facilities, it was indispensable to aggregate

Table 4.1: Variable names, units, and detailed description

Variable	Abbreviation	Unit	Description
Log_RSEI	RSEI		Natural logarithm of (1+ RSEI score)
Pop_dens	Pop	population per km2	Population density
Black_Pc	Black	%	The percentage of Black or African American population
Asian_Pc	Asian	%	The percentage of Asian population
Others_Pc	Other	%	The percentage of other races
Hispanic_Pc	Hisp	%	The percentage of Hispanic or Latino
Median_household_income_k	income	thousand dollars	Median household income
Unemploy_Pc	Unemp	%	The percentage of unemployment, which is measured by the percent of the labor force that are unemployed
Vacant_Pc	Vac	%	The percentage of vacant housing
Owner_Pc	Own	%	The percentage of owner-occupied housing
Mortgage_Pc	mort	%	The percent of housing units with a mortgage
Not_High_Pc	NH	%	The percentage of adults without high school diploma (for population 25 years and over)

the information to the census block group level. We utilized ArcGIS Map to merge the TRI facility-level data to the block group level based on the GEOID of each block group. The block group shapefile of the City of Houston was retrieved from the United States Census Bureau TIGER (Topologically Integrated Geographic Encoding and Referencing)/Line Shapefiles: Block Groups [149] and the COHGIS (The City of Houston Geographic Information System) Open Data Portal [150].

Table 4.2: Statistical Summary of the variables

	nbr.val	min	max	range	sum	median	mean	SE.mean	CI.mean.0.95	var	std.dev	coef.var
year	13447	2011	2017	6	27082258	2014	2014	0.02	0.03	4.00	2.00	0.00
Log_RSEI	13447	0	17.02	17.02	5004.10	0.00	0.37	0.02	0.03	3.64	1.91	5.13
Pop_dens	13447	0	379.24	379.24	346516.43	20.00	25.77	0.22	0.43	639.34	25.29	0.98
Black_Pc	13447	0	100.00	100.00	280303.90	11.04	20.85	0.21	0.41	596.11	24.42	1.17
Asian_Pc	13447	0	74.84	74.84	88668.92	2.94	6.59	0.08	0.16	85.64	9.25	1.40
Others_Pc	13447	0	94.15	94.15	152913.79	8.04	11.37	0.10	0.19	129.01	11.36	1.00
Hispanic_Pc	13447	0	100.00	100.00	547297.25	33.65	40.70	0.24	0.48	805.70	28.38	0.70
Median.household_income_k	13447	0	249.95	249.95	825521.10	51.18	61.39	0.35	0.68	1641.25	40.51	0.66
Unemploy_Pc	13447	0	61.27	61.27	98731.13	5.86	7.34	0.06	0.11	42.17	6.49	0.88
Vacant_Pc	13447	0	73.80	73.80	135186.83	8.32	10.05	0.08	0.15	78.44	8.86	0.88
Owner_Pc	13447	0	100.00	100.00	746986.04	60.38	55.55	0.26	0.50	888.80	29.81	0.54
mortgage_Pc	13447	0	728.57	728.57	735406.99	58.70	54.69	0.24	0.46	744.82	27.29	0.50
Not_High_Pc	13447	0	199.72	199.72	1057520.84	82.98	78.64	0.17	0.32	366.39	19.14	0.24

nbr.val means the number of variables; SE.mean means standard error of the mean; CI means the confidence interval; std.dev means standard deviation; coef.var means the coefficient of variation.

4.3 Results and Discussion

4.3.1 Analysis of estimation results

Based on Equation 4.3, we estimate the correlation between racial and ethnic composition, socioeconomic indicators, and toxic chemical risk. The estimates reveal that the optimal bandwidth amounts to 1345, which determines the size of the local neighborhood surrounding each point in space and time for which the regression model is computed. For the assessment of the model's goodness of fit, the results indicate that the AICc value reaches 53559.6, whereas the R^2 and R^2 Adjusted values amount to 0.1615 and 0.1609, respectively. The Spatio-temporal Distance Ratio (STDR) stands at 0.2688, implying that the spatial and temporal scales of the data are relatively minor compared to the variability observed in the response variable. This finding suggests that the GTWR model may prove suitable for the data, as it can capture the local spatial and temporal patterns present in the data.

To assess the significance of the association between each independent variable and the dependent variable, a t-test was conducted and detailed in Appendix B. The findings indicate that the explanatory variables that have a statistically significant relationship with the outcome variable at a significance level of 0.5 are as follows: Intercept, Pop_dens, Black_Pc, Asian_Pc, Other_Pc, Hisp_Pc, Unemployment, Owner_Pc, and Not_High_Pc. A multicollinearity assessment was performed to ascertain the presence of a significant degree of inter-correlation among the predictor variables (see Appendix B). The examination of the variance inflation factor (VIF) values indicated that all variables demonstrated VIF values below 10, thus indicating the absence of highly correlated predictor variables.

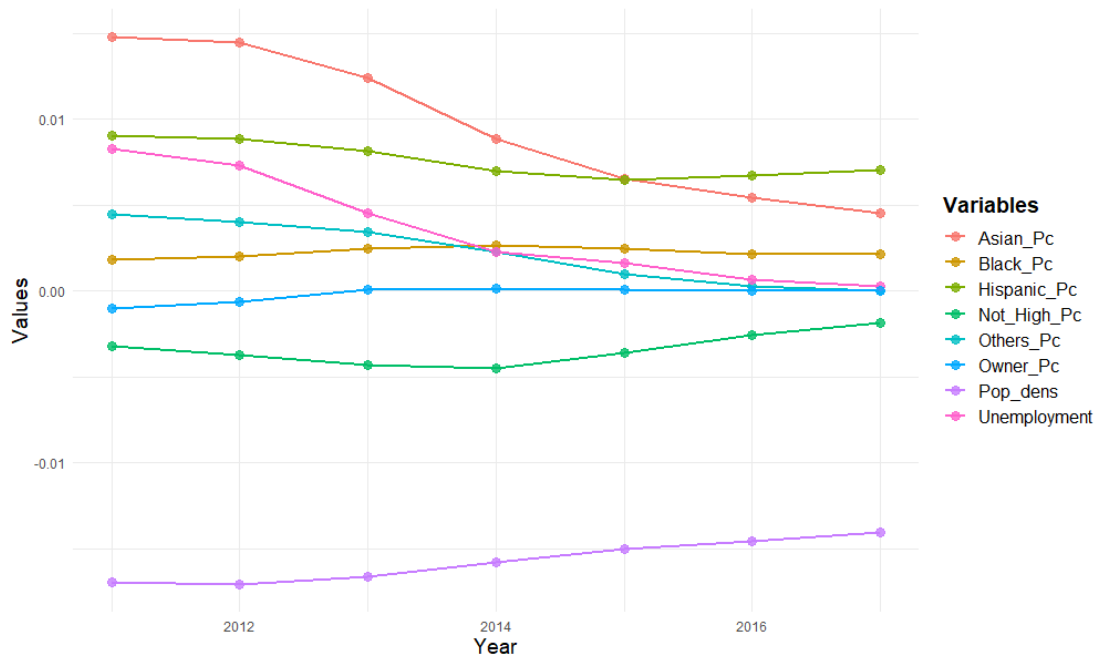


Figure 4.6: Temporal variation of the parameter coefficients

The present study presents a graphical analysis of the temporal variation of parameter estimates obtained through the Geographically and Temporally Weighted Regression (GTWR) model for the years 2011-2017. Specifically, Figure 4.6 displays the average trend of these parameter estimates over the aforementioned period, whereas Figure 4.7 illustrates the time series change trend of all the GTWR estimation coefficients.

Figure 4.6 reveals that population density exhibits a negative association with toxic risk, as evidenced by the negative relationship between the two variables. Moreover, the estimates suggest that the magnitude of this relationship increased between 2011 and 2017. Similarly, the percentage of adults without a high school diploma, used as an education indicator, is also found to be negatively related to toxic risk during the study period. However, the temporal trend of the parameter estimates indicates that the magnitude of this relationship declined from 2011 to 2014, followed by an increase from 2014 to 2017. The housing indicator, specifically the

percentage of owner-occupied housing, exhibits an average negative association with toxic risk. The study findings indicate that the percentage of minorities, which includes Asian, Black or African American, and other ethnic groups, shows an average positive correlation with toxic risk. Specifically, the parameter estimates for Asian_Pc and Others_Pc demonstrate a consistent decreasing trend between 2011 and 2017. The estimates for the variable Black_Pc demonstrate an increasing trend from 2011 to 2014, followed by a stable pattern from 2014 to 2017. Regarding the ethnicity indicator, the percentage of Hispanic or Latino individuals demonstrates a decreasing positive association with toxic risk from 2011 to 2015. However, the average parameter estimates have been increasing since 2016.

Figure 4.7 presents the GTWR regression coefficients for the years 2011 to 2017, representing the complete time series rather than just the average temporal trend. From this figure, it is apparent that although the average parameter estimates for variables Not_High_Pc and Owner_Pc indicated negative relationships with toxic risk (as depicted in Figure 4.6), the relationships become mixed when considering all coefficients. A similar pattern of mixed results is observed for the relationship between unemployment and toxic risk, with average coefficients indicating a positive association. Furthermore, the variables Black_Pc and Other_Pc exhibit mixed results from 2011 to 2014. These mixing outcomes indicate the presence of spatial-temporal heterogeneity in the distribution of coefficient estimates and emphasize the significance of conducting spatial-temporal analyses to provide targeted guidelines for reducing toxic risk. In the next section, a detailed analysis of spatiotemporal heterogeneity in the influence of socioeconomic development and racial/ethnic composition on toxic risk is conducted to uncover the variations across regions during the study period.

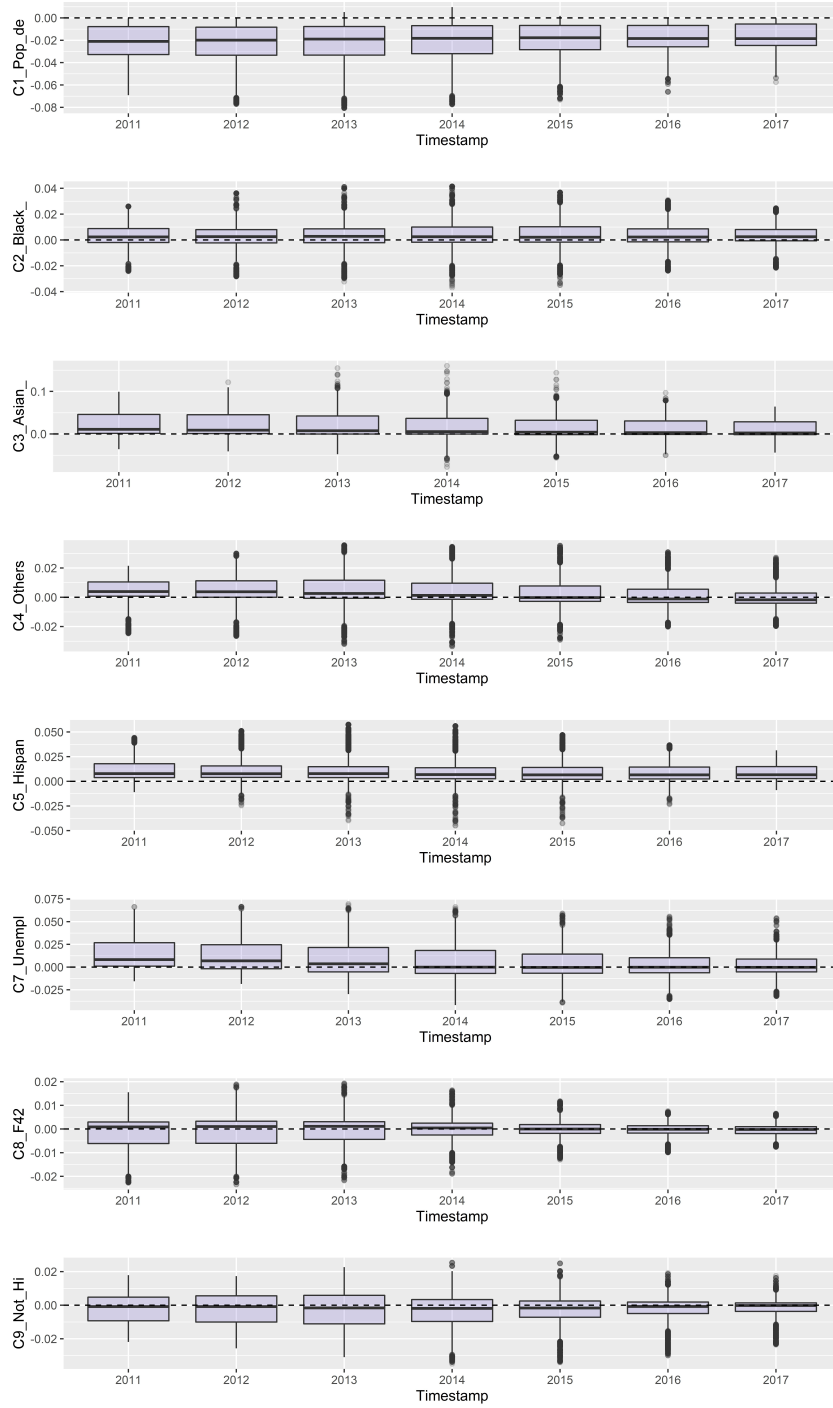


Figure 4.7: Time series change trend of GTWR regression coefficients

4.3.2 Spatiotemporal heterogeneity distribution of coefficients

This section presents an in-depth spatiotemporal investigation into the associations between racial/ethnic and socioeconomic characteristics and toxic risk. The spatial distributions of each indicator from 2011 to 2017 are displayed in Appendix B. Within this section, we have specifically chosen to showcase the spatiotemporal trends of each explanatory variable by focusing on the results from 2011, 2014, and 2017.

Regarding the relationship between racial minorities and toxic risk, the results of the percentage of Black or African American are shown in Figure 4.8 and B.4. The geographical results indicate that for eastern Houston and part of western Houston, the percentage of the Black population tends to have a negative relationship with toxic risk, which indicates that the block groups with a high percentage of Black or African American people are less likely to be exposed to toxic risk. For the majority of other areas of Houston, the percentage of the Black population tends to have positive associations, suggesting that densely concentrated racial minority groups are more likely to be exposed to toxic risk.

Notable temporal findings indicate that the coefficient estimates exhibit an upward trend for most regions of Houston, Texas, between 2011 and 2014, followed by a decline from 2014 to 2017. This trend is particularly evident in the northwestern region of Houston. As illustrated in Figure 4.8a, the coefficient estimates range from 0 to 0.03. A coefficient of 0.03 implies that, while controlling for all other variables, a one percentage point increase in the share of the Black population within the block group results in an approximate 3% increase in toxic risk. In 2014, the coefficient estimates range from 0.02 to 0.05 (see Figure 4.8b), which means that while holding all other variables constant, a one percentage point increase in the share of the Black population

results in a 2% to 5% increase in toxic risk. From 2014 to 2017, the coefficient estimates declined for the majority of Houston (see Figure 4.8c).

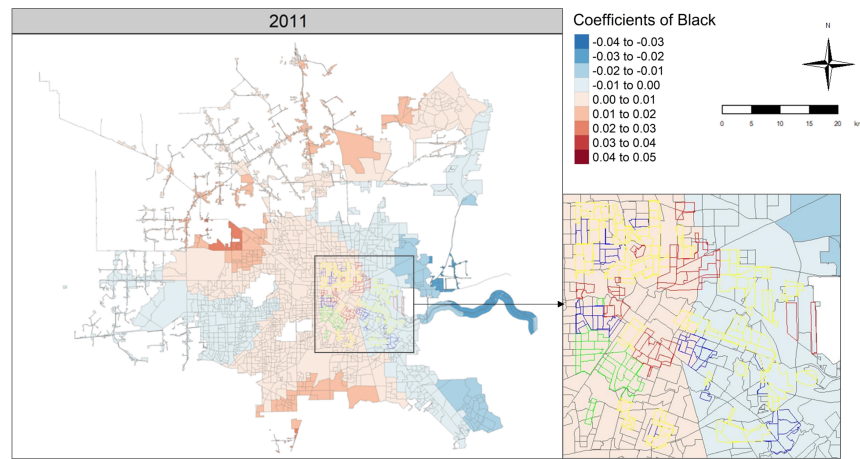
For the redlining areas in Houston, we can see that there are mixing positive and negative associations between the percentage of Black or African Americans and toxic risk. Specifically, in 2011 and 2017, areas designated as "Hazardous" tend to have positive coefficients, indicating a positive relationship between the percentage of Black or African Americans and toxic risk. However, in 2014, these areas tend to have negative coefficients, suggesting a negative relationship between the two variables. The same pattern is observed in the "Best" areas, which are characterized by the least risks for lenders and insurers. These findings suggest that the relationship between race and toxic risk may be complex in areas with redlining history. Further research is needed to better understand the underlying factors driving these patterns and their implications for environmental justice.

A possible mechanism for temporal environmental justice is Neighborhood Change, that is, the racial composition of neighborhoods. The Classical Invasion-Succession Model [151] is a theoretical framework used to describe the process of neighborhood change over time based on the concept of ecological succession. The process starts with Black residents moving into predominantly White neighborhoods, which can create tensions and conflicts. However, over time, the minority group becomes more established and builds social networks and institutions, leading to a gradual replacement of the existing White population. Therefore, the white population has been gradually shifting to more urban and hence more polluted areas [152]. Another explanation would suggest that areas with a higher proportion of African Americans may have experienced larger reductions in pollution levels over time compared to other areas. Currie et al. (2023) [153] proposed that the Clean Air Act (CAA) has probably had a considerable impact in lessening the

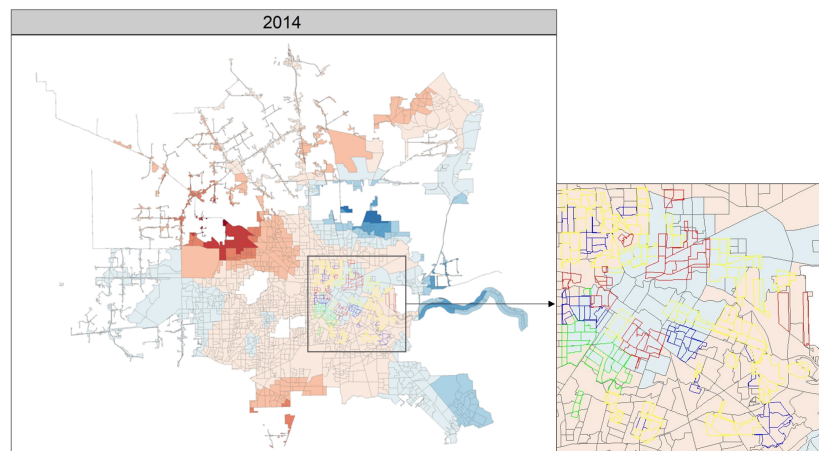
differences in air pollution exposure between different racial groups. The authors argued that the law's deliberate focus on cleaning up the most heavily polluted areas, where African Americans were more likely to reside, has been instrumental in achieving this. As a result, despite not being explicitly designed for this purpose, the CAA has played a role in reducing environmental disparities among racial groups in the United States.

With respect to other racial minorities, such as Asians and other ethnic groups, the findings are depicted in Figures 4.9 and 4.10. In the eastern and southeastern regions, the percentage of Asians exhibits a negative association with toxic risk in 2011, 2014, and 2017, which is comparable to the results observed for the percentage of the Black population. This suggests that block groups with a higher proportion of Asians tend to encounter a lower level of toxic risk. Conversely, in most other regions, the percentage of Asians is inversely related to toxic risk, indicating that block groups with a higher percentage of Asians are exposed to greater levels of toxic risk. Notably, the observed coefficients for the percentage of Asians demonstrate a declining trend from 2011 to 2017, which suggests evidence of environmental justice improvement in Houston.

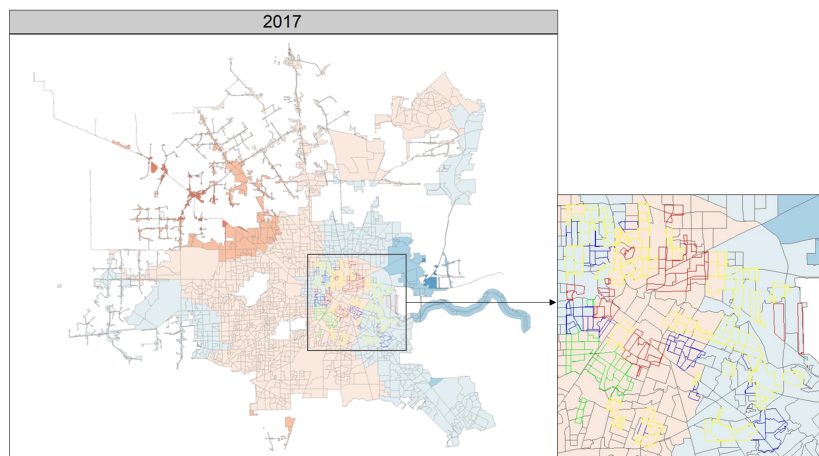
Other racial minority groups, including American Indian and Alaska Native, Native Hawaiian and Other Pacific Islander, and other groups, exhibit positive associations with toxic risk in 2011 and 2014. However, in 2017, negative associations are evident between the percentage of Other racial minorities and toxic risk for the majority of Houston, except for the northeastern areas. Notably, within the redlining area, block groups with a higher percentage of Other racial minorities tend to be exposed to elevated levels of toxic risk from 2011 to 2017.



(a) 2011

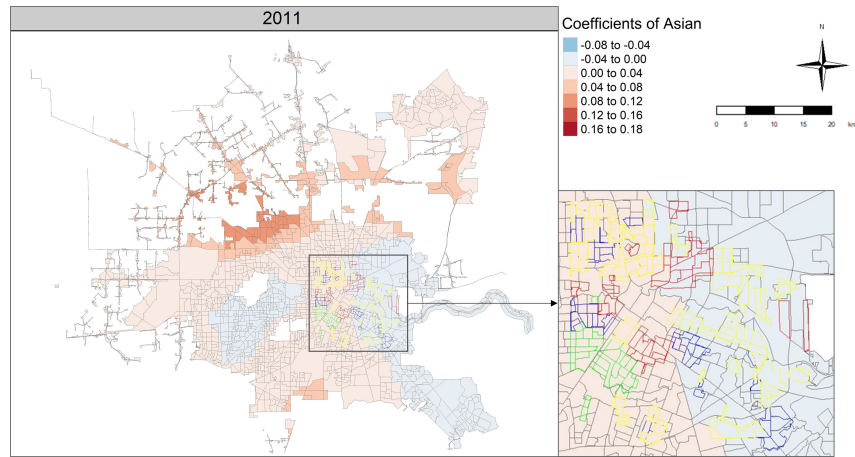


(b) 2014

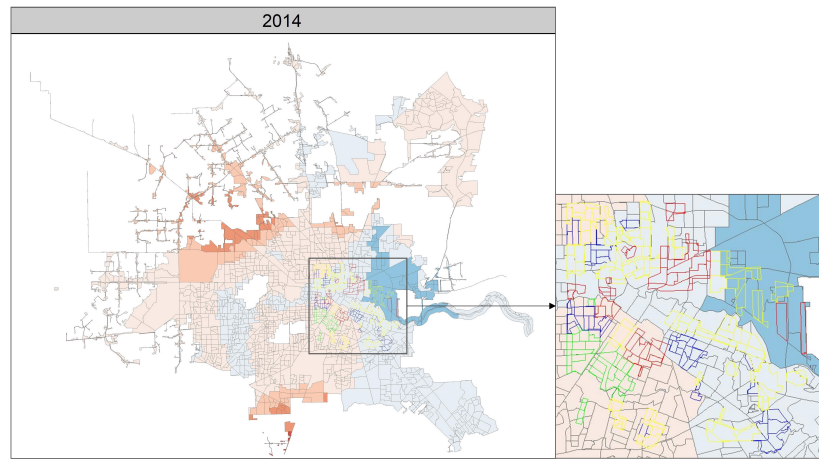


(c) 2017

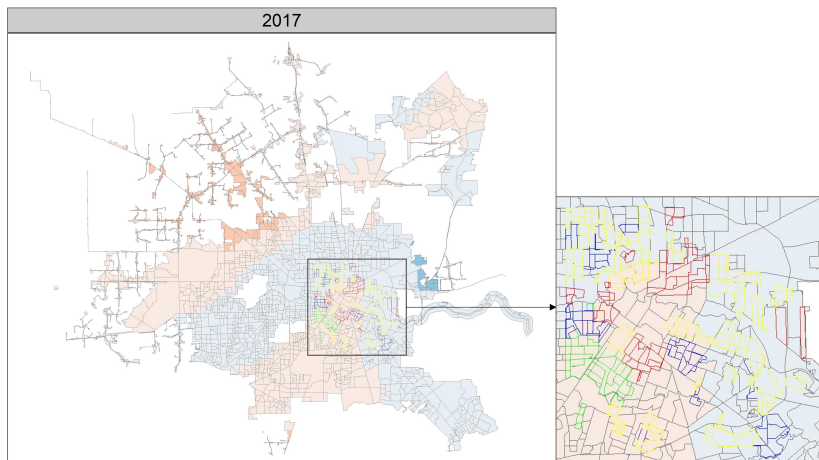
Figure 4.8: The spatiotemporal distribution of regression coefficient estimates of Black_Pc
 Note: Green boundary means "Best", blue boundary means "Still Desirable", yellow boundary means "Definitely Declining", and Red boundary means "Hazardous"



(a) 2011

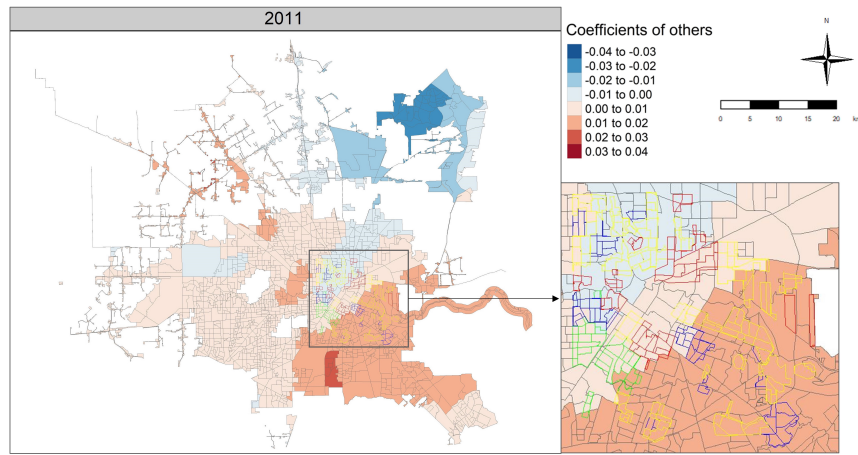


(b) 2014

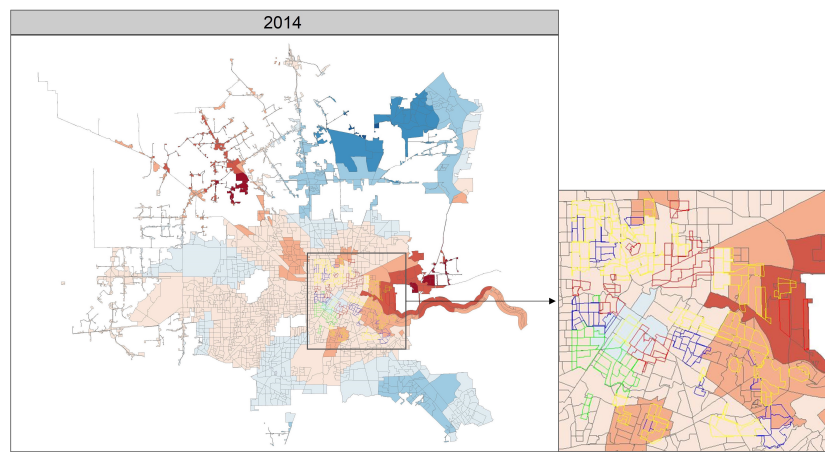


(c) 2017

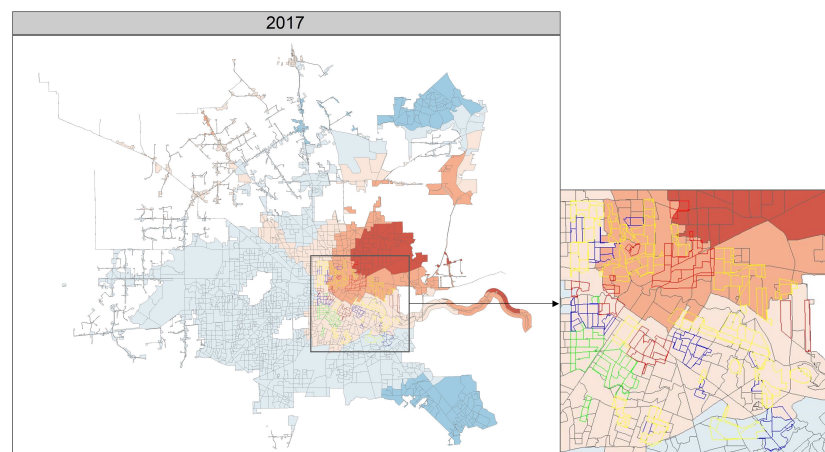
Figure 4.9: The spatiotemporal distribution of regression coefficient estimates of Asian_Pc
 Note: Green boundary means "Best", blue boundary means "Still Desirable", yellow boundary means "Definitely Declining", and Red boundary means "Hazardous"



(a) 2011



(b) 2014

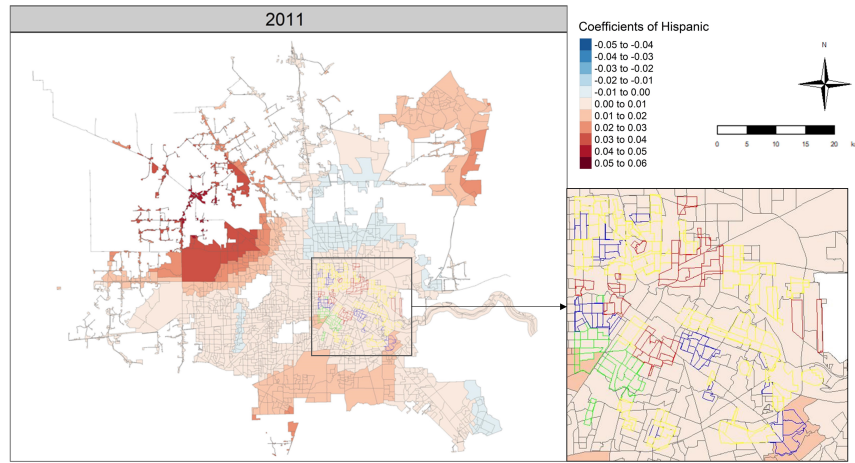


(c) 2017

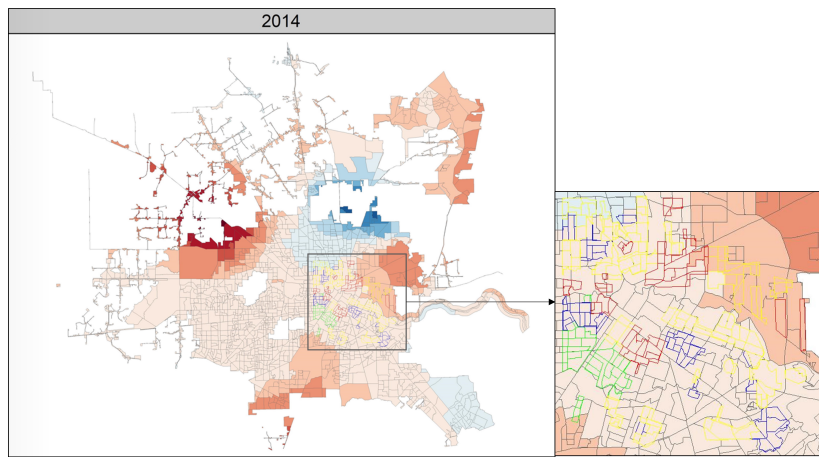
Figure 4.10: The spatiotemporal distribution of regression coefficient estimates of Others.Pc
 Note: Green boundary means "Best", blue boundary means "Still Desirable", yellow boundary means "Definitely Declining", and Red boundary means "Hazardous"

The coefficient estimates for the ethnic composition are presented in Figure 4.11. As depicted in the figure, the percentage of Hispanic or Latino individuals is positively associated with toxic risk across the majority of areas, with the exception of the area surrounding Aldine, Houston. This indicates that block groups with a higher proportion of Hispanic or Latino individuals in the vicinity of Aldine, Houston, are exposed to lower levels of toxic risk. However, in the majority of other areas, block groups with a higher percentage of Hispanic or Latino individuals are exposed to greater levels of toxic risk. Notably, within redlining areas, the percentage of Hispanic or Latino individuals shows a positive association with toxic risk from 2011 to 2014, regardless of the assigned color codes. The coefficient estimates for the percentage of Hispanic or Latino individuals within the redlining areas ranged from 0 to 0.02. Holding all other variables constant, a 1% increase in the proportion of Hispanic or Latino individuals within the redlining areas would result in a maximum increase in the toxic risk of about 2%, based on the coefficient estimates obtained.

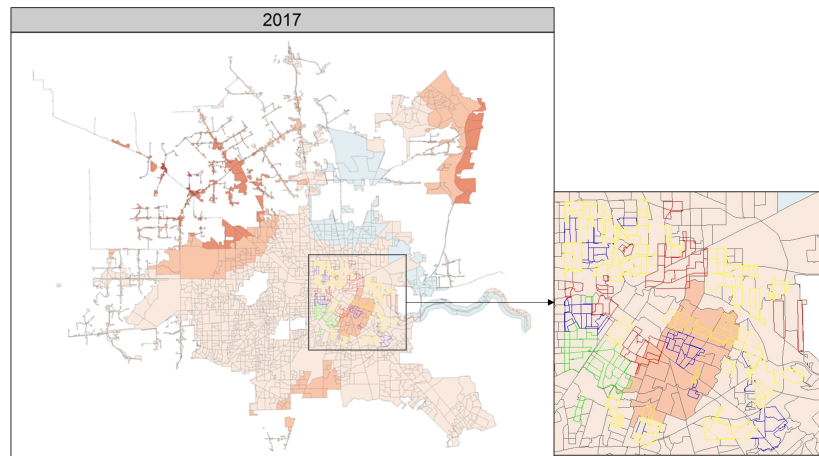
Between 2011 and 2014, the coefficient estimates in the northwestern region exhibit an increasing trend. Specifically, in 2014, the coefficients for the percentage of Hispanic or Latino individuals in the northwestern region ranged from 0.02 to 0.06. This indicates that with other variables held constant, a 1% increase in the percentage of Hispanic or Latino individuals would lead to an increase in toxic risk ranging from 2% to 6%. However, between 2014 and 2017, the coefficient estimates in the northwestern region demonstrate a decreasing trend. In 2017, the coefficients for the percentage of Hispanic or Latino individuals ranged from 0 to 0.03. Therefore, holding all other variables constant, a 1% increase in the proportion of Hispanic or Latino individuals would lead to a maximum increase in the toxic risk of 3%.



(a) 2011



(b) 2014



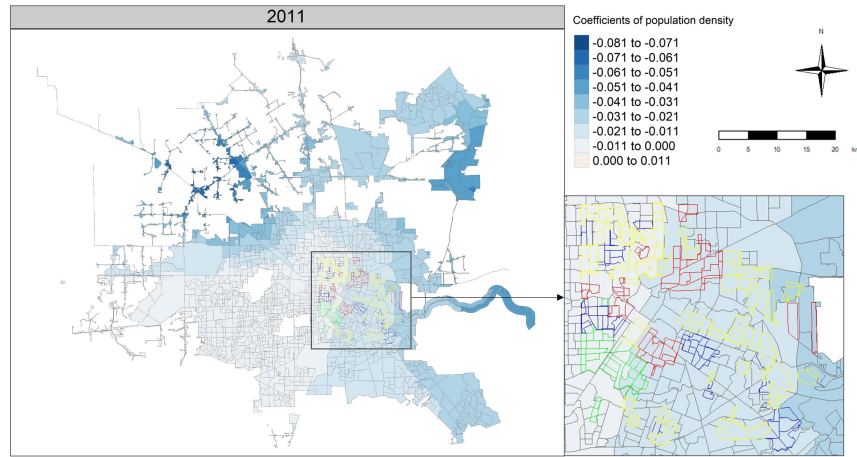
(c) 2017

Figure 4.11: The spatiotemporal distribution of regression coefficient estimates of Hispanic_Pc
 Note: Green boundary means "Best", blue boundary means "Still Desirable", yellow boundary means "Definitely Declining", and Red boundary means "Hazardous"

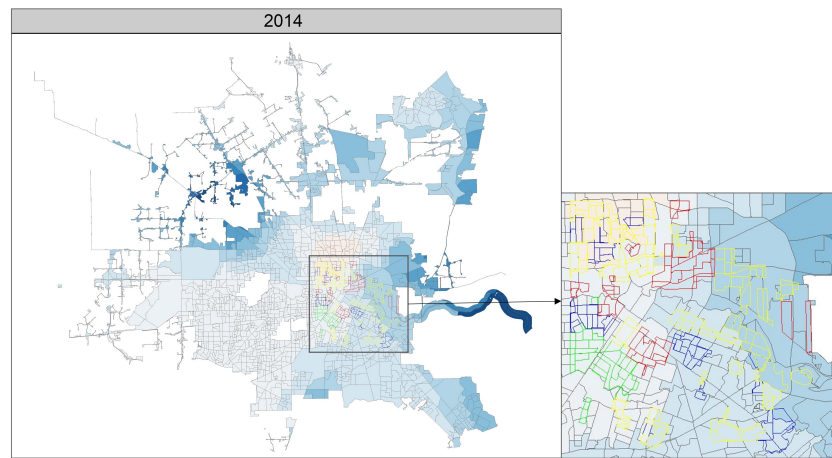
An additional significant finding is that population density exhibits a consistently negative association with toxic risk across the entirety of Houston, including the redlining areas. From Figure 4.12, we can see that the coefficient estimates of the population density ranged from -0.081 to 0 during the period 2011 to 2017. This finding suggests that block groups characterized by high population density are generally associated with a lower level of toxic risk exposure. More specifically, the coefficient estimates for the northern, northeastern, and eastern regions are comparatively smaller than those for other regions. This indicates that when controlling for all other variables, a one-unit increase in population density is associated with a greater reduction in toxic risk in these regions compared to other regions.

The results coincide with the previous Chapter and also with previous studies. In environmental justice analysis, population density is frequently used as a control variable. For example, Li et al. [87] reported a negative correlation between excess emissions and population density. Gilbert and Chakraborty [91] also found a negative association between population density and cancer risk in Jacksonville, but an opposite effect in Tampa, indicating the relationship varies across different areas.

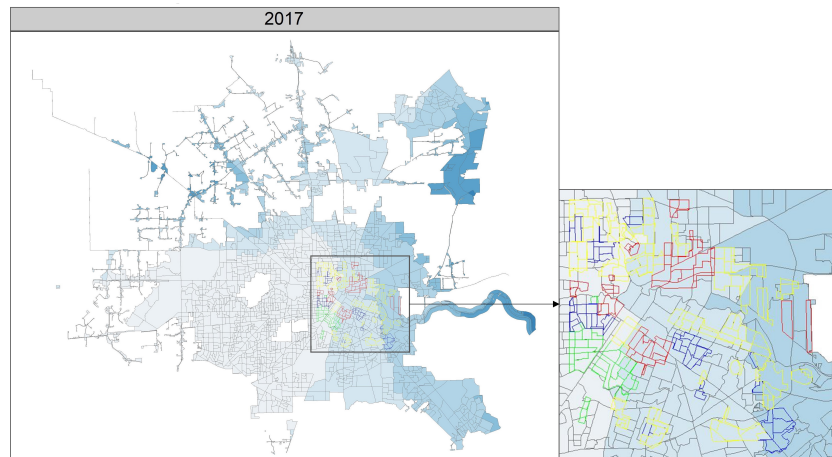
The inverse correlation between population density and toxic chemical risk can be attributed to several factors. Firstly, areas with high population densities tend to have less industrial activity, resulting in fewer sources of toxic chemical exposure [154]. Secondly, areas with high population densities often have more robust environmental regulations, limiting the use of toxic chemicals [155, 156]. Lastly, waste management systems tend to be better in areas with high population densities, reducing the risk of exposure to toxic chemicals [157]. The above factors collectively contribute to the negative relationship between population density and toxic chemical risk, underscoring the need for effective environmental regulation and management.



(a) 2011



(b) 2014

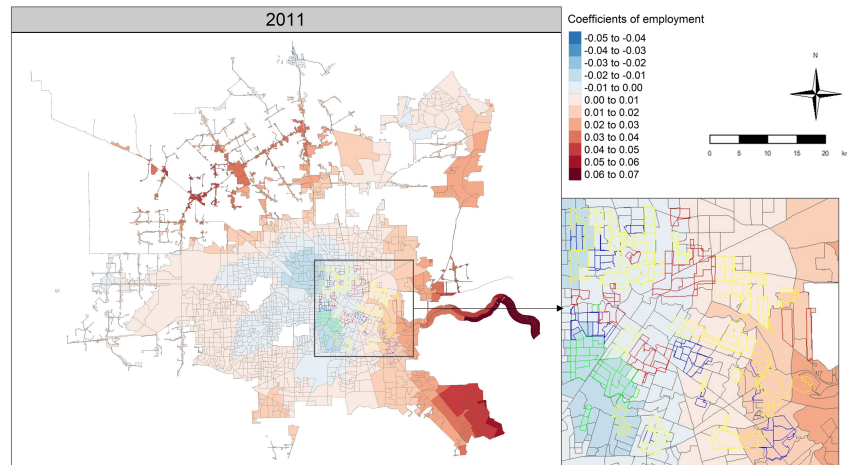


(c) 2017

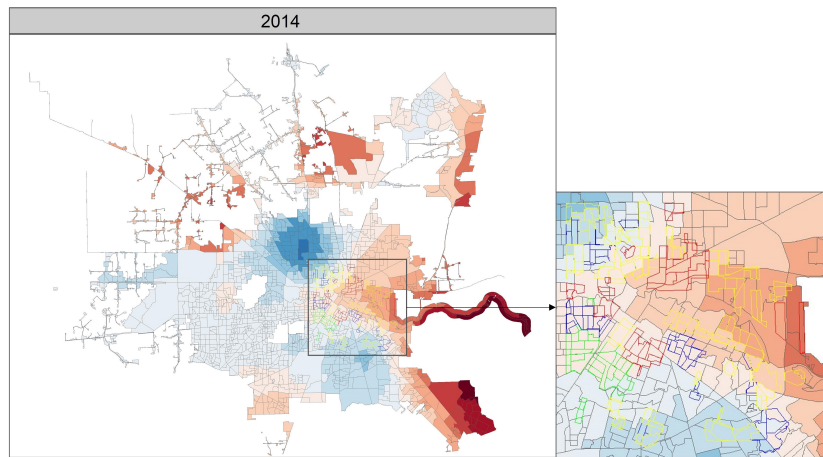
Figure 4.12: The spatiotemporal distribution of regression coefficient estimates of Pop_Dens
 Note: Green boundary means "Best", blue boundary means "Still Desirable", yellow boundary means "Definitely Declining", and Red boundary means "Hazardous"

Figure 4.13 through 4.15 present the coefficient estimates that demonstrate the associations between toxic risk and various socioeconomic characteristics, such as the percentage of unemployment, the percentage of owner-occupied housing, and the percentage of adults without high school diplomas.

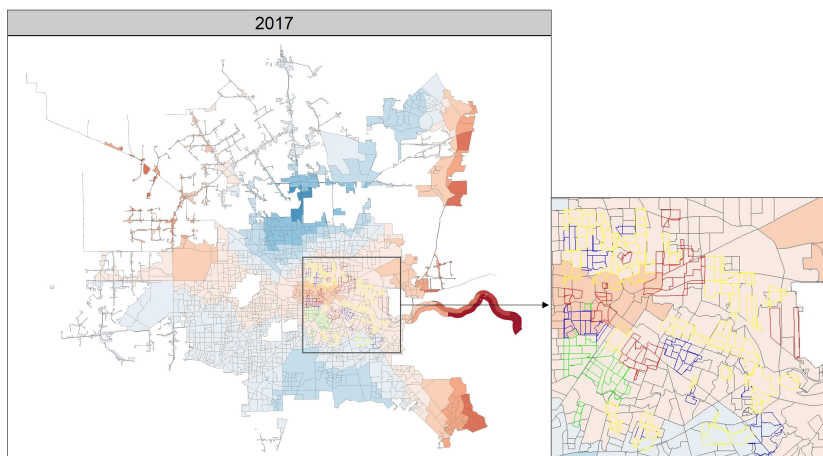
While the average coefficient estimates for the percentage of unemployment appear to be positive between 2011 and 2017, as depicted in Figure 4.6, the spatiotemporal distribution of these estimates reveals both positive and negative associations across regions and over time (Figure 4.13). A noteworthy observation is that in the eastern region, particularly along the Buffalo Bayou river, the percentage of unemployment demonstrates a positive relationship with toxic risk in 2011, 2014, and 2017. The coefficient estimates for the percentage of unemployment in these areas reached 0.08, indicating that when controlling for all other variables, a one-percentage-point increase in unemployment within a block group is associated with an up to 8.33% increase in toxic risk. The figures for 2011 and 2014 (Figures 4.13a and 4.13b) demonstrate that in central Houston, there are negative associations between the percentage of unemployment and toxic risk, indicating that block groups with a high percentage of unemployment are generally associated with low levels of toxic risk. Additionally, the figure for 2017 (Figure 4.13c) shows that negative associations exist between the percentage of unemployment and toxic risk in the southern, southwestern, and northern regions. In the redlining areas, on the other hand, the associations between the percentage of unemployment and toxic risk are positive, meaning that high levels of unemployment within block groups in these areas are generally associated with higher levels of toxic risk.



(a) 2011



(b) 2014

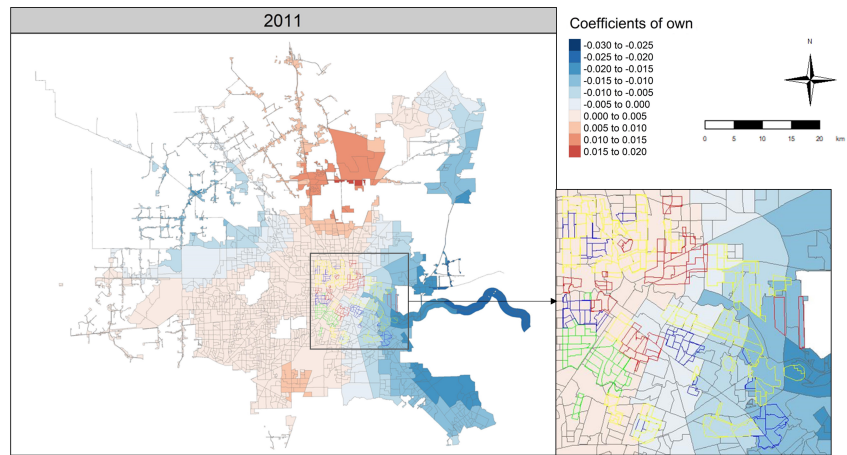


(c) 2017

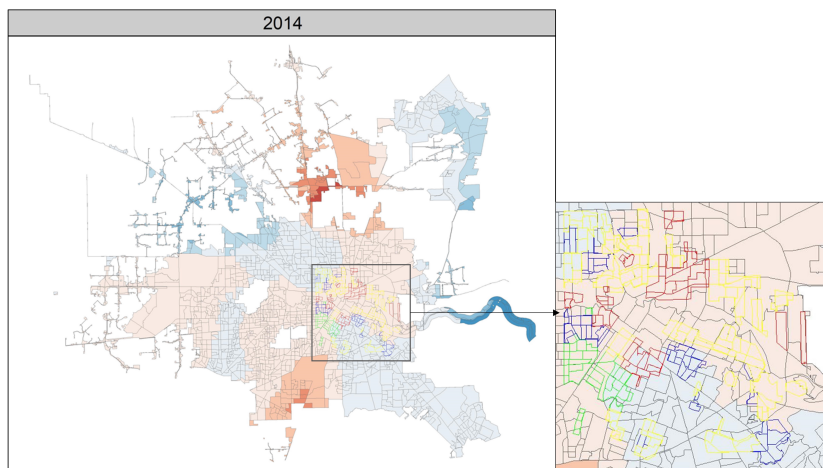
Figure 4.13: The spatiotemporal distribution of regression coefficient estimates of Unemploy_Pc
 Note: Green boundary means "Best", blue boundary means "Still Desirable", yellow boundary means "Definitely Declining", and Red boundary means "Hazardous"

In terms of the percentage of owner-occupied housing (see Figure 4.14), the findings indicate that in both 2011 and 2014, the distribution of spatial coefficient estimates follows a comparable trend. Specifically, in the eastern, southeastern, and northwestern regions, there exists an inverse correlation between the percentage of owner-occupied housing and the prevalence of toxic risk (Figures 4.14a and 4.14b). This suggests that block groups with a greater concentration of owner-occupied housing tend to be associated with lower exposure to toxic risk. In contrast, positive associations are observed between the percentage of owner-occupied housing and toxic risk in other geographical areas. In 2017, a negative association was found between the proportion of owner-occupied housing and toxic risk in the northeastern, northeastern, and southeastern regions, while predominantly positive relationships were observed in redlining areas (Figure 4.14c).

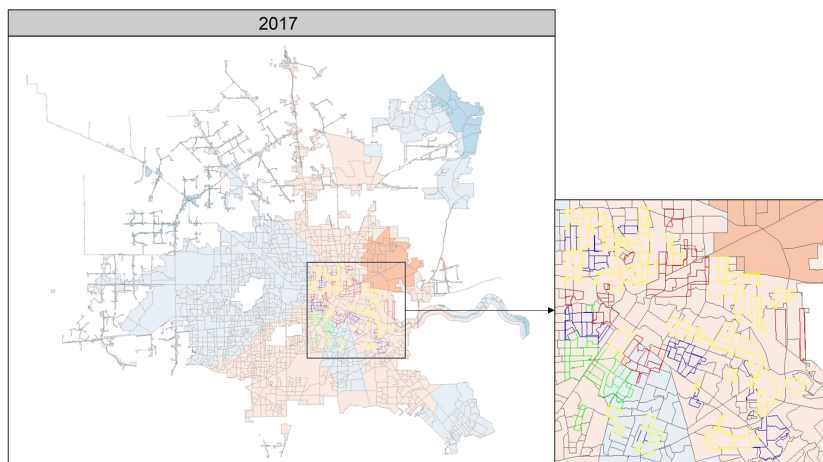
Regarding the education variable, which pertains to the percentage of adults without high school diplomas, the findings demonstrate a heterogeneous pattern of associations with toxic risk, encompassing both positive and negative relationships Figure 4.15. Spanning from 2011 to 2017, the percentage of adults without high school diplomas exhibits negative correlations with toxic risk in the northwestern, northeastern, and eastern areas of Houston. Conversely, predominantly positive associations are observed between the education variable and toxic risk in the northern, central, and southwestern regions of Houston. In redlining areas, particularly those block groups designated with the color code red, a tendency is observed for a positive correlation between the percentage of adults without high school diplomas and toxic risk. This indicates that in these areas, block groups with a higher concentration of adults without high school diplomas tend to exhibit greater exposure to toxic risk.



(a) 2011

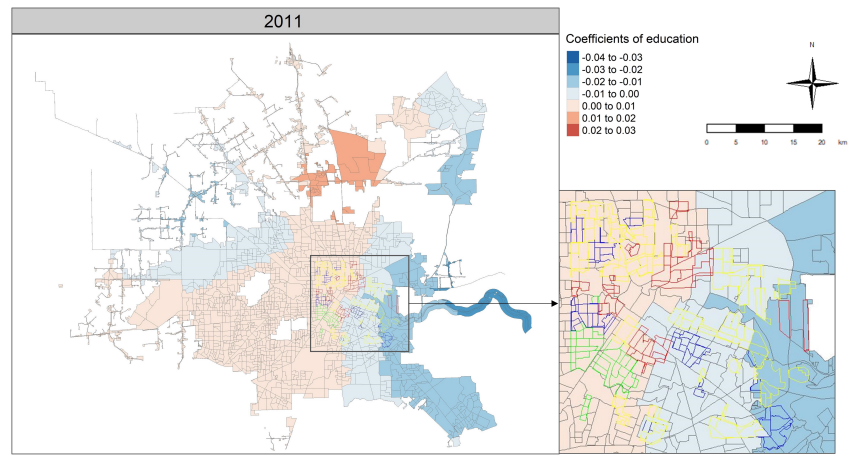


(b) 2014

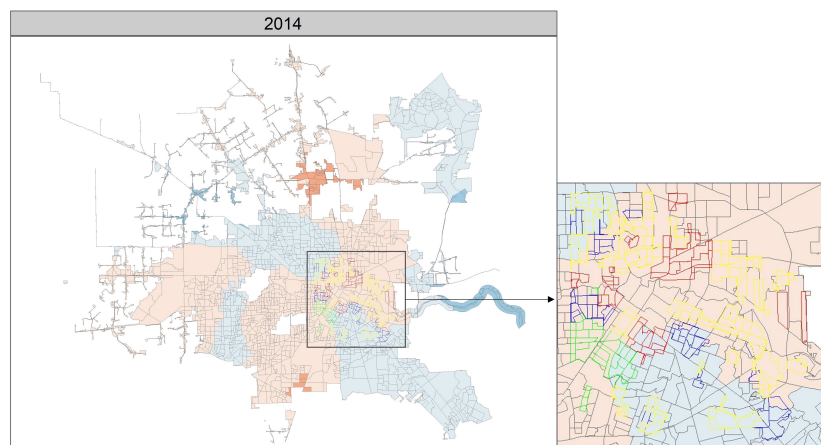


(c) 2017

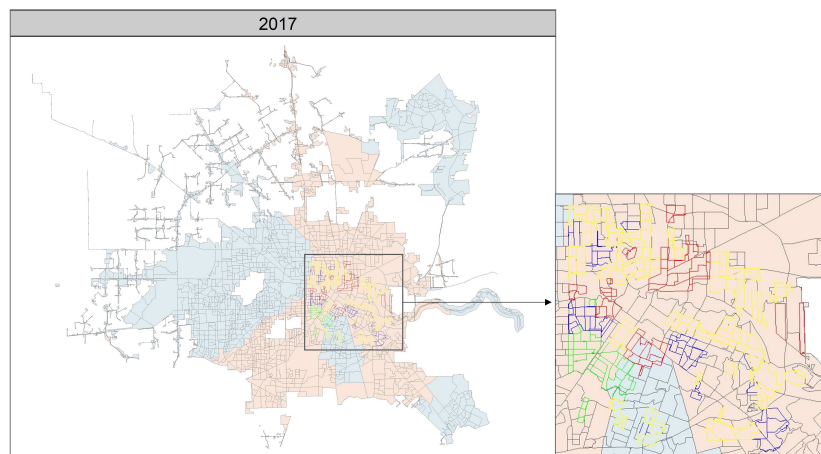
Figure 4.14: The spatiotemporal distribution of regression coefficient estimates of Owner.Pc
 Note: Green boundary means "Best", blue boundary means "Still Desirable", yellow boundary means "Definitely Declining", and Red boundary means "Hazardous"



(a) 2011



(b) 2014



(c) 2017

Figure 4.15: The spatiotemporal distribution of regression coefficient estimates of Not_High_Pc
 Note: Green boundary means "Best", blue boundary means "Still Desirable", yellow boundary means "Definitely Declining", and Red boundary means "Hazardous"

4.4 Conclusions and policy recommendations

The chapter utilized the Geographical and Temporal Weighted Regression (GTWR) model to examine the spatiotemporal variability in the associations between indicators of race/ethnicity and socioeconomic status with toxic risk. This model is capable of simultaneously accounting for spatial autocorrelation across regions and temporal decay effects. The findings indicate that, with the exception of the eastern and western regions of Houston, the proportion of Black residents tends to be positively associated with toxic risk, implying that densely populated racial minority communities are at higher risk of exposure to harmful pollutants. Temporally, the coefficient estimates display an increasing trend in most areas of Houston, Texas, between 2011 and 2014, followed by a decrease from 2014 to 2017. Furthermore, the proportion of Asian residents exhibits a notable reduction in the coefficient estimates from 2011 to 2017, indicating the environmental justice improvement in Houston. The results also suggest that the proportion of individuals identifying as Hispanic or Latino is positively correlated with toxic risk in most areas, including redlining areas, with the notable exception of the region surrounding Aldine, Houston spanning from 2011 to 2017. Regarding the socioeconomic indicators, population density demonstrates a consistently negative association with toxic risk throughout Houston, encompassing the redlining areas. Conversely, other variables such as employment, education, and housing tend to exhibit a combination of positive and negative associations with toxic risk.

This study focused on the block group level environmental inequality. The results provided important information on the spatiotemporal heterogeneity of the various relationships between racial/ethnic and socioeconomic characteristics and environmental negative outcomes. The findings of this study might provide information on policy interventions aimed at improving environmental

conditions in disadvantaged communities to be developed and implemented. These interventions could address environmental inequality and social injustice by taking into account the complex and multifaceted relationships between race, ethnicity, socioeconomic status, and the incidence of adverse environmental outcomes at the block group level. To achieve this goal, collaboration and cooperation between local organizations, such as the Texas Environmental Justice Advocacy Service, and the Texas Commission on Environmental Quality, as well as public officials and industry stakeholders might be important. A concerted effort might be required to identify and prioritize the needs of communities and develop policies that are tailored to address their unique circumstances. By taking a holistic approach and working together, the whole city may promote environmental justice, reduce inequality, and improve the health and well-being of communities that are most vulnerable to environmental harm.

The results of this study could provide valuable insights into enriching the practice of Pollution prevention (P2), which refers to any practice that reduces, eliminates, or prevents the creation of pollution and its release into the natural environment [158]. This chapter has offered comprehensive information regarding the quantitative assessment of the correlations between racial/ethnic and socioeconomic determinants and toxic chemical exposure risk. The findings of this study could augment the quantitative perspectives of the Pollution Prevention (P2) Environmental Justice Facility Mapping Tool. This information could add quantitative perspectives to the P2 Environmental Justice Facility Mapping Tool, a TRI program that is designed to detect and delineate underserved communities with environmental justice (EJ) concerns based on a demographic index. By incorporating the quantitative measures of the relationships between toxic chemical exposure risk and race/ethnicity and socioeconomic status, the P2 tool can provide a more nuanced understanding of the environmental injustices that may be pervasive in underserved communities.

This could allow policymakers and stakeholders to implement more effective and targeted measures to mitigate and prevent environmental harm in these communities.

The temporal trend of environmental justice development over time represents another vital perspective provided by this study. This perspective underscores the dynamic and complex nature of the relationships between environmental justice over time. By analyzing the temporal trends of environmental justice development, policymakers may identify the causes and consequences of environmental injustice and design interventions that are tailored to address the evolving needs of vulnerable communities over time. Furthermore, this perspective can inform the development of more comprehensive and effective policies that account for the intricate interplay between environmental justice and the temporal dynamics of social, economic, and political factors that shape the environmental landscape. For example, the emergence of racial environmental justice disparities could potentially result from the demographic shifts occurring within neighborhoods with varying racial compositions. Therefore, it might be helpful that government officials remain cognizant of these dynamic trends to facilitate the development of appropriate guidelines that reflect the current context.

Although the GTWR model can help provide information on the temporal and spatial effects of racial/ethnic and socioeconomic on toxic risk, there are several limitations that need further studies. Lack of interpretation is one main limitation of the GTWR model. The bandwidth for the GTWR model encompasses both the spatial and temporal dimensions, making it more challenging to interpret intuitively than a single spatial bandwidth, such as distance or the number of neighborhoods. Furthermore, although spatiotemporal matrix weights have been integrated into the GTWR model to enhance the fitting effect, the heterogeneous variable weights have not been adequately considered. This disregard for the potential disparity between each variable

and its corresponding spatiotemporal variation ultimately impacts the model's overall fitting effect [159]. To mitigate the impact of high collinearity among explanatory variables on the model's operation and subsequent analysis results, several variables were excluded prior to model deployment. However, this exclusion can potentially restrict the pool of candidate explanatory variables and introduce bias to the analysis results [160].

Chapter 5: Conclusions

5.1 Summary of major findings and synthesis of contributions

The study of distributive environmental justice examines whether particular racial/ethnic or socioeconomic groups bear a disproportionate burden of negative environmental consequences. This has raised increasing concerns among various stakeholders, including the central government, local government, industrial facilities, and residents. Toxic chemical releases have adverse impacts not only on climate change but also on ecosystems and human health. Analyzing the interregional transfer of toxic chemical releases can yield information on the distribution of responsibility for pollution reduction in a more equitable and just manner, while also taking into account the consumers of these toxic chemicals. Moreover, comprehending the spatial heterogeneity of the relationship between environmental hazards and socioeconomic indices is crucial for understanding the resulting health outcomes at different spatial scales and locations. This, in turn, enables policymakers to develop targeted interventions at the local level that focus on vulnerable groups. By exploring spatiotemporal heterogeneity in the influence of socioeconomic development on toxic risk, it may be possible to uncover variations across regions and over time, thereby facilitating the development of effective solutions for pollution reduction.

Although the existing literature on this topic is expanding, this dissertation aims to address the gaps in current research by examining the variations in environmental inequality across

spatial scales and from both consumption- and production-based perspectives. Previous research mainly focused on TCR in the US from the production perspective, analyzing pollution exposure across regions and demographic groups, and there has been an absence of research analyzing the toxic chemical release transfer embodied in interregional trade and disproportionate exposure of consumption-based TCR faced by different income groups across states in the US. In Chapter 2, we aim to bridge these gaps by first evaluating regional consumption-based assessments of pollution inequalities and the outsourcing of embodied TCR across states in the US. We then proceed to estimate the TCR footprint of various income groups at the state level. Furthermore, while there is ample evidence documenting racial-ethnic and socioeconomic disparities in exposure to environmental hazards in the United States, there has been a dearth of research investigating whether the relationship between negative environmental outcomes and racial, ethnic, or socioeconomic indicators remains consistent across various spatial scales within the US. Chapter 3 filled this gap by utilizing the MGWR model to examine the spatial non-stationarity in the associations between toxic chemical hazard risk and community characteristics, such as race/ethnicity and socioeconomic status, at the census block group level for the year 2017. To the best of my knowledge, no prior study has investigated the relationship between socioeconomic indicators and toxic risk in the United States using a localized model that incorporates a temporal dimension. Therefore, Chapter 4 employed the geographical and temporal weighted regression (GTWR) model to examine the spatiotemporal heterogeneity of the impact of racial disparities and socioeconomic development on toxic risk in Houston to answer the question of whether pollution inequality has improved or worsened over time.

The findings of Chapter 2 indicate that the production of TCR is unevenly distributed across regions in the US, with particular concentration in certain central and southern states, as well as

the Great Lakes Region. Examples of these states include Texas, Indiana, Ohio, among others. This disparity can be partially explained by the primary sectors present in each region. In contrast, the distribution of consumption-based TCR is relatively more equal across regions in the US. The coastal states primarily serve as net importers of embodied TCR, suggesting that they should bear a greater responsibility for reducing TCR under the consumption-based principle as opposed to the production-based principle. Household final demand across all income groups is responsible for the largest portion of embodied TCR, and the consumption of the top 35% of households alone contributes to over 50% of the total TCR in the US. This set of data can provide valuable insight for the government and other stakeholders in developing mitigation policies and actions that address both production-based and consumption-based perspectives.

Chapter 3 finds that vulnerable racial minorities, as measured by the percentage of the Black or Asian population in a census block group, exhibit a consistent positive correlation with toxic risk across Texas, suggesting that regardless of location, these groups face a high level of exposure to toxic chemicals. In contrast, the relationship between toxic risk and socio-economic variables is not constant throughout the state, displaying regional variations in spatial patterns. Specifically, income is significantly negatively associated with toxic risk in the Dallas-Fort Worth-Arlington Metropolitan Statistical Area, indicating that low-income communities are concentrated in areas with higher exposure to toxic chemical risk. The proportion of the population without high school diplomas and the unemployment rate both exhibit positive relationships with toxic risk in the eastern region of Texas, indicating that low socio-economic status groups are concentrated in areas with higher exposure to toxic chemicals in that region. The percentage of vacant housing has a negative association with toxic risk around the city of Beaumont but is positively correlated with toxic risk around the cities of Corpus Christi and Houston. The

relationship between toxic risk and the population share of Hispanic/Latino minorities varies across the state and is only significantly positive in the eastern areas of Texas. Finally, population density has a locally varying relationship with toxic risk, with a significantly negative correlation observed in the northern and eastern regions of Texas.

Chapter 4 findings reveal that, except for the eastern and western regions of Houston, the density of Black residents tends to be positively correlated with toxic risk, indicating that densely populated communities of racial minorities are at higher risk of exposure to harmful pollutants. Temporally, the coefficient estimates indicate an increasing trend in most areas of Houston, Texas, between 2011 and 2014, followed by a decrease from 2014 to 2017. Furthermore, the density of Asian residents exhibits a notable decline in the coefficient estimates from 2011 to 2017, suggesting an improvement in environmental justice in Houston. The results also suggest that the density of individuals identifying as Hispanic or Latino is positively associated with toxic risk in most areas, including redlining areas, with the exception of the region surrounding Aldine, Houston spanning from 2011 to 2017. In terms of socioeconomic indicators, population density consistently demonstrates a negative association with toxic risk throughout Houston, encompassing the redlining areas. However, other variables such as employment, education, and housing tend to exhibit a mix of positive and negative associations with toxic risk.

5.2 Limitations and future studies

This dissertation solely concentrated on the disparities in environmental inequity across temporal dimensions and spatial magnitudes. Starting from the 1900s, the climate justice and energy justice movements garnered increasing recognition, owing to their more expansive objective

of facilitating a just transition. Just transition is a transition away from the fossil-fuel economy to a new economy that provides “dignified, productive, and ecologically sustainable livelihoods; democratic governance; and ecological resilience” [161]. In order to realize a just transition, it is essential to acquire a holistic comprehension of the disparities in environmental justice, climate justice, and energy justice across both spatial and temporal dimensions. In future studies, endeavors should adopt a more comprehensive approach to inform policy initiatives aimed at mitigating inequitable environmental and climate impacts experienced by disadvantaged communities with low socioeconomic status and communities of color. Additionally, such studies should advocate for the equitable provision of clean and affordable energy access for all.

For Chapter 2, there are some limitations that should be addressed in future studies. First, the toxic release inventory only includes the number of TCR but does not include concentration information of chemical substances in air, water, and land. In future studies, the concentration of the toxic chemicals should be estimated in order to achieve more accurate results. Second, TRI does not include all industrial sectors and does not include all facilities in these sectors. TRI only includes large facilities in manufacturing, metal mining, electric power generation, chemical manufacturing, and hazardous waste treatment that meet TRI reporting criteria. Third, this paper does not study the temporal dimensions of pollution inequality. When years of MRIO tables are available, the analysis of embodied TCR trends over time can be conducted and the decomposition analysis can be proposed to investigate the driving forces of changes.

Chapter 3 could be strengthened by addressing some limitations. First, this chapter only used the RSEI score as a measure of toxic risk, which may present a certain extent of uncertainty. RSEI incorporates information from EPA’s TRI database, which tracks certain toxic chemical releases and waste management activities at federal facilities and larger industrial facilities across

the United States and its territories [112]. Other exposures, for example, exposure to transportation and occupational exposure, are needed to be considered in future studies. Second, this chapter conducted a cross-sectional spatial analysis only on the relationships between toxic chemical risks and the distributions of socioeconomic status, and racial and ethnic minorities, further research needs to be done for the spatial-temporal features of environmental inequality across varying geographical scales. Furthermore, there might be an endogeneity problem due to the reverse causality between income and environmental degradation [114]. One way to address this issue is by employing instrumental variables that are correlated with the income variable but not the error term. This enables the estimation of the causal effect of the income variable on the environmental variable with greater consistency. Finally, this chapter used toxic risk data from the RSEI model, which contains the toxic chemical releases, the chemical's toxicity, and human exposure. However, toxic chemical releases have adverse impacts on people's health, from acute poisonings and events to chronic effects like cancer and non-cancer potential [2, 7]. The health impacts of toxic chemical releases need to be considered in future analyses of environmental inequality studies.

While Chapter 4 provides valuable insights into the spatial and temporal analysis of environmental justice studies, there are several limitations that should be addressed in future research. One limitation of this chapter is that it only considers data from the years 2011 to 2017, due to data availability constraints. To extend the study period and investigate more recent trends in environmental justice, it would be necessary to analyze data from additional years. Additionally, the global COVID-19 pandemic has significantly impacted economic and social activities, which may have important implications for environmental injustice. Examining the effects of the pandemic on environmental justice would provide valuable insights and perspectives for future research.

Secondly, the present chapter only examines the associations between racial, ethnic, and socioeconomic attributes and toxic risk, without delving into the underlying mechanisms driving the temporal changes in such disparities. To address this gap, future research should adopt a causal approach, such as employing a difference-in-differences (DiD) model, to investigate the determinants that contribute to the reduction or exacerbation of pollution gaps between groups differentiated by race or socioeconomic status over time. Another limitation of this chapter is that it focuses on only one city, Houston, Texas, which may limit the generalizability of the findings. In future research, it would be beneficial to conduct studies in multiple cities with redlining practices to analyze the spatial and temporal heterogeneity of environmental inequality. By doing so, more precise guidelines for reducing toxic risk could be developed. What's more, the release of toxic chemicals has significant implications for human health [2, 67]. Future research should address the health consequences of such releases, through examination of their spatial and temporal effects on diverse demographic groups. Policymakers can use this analysis to develop targeted and adaptable guidelines.

Appendix A: Appendix for Chapter 3

A.1 The OLS model results

Table A.1: The OLS model results for the spatial stationarity of parameters

	Coefficient estimate	STD	t value	Coefficient estimate	STD	t value	Coefficient estimate	STD	t value	Coefficient estimate	STD	t value
Intercept	0.2684***	0.0237	11.315	0.4228***	0.0733	5.769	0.4114***	0.0541	7.604	0.5352***	0.0854	6.268
BLACK	0.003405***	0.0007	5.023	0.002774***	0.0007	3.848	0.0019***	0.0007	2.592	0.0022***	0.0007	3.021
ASIAN	0.0027**	0.0016	1.772	0.00351**	0.0016	2.214	0.0030*	0.0016	1.923	0.0027*	0.0016	1.688
OTHERS	0.002405**	0.0012	2.023	0.002203*	0.0012	1.850	0.0016	0.0012	1.374	0.0019	0.0012	1.605
HISPANIC	0.001322***	0.0004	3.054	0.000809*	0.0005	1.704	-0.0006	0.0006	-0.995	0.0009*	0.0005	1.826
POP	-0.0001***	0.0000	-14.409	-9.66e-05***	0.0000	-14.428	-0.0001***	0.0000	-14.380	-0.0001***	0.0000	-14.266
INCOME				-9.72e-07***	0.0000	-2.649				-3E-07	0.0000	-0.721
INCOME2												1.27E-06
GINI				-0.1631	0.1396	-1.168				-0.2423	0.1430	-1.695
OWNER							-0.0018***	0.0005	-3.582	-0.0019***	0.0006	-3.305
VACANT							-0.0004	0.0011	-0.326	0.0005	0.0011	0.440
NHSD							0.0045***	0.0010	4.275			
UNEMP							0.0014	0.0022	0.641	0.0021	0.0022	0.953
												0.0022
												1.114
n	15154			14143			15154			15154		
R square	0.01413			0.01463			0.01638			0.01541		
AIC	52655.8			52652.17			52629.13			52646.09		

Note: Signif. codes: 0.001 '***', 0.01 '**', 0.05 '*', 0.1 '+', 1. ' '.

A.2 Testing regression assumptions

A.2.1 Distribution of the dependent variable

We plotted the distribution of the dependent variable in both the original and logged formats.

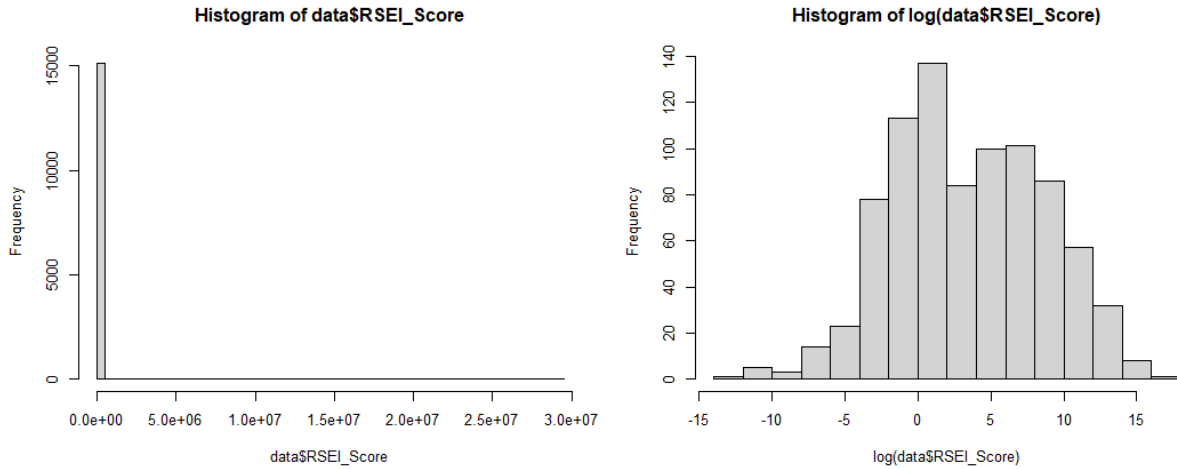


Figure A.1: Distribution of dependent variable

The dependent variable is highly skewed (Figure A.1, left), so in our paper, we used logged format (Figure A.1, right), which is roughly bell-shaped.

A.2.2 Linear relationship

In Figure A.2, the first plot “Residuals vs Fitted” is used to check the assumptions of linearity. If the residuals are spread equally around a horizontal line without distinct patterns (red line is approximately horizontal at zero), that is a good indication of having a linear relationship . In this figure, the residuals are spread equally around the red line, indicating a linear relationship.

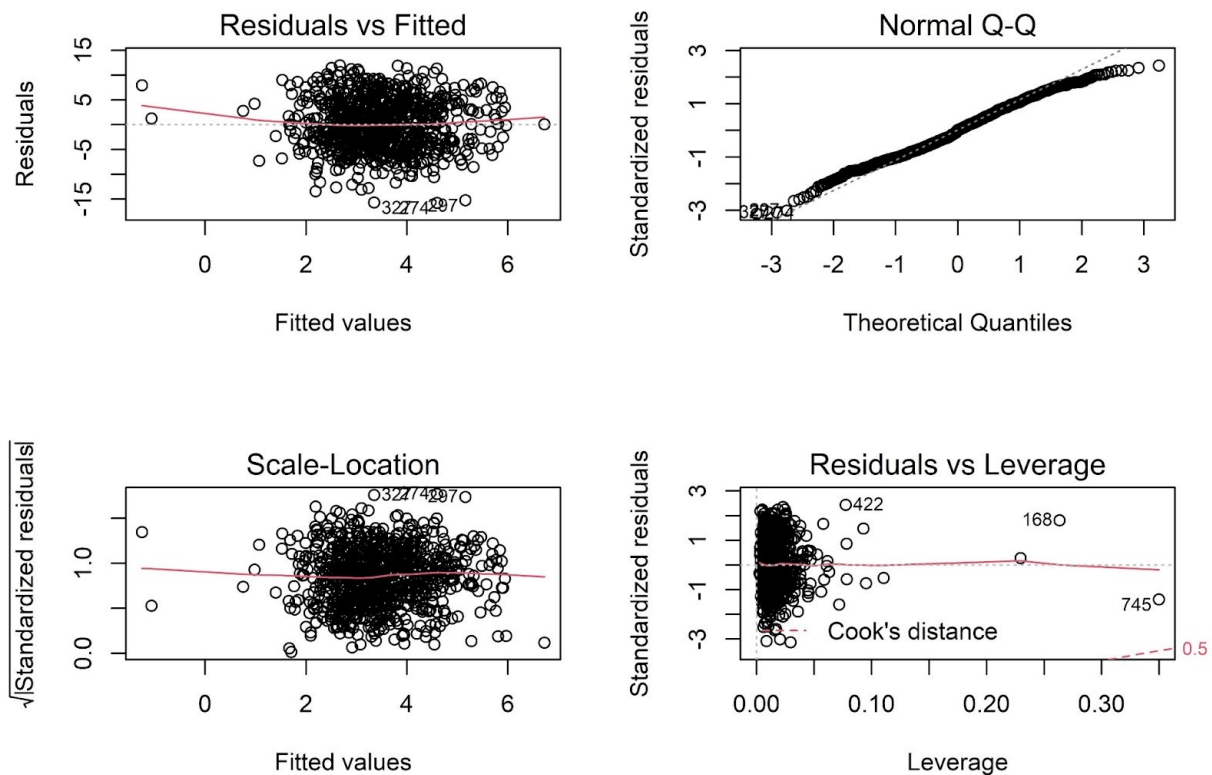


Figure A.2: Regression diagnostics plots

Also, we could use `crPlots` function from `car` package in R to plot partial-residual plots.

Evaluate Nonlinearity

component + residual plot aka partial-residual plots

`library(car)`

`crPlots(simple_model)`

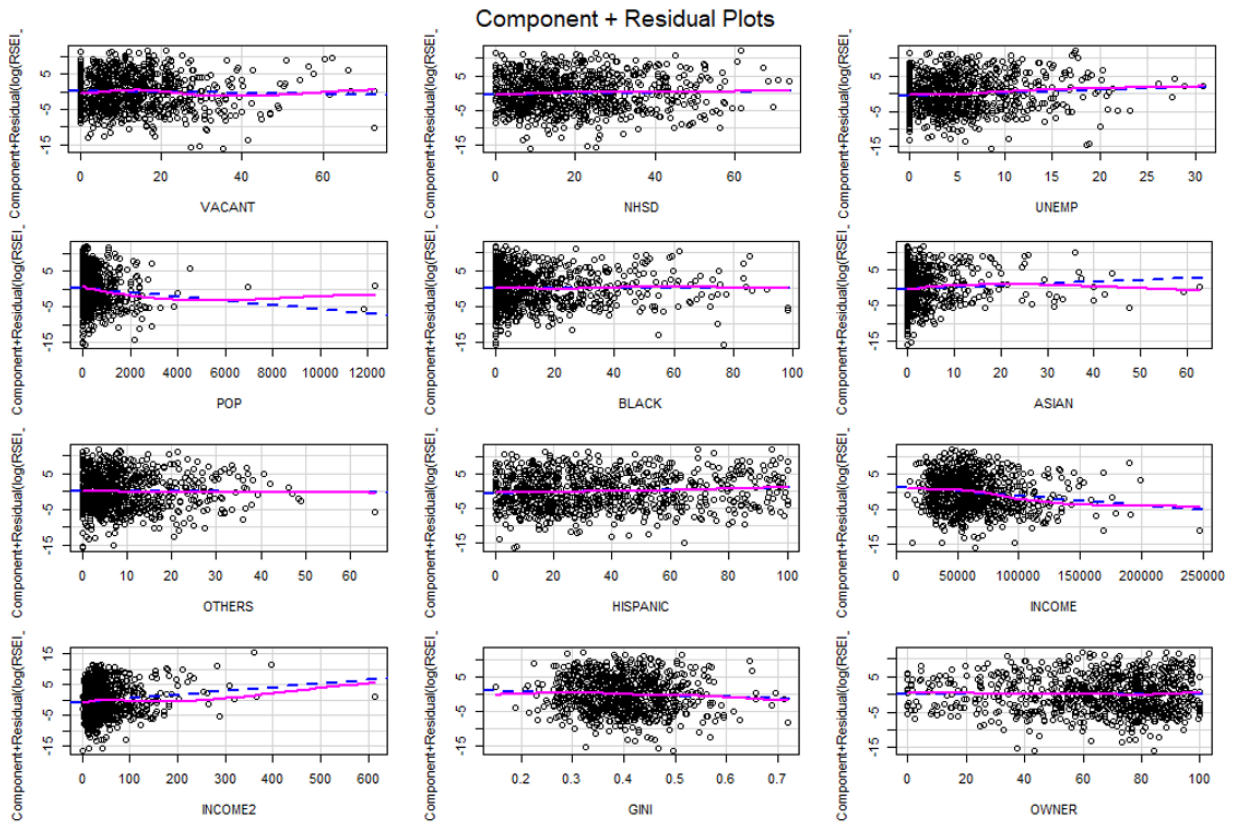


Figure A.3: Partial-residual plots

The blue line shows the expected residuals if the relationship between the predictor and response variable was linear. The pink line shows the actual residuals. The two lines of all the figures are similar, which means linear relationships.

A.2.3 Normality of residuals

In Figure A.2 “Residual vs. Fitted”, because the residuals are close to zero and the trend line is relatively flat (first plot), we can assume that the residuals follow normal distribution. Figure A.2 “Normal Q-Q” is used to check the normality of residuals assumption. If the majority of the residuals follow the straight dashed line, then the assumption is fulfilled. In our Figure, most of the points are located on the diagonal line, showing the normality of residuals.

Also, a histogram of residuals is plotted to show the distribution of residuals.

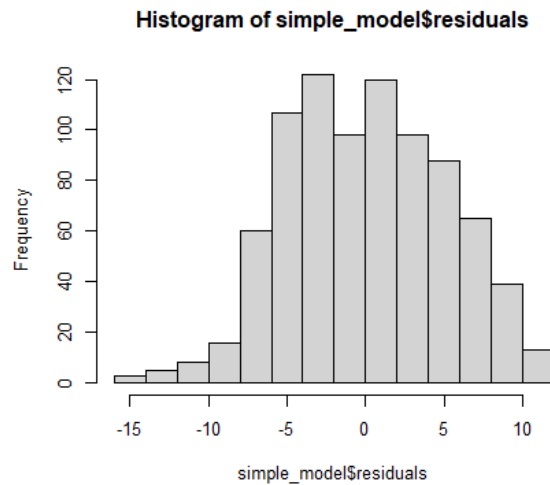


Figure A.4: Histogram of residuals

The histogram is close to bell-shaped and not radically skewed, and it is centered to zero, indicating a normal distribution of residuals.

A.2.4 Testing the Homoscedasticity Assumption

Figure A.2 “Scale-Location” is used to check the homoscedasticity of residuals (equal variance of residuals). If the residuals are spread randomly and the see a horizontal line with equally (randomly) spread points, then the assumption is fulfilled. In this figure, the residuals are spread randomly along the horizontal line, indicating the homoscedasticity of residuals.

We also performed a Breusch-Pagan Test using `bptest` function from `lmtest` package.

```
load lmtest library
```

```
library(lmtest)
```

```
lmtest::bptest(simple_model)
```

studentized Breusch-Pagan test

data: simple_model

BP = 8.3262, df = 12, p-value = 0.7591

The test statistic is 8.3262 and the corresponding p-value is 0.7591. Since the p-value is not less than 0.05, we fail to reject the null hypothesis. We do not have sufficient evidence to say that heteroscedasticity is present in the regression model.

A.2.5 Multicollinearity

In order to conduct multicollinearity test, we need to first assume there is no multicollinearity, and fit the linear model as follow:

```
Call:
lm(formula = Log_RSEI_score ~ POP + BLACK + ASIAN + OTHERS +
    HISPANIC + INCOME + INCOME2 + GINI + OWNER + VACANT + NHSD +
    UNEMP, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-0.8220 -0.3355 -0.2511 -0.1337  16.6753

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.526e-01  1.039e-01   3.394  0.00069 ***
POP          -1.073e-04  7.422e-06 -14.457 < 2e-16 ***
BLACK        2.154e-03  7.524e-04   2.863  0.00420 **
ASIAN        2.494e-03  1.609e-03   1.550  0.12124
OTHERS       1.364e-03  1.198e-03   1.139  0.25490
HISPANIC     -4.982e-04  5.880e-04  -0.847  0.39689
INCOME       3.378e-06  1.234e-06   2.737  0.00620 **
INCOME2     -1.372e-03  5.020e-04  -2.732  0.00630 **
GINI        -2.005e-01  1.453e-01  -1.380  0.16769
OWNER       -2.440e-03  5.813e-04  -4.197  2.72e-05 ***
VACANT       4.838e-04  1.132e-03   0.428  0.66899
NHSD        5.586e-03  1.135e-03   4.920  8.75e-07 ***
UNEMP       2.254e-03  2.214e-03   1.018  0.30869
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.373 on 15141 degrees of freedom
Multiple R-squared:  0.01712,    Adjusted R-squared:  0.01634
F-statistic: 21.98 on 12 and 15141 DF,  p-value: < 2.2e-16
```

Figure A.5: Lineal model results assuming no multicollinearity

From Figure A.5, the R-square is 0.1634, and F-value is high enough implying that all the explanatory variables together significantly explain Log_RSEI_score. However, 6 of the total variables are not significant.

To check the multicollinearity problem, we first used pair-wise correlation to test it. First look at the pair-wise correlation among the explanatory variables (Figure A.6). The correlation matrix shows that the pair-wise correlation among all the explanatory variables are not very high, except for the pairs INCOME-INCOME2 (0.946, structural multicollinearity), NHSD-HISPANIC (0.702, data multicollinearity), and NHSD-INCOME (0.535, data multicollinearity).

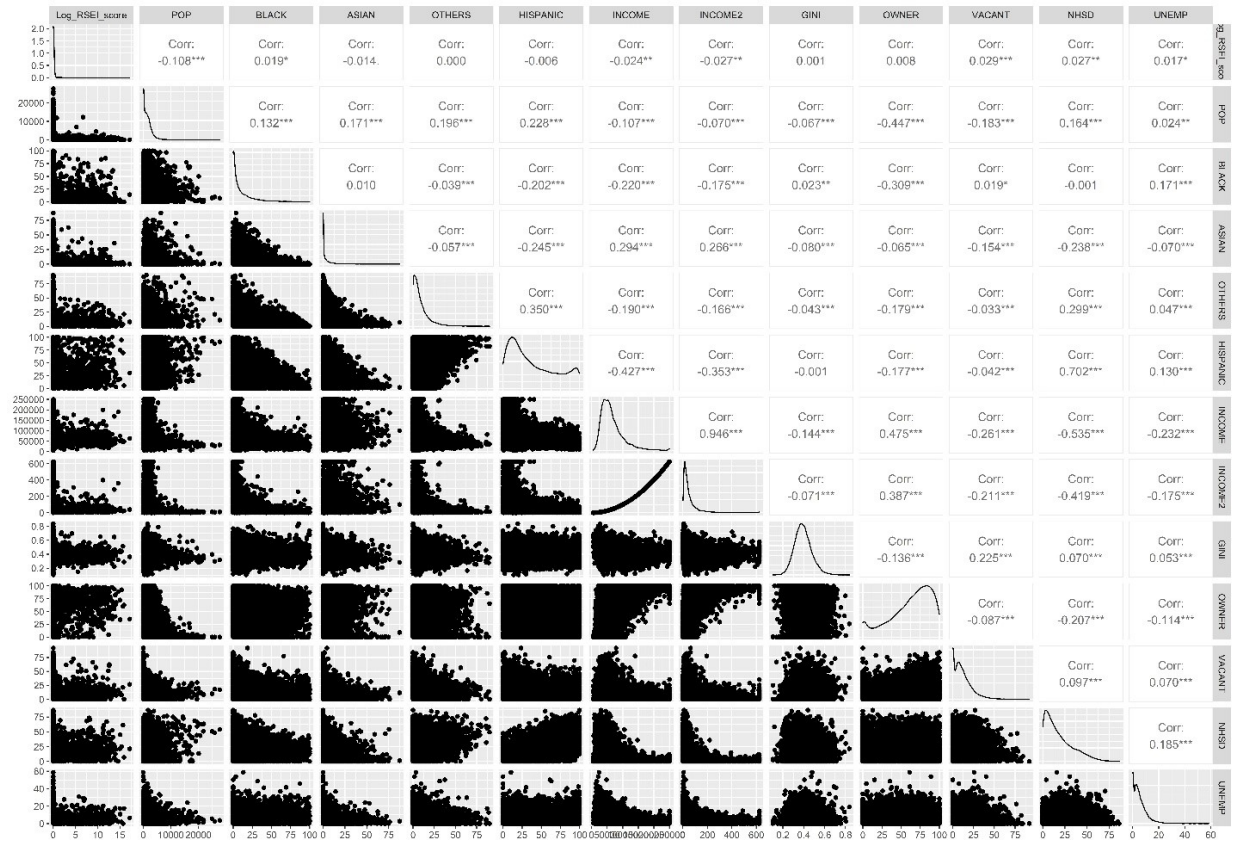


Figure A.6: Pair-wise correlation

Again, by looking at the partial correlation coefficient matrix among the variables (Figure A.7), it's also clear that the partial correlation between INCOME-INCOME2, NHSD-HISPANIC,

and NHSD-INCOME are quite high.

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13
1	1.00000000	-0.11668768	0.02326210	0.012593037	0.009252662	-0.006884975	0.02224066	-0.02219672	-0.011212262	-0.034091131	0.003474591	0.039950499	0.008272911
2	-0.116687685	1.00000000	0.08372814	0.141458518	0.070960636	0.156003006	0.06148593	-0.01455975	-0.072779511	-0.405489446	-0.148844434	0.042463585	-0.020081574
3	0.023262097	0.08372814	1.00000000	-0.043245477	-0.022351471	-0.347881426	-0.16886629	0.11463967	-0.060410482	-0.163664039	-0.101886579	0.065029935	0.134922737
4	0.012593037	0.14145852	-0.04324548	1.000000000	-0.004722851	-0.143948990	0.11949231	-0.04696794	-0.032822594	-0.163762932	-0.063217353	0.014082165	0.017636448
5	0.009252662	0.07096064	-0.02235147	-0.004722851	1.000000000	0.161657079	0.04352415	-0.04467209	-0.046845556	-0.088534055	-0.007858513	0.086782095	-0.002066004
6	-0.006884975	0.15600301	-0.34788143	-0.143948990	0.161657079	1.000000000	-0.09968643	0.04812409	-0.050985163	-0.009469043	-0.170226576	0.536063627	0.040795904
7	0.022240662	0.06148593	-0.16886629	0.119492313	0.043524145	-0.099686432	1.000000000	0.93493136	-0.172294042	0.341067462	-0.182638379	-0.348228027	-0.133011079
8	-0.022196722	-0.01455975	0.11463967	-0.046967941	-0.044672092	0.048124087	0.93493136	1.000000000	0.178584732	-0.211581404	0.119558924	0.270762784	0.105714651
9	-0.011212262	-0.07277951	-0.06041048	-0.032822594	-0.046845556	-0.050985163	-0.17229404	0.17858473	1.000000000	-0.090825330	0.135943442	-0.003297866	0.003045744
10	-0.034091131	-0.40548945	-0.16366404	-0.163762932	-0.088534055	-0.009469043	0.34106746	-0.21158140	-0.090825330	1.000000000	-0.028393796	0.147392659	0.039008746
11	0.003474591	-0.14884443	-0.10188658	-0.063217353	-0.007858513	-0.170226576	-0.18263838	0.11955892	0.135943442	-0.028393796	1.000000000	0.041715846	0.006862803
12	0.039950499	0.04246358	0.06502994	0.014082165	0.086782095	0.536063627	-0.34822803	0.27076278	-0.003297866	0.147392659	0.041715846	1.000000000	0.019385007
13	0.008272911	-0.02008157	0.13492274	0.017636448	-0.002066004	0.040795904	-0.13301108	0.10571465	0.003045744	0.039008746	0.006862803	0.019385007	1.000000000

Figure A.7: Partial correlation coefficient matrix

The pair-wise correlation has shown the basic results of multicollinearity and indicated that the possible source if multicollinearity might be INCOME-INCOME2, NHSD-HISPANIC, and NHSD-INCOME relationships. However, pair-wise correlation between the explanatory variables may be considered as the sufficient, but not the necessary condition for the multicollinearity. Next, we used Farrar – Glauber Test (F-G test) to deal with the problem of multicollinearity (<https://datascienceplus.com/multicollinearity-in-r/>). For overall diagnostic checking, the results can be seen as follow:

```
call:
omcdiag(mod = model)

Overall Multicollinearity Diagnostics

Determinant |X'X|:      MC Results detection
Farrar Chi-Square:  76260.5998      1
Red Indicator:      0.2549          0
Sum of Lambda Inverse:  44.8907      0
Theil's Method:      4.6492          1
Condition Number:      34.0148      1

1 --> COLLINEARITY is detected by the test
0 --> COLLINEARITY is not detected by the test
```

Figure A.8: Overall diagnostic checking results

The value of the determinant is very small (0.0065) and the Chi-Square is very high (Figure A.8), indicating that the detection of collinearity by the overall test.

Besides, for individual diagnostic checking, the results can be seen as follow:

```
Call:
lmcdiag(mod = model)

All Individual Multicollinearity Diagnostics Result
```

	VIF	TOL	Wi	Fi	Leamer	CVIF	Klein	IND1	IND2
POP	1.4617	0.6841	635.5570	699.1589	0.8271	1.4593	1	5e-04	0.7835
BLACK	1.3458	0.7430	476.0302	523.6678	0.8620	1.3436	1	5e-04	0.6374
ASIAN	1.2298	0.8132	316.3050	347.9585	0.9017	1.2277	1	6e-04	0.4635
OTHERS	1.1797	0.8477	247.3379	272.0896	0.9207	1.1777	1	6e-04	0.3778
HISPANIC	2.5164	0.3974	2087.4009	2296.2926	0.6304	2.5122	1	3e-04	1.4948
INCOME	16.9283	0.0591	21925.9850	24120.1764	0.2430	16.9002	1	0e+00	2.3340
INCOME2	12.3520	0.0810	15626.5808	17190.3741	0.2845	12.3315	1	1e-04	2.2798
GINI	1.1249	0.8890	171.8626	189.0613	0.9429	1.1230	1	6e-04	0.2753
OWNER	1.8951	0.5277	1232.1086	1355.4089	0.7264	1.8919	1	4e-04	1.1716
VACANT	1.2251	0.8162	309.8958	340.9079	0.9035	1.2231	1	6e-04	0.4558
NHSD	2.5320	0.3949	2108.8791	2319.9202	0.6284	2.5278	1	3e-04	1.5009
UNEMP	1.1000	0.9091	137.6241	151.3965	0.9535	1.0982	1	7e-04	0.2255

```

1 --> COLLINEARITY is detected by the test
0 --> COLLINEARITY is not detected by the test

ASIAN , OTHERS , HISPANIC , GINI , VACANT , UNEMP , coefficient(s) are non-significant m
ay be due to multicollinearity

R-square of y on all x: 0.0171

* use method argument to check which regressors may be the reason of collinearity
=====

```

Figure A.9: Individual diagnostic checking results

The degree of freedom is (12, 15144), the theoretical value of F at 5% level of significance is 1.75. The F-statistic for the variable INCOME and INCOME2 is very high (16.9283 and 12.3520), followed by NHSD (2.5320) and HISPANIC (2.5164) (Figure A.9).

Finally, in order to exam the pattern of multicollinearity, we conducted t-test for correlation coefficient.

<code>> pcor(X, method = "pearson")</code>														
sestimate														
Log_RSEI_score	0.00000000	-0.11668768	0.02326210	0.012593037	0.009252662	-0.006884975	0.02224066	-0.02219672	-0.01121262	-0.034091131	0.003474591	0.039950499	0.008272911	
POP	-0.11668768	1.00000000	0.08372814	0.141458318	0.070960636	0.156003006	0.06148593	-0.01455975	-0.072779511	-0.405489446	-0.148844434	0.042463585	-0.020081574	
BLACK	0.02326210	0.08372814	1.00000000	-0.043245477	-0.02331471	-0.143948990	0.11463967	-0.060410482	-0.163664039	-0.101886579	0.065029935	-0.134022737		
ASIAN	0.012593037	0.14145832	-0.04324548	1.000000000	-0.004722851	-0.143948990	0.11949231	-0.04696794	-0.032822594	-0.163762932	-0.063217353	0.014082165	0.017636448	
OTHERS	0.009252662	0.07096064	-0.0233147	-0.004722851	1.000000000	0.161657079	0.04352415	-0.04467209	-0.046845556	-0.088534055	-0.007858513	0.086782095	-0.002066004	
HISPANIC	-0.006884975	0.15600301	-0.14394899	-0.143948990	0.161657079	1.000000000	-0.09968643	0.04812409	-0.050985163	-0.009469043	-0.170226576	0.536063627	0.040795904	
INCOME	0.022240662	0.06148593	-0.16366404	0.119492313	0.043524145	-0.099686432	1.000000000	0.93493136	-0.172294042	0.341067462	-0.182638779	-0.348228027	-0.133011078	
INCOME2	-0.022196722	-0.01455975	0.11463967	-0.046967941	-0.044672092	0.048124087	0.93493136	1.000000000	0.178584732	-0.211581404	0.119558924	0.270762784	0.105714651	
GINI	-0.011212622	-0.07277951	-0.06041048	-0.032822594	-0.046845556	-0.050985163	-0.17229404	0.17858473	1.000000000	-0.090825330	0.135943442	-0.003297866	0.003045744	
OWNER	-0.034091131	-0.40548945	-0.16366404	-0.163762932	-0.088534055	-0.009469043	0.34106746	-0.21158140	-0.090825330	1.000000000	-0.028393796	0.147392659	0.39008746	
VACANT	0.003474591	-0.14884443	-0.10188658	-0.06502994	-0.007858513	-0.170226576	0.536063627	-0.34822803	0.27076278	-0.003297866	1.000000000	0.041715846	0.006862803	
NHSD	0.039950499	0.04246358	0.06502994	0.014082165	0.086782095	0.536063627	-0.34822803	0.27076278	0.147392659	0.041715846	1.000000000	0.019385007	1.000000000	
UNEMP	0.008272911	-0.02008157	0.13402274	0.017636448	-0.002066004	0.040795904	-0.13301108	0.10571465	0.003045744	0.039008746	0.006862803	0.019385007	1.000000000	
sp.value														
Log_RSEI_score	0.00000000	4.657057e-47	4.200315e-03	1.212391e-01	2.548975e-01	3.968927e-01	6.200616e-03	6.303321e-03	1.676865e-01	2.716438e-05	6.689872e-01	8.754830e-07	3.086912e-01	
POP	4.657057e-47	0.00000000	0.686649e-25	1.590637e-68	2.280146e-18	0.094443e-83	3.652424e-14	7.319304e-02	3.042180e-19	0.000000e+00	9.509831e-76	1.719444e-07	1.346524e-02	
BLACK	4.200315e-03	0.686649e-25	0.000000e+00	1.016746e-07	5.948273e-03	0.000000e+00	2.921049e-97	1.796795e-45	1.006029e-13	2.130306e-91	3.105588e-36	1.145215e-15	1.890986e-62	
ASIAN	1.212391e-01	1.590637e-68	1.016746e-07	0.000000e+00	5.611497e-01	6.433840e-71	2.813462e-49	7.362351e-09	5.349644e-05	1.654959e-91	6.890298e-15	8.312225e-02	2.998592e-02	
OTHERS	2.548975e-01	2.280146e-18	5.948273e-03	5.611497e-01	0.000000e+00	3.457481e-89	8.412818e-08	3.808685e-08	8.052334e-09	9.744442e-28	3.355444e-01	1.035047e-26	7.993292e-01	
HISPANIC	3.968927e-01	0.094443e-83	0.000000e+00	6.433840e-71	3.457481e-89	0.000000e+00	9.431685e-35	3.124025e-09	3.437487e-10	2.439538e-01	7.954905e-99	0.000000e+00	5.116456e-07	
INCOME	6.200616e-03	3.652424e-14	2.921049e-97	2.813462e-49	8.412818e-08	9.431685e-35	0.000000e+00	0.000000e+00	3.138968e-101	0.000000e+00	1.016907e-113	0.000000e+00	9.945705e-61	
INCOME2	6.303321e-03	7.319304e-02	1.796795e-45	7.362351e-09	3.808685e-08	1.224025e-09	0.000000e+00	0.000000e+00	9.854173e-109	8.024547e-153	2.488218e-49	1.115685e-252	6.838656e-39	
GINI	1.676865e-01	3.042180e-19	1.006029e-13	5.349644e-05	8.052334e-09	3.437487e-10	3.138968e-101	9.854173e-109	0.000000e+00	4.127701e-29	2.226344e-63	6.848949e-01	7.078317e-01	
OWNER	2.716438e-05	0.000000e+00	2.130306e-91	1.654959e-91	9.744442e-28	2.439538e-01	0.000000e+00	8.024547e-153	4.127701e-29	0.000000e+00	4.750066e-04	2.679892e-74	1.573034e-06	
VACANT	6.689872e-01	9.509831e-76	3.105588e-36	6.890298e-15	3.355444e-01	7.954905e-99	1.016907e-113	2.488218e-49	2.226344e-63	4.750066e-04	0.000000e+00	2.817931e-07	3.984150e-01	
NHSD	8.754830e-07	1.719444e-07	1.145215e-15	8.312225e-02	1.035047e-26	0.000000e+00	0.000000e+00	1.115685e-252	6.848949e-01	2.679892e-74	2.817931e-07	0.000000e+00	1.705657e-02	
UNEMP	3.086912e-01	1.346524e-02	1.890986e-62	2.998592e-02	7.993292e-01	5.116456e-07	9.945705e-61	6.838656e-39	7.078317e-01	1.573034e-06	3.984150e-01	1.705657e-02	0.000000e+00	
sstatistic														
Log_RSEI_score	0.00000000	-14.457037	2.863147	1.5496807	1.1385774	-0.8472078	2.737363	-2.731952	-1.3797418	-4.197311	0.4275467	4.9197874	1.0180064	
POP	-14.4570371	0.0000000	10.338949	17.583109	8.7336861	-19.4339157	7.580109	-1.791750	-8.9792419	-54.583736	-18.5214420	5.2298091	-2.4715114	
BLACK	2.8631472	10.338949	0.000000	-5.326286	-2.7510083	-45.6582637	-21.081543	14.199888	-7.4470367	-20.413917	-12.6026020	8.0188269	16.7532852	
ASIAN	1.5496807	17.583109	-5.326286	0.0000000	-0.5811474	-17.8991641	14.809490	-5.785733	-4.0409572	-20.426592	-7.7944082	1.7329650	2.1704809	
OTHERS	1.1385774	8.733686	-2.751008	-0.5811474	0.0000000	20.1568277	5.360672	-5.502339	-5.7706235	-10.936954	-0.9670102	10.7188691	-0.2542198	
HISPANIC	-0.8472078	19.433916	45.658264	-17.8991641	20.1568277	0.0000000	-12.327698	5.9284787	-6.2818318	-1.165206	-21.2564093	78.1375529	5.0240684	
INCOME	2.7373633	7.580109	-21.081543	14.8094904	-5.3606725	-12.3276984	0.0000000	324.218927	-21.5226259	44.644892	-22.8578942	-45.7100278	-16.5135804	
INCOME2	-2.7319524	-1.791750	14.199888	-5.7857325	-5.5023390	5.9284787	324.218927	0.0000000	22.3336557	-26.637904	14.8178657	34.6098403	13.0813592	
GINI	-1.3797418	-8.979242	-7.447037	-0.4049572	-5.7706235	-6.2818318	-21.522426	22.333656	0.0000000	-11.222329	16.8844182	-0.4038006	0.3747768	
OWNER	-4.1973115	-54.583736	-20.413917	-0.2465922	-10.9369545	-1.1652064	44.644892	-26.637904	-11.2223290	0.0000000	-3.4952310	18.3367585	4.8036345	
VACANT	0.4275467	-18.521442	-12.602602	-7.7944082	-0.9670102	-21.2564093	-22.857894	14.817866	16.8844182	-3.495231	0.0000000	5.1375557	0.8444793	
NHSD	4.9197874	5.229809	8.018827	1.7329650	10.7188691	78.1375529	-45.710028	34.609840	-0.4038006	18.336759	5.1375557	0.0000000	2.3857496	
UNEMP	1.0180064	-2.471511	16.755285	2.1704809	-0.2542198	5.0240684	-16.513580	13.081359	0.3747768	4.803634	0.8444793	2.3857496	0.0000000	

Figure A.10: T-test for correlation coefficient results

Note: n=15154, gp=11, method="Pearson".

Again, the high partial correlation between “INCOME-INCOME2”, “NHSD-HISPANIC”, and “NHSD-INCOME” are found to be highly statistically significant (Figure A.10).

For remission measure: exclude INCOME2 and NHSD.

A.2.6 Moran’s I analysis for dependent variable

Hypothesis: the RSEI score values are randomly distributed across block groups following a completely random process. There are two methods for testing this hypothesis: an analytical method and a Monte Carlo method.

Analytical method

To run the Moran’s I analysis using the analytical method, we used the `moran.test` function.

```
moran.test(map_and_data$Log_RSEI_score, lw, alternative="less")
```

Moran I test under randomisation

data: map_and_data\$Log_RSEI_score

weights: lw

Moran I statistic standard deviate = 22.247, p-value < 2.2e-16

alternative hypothesis: greater

sample estimates:

<i>Moran I statistic</i>	<i>Expectation</i>	<i>Variance</i>
<i>1.024675e-01</i>	<i>-6.325111e-05</i>	<i>2.124033e-05</i>

P value is less than 2.2e-16, therefore, we reject the null, meaning the RSEI score values are strongly clustered.

Monte Carlo Method

The analytical approach to Moran's I analysis benefits from being fast. But it may be sensitive to irregularly distributed polygons. A safer approach to hypothesis testing is to run an MC simulation using the `moran.mc()` function.

```
MC <- moran.mc(map_and_data$Log_RSEI_score, lw, nsim=999, alternative="greater")
```

MC

Monte-Carlo simulation of Moran I

data: map_and_data\$Log_RSEI_score

weights: lw

number of simulations + 1: 1000

statistic = 0.10247, observed rank = 1000, p-value = 0.001

alternative hypothesis: greater

The MC simulation generates a very small p-value, 0.001.

Plot the Null distribution (note that this is a density plot instead of a histogram)

plot(MC)

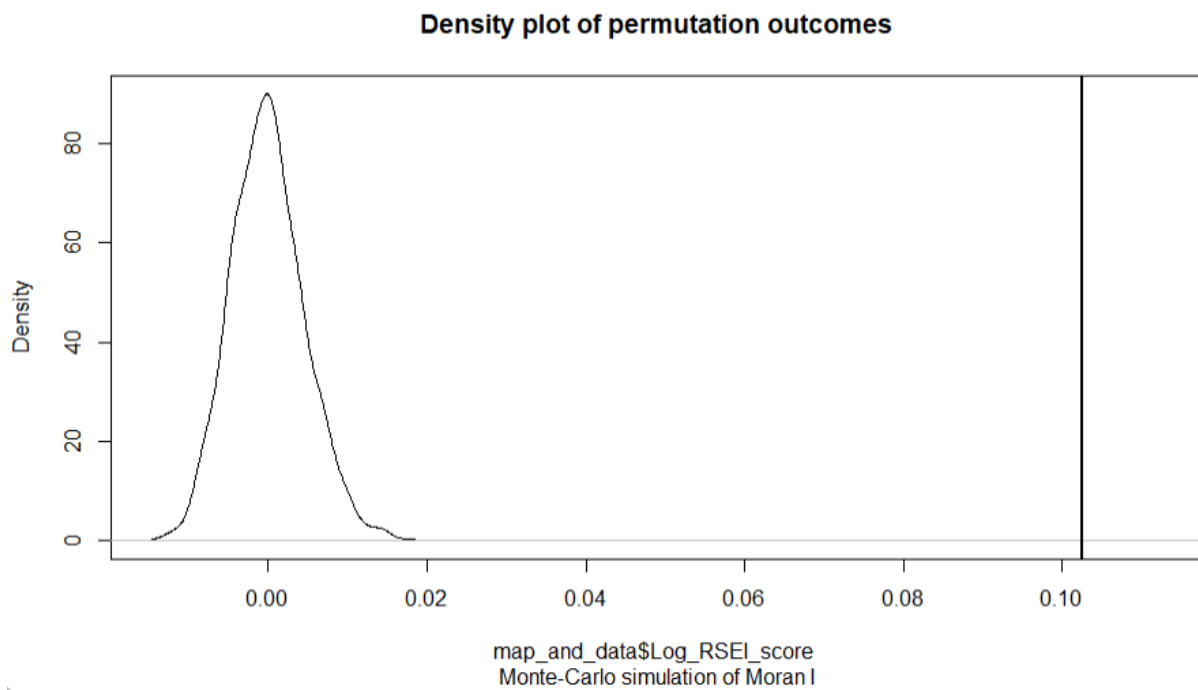


Figure A.11: Monto-Carlo simulation of Moran's I

Our observed statistic, 0.10247, falls way to the right of the distribution suggesting that the RSEI score values are clustered (a positive Moran's I value suggests clustering whereas a negative Moran's I value suggests dispersion).

A.2.7 Statistical tests for MGWR results

We used Monte Carlo simulation of Moran's I to determine if there is spatial non-stationary of MGWR's coefficients. Our observed statistics falls way to the right of the distribution suggesting

that the coefficient values are clustered.

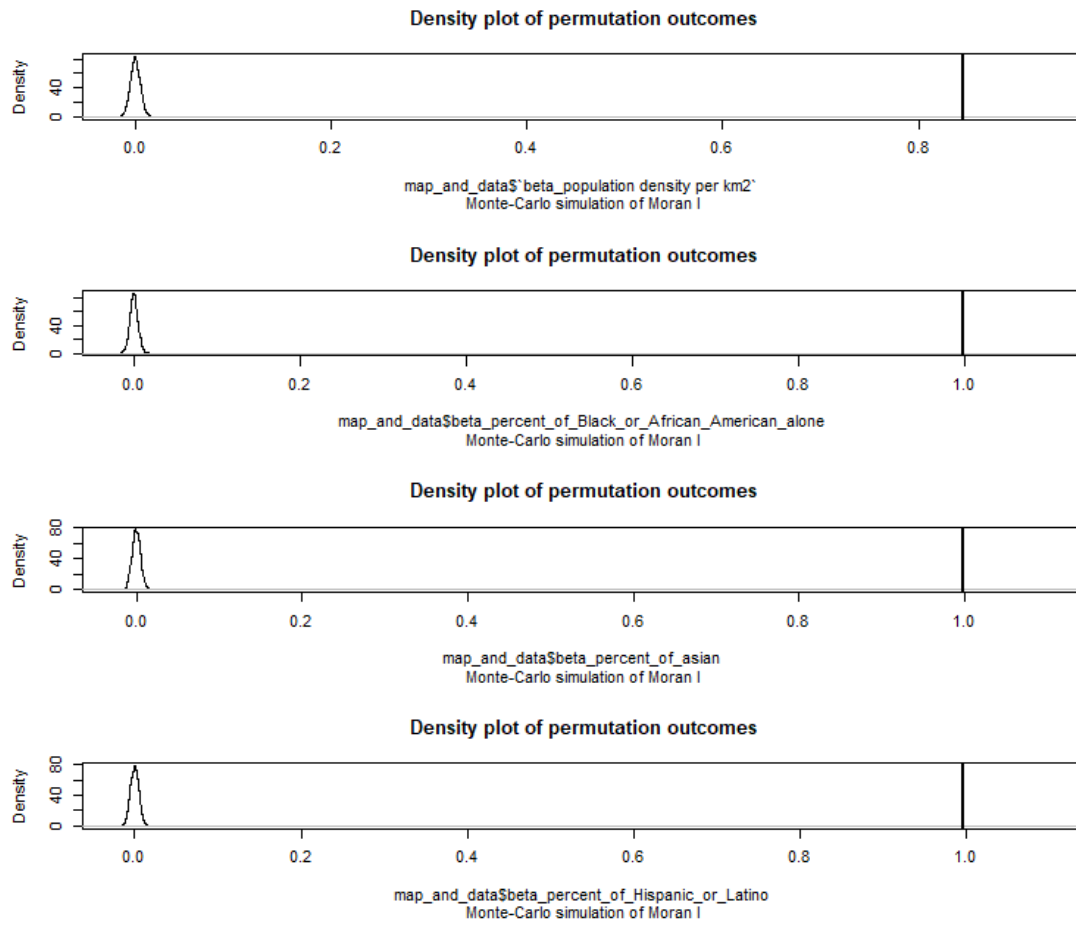


Figure A.12: Monto-Carlo simulation of Moran's I for Model 1 coefficients

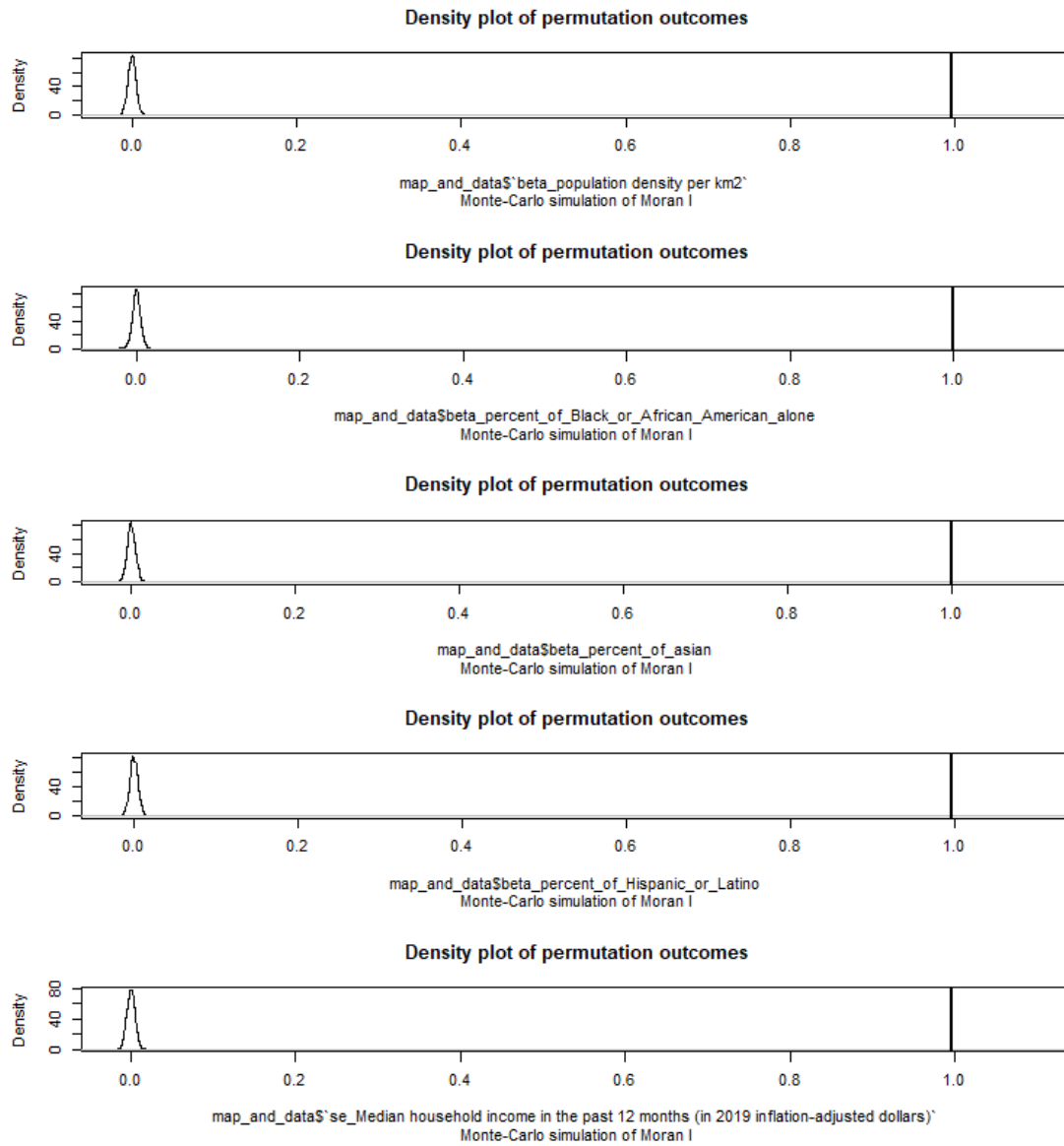


Figure A.13: Monto-Carlo simulation of Moran's I for Model 2 coefficients

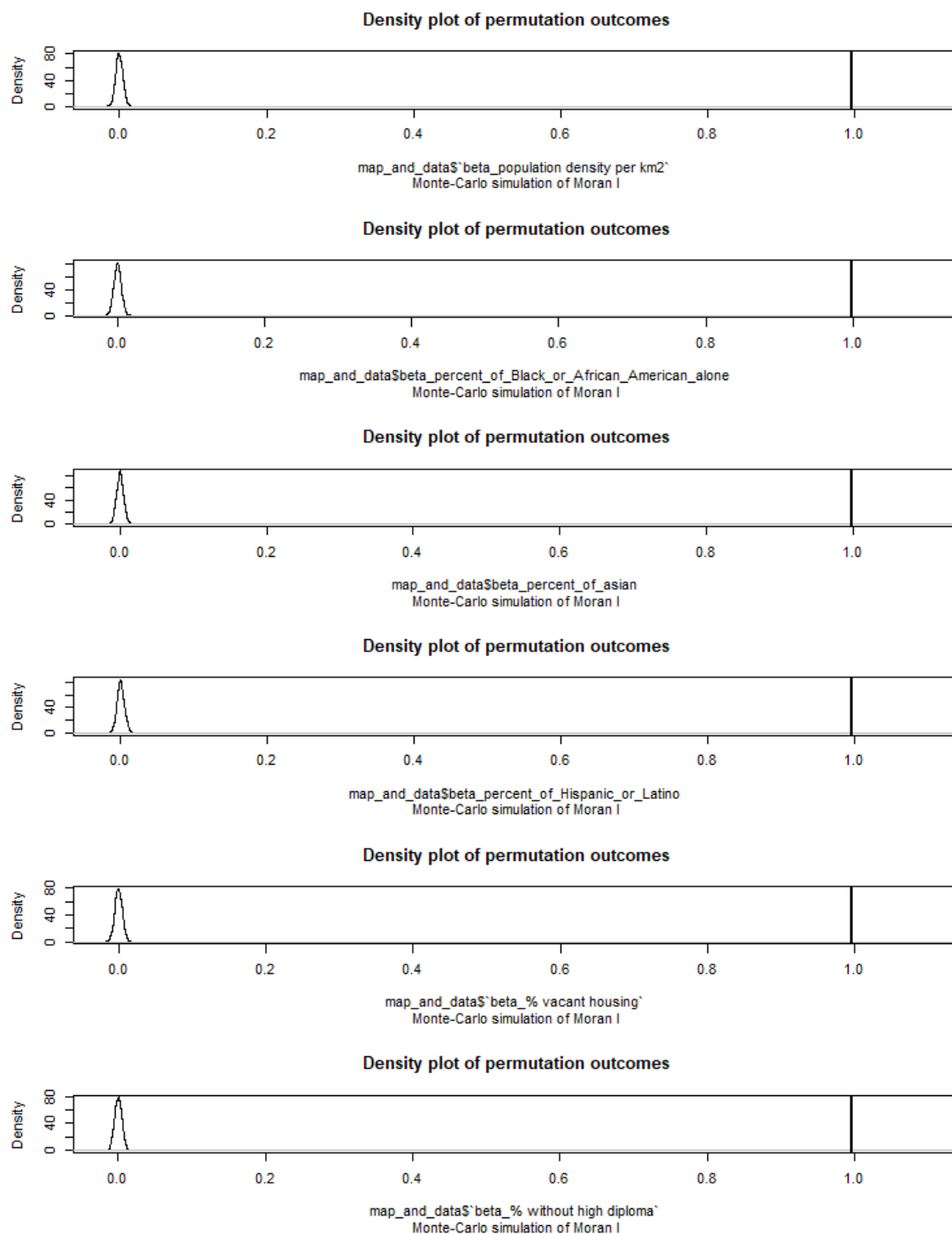


Figure A.14: Monto-Carlo simulation of Moran's I for Model 3 coefficients

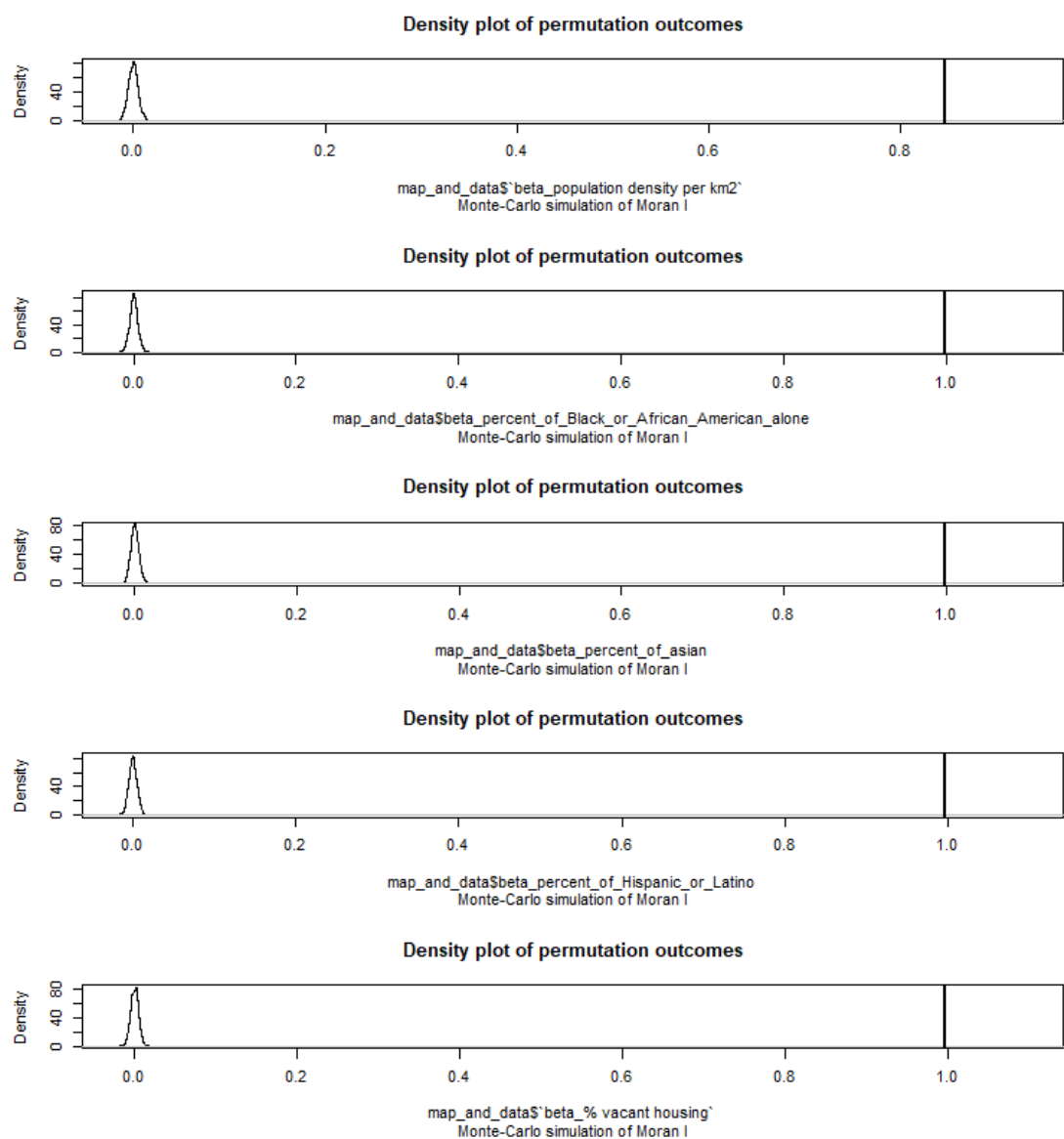


Figure A.15: Monto-Carlo simulation of Moran's I for Model 4 coefficients

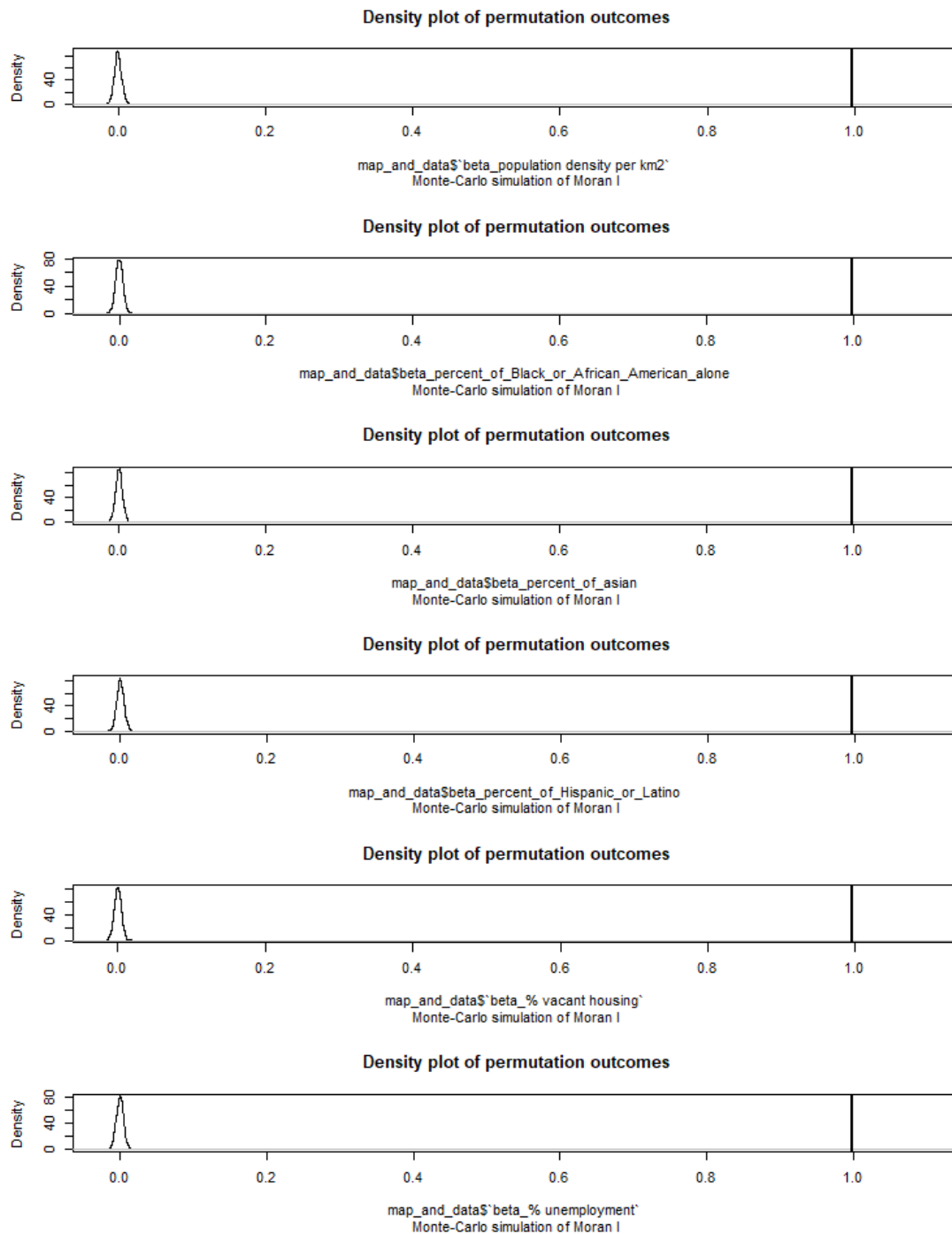


Figure A.16: Monto-Carlo simulation of Moran's I for Model 5 coefficients

A.3 Parameter estimates for all models

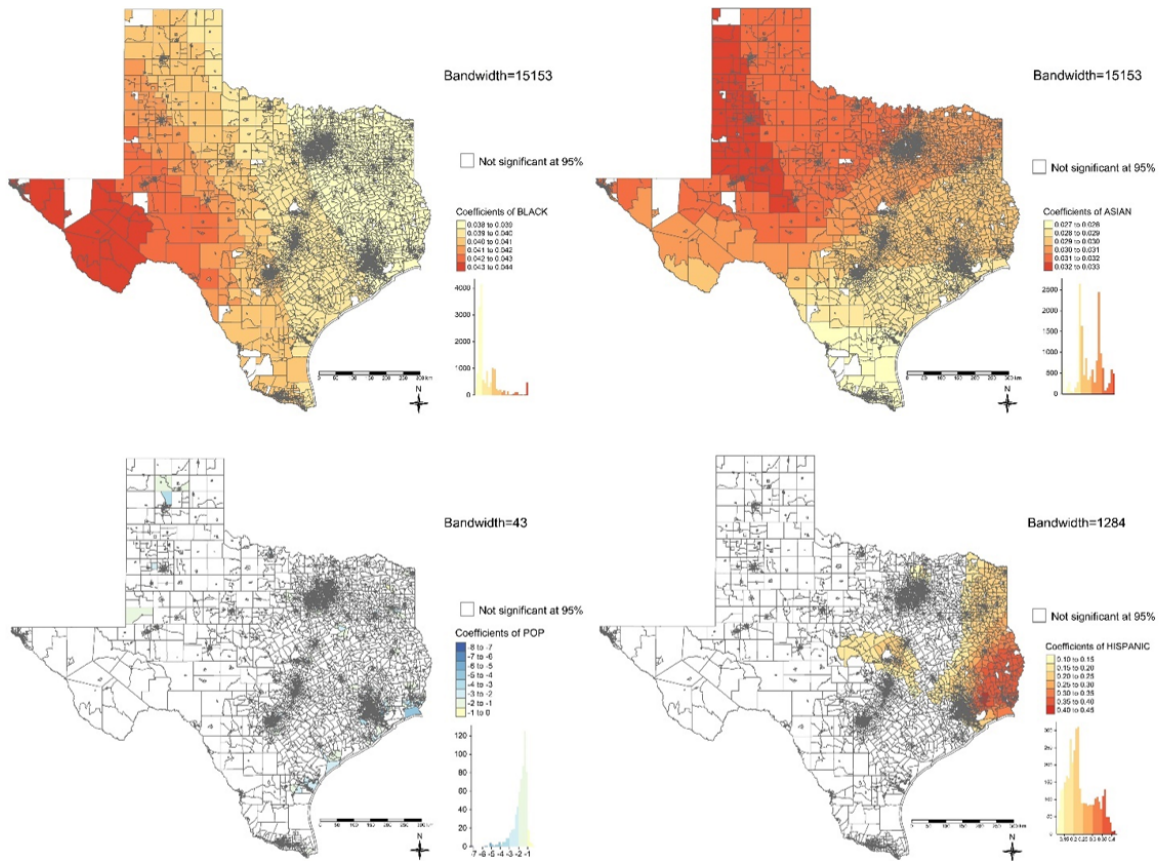


Figure A.17: parameter estimates for model 1

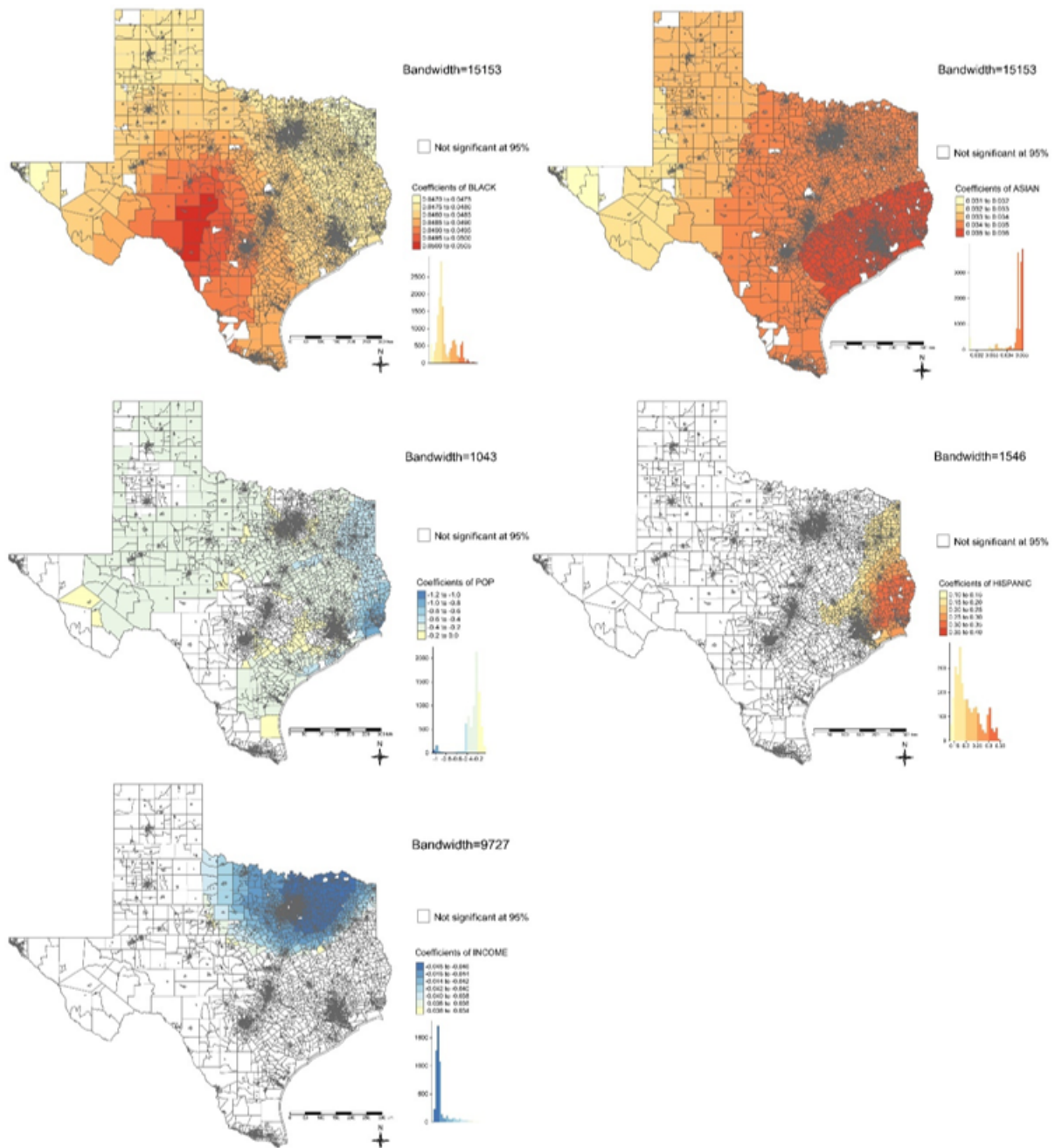


Figure A.18: parameter estimates for model 2

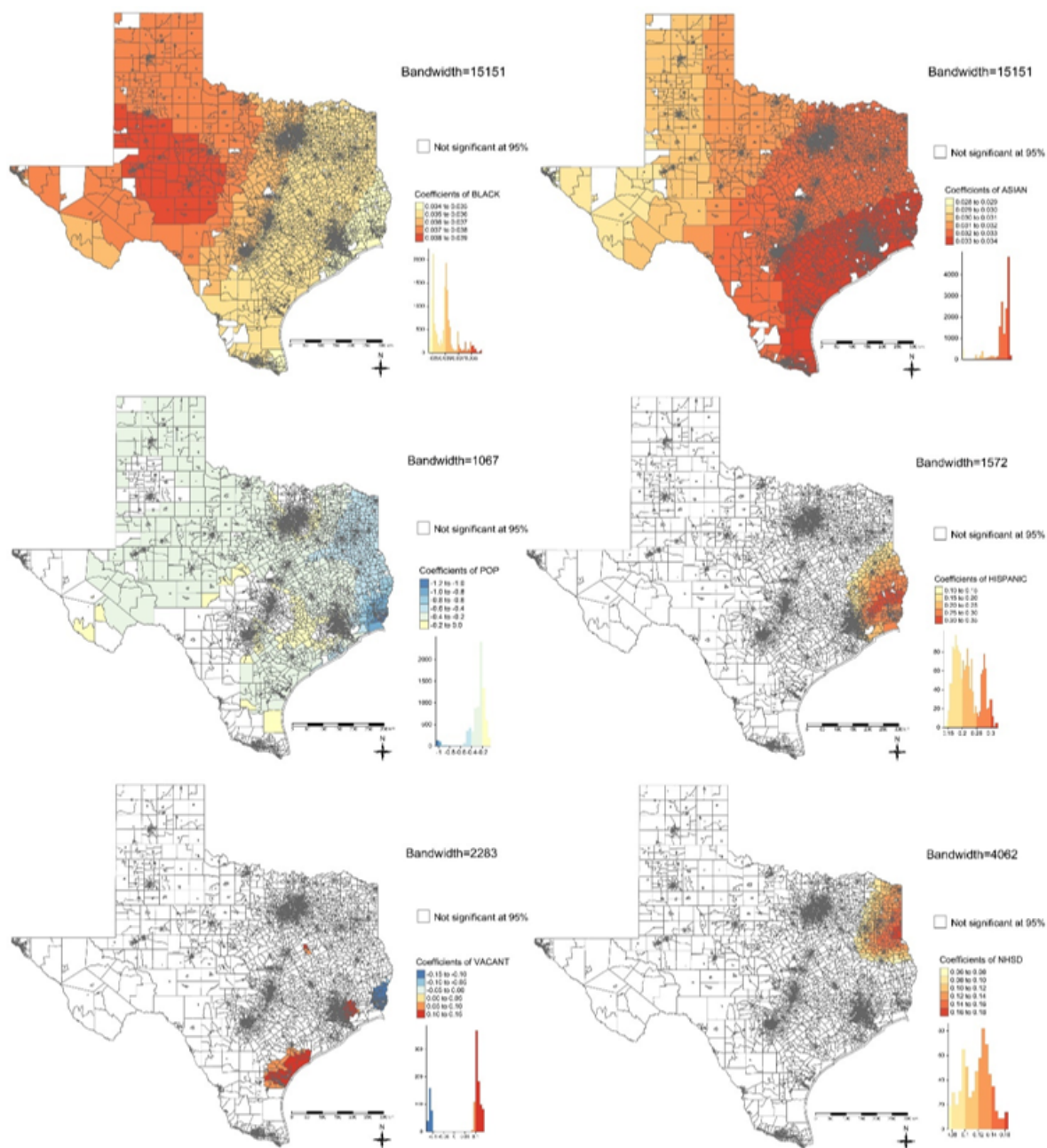


Figure A.19: parameter estimates for model 3

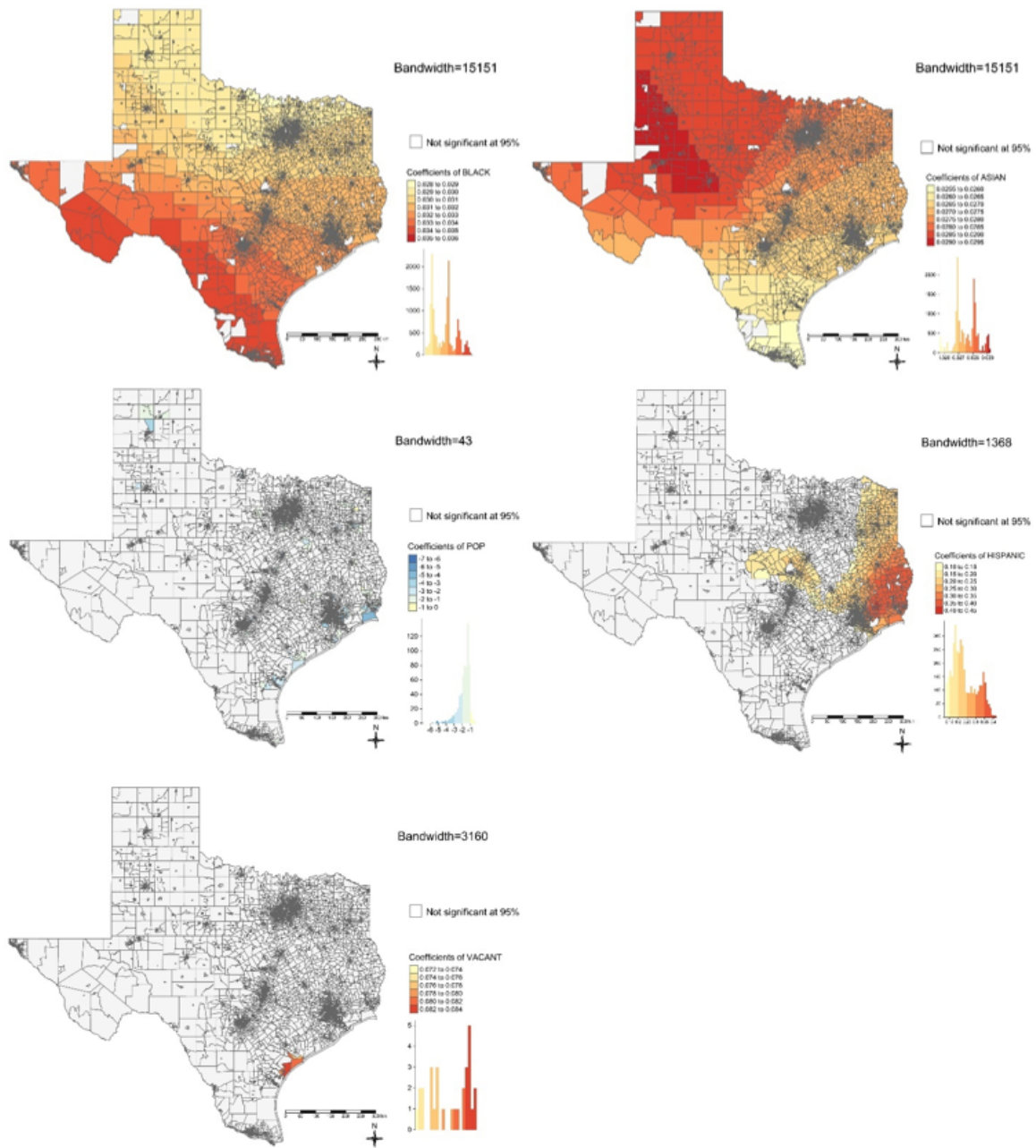


Figure A.20: parameter estimates for model 4

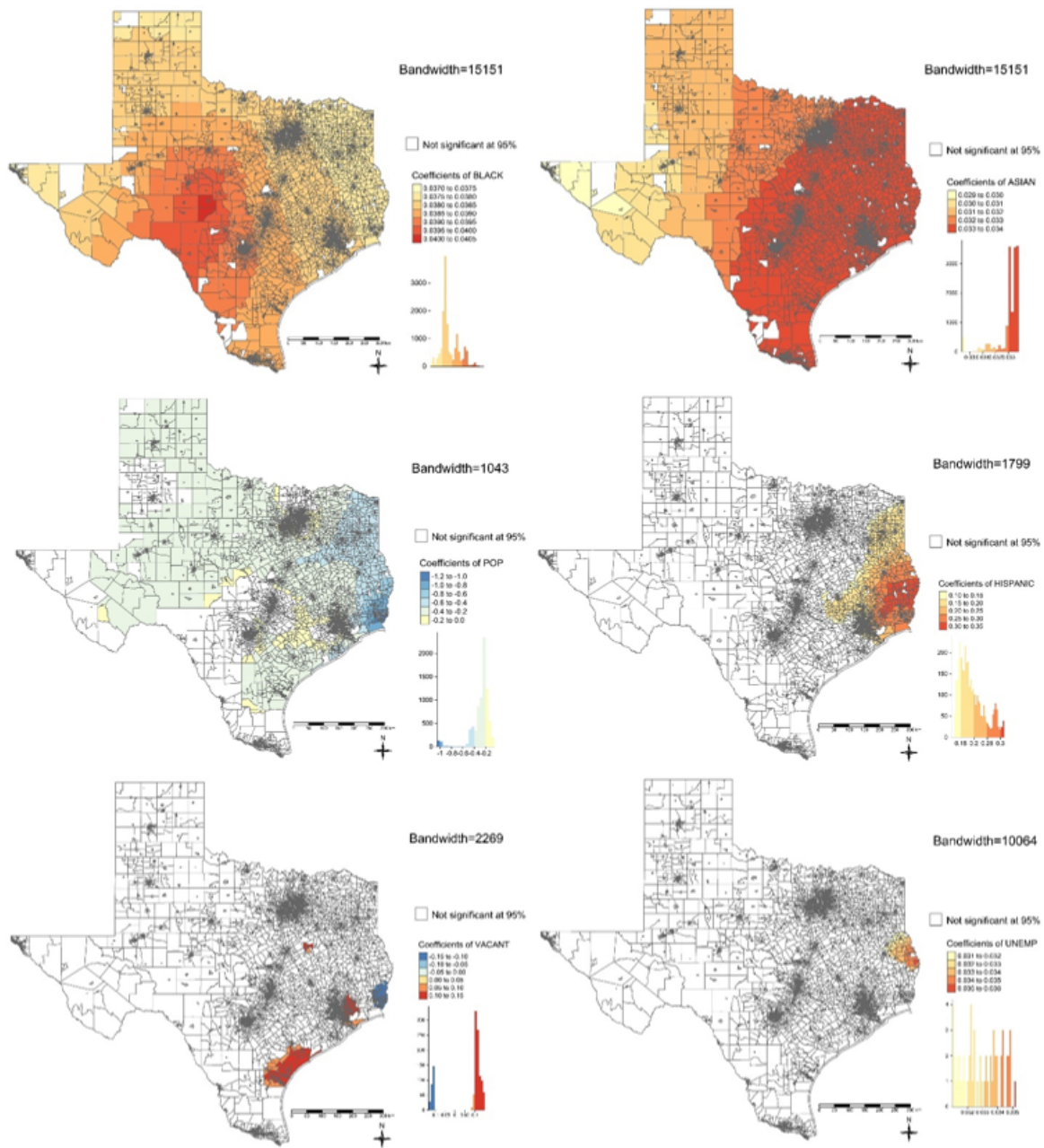


Figure A.21: parameter estimates for model 5

Appendix B: Appendix for Chapter 4

B.1 Statistical test

B.1.1 t-test for significance

We used the t-test to evaluate whether each predictor in the model has a significant relationship with the outcome variable. The R code is shown below:

```
model <- lm(RSEI ~ Pop + Black + Asian + Other + Hisp + income + Unemp + Vac +  
Own + mort + NH, data = data)  
coef_test <- -coef(summary(model))[, "Pr(> |t|)"]  
coef_test  
significant_predictors <- -names(coef_test[coef_test < 0.5])  
significant_predictors
```

The code produced a list of names of the predictor variables that are significantly related to the outcome variable. These are the predictor variables whose p-values are less than the chosen significance level (0.5): Intercept, Pop, Black, Asian, Other, Hisp, Unemp, Own, and NH.

B.1.2 Test for multicollinearity

This test evaluates whether there is a high degree of correlation among the predictor variables in the model, which can lead to inaccurate regression coefficients. One popular test for this is the variance inflation factor (VIF).

Variables	Pop	Black	Asian	Other	Hisp	Unemp	Own	NH
VIF	1.1848	2.1621	1.6796	1.2431	5.9174	1.1636	1.2806	3.6353

As a general rule of thumb, VIF values above 10 are considered very high and indicate that multicollinearity is a severe problem in the regression model. In this case, All the remaining variables have VIF values less than 10, which suggests that there are no highly correlated predictor variables.

B.1.3 Regression diagnostics tests

In Figure B.1, specifically in the "Residual vs. Fitted" plot, the proximity of residuals to zero and the relatively flat trend line suggests that the residuals conform to a normal distribution. To verify the normality assumption, Figure B.1's "Normal Q-Q" plot is employed. The validity of the normality assumption is established when the bulk of the residuals aligns along the solid diagonal line. As illustrated in our figure, the majority of the residuals are positioned on this line, confirming the normality of residuals.

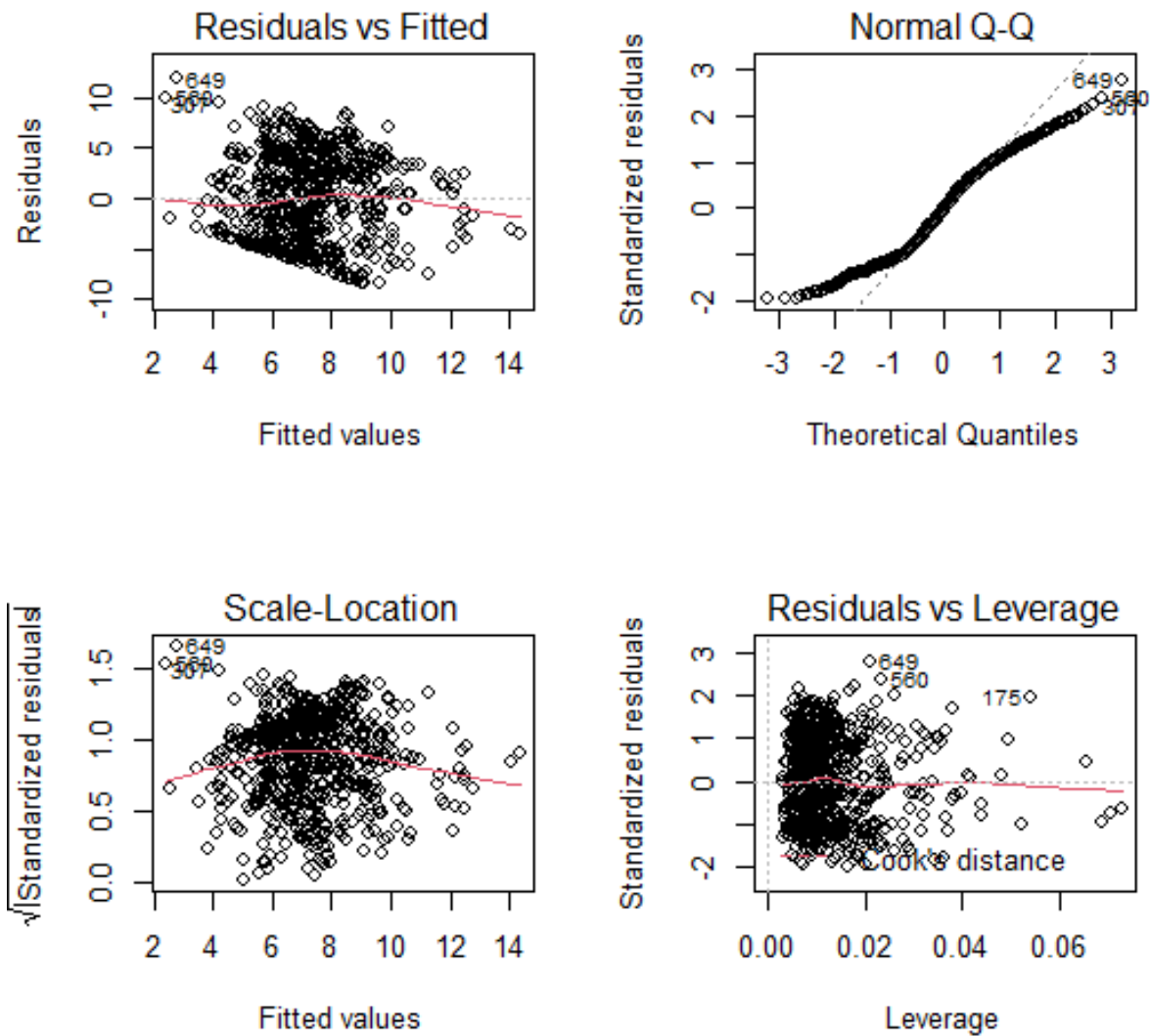


Figure B.1: Regression diagnostics plots

In Figure B.1, the "Residuals vs. Fitted" plot can be utilized to assess the assumption of linearity. In this figure, the residuals are spread equally around the red line, thereby corroborating the presence of a linear relationship. Furthermore, in Figure B.2, the blue line represents the anticipated residuals under the assumption of a linear relationship between the predictor and response variable. Meanwhile, the pink line represents the actual residuals. The similarity

between the two lines in all the figures suggests a linear relationship.

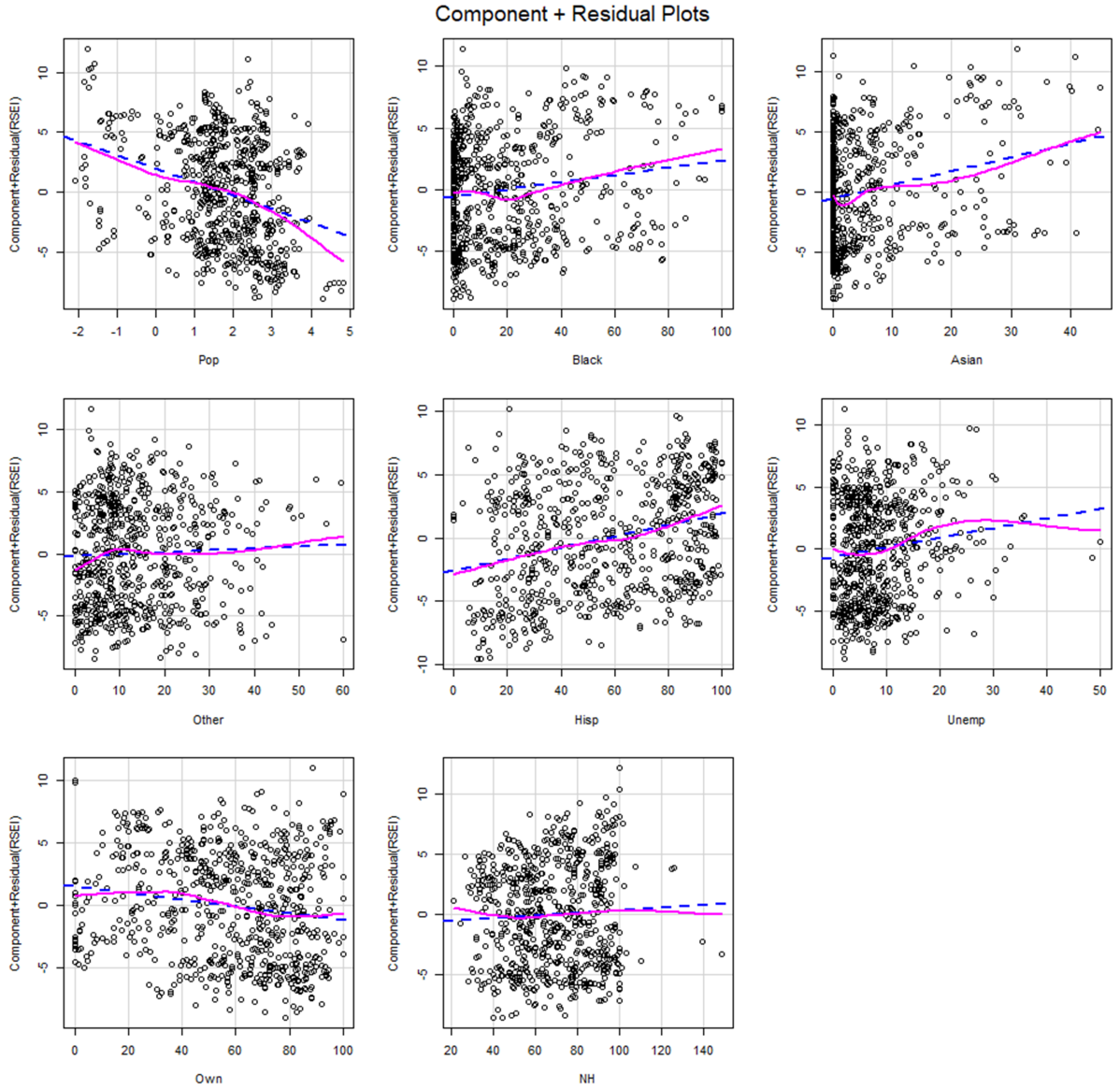


Figure B.2: Test of Linearity

In Figure B.1, the "Scale-Location" plot is employed to verify the homoscedasticity assumption. The presence of a horizontal line with uniformly distributed (randomly spread) points signifies that the residuals are homoscedastic. As observed in this figure, the residuals are randomly scattered along the horizontal line, indicating the homoscedasticity of residuals.

B.2 All results

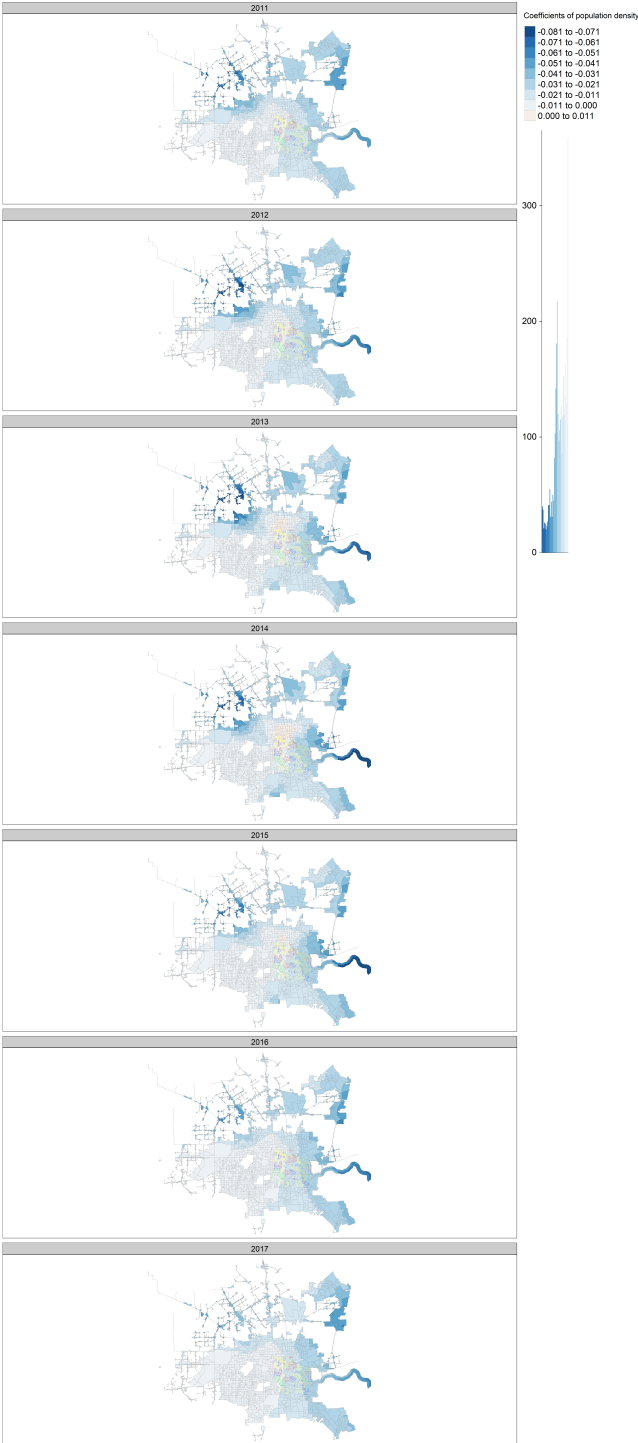


Figure B.3: The spatiotemporal distribution of regression coefficient estimates of population density 2011-2017



Figure B.4: The spatiotemporal distribution of regression coefficient estimates of Black_Pc 2011-2017

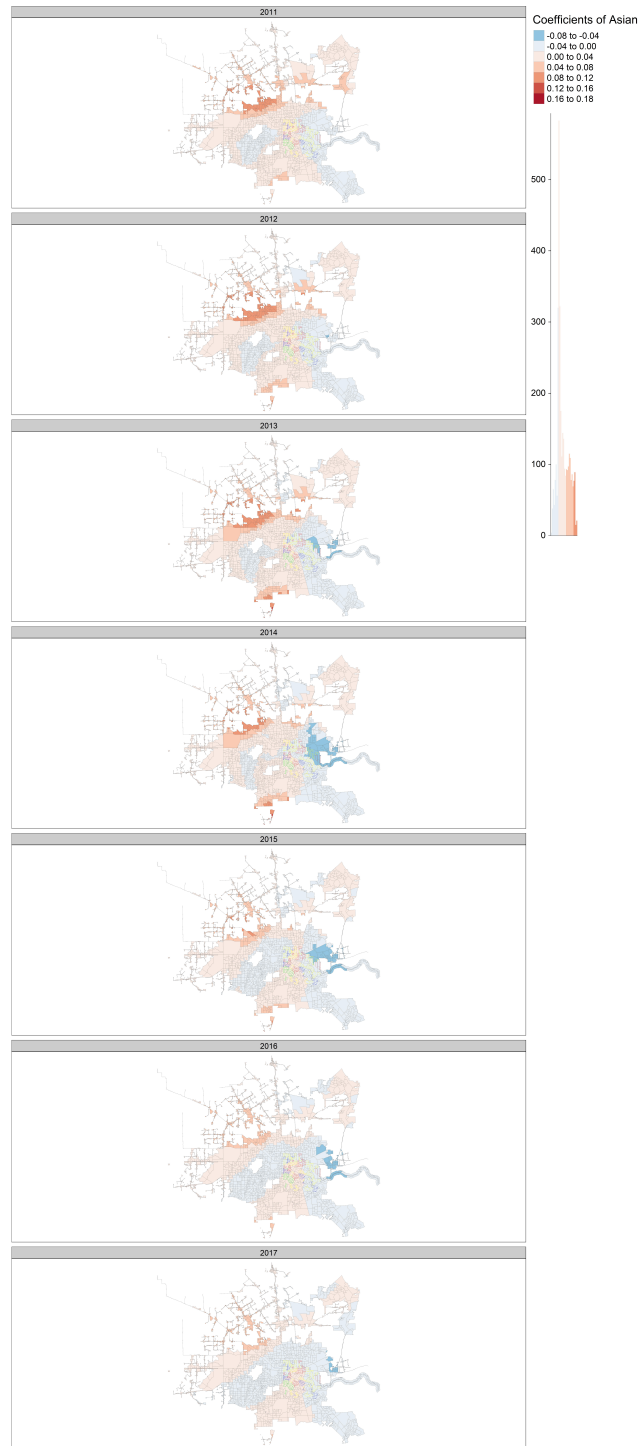


Figure B.5: The spatiotemporal distribution of regression coefficient estimates of Asian_Pc 2011-2017

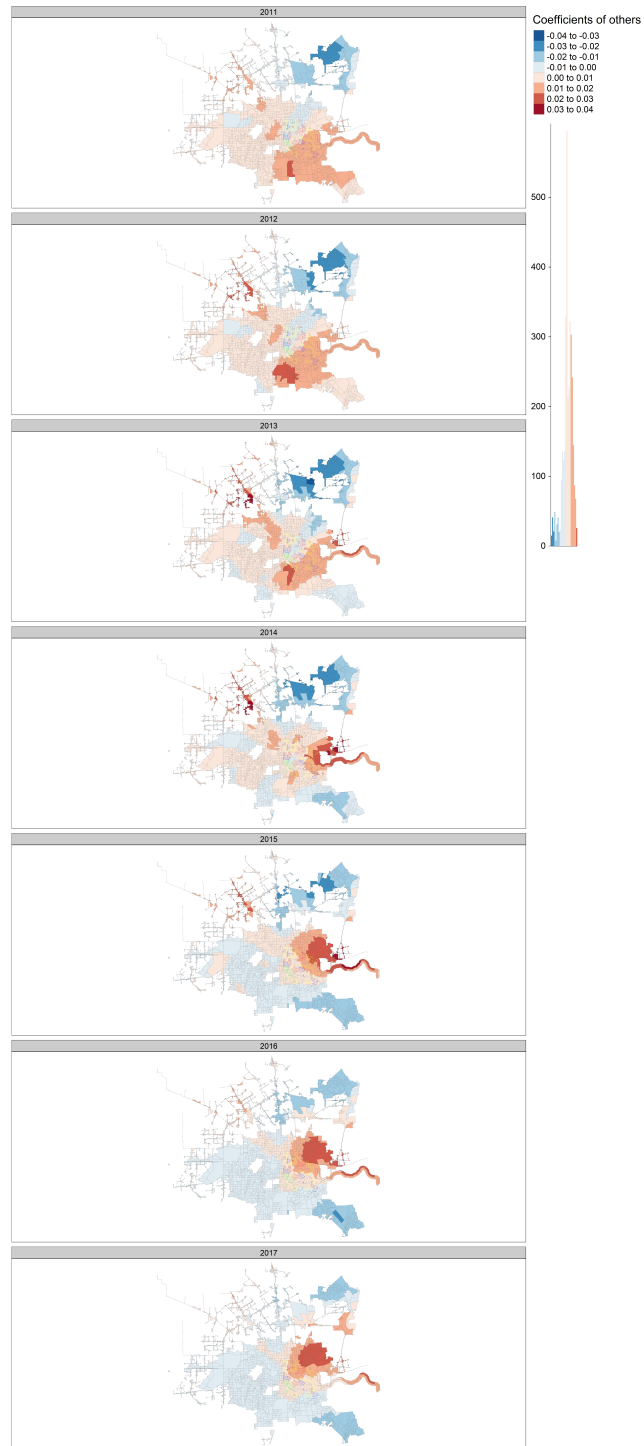


Figure B.6: The spatiotemporal distribution of regression coefficient estimates of Others_Pc 2011-2017

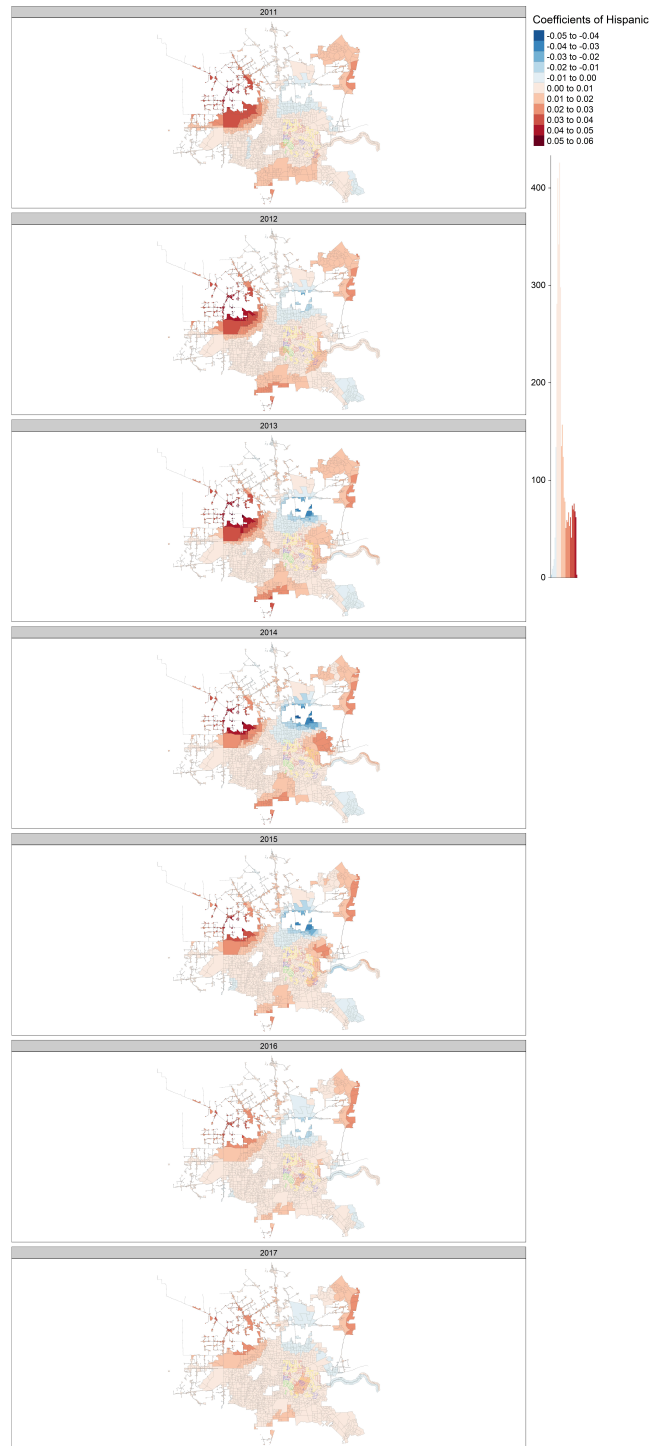


Figure B.7: The spatiotemporal distribution of regression coefficient estimates of Hispanic_Pc 2011-2017



Figure B.8: The spatiotemporal distribution of regression coefficient estimates of employment 2011-2017

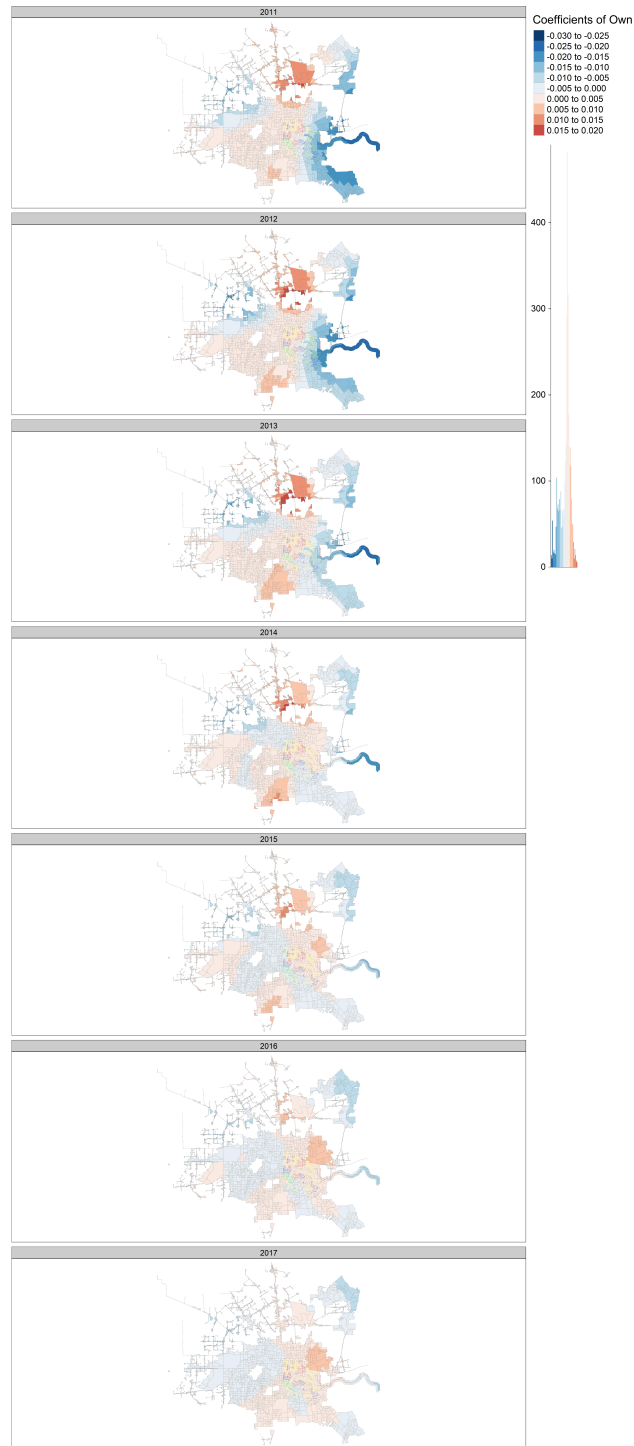


Figure B.9: The spatiotemporal distribution of regression coefficient estimates of housing 2011-2017

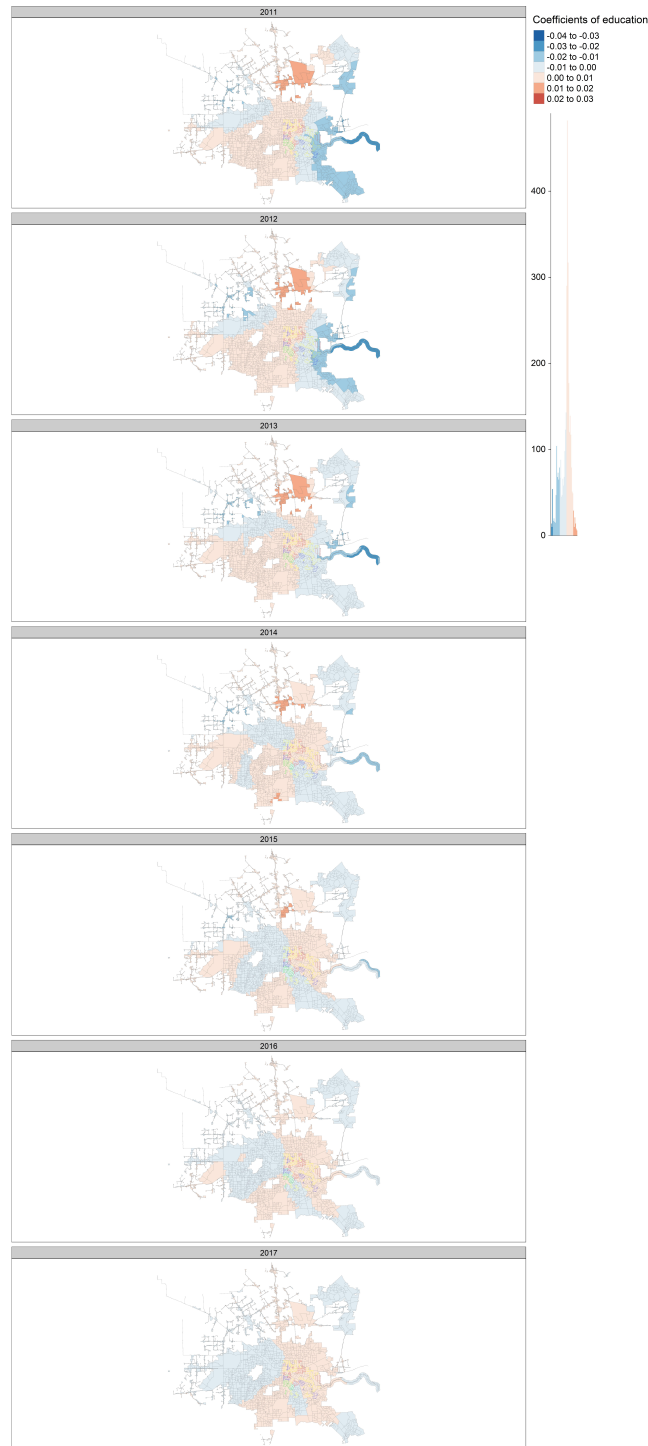


Figure B.10: The spatiotemporal distribution of regression coefficient estimates of education 2011-2017

Bibliography

- [1] SCL Koh, T Ibn-Mohammed, Adolf Acquaye, Kuishuang Feng, IM Reaney, Klaus Hubacek, H Fujii, and Khaled Khatab. Drivers of us toxicological footprints trajectory 1998–2013. *Scientific reports*, 6(1):1–10, 2016.
- [2] John M Balbus, Alistair BA Boxall, Richard A Fenske, Thomas E McKone, and Lauren Zeise. Implications of global climate change for the assessment and management of human health risks of chemicals in the natural environment. *Environmental toxicology and chemistry*, 32(1):62–78, 2013.
- [3] Odd Godal and Jan Fuglestad. Testing 100-year global warming potentials: impacts on compliance costs and abatement profile. *Climatic Change*, 52(1-2):93–127, 2002.
- [4] Conference of the Parties. Report of the Conference of the Parties on its third session, held at Kyoto from 1 to 11 December 1997 Addendum Part Two: Action taken by the Conference of the Parties at its third session. *Methodological issues related to the Kyoto Protocol*, pages 1–60, 1998.
- [5] PM Glibert, CJ Madden, W Boynton, D Flemer, C Heil, and J Sharp. Nutrients in estuaries: a summary report of the national estuarine experts workgroup (2005–2007). *US EPA report*, 2010.
- [6] J. Kaiser. Hazards of particles, PCBs focus of Philadelphia meeting. *Science*, 288(5465):424–425, 2000.
- [7] Seong Rin Lim, Carl W. Lam, and Julie M. Schoenung. Quantity-based and toxicity-based evaluation of the U.S. Toxics Release Inventory. *Journal of Hazardous Materials*, 178(1-3):49–56, 2010.
- [8] US Environmental Protection Agency. Toxics Release Inventory (TRI) Program, 2022.
- [9] U.S. Department of Health HHS and Human Services. National Toxic Substance Incidents Program (NTSIP) Biennial Report 2013-2014. Technical report, Agency for Toxic Substances and Disease Registry, 2014.

- [10] Matthew A. Cole, Robert J.R. Elliott, and Khemrutai Khemmarat. Local exposure to toxic releases: Examining the role of ethnic fractionalization and polarisation. *Ecological Economics*, 93:249–259, 2013.
- [11] Minnesota Pollution Control Agency. Disproportionate impacts in Minnesota, 2021.
- [12] US Environmental Protection Agency. Regional Profiles, 2021.
- [13] Jayajit Chakraborty, Timothy W. Collins, Sara E. Grineski, Marilyn C. Montgomery, and Maricarmen Hernandez. Comparing Disproportionate Exposure to Acute and Chronic Pollution Risks: A Case Study in Houston, Texas. *Risk Analysis*, 34(11):2005–2020, 2014.
- [14] Bin Zou, Fen Peng, Neng Wan, Keita Mamady, and Gaines J. Wilson. Spatial cluster detection of air pollution exposure inequities across the United States. *PLoS ONE*, 9(3):e91917, 2014.
- [15] Kerry Ard. Trends in exposure to industrial air toxins for different racial and socioeconomic groups: A spatial and temporal examination of environmental inequality in the U.S. from 1995 to 2004, 2015.
- [16] Klara Zwickl and Mathias Moser. Informal environmental regulation of industrial air pollution: Does neighborhood inequality matter? (15330):1–23, 2015.
- [17] Klara Zwickl, Michael Ash, and James K. Boyce. Regional variation in environmental inequality: Industrial air toxics exposure in U.S. cities. *Ecological Economics*, 107:494–509, 2014.
- [18] Maninder P.S. Thind, Christopher W. Tessum, Inês L. Azevedo, and Julian D. Marshall. Fine Particulate Air Pollution from Electricity Generation in the US: Health Impacts by Race, Income, and Geography. *Environmental Science and Technology*, 53(23):14010–14019, 2019.
- [19] Manfred Lenzen, Joy Murray, Fabian Sack, and Thomas Wiedmann. Shared producer and consumer responsibility - Theory and practice. *Ecological Economics*, 61(1):27–42, 2007.
- [20] Jesper Munksgaard and Klaus Alsted Pedersen. CO2 accounts for open economies: Producer or consumer responsibility? *Energy Policy*, 29(4):327–334, 2001.
- [21] Glen P. Peters. From production-based to consumption-based national emission inventories. *Ecological Economics*, 2008.
- [22] Arjen Y. Hoekstra and Thomas O. Wiedmann. Humanity’s unsustainable environmental footprint. *Science*, 344(6188):1114–1117, 2014.
- [23] Thomas Wiedmann and Manfred Lenzen. Environmental and social footprints of international trade. *Nature Geoscience*, 11(5):314–321, 2018.

- [24] Christopher W. Tessum, Joshua S. Apte, Andrew L. Goodkind, Nicholas Z. Muller, Kimberley A. Mullins, David A. Paoletta, Stephen Polasky, Nathaniel P. Springer, Sumil K. Thakrar, Julian D. Marshall, and Jason D. Hill. Inequity in consumption of goods and services adds to racial-ethnic disparities in air pollution exposure. *Proceedings of the National Academy of Sciences of the United States of America*, 116(13):6001–6006, 2019.
- [25] Janis Brizga, Kuishuang Feng, and Klaus Hubacek. Household carbon footprints in the Baltic States: A global multi-regional input–output analysis from 1995 to 2011. *Applied Energy*, 189:780–788, 2017.
- [26] Thomas Wiedmann. A review of recent multi-region input-output models used for consumption-based emission and resource accounting. *Ecological Economics*, 69(2):211–222, 2009.
- [27] Thomas Wiedmann, Manfred Lenzen, Karen Turner, and John Barrett. Examining the global environmental impact of regional consumption activities - Part 2: Review of input-output models for the assessment of environmental impacts embodied in trade. *Ecological Economics*, 61(1):15–26, 2007.
- [28] Chris Brunsdon, A. Stewart Fotheringham, and Martin E. Charlton. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, 28(4):281–298, 1996.
- [29] Jeremy L Mennis and Lisa Jordan. The Distribution of Environmental Equity: Exploring Spatial Nonstationarity in Multivariate Models of Air Toxic Releases. *Annals of the Association of American Geographers*, 95(2):249–268, 2008.
- [30] Abolfazl Mollalo, Behzad Vahedi, and Kiara M. Rivera. GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Science of the Total Environment*, 728:138884, 2020.
- [31] Hannah Aoyagi and Oladele A. Ogunseitan. Toxic releases and risk disparity: A spatiotemporal model of industrial ecology and social empowerment. *International Journal of Environmental Research and Public Health*, 12(6):6300–6318, 2015.
- [32] Yoo Min Park and Mei Po Kwan. Understanding racial disparities in exposure to traffic-related air pollution: Considering the spatiotemporal dynamics of population distribution. *International Journal of Environmental Research and Public Health*, 17(3), 2020.
- [33] Bo Huang, Bo Wu, and Michael Barry. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science*, 24(3):383–401, 2010.
- [34] A. Stewart Fotheringham, Ricardo Crespo, and Jing Yao. Geographical and Temporal Weighted Regression (GTWR). *Geographical Analysis*, 47(4):431–452, 2015.

- [35] Kerry Ard and Kevin Smiley. Examining the Relationship Between Racialized Poverty Segregation and Hazardous Industrial Facilities in the U.S. Over Time. *American Behavioral Scientist*, 66(7):974–988, 2022.
- [36] Wei Zhang, Yu Liu, Kuishuang Feng, Klaus Hubacek, Jinnan Wang, Miaomiao Liu, Ling Jiang, Hongqiang Jiang, Nianlei Liu, Pengyan Zhang, Ying Zhou, and Jun Bi. Revealing Environmental Inequality Hidden in China’s Inter-regional Trade. *Environmental Science and Technology*, 52(13):7171–7181, 2018.
- [37] Zhan Ming Chen, Stephanie Ohshita, Manfred Lenzen, Thomas Wiedmann, Magnus Jiborn, Bin Chen, Leo Lester, Dabo Guan, Jing Meng, Shiyun Xu, Guoqian Chen, Xinye Zheng, Jin Jun Xue, Ahmed Alsaedi, Tasawar Hayat, and Zhu Liu. Consumption-based greenhouse gas emissions accounting with capital stock change highlights dynamics of fast-developing countries. *Nature Communications*, 9(1), 2018.
- [38] S. C.L. Koh, T. Ibn-Mohammed, A. Acquaye, K. Feng, I. M. Reaney, K. Hubacek, H. Fujii, and K. Khatab. Drivers of U.S. toxicological footprints trajectory 1998-2013. *Scientific Reports*, 6(December):1–10, 2016.
- [39] Hidemichi Fujii, Shunsuke Okamoto, Shigemi Kagawa, and Shunsuke Managi. Decomposition of toxicity emission changes on the demand and supply sides: Empirical study of the US industrial sector. *Environmental Research Letters*, 12(12), 2017.
- [40] Keisuke Nansai, Susumu Tohno, Satoru Chatani, Keiichiro Kanemoto, Midori Kurogi, Yuta Fujii, Shigemi Kagawa, Yasushi Kondo, Fumiya Nagashima, Wataru Takayanagi, and Manfred Lenzen. Affluent countries inflict inequitable mortality and economic loss on Asia via PM2.5 emissions. *Environment International*, 134(October 2019), 2020.
- [41] Fahui Wang and Yvette C Feliberty. Geospatial Techniques in Urban Hazard and Disaster Analysis. In *Geospatial Techniques in Urban Hazard and Disaster Analysis*, pages 157–177. 2010.
- [42] Lara P. Clark, Dylan B. Millet, and Julian D. Marshall. National patterns in environmental injustice and inequality: Outdoor NO2 air pollution in the United States. *PLoS ONE*, 9(4):1–8, 2014.
- [43] Amanda T. Charette, Mary B. Collins, and Jaime E. Mirowsky. Assessing residential socioeconomic factors associated with pollutant releases using EPA’s Toxic Release Inventory. *Journal of Environmental Studies and Sciences*, 11(2):247–257, 2021.
- [44] Jacob Bor, Gregory H. Cohen, and Sandro Galea. Population health in an era of rising income inequality: USA, 1980–2015. *The Lancet*, 389(10077):1475–1490, 2017.
- [45] Congressional Research Service. The U.S. Income Distribution: Trends and Issues. 2021.
- [46] Andrew Jorgenson, Juliet Schor, and Xiaorui Huang. Income Inequality and Carbon Emissions in the United States: A State-level Analysis, 1997–2012. *Ecological Economics*, 134:40–48, 2017.

- [47] Kuishuang Feng, Klaus Hubacek, and Kaihui Song. Household carbon inequality in the U.S. *Journal of Cleaner Production*, 278:123994, 2021.
- [48] Jingquan Chen, Xinyan Lian, Hanning Su, Ziyang Zhang, Xiaojun Ma, and Baishu Chang. Analysis of China’s carbon emission driving factors based on the perspective of eight major economic regions. *Environmental Science and Pollution Research*, 28(7):8181–8204, 2021.
- [49] Adrian Foong, Prajal Pradhan, Oliver Frör, and Jürgen P. Kropp. Adjusting agricultural emissions for trade matters for climate change mitigation. *Nature Communications*, 13(1), 2022.
- [50] Lulu Lian, Tao Huang, Xianmin Ke, Zaili Ling, Wanyanhan Jiang, Zhanxiang Wang, Shijie Song, Jixiang Li, Yuan Zhao, Hong Gao, Shu Tao, Junfeng Liu, and Jianmin Ma. Globalization-Driven Industry Relocation Significantly Reduces Arctic PAH Contamination. *Environmental Science and Technology*, 56(1):145–154, 2022.
- [51] Glen P. Peters and Edgar G. Hertwich. Production factors and pollution embodied in trade: Theoretical development. *Industrial Ecology Programme (IndEcol). Working Papers*, 5(5):2004, 2004.
- [52] Kjartan Steen-Olsen, Richard Wood, and Edgar G. Hertwich. The Carbon Footprint of Norwegian Household Consumption 1999–2012. *Journal of Industrial Ecology*, 20(3):582–592, 2016.
- [53] IMPLAN Group. LLC. IMPLAN Data Overview & Sources. Technical report, 2021.
- [54] US Environmental Protection Agency. Regional Profile for EPA Region 6, 2021.
- [55] US Environmental Protection Agency. EPA Proposes “Good Neighbor” Plan to Cut Smog Across Much of the United States — US EPA, 2022.
- [56] US Environmental Protection Agency. Introduction to the Toxics Release Inventory (TRI) for Industrial and Federal Facilities What is the TRI ? Technical Report July, 2022.
- [57] United States Census Bureau. American Community Survey - Data Profiles 2017, 2017.
- [58] Candi Clouse. Household Income Events – IMPLAN - Support, 2022.
- [59] US Environmental Protection Agency. Emergency Planning and Community Right-to-Know Act (EPCRA), 2021.
- [60] US Environmental Protection Agency. TRI National Analysis, 2021.
- [61] Statista. GDP by industry California 2021 — Statista, 2022.
- [62] Forum of Federation. United States of America, 2021.
- [63] Estelle Sommeiller and Mark Price. The new gilded age Income inequality in the U.S. by state, metropolitan area, and county. Technical report, Economic Policy Institute, Washington, DC, 2018.

- [64] Statista. Population density in the U.S., by state 2020 Published by Statista Research Department, Jan 21, 2021 In 2020, Washington, D.C. had the highest population density in the United States, with 11,686 people per square mile. As a whole, there were 93 resident, 2021.
- [65] Ronald E Miller and Peter D Blair. *Input-output analysis: foundations and extensions*. Cambridge university press, 2009.
- [66] George D. Haddow, Jane A. Bullock, and Damon P. Coppola. *Introduction to Emergency Management*. Elsevier, jan 2020.
- [67] Patricia M; Gilbert, Christopher J.; Madden, Walter; Boynton, David; Flemer, Cynthia; Heil, and Jonathan Sharp. Nutrients in estuaries: a summary report of the National Estuarine Experts Workgroup (2005–2007). *National Estuarine Experts*, (November):188, 2010.
- [68] Liton Chakraborty, Horatiu Rus, Daniel Henstra, Jason Thistlethwaite, Andrea Minano, and Daniel Scott. Exploring spatial heterogeneity and environmental injustices in exposure to flood hazards using geographically weighted regression. *Environmental Research*, 210(February):112982, 2022.
- [69] FEMA. What do we call Ourselves? Hazards Risk Managers of Course, 2022.
- [70] Lubka Tchankova. Risk identification – basic stage in risk management. *Environmental Management and Health*, 13(3):290–297, aug 2002.
- [71] Barry E Flanagan, Edward W Gregory, / Atsdr, Elaine J Hallisey, Janet L Heitgerd, Cdc / Nchhstp, Brian Lewis, Barry E ; Flanagan, Edward W ; Gregory, Elaine J ; Hallisey, and Janet L ; Heitgerd. A Social Vulnerability Index for Disaster Management. *Journal of Homeland Security and Emergency Management*, 8, 2011.
- [72] Sapam Ranabir Singh, Mohammad Reza Eghdami, and Sarbjeet Singh. The Concept of Social Vulnerability: A Review from Disasters Perspectives. *International Journal of Interdisciplinary and Multidisciplinary Studies (IJIMS)*, 1(6):71–82, 2014.
- [73] Alice Fothergill and Lori A Peek. Poverty and Disasters in the United States: A Review of Recent Sociological Findings. *Natural Hazards*, 32:89–110, 2004.
- [74] Michel Masozera, Melissa Bailey, and Charles Kerchner. Distribution of impacts of natural disasters across income groups: A case study of New Orleans. *Ecological Economics*, 63(2-3):299–306, aug 2007.
- [75] Farin Fatemi, Ali Ardalan, Benigno Aguirre, Nabiollah Mansouri, and Iraj Mohammadfam. Social vulnerability indicators in disasters: Findings from a systematic review. *International Journal of Disaster Risk Reduction*, 22:219–227, jun 2017.
- [76] James R. Elliott and Jeremy Pais. Race, class, and Hurricane Katrina: Social differences in human responses to disaster. *Social Science Research*, 35(2):295–321, jun 2006.

- [77] Eric Tate, Md Asif Rahman, Christopher T. Emrich, and Christopher C. Sampson. Flood exposure and social vulnerability in the United States. *Natural Hazards*, 106(1):435–457, mar 2021.
- [78] William Bowen. An Analytical Review of Environmental Justice Research: What Do We Really Know? *Environmental Management*, 29(1):3–15, 2002.
- [79] Robert J. Brulle and David N. Pellow. ENVIRONMENTAL JUSTICE: Human Health and Environmental Inequalities. *Annual Review of Public Health*, 27(1):103–124, 2006.
- [80] Michelle L. Bell and Keita Ebisu. Environmental inequality in exposures to airborne particulate matter components in the United States. *Environmental Health Perspectives*, 120(12):1699–1704, 2012.
- [81] Sacoby M. Wilson, Herb Fraser-Rahim, Edith Williams, Hongmei Zhang, Lashanta Rice, Erik Svendsen, and Winston Abara. Assessment of the distribution of toxic release inventory facilities in metropolitan Charleston: An environmental justice case study. *American Journal of Public Health*, 102(10):1974–1980, 2012.
- [82] Sirisha C Naidu, Panayiotis T Manolakos, and Thomas E Hopkins. Environmental Justice in Ohio. *Review of Radical Political Economics*, 45(3):384–399, 2012.
- [83] Kerry Ard. Trends in exposure to industrial air toxins for different racial and socioeconomic groups: A spatial and temporal examination of environmental inequality in the U.S. from 1995 to 2004. *Social Science Research*, 53(2015):375–390, 2015.
- [84] Timothy W. Collins and Sara E. Grineski. Environmental Injustice and Religion: Outdoor Air Pollution Disparities in Metropolitan Salt Lake City, Utah. *Annals of the American Association of Geographers*, 109(5):1597–1617, 2019.
- [85] Ryan Johnson, Kim Ramsey-White, and Christina H. Fuller. Socio-demographic differences in toxic release inventory siting and emissions in metro Atlanta. *International Journal of Environmental Research and Public Health*, 13(8):747, 2016.
- [86] Nicole Kravitz-Wirtz, Kyle Crowder, Anjum Hajat, and Victoria Sass. The Long-term Dynamics of Racial/Ethnic Inequality in Neighborhood Air Pollution Exposure, 1990–2009. *Du Bois Review*, 13(2):237–259, 2016.
- [87] Zhengyan Li, David M. Konisky, and Nikolaos Zirotiannis. Racial, ethnic, and income disparities in air pollution: A study of excess emissions in Texas. *PLoS ONE*, 14(8):1–16, 2019.
- [88] Ronald Pope, Jianguo Wu, and Christopher Boone. Spatial patterns of air pollutants and social groups: a distributive environmental justice study in the phoenix metropolitan region of USA. *Environmental Management*, 58(5):753–766, 2016.
- [89] Jason M Post, Perry L Carter, Cynthia L Sorresen, and Spatial Environmental. Spatial Environmental Inequality in Lubbock, Texas. *Current Research Journal of Social Sciences and Humanities*, 01(1):01–12, 2018.

- [90] Meng Li, Yuning Gao, and Shenglong Liu. China's energy intensity change in 1997–2015: Non-vertical adjusted structural decomposition analysis based on input-output tables. *Structural Change and Economic Dynamics*, 53:222–236, 2020.
- [91] Angela Gilbert and Jayajit Chakraborty. Using geographically weighted regression for environmental justice analysis: Cumulative cancer risks from air toxics in Florida. *Social Science Research*, 40(1):273–286, 2011.
- [92] Rehana Shrestha, Johannes Flacke, Javier Martinez, and Martin Van Maarseveen. Environmental Health Related Socio-Spatial Inequalities: Identifying "Hotspots" of Environmental Burdens and Social Vulnerability. *International Journal of Environmental Research and Public Health*, 13(7):691, 2016.
- [93] Gordon Walker. Beyond Distribution and Proximity: Exploring the Multiple Spatialities of Environmental Justice. *Antipode*, 41(4):614–636, 2009.
- [94] Maggie Chamberlain. *A quantitative analysis of environmental inequality in Houston, Texas*. Master dissertation, Texas State University, San Marcos, Texas., 2017.
- [95] Sara E. Grineski, Timothy W. Collins, and Hector A. Olvera. Local Variability in the Impacts of Residential Particulate Matter and Pest Exposure on Children's Wheezing Severity: A Geographically Weighted Regression Analysis of Environmental Health Justice. *Population and Environment*, 37(1):22–43, 2015.
- [96] Chunrong Jia, Wesley James, and Satish Kedia. Relationship of racial composition and cancer risks from air toxics exposure in memphis, Tennessee, USA. *International Journal of Environmental Research and Public Health*, 11(8):7713–7724, 2014.
- [97] Kim M Lersch and Timothy C Hart. Sociological Spectrum Mid-South Sociological Association Environmental Justice, Lead, and Crime: Exploring the Spatial Distribution and Impact of Industrial Facilities in Hillsborough County, Florida. *Sociological Spectrum*, 34(1):1–21, 2014.
- [98] Seong Nam Hwang and Venu Srivastav Kannegalla. Relationships between Potential Risk of Hazardous Material Releases and Socio-Economic Demographic Characteristics Using Geospatial Analysis. *Journal of Geoscience and Environment Protection*, 09(07):62–85, 2021.
- [99] Jun Tu, Wei Tu, and Stuart H. Tedders. Spatial variations in the associations of term birth weight with ambient air pollution in Georgia, USA. *Environment International*, 92-93:146–156, 2016.
- [100] Maryia Bakhtsiyarava and Raphael J. Nawrotzki. Environmental inequality and pollution advantage among immigrants in the United States. *Applied Geography*, 81:60–69, 2017.
- [101] A. Stewart Fotheringham, Wenbai Yang, and Wei Kang. Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers*, 107(6):1247–1265, 2017.

- [102] Taylor M. Oshan, Ziqi Li, Wei Kang, Levi J. Wolf, and A. Stewart Fotheringham. MGWR: A python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS International Journal of Geo-Information*, 8(6):269, 2019.
- [103] Hanchen Yu, A. Stewart Fotheringham, Ziqi Li, Taylor Oshan, Wei Kang, and Levi John Wolf. Inference in Multiscale Geographically Weighted Regression. *Geographical Analysis*, 52(1):87–106, 2020.
- [104] Hirotogu Akaike. Information theory and an extension of the maximum likelihood principle. *Selected papers of hirotogu akaike*, pages 199–213, 1998.
- [105] Zahratu Shabrina, Boyana Buyuklieva, and Matthew Kok Ming Ng. Short-Term Rental Platform in the Urban Tourism Context: A Geographically Weighted Regression (GWR) and a Multiscale GWR (MGWR) Approaches. *Geographical Analysis*, pages 1–22, 2020.
- [106] D Anderson and K Burnham. Model selection and multi-model inference. *Second. NY: Springer-Verlag*, 63(2020):10, 2004.
- [107] de Leon Arnoldo and A. Calvert Robert. Segregation, aug 2021.
- [108] Daniel Aaronson, Daniel A. Hartley, and Bhashkar Mazumder. The Effects of the 1930s Holc 'Redlining' Maps. *Ssrn*, 13(4):355–392, 2020.
- [109] Erin Douglas. EPA chief to visit Texas neighborhoods plagued by pollution — The Texas Tribune, 2021.
- [110] Bruce Mitchell and Juan Franco. HOLC "Redlining" maps: The persistent structure of segregation and economic inequality. Technical report, Natioanl Community Reinvestment Coalition, Washington, DC, 2018.
- [111] Andy Olin. Texas economy ranks near the top for racial equality in a nation with pervasive wealth inequality — The Kinder Institute for Urban Research, 2020.
- [112] US Environmental Protection Agency. EPA 's Risk -Screening Environmental Indicators (RSEI) Methodology. Technical Report Version 2.3.8, US Environmental Protection Agency, Washington, DC, 2019.
- [113] Bruna Galobardes, Mary Shaw, Debbie A Lawlor, John W Lynch, and George Davey Smith. Indicators of socioeconomic position (part 1). *J Epidemiol Community Health*, 60:7–12, 2006.
- [114] C-Y Cynthia Lin and Zachary D Liscow. Endogeneity in the environmental kuznets curve: an instrumental variables approach. *American Journal of Agricultural Economics*, 95(2):268–274, 2013.
- [115] Jianli Wu and Yue Pu. Air pollution, general government public-health expenditures and income inequality: Empirical analysis based on the spatial durbin model. *Plos one*, 15(10):e0240053, 2020.

- [116] United States Census Bureau. Houston city, Texas - Census Bureau Profile, 2022.
- [117] Taylor M. Oshan, Jordan P. Smith, and A. Stewart Fotheringham. Targeting the spatial context of obesity determinants via multiscale geographically weighted regression. *International Journal of Health Geographics*, 19(1), 2020.
- [118] Kerry Ard. Trends in exposure to industrial air toxins for different racial and socioeconomic groups: A spatial and temporal examination of environmental inequality in the U.S. from 1995 to 2004. *Social Science Research*, 53(2015):375–390, 2015.
- [119] Paul Mohai and Robin Saha. Which came first, people or pollution? Assessing the disparate siting and post-siting demographic change hypotheses of environmental injustice. *Environmental Research Letters*, 10(11), 2015.
- [120] Gilbert C. Gee and Chandra L. Ford. Structural racism and health inequities: Old Issues, New Directions. *Du Bois Review*, 8(1):115–132, 2011.
- [121] Jennifer Zanoni. Rural Texas. Technical report, Geography Division U.S. Census Bureau, 2018.
- [122] Rainald Borck and Philipp Schrauth. Population density and urban air quality. *Regional Science and Urban Economics*, 86(October 2020):103596, 2021.
- [123] Paul Mohai, Paula M. Lantz, Jeffrey Morenoff, James S. House, and Richard P. Mero. Racial and socioeconomic disparities in residential proximity to polluting industrial facilities: evidence from the Americans’ Changing Lives Study. *American journal of public health*, 99 Suppl 3:649–656, 2009.
- [124] Ziqi Li and A Stewart Fotheringham. Computational improvements to multi-scale geographically weighted regression. *International Journal of Geographical Information Science*, 34(7):1378–1397, 2020.
- [125] Hiroko Tabuchi and Nadja Popovich. People of Color Breathe More Hazardous Air. The Sources Are Everywhere. - The New York Times, 2021.
- [126] Douglas S Massey, Jonathan Rothwell, and Thurston Domina. The changing bases of segregation in the united states. *The Annals of the American Academy of Political and Social Science*, 626(1):74–90, 2009.
- [127] John Hope Franklin. History of racial segregation in the united states. *The Annals of the American Academy of Political and Social Science*, 304(1):1–9, 1956.
- [128] David K Fremon. *The Jim Crow laws and racism in American history*. Enslow Publishers Berkeley Heights, NJ, USA., 2000.
- [129] Daniel Aaronson, Daniel Hartley, and Bhashkar Mazumder. The effects of the 1930s holc” redlining” maps. *American Economic Journal: Economic Policy*, 13(4):355–92, 2021.
- [130] US Environmental Protection Agency. Air Quality - National Summary — US EPA, 2022.

- [131] Center on Budget and Policy Priorities. Income inequality has grown in Texas, 2023.
- [132] Megan Kirby-McGregor, Chen Chen, Hong Chen, Tarik Benmarhnia, and Jay S. Kaufman. Inequities in ambient fine particulate matter: A spatiotemporal analysis in Canadian communities. *Science of the Total Environment*, 858(October 2022):159766, 2023.
- [133] Tobias Rüttenauer and Henning Best. Consistent inequality across Germany? Exploring spatial heterogeneity in the unequal distribution of air pollution. *Research Handbook on Environmental Sociology*, (2021):41–61, 2021.
- [134] Cindy M. Padilla, Wahida Kihal-Talantikite, Verónica M. Vieira, Philippe Rossello, Geraldine Le Nir, Denis Zmirou-Navier, and Severine Deguen. Air quality and social deprivation in four French metropolitan areas-A localized spatio-temporal environmental inequality analysis. *Environmental Research*, 134:315–324, 2014.
- [135] Jing Ma, Bochu Liu, Gordon Mitchell, and Guanpeng Dong. A spatial analysis of air pollution and environmental inequality in beijing, 2000–2010. *Journal of Environmental Planning and Management*, 62(14):2437–2458, 2019.
- [136] Lara P. Clark, Dylan B. Millet, and Julian D. Marshall. Changes in transportation-related air pollution exposures by race-ethnicity and socioeconomic status: Outdoor nitrogen dioxide in the United States in 2000 and 2010. *Environmental Health Perspectives*, 125(9):1–10, 2017.
- [137] Anna Rosofsky, Jonathan I. Levy, Antonella Zanobetti, Patricia Janulewicz, and M. Patricia Fabian. Temporal trends in air pollution exposure inequality in Massachusetts. *Environmental Research*, 161(May 2017):76–86, 2018.
- [138] Sijing Liu, Nannan He, Yuting Shi, and Guoqi Li. The roles logistics agglomeration and technological progress play in air pollution — New evidence in sub-regions of Chongqing, China. *Journal of Cleaner Production*, 317(April):128414, 2021.
- [139] Abdulrahman Jbaily, Xiaodan Zhou, Jie Liu, Ting Hwan Lee, Leila Kamareddine, Stéphane Verguet, and Francesca Dominici. Air pollution exposure disparities across US population and income groups. *Nature*, 601(7892):228–233, 2022.
- [140] Jonathan Colmer, Ian Hardman, Jay Shimshack, and John Voorheis. Disparities in PM 2.5 air pollution in the United States. *Science*, 369(6503):575–578, 2020.
- [141] Mercedes A. Bravo, Joshua L. Warren, Man Chong Leong, Nicole C. Deziel, Rachel T. Kimbro, Michelle L. Bell, and Marie Lynn Miranda. Where Is Air Quality Improving, and Who Benefits? A Study of PM2.5 and Ozone Over 15 Years. *American Journal of Epidemiology*, 191(7):1258–1269, 2022.
- [142] Timothy W. Collins, Sara E. Grineski, Yasamin Shaker, and Casey J. Mullen. Communities of color are disproportionately exposed to long-term and short-term PM2.5 in metropolitan America. *Environmental Research*, 214(P4):114038, 2022.

- [143] John K. Kodros, Michelle L. Bell, Francesca Dominici, Christian L'Orange, Krystal J. Godri Pollitt, Scott Weichenthal, Xiao Wu, and John Volckens. Unequal airborne exposure to toxic metals associated with race, ethnicity, and segregation in the USA. *Nature Communications*, 13(1), 2022.
- [144] Waldo R Tobler. A computer movie simulating urban growth in the detroit region. *Economic geography*, 46(sup1):234–240, 1970.
- [145] National Archives. The Emancipation Proclamation, 2023.
- [146] National Archives. Civil Rights Act (1964), 2023.
- [147] US Environmental Protection Agency. EasyRSEI Dashboard v2.3.10, 2023.
- [148] United States Census Bureau. American Community Survey Summary File, 2022.
- [149] United States Census Bureau. 2017 TIGER/Line® Shapefiles, 2023.
- [150] COHGIS Open Data Portal. HOUSTON CITY LIMIT, 2020.
- [151] Robert J Bursik Jr and Jim Webb. Community change and patterns of delinquency. *American journal of sociology*, 88(1):24–42, 1982.
- [152] Elizabeth A. Richardson, Niamh K. Shortt, and Richard J. Mitchell. The mechanism behind environmental inequality in Scotland: Which came first, the deprivation or the landfill? *Environment and Planning A*, 42(1):223–240, 2010.
- [153] Janet Currie, John Voorheis, and Reed Walker. What Caused Racial Disparities in Particulate Exposure to Fall? New Evidence from the Clean Air Act and Satellite-Based Measures of Air Quality. *American Economic Review*, 113(1):71–97, 2023.
- [154] Christopher Jones and Daniel M Kammen. Spatial distribution of us household carbon footprints reveals suburbanization undermines greenhouse gas benefits of urban population density. *Environmental science & technology*, 48(2):895–902, 2014.
- [155] May A Massoud, Akram Tarhini, and Joumana A Nasr. Decentralized approaches to wastewater treatment and management: applicability in developing countries. *Journal of environmental management*, 90(1):652–659, 2009.
- [156] Petter Naess. Urban planning and sustainable development. *European Planning Studies*, 9(4):503–524, 2001.
- [157] Michael D LaGrega, Phillip L Buckingham, and Jeffrey C Evans. *Hazardous waste management*. Waveland Press, 2010.
- [158] US Environmental Protection Agency. Pollution Prevention (P2) — US EPA, 2023.
- [159] Menglu Zhao, Huijuan Wang, Jiaxun Sun, Rui Tang, Boya Cai, Xiaoyong Song, Xinmiao Huang, Jian Huang, and Zhengqiu Fan. Spatio-temporal characteristics of soil cd pollution and its influencing factors: A geographically and temporally weighted regression (gtwr) method. *Journal of Hazardous Materials*, 446:130613, 2023.

- [160] Donald E Farrar and Robert R Glauber. Multicollinearity in regression analysis: the problem revisited. *The Review of Economic and Statistics*, pages 92–107, 1967.
- [161] Xinxin Wang and Kevin Lo. Just transition: A conceptual review. *Energy Research & Social Science*, 82:102291, 2021.