THE SOCIAL DETERMINANTS OF HEPATITIS C – A CASE STUDY OF THE UNITED STATES

BY.
NICKLAS MATTI BRYDER

Lund University

Department of Human Geography

Bachelor thesis, spring 2018

Supervisor: Josephine Rekers

Word count: 16847

ABSTRACT

The social determinants of health are important to understand for the development of a world of equal opportunities in regards of life quality and length. This thesis aimed to contribute into such research using quantative methodology on Hepatitis C infection rates. Correlations were found for multiple social factors, but the study was unable to significantly work these results into a regression. Only three variables were significant, poverty, incarceration rates and African American or Black identity. Broadly this thesis highlights the need for a more holistic view on the social circumstances and to take into consideration multicollinearity when constructing a study. Furthermore, it highlights the need to use right scale when performing geographical research.

CONTENTS

1. Introduction 5
1.1. Problem Formulation 5
1.2. Research Questions and Aims 7
2. Conceptual Framework 9
2.1. The Social Determinants of Health 9
2.1.1. Financial Situation 10
2.1.2. Healthcare Access 11
2.1.3. Education 12
2.1.4. Community 13
2.1.5. The current Focus of the SDH and the Case for Expansion 14
2.2. The Risk Factors for Hepatitis C 15
2.3. The Importance of Social Factors 17
2.3.1. Intravenous Drug Use 18
2.3.2. Unsafe healthcare settings 19
2.3.3. Contaminated Tools 21
2.4. Does Geography Matter? 21
2.4.1. Clustering of Social Factors 21
2.4.2. HCV and the Spatial Dimension 23
2.5. Conclusion 24
3. Methodology 26
3.1. Data Collection 26
3.2. Research Design 27
3.3. Methodological Theory 28
3.3.1. Limitations 30
3.3.2. Hypothesis 31

4. Results 33
4.1. Correlation Analysis 33
4.2. Regression Analysis 1 33
4.3. Regression Analysis 2 36
4.4. Moran's I 36
4.5. Geographically Weighted Regression 36
5. Discussion 41
5.1. Findings 41
5.2. Alternative Research 43
6. Conclusion 45
7. References 46
Appendix 1: Table of Residuals 54
Appendix 2: Full correlation table between all variables 55

1. INTRODUCTION

The following chapter will provide a problem formulation, a set of research aims and questions.

1.1. PROBLEM FORMULATION

Hepatitis C (HCV) has surpassed HIV/AIDS as the deadliest infectious disease in the United States in recent years (Bakalar, 2012). In the country an estimated three million people are currently living with the disease and the numbers are growing by approximately 17000 cases every year (HHS.gov, 2018). The Centre for Disease Control (CDC) has created an individual division to map the spread of the disease and slow its growth (CDC, 2016). Yet there are still long ways to go and new strategies to target the disease are required.

The disease is caused by a single stranded RNA virus of the *Flaviviridae* family and was first extensively characterized in 1988. It comes in six different strains with multiple subtypes, with strain 1, the first discovered, resulting in around 60% of all cases (Nakano et al, 2011). The disease causes its eponymous condition, viral hepatitis, and has risen to become the worlds largest cause of liver transplants. If left untreated the disease has the potential to cause cirrhosis, hepatocellular carcinoma and a variety of extrahepatic conditions. While potentially fatal the disease has a relatively low case mortality rate[1] with only approximately 3 out of every 1.000 of infected individuals dying as a result of the virus. However due to the high prevalence of the disease such mortality rates still have the possibility to kill thousands in the years to come.

The disease has the potential to cause both an acute and chronic infection, with the majority of cases being chronic. This means that the majority of cases set in slowly, with individuals being able to live years, if not decades with the disease before symptoms set in (Nelson et al., 2011), also giving infected individuals time to transmit the disease to others.

HCV can be cured and 95% of all patients with HCV are treated back to full health[2] through the usage of anti-viral medication and 25% of all acute patients clear even without the use of medication. Medication however is expensive and according to Henry (2018) the cost of treating all infections in the United States currently present would be approximately 310 billion USD. This cost is expected to rise as the disease prevalence rises and as current

[1] Case mortality rate is defined as the amounts of deaths per 100.000 infected individuals
[2] Full health is defined as no viral RNA detected three months after a viral suppression treatment, also referred to as a sustained virologic response.

patterns of increasing healthcare and medication cost continue (Henry, 2018). Beyond this the disease cause non-fatal health conditions, such as increased risk of Type II Diabetes, anemia and digestive issues (Healthline, 2018), upon affected individuals, having the potential to reduce the quality of their everyday life. As such, even with minimal symptom the disease has the potential to place massive burdens on both individuals and healthcare systems if left untargeted for mitigation.

The risk factors for the disease has been known for over two decades. The virus has a parenteral mode of transmission, meaning that it transmits outside of the alimentary tract, by for example intravenous, intramuscular or subcutaneous exposure (Denniston et al, 2014). It was early on shown to have a high prevalence rate among intravenous drug users, transmitted through shared needles, and amongst medical patents who had received blood transfusion before the identification of the virus (Denniston et al, 2014). These factors, and others, have provided deep insight into the personal decisions and risk factors which may cause an HCV infection.

Despite the extent of previously mentioned research there is little knowledge regarding how social circumstances may influence the spread of the virus. While social factors have no ability to directly cause a viral infection, they have the possibility to increase the likelihood of individuals exposure to the direct risk factors. In the modern day where the social determinants of health are becoming more pronounced (Braveman and Gottlieb, 2014) it is problematic that one of the worlds growing infectious diseases is potentially poorly understood from such angles. Understanding how social factors can influence HCV infection risk is an important tool for efficient and well-planned risk prevention. Due to the high variability of its strains and capacity to rapidly mutate there is no vaccine against HCV (Yu and Chiang, 2010). As such, the only tool for long term prevention is to mitigate the behaviors which increases infection risk. If social conditions having the potential to cluster in spatial dimensions and since healthcare policy in the United States is largely being decided on a state level, an optimal starting point to assess the relationship between HCV and social factors is through a spatial analysis. Extensive targeting to reduce viral presence has the potential to achieve similar goals as vaccination attempts, massively lowering the prevalence of the disease. Contributing to such goals have the potential to massively reduce social, economic and medical burdens. Contributing into growing research regarding the social determinants of health, this thesis will set out to explore those social circumstances.

1.2. RESEARCH QUESTIONS AND AIMS

The aim of this thesis is to understand two primary questions;

How does social factors influence the risk of Hepatitis C infection?
Which social factors are the most important to estimate HCV infection and why?

The thesis will use state level data in the United States and as such is partially place specific to the country. However, the underlying logic should be applicable to other countries with similar rates of HCV and similar social conditions.

The two research questions provide a baseline for contributing into both policy recommendation on the national and regional level and for suggesting further research which can be conducted to deepen our knowledge of the disease. Lastly it aims to contribute into a growing field of geographically driven social disease research to understand more broadly how space impact people's health.

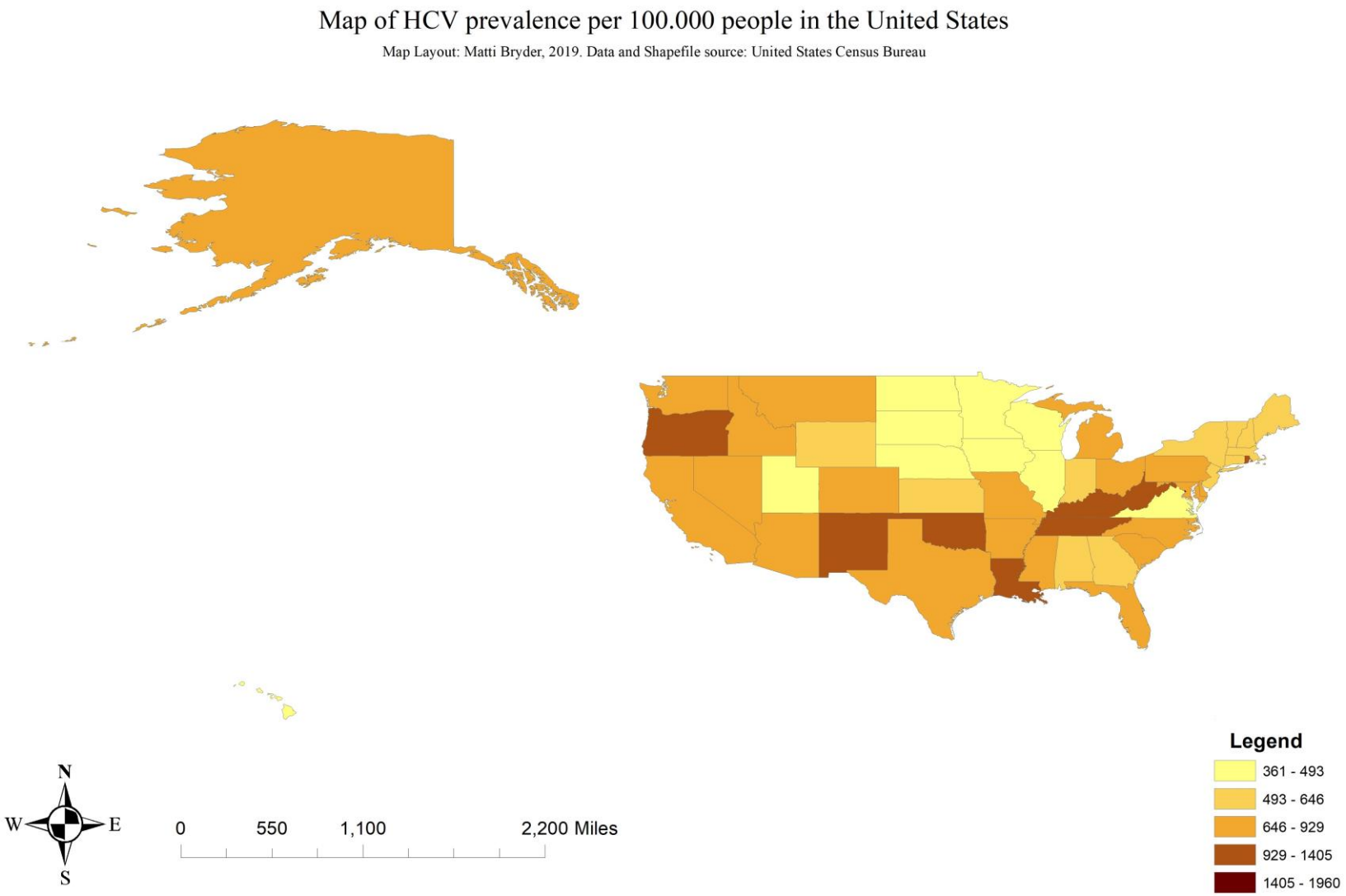


Figure 1: Map of HCV prevalence in the United States

2. CONCEPTUAL FRAMEWORK

This chapter will detail the risk factors for HCV, provide overview of theories which link those risk factors to social conditions, provide overview of theories which link those social conditions to geography and lastly present a conclusive theory regarding the answers to the research questions. Furthermore, this chapter will contain all relevant definitions not presented in the introduction.

2.1. THE SOCIAL DETERMINANTS OF HEALTH

The social determinants of health (SDH), are the social and economic factors which influence health statuses (Braveman and Gottlieb, 2014). There have been studied in some form at least since the 1840s with works done by Engel (1845) on the conditions of the British working class and Virchow (approx. 1840, most the writings were republished in 1985) on what is modern day Poland. In recent years they have come to the forefront of medical research and policy implementation across the globe, with the emergence of data that imply stark differences of health status between socioeconomic groups. Whereas the primary combatant of disease used to be innovation and technology it has now instead become the social and planning arm of states and international organizations. The very idea that an individual's social place has a non-consensual impact upon his or her health status is a question of social justice and a failure of healthcare systems, states and communities. Furthermore, with globalization potentially causing inequalities across multiple dimension to increase it is important to understand what the impacts are.

The SDH are perhaps the greatest concern for people engaged in medical questions in the modern world. In the United States Woolf et al (2007) concluded that more years of life were lost due to social inequalities than medical advances saved during the period between 1998 and 2002. While challenging to modern narratives it is a confirmation of McKeown's thesis (1955) which claimed that the health improvement across the Western world during the last century had more to do with economic equity than medicines or technology.

It is important to remember that, being poor or disenfranchised does not induce disease itself, especially not infectious disease. Instead social position influences the everyday situations and experiences individuals are exposed to, and it is those which, both in the short term and the long term, causes individuals to become sick. The studies of the SDH are studies of

"causes of causes" as it is commonly expressed in health literature (Braveman & Gottlieb, 2014)[3]. As such when dealing with issues of SDH it is important to establish knowledge in both the natural determinants of disease, such as pathogens or drug consumption, and the social circumstances which cause exposure to those natural causes.

To start analysing the SDH it is important to distinguish social factors from natural factors in order to not trap oneself in ecological fallacies. Social factors are those which are created by human social, political and economic structures, as well as historical circumstance, and as such are subject to change (WHO 2, 2018). These factors have the capacity to affect virtually all situations in our lives, except biological ones. As such differences between groups can be natural in that they are influenced by for example genetics or an aging population. The fact elderly people die more often than young ones or that healthy women live longer than healthy men are not products of the SDH even if they are quantitively able to be placed in certain groups. Moreover, with genetics and pathogens being spread in local spaces, certain diseases and health outcomes are geographically constrained without social influences. As such the SDH are not only correlational indicators but need to have a causality attached to the specific social factor. The following segment will provide a literary review over the commonly studied social variables often linked to differences in health outcomes.

2.1.1. FINANCIAL SITUATION

The financial situation of individuals both directly and indirectly affects their health. In the United States poverty has been shown to directly influence the stress level of individuals which measurable impact the risk for myocardial infarctions, heart-wall abnormalities and even sudden death (Dimsdale, 2008). This stress arises from a variety of uncertainties regarding the future of the self and of potential family. Stress is also a determinant for unhealthy behaviours such as tobacco and alcohol consumption in attempts to relieve them of their condition for a while (Thoits, 2012). Moreover, financial strain is often a contribution to family issues which deepen the risk for unhealthy stress (Dimsdale, 2008). Indirectly financial situation is a metric for accesses to essential goods and resources. This is partially location specific and its significance depends on regulatory matters. In the United States, with low reimbursement policies for medicines there is a discrepancy between economic classes on access to them. The same holds true for any good which has an economic value and is left victim to free market price mechanisms. This is due to that goods that have medical value

[3] In reality the SDH studies long linkages of causes, making maps of social determinants until the connection to the natural determinant is found.

often have a high demand yet not a high supply. Furthermore, economic situations influence decisions regarding what risks that need to be taken in order for improvement. Poor individuals are at higher likelihood to take jobs which pose a direct risk to their health and which offer low external benefits, such as insurance, due to higher desperation for financial reward. Lastly poverty has been shown to be a direct determinant of individuals' mental health status, low income is linked to lower self-worth and increased risk of the development of depression (Saraceno and Barbui, 1997).

2.1.2. HEALTHCARE ACCESS

Quality healthcare is essential for most people to live healthy lives. Healthcare encompasses all services that directly serve to diagnose and improve physiological function. This need is rooted in that human physiology is complex and to improve it specialized education and resources are needed.

Access to healthcare can be come in three primary forms, financial access, geographical access and cultural-community access. Financial access includes both the ability to directly pay healthcare systems and the ability to pay for insurance. Geographical access is the availability of relevant services, clinics or insurance offers, within specific spaces. Cultural-community access encompasses relevant trust in the healthcare system from the patient, which may arise from a variety of factors such as poor performance or historical discrimination. These dimensions have often been found to be overlapping (Kullgren et al, 2012)

Lack of healthcare access, due to financial means, measured by affordable insurance, were found by Woolhandler and Himmelstein (2017) in a review to increase the risk of premature death by 26% in the United States. This is due to a set of factors; the first is the inability to access essential medicines or treatments when under lethal acute conditions, the second is lack of treatment options for chronic conditions and the third is a general degradation of health behaviour. It has been highlighted that the healthcare system therefore is not only essential for disease treatment but for the long-term access to relevant knowledge regarding health behaviour which increase life expectancy. Furthermore, access to insurance decreases the longer individuals have been uninsured due to heightened health risks. The two most

notable affects[4] regarding chronic health issues, caused by behaviour, have been highlighted as hypertension and hyperlipidaemia.

Non-financial barriers share all of these conditions when they result in a complete barring from the healthcare system. However, both geographical and cultural access, which does not correlate with affordability often, only result in a partial lack of access. Regarding both of these factors African-American groups have been studied to a high degree. Due to a history of discrimination and segregation African-American are much more likely to both be geographically removed from healthcare systems and receive lower-quality treatment than non-Hispanic whites. Partial barriers to healthcare access have shown similar negative results as full barriers but with some lower impacts across certain facets, especially when dealing with acute medications. With the implementation of the ACA in 2014 the racial discrepancies in geographical access have been reduced and it is still not fully understood what end impacts it will have (Chen et al, 2016).

2.1.3. EDUCATION

Education has been recorded as being strongly correlated to health outcomes for several decades and has been a key focus point in health policy. While certainly possessing unique benefits studies, which focus on education have recently come under scrutiny for a variety of non-causal linkages and ecological fallacies. Being strongly interlinked with a set of other factors such as financial situation, ethnic belonging and geography the origin points for many of previously linked educational benefits have been debated. With education being so strongly linked to a set of factors which themselves it is important to analyse the underlying causal linkages and not only correlational data.

Regardless, education has a set of powerful health influencing factors that must not be ignored. The first of these is educations impact on potential economic conditions, leading into the financial aspects of health. Education and income are a consistent feedback loop and one of the roots of cross-generational poverty and health inequalities within liberal market economies. Education serves by providing access to high-income employment and healthy spaces. Furthermore, higher income and more equipped living spaces allow better access to education and increases the likelihood of children to themselves get access to improved lives.

[4] A varying set of conditions are have been claimed to be affected by insurance status in recent studies, however many of them suffer from poor statistical power and non-significant results (Woolhandler and Himmelstein, 2017).

Those who break this cycle often by becoming first generation secondary school graduates or achieve unusually high secondary school results have been recorded to achieve unique health benefits. The health benefits to first generation tertiary students are even more pronounced in regard to their physical health[5]. The first reason to this is linked to increased income and the second is less understood but is premised on improved health behaviours. These benefits have been highlighted by case studies regarding individuals and groups that have broken traditional norms of having educational status correlated to financial status, and vice versa (McFadden, 2016). The basic idea is that education brings out better behaviour in individuals in relation to their health. Knowledge regarding such subjects as dietary planning (Hiza et al, 2013), physical exercise and the dangers of tobacco (Escobedo & Peddicord, 1996) and alcohol consumption (Assari and Lankarani, 2016), provide the basis for a healthy long-term lifestyle.

2.1.4. COMMUNITY

The ecological models of health behaviours described in the later 1990s and early 2000s provided the basis for community and health interactions. They proposed that health behaviours were directly influenced by the social environments which individuals operated within. Community factors have the capacity to influence different groups, with relatively similar economic, educational and geographic situations, to different health outcomes (Goodman et al, 2014). This is done through a set of mechanisms and work both to the advantage and disadvantage of different groups.

By being placed in a community which have specific knowledges regarding health behaviours individuals are more likely to passively consume that knowledge or the behaviours which it leads to. This is used to, for example, further explain the positive aspects of education, where even if an individual does not consume the relevant knowledge themselves, they are more likely to passively adjust to it due to their exposure to others who have (Merzel & D'Afflitti, 2003). Similarly, if negative behaviours, such as smoking and alcohol consumption, become accepted within a community they are more likely to spread to new members as they integrate (Shelly et al, 2008). This can be caused either by the inflow of false information regarding the danger of these behaviours or by passive actions.

[5] First generation tertiary education has been shown to have more mental health issues compared to the average population due to increased pressure and perceptions of inferiority. Furthermore, first generation tertiary students are the highest group likely to drop out of their education, contributing to a feeling of failure and disappointment (McFadden, 2016).

While knowledge spill over is relatively understood in the sense that interactions between people allows them to learn new things, passive spill over is less understood. Moreover, different groups of people are more or less likely to be influenced by passive effects. A theory which has been brought forth by paediatric studies (Gerrard et al, 2005) is that subconscious positive associations between behaviours and important members of their community increases their risks. This is used to explain why children of all social classes are more likely to consume alcohol and tobacco if it is common amongst their parents but can also be used to describe the importance of community figureheads in general.

Community based health policies have become common in the last two decades with mixed results. While the successes have certainly highlighted that the field is promising for achieving beneficial health outcomes it has also highlighted a need to think holistically for each intervention (Mezel and D'Afflitti, 2003). Each community may differ in the underlying cause for its community behaviour. For certain communities it may be relevant to input the relevant knowledge and for others it may be important to bridge social interaction in order to ensure that the relevant knowledge can reach the intended individuals. Furthermore, it is important to understand the context of each community, what may from the outside be considered a single community may possess set of informal sub-communities and cross community interactions that can influence the results of an intervention. Lastly, new modes of communication and transportation allows for new dynamics in community creation and questions traditional approaches, how to influence communities across the internet and the different communities of the workplace compared to the home needs to be tackled in future medical policy.

2.1.5. THE CURRENT FOCUS OF THE SDH AND THE CASE FOR EXPANSION

In the United States rising levels of non-communicable diseases[6] has heavily influenced the empirical focus of medical resources overall. The research on the SDH have followed these patterns (Rosenberg, 2013). Social and economic dimensions have been successfully characterized as powerful influences upon individual's health related behaviours and accesses to official medical services. Moreover, poverty and discrimination have in themselves been proven to be powerful direct influences on individual's health, leading to higher rates of cardiovascular disease and mental health issues (Mays et al, 2007). Furthermore, economic

[6] Non-communicable diseases are those which are unable to spread between individuals and are acquired due to behavioral or genetic reasons. Common examples are various forms of cardiovascular, diabetes and various forms of cancer.

and social instability have been linked to dramatic increases in risky health behaviours such as alcohol and drug abuse (Leon, 1998).

Despite this understanding, improvements have been hard to come by and many factors are continuing to grow in influence. The perhaps strongest influence on life-expectancy is education, which has been known since the early 1990s, yet the divides between groups of different educations are growing and the different access to education between socioeconomic groups are increasing (Boschma, 2016; Olshansky et al, 2012). Beyond that intergenerational social mobility is becoming harder to achieve, with many children forced to endure the same social and economic conditions as their parents (Chetty et al, 2014).

While the SDH have become standard concepts in health policy across Africa, Latina America, Southeast Asia, Europe and Canada it is worrying how slow the United States is in its adaptation. It is perhaps the challenge to modern forms of technocapitalism or the decentralized political landscape which is the cause. Regardless of the reasons it is high time that the country adopts novel forms of policy implementation regarding health.

The focus of the research that exists is naturally on non-communicable diseases due to the heavy burden they place upon the healthcare system and society at large. The advancement of medical sciences in the last century have dramatically diminished the influence of infectious diseases[7]. Yet despite this there remains diseases for which there are little to no preventive measures to be taken and, especially in the United States, medication is expensive. The disregard for these diseases is not only potentially dangerous for the lives of individuals but also a question of social equity. If certain diseases are more common in disenfranchised groups and if those groups are left with heavier health and financial burdens, that is a display of injustice.

2.2. THE RISK FACTORS FOR HEPATITIS C

HCV is primarily spread through blood-to-blood contact with lower risks of it being transmitted vertically[8] or sexually. Furthermore, the onset of the chronic form of the disease has been shown to be accelerated by alcohol consumption.

For blood-to-blood contact three main factors have shown heavy correlation to HCV infection; intravenous drug use, blood transfusion and unsafe therapeutic injections (Shepard

[7] Infectious diseases are those that transmit between individuals. Common examples are AIDS, all Hepatitis forms, and Tuberculosis. The perhaps most common, is the flu.
[8] Vertically means spread from mother to child during pregnancy.

et al, 2005). Additionally, the sharing of penetrative equipment for non-medical use, such as piercing and tattooing equipment have been shown to spread the virus through blood contact (Tholme and Holmberg, 2012). The weight of these factors' dependants on geographical location and it is often claimed that intravenous drug use is the primary infection driver in developed countries while blood transfusions and unsafe therapeutic injections are dominant in the developing world (Shepard et al, 2005).

Intravenous drug use is a risk factor due to the practice of sharing and reusing injection needles during consumption. Traces of blood and HCV can remain for relatively long periods on injection equipment. Furthermore, drug users have a higher prevalence of HIV, which has been found to increase infection rates and lower liver function which accelerates the onset of the disease. Lower liver function has been traced to both drug usage and to interlinkages between drug abuse and alcohol abuse.

Blood transfusion can cause an infection due to being transfused with infected blood and remaining viral presence on and in transfusion equipment. While not a transfusion, dialysis procedures have also been linked to higher HCV prevalence and to multiple local outbreak of HCV. In modern medical systems blood is tested for HCV and the rates of occurring due to infected blood is virtually non-existent in new patients in the developed world. Despite advances in medical hygiene infection due to improperly sanitized equipment still occurs in both the developed and developing world, though at different rates.

Unsafe therapeutic injections have recently been recorded to be a possible larger cause than previously thought, at least in the developed world. A study by Denniston et al (2014) found that minimally several hundred cases of HCV infection in the United States and Europe during 1998-2008 were due to improper injections in healthcare settings, occurring on both hospital and non-hospital healthcare settings.

No reported case of HCV has occurred in professional settings due to either piercing or tattooing. However, in private and institutionalized settings the risk has been recorded as significant (Tholme and Holmberg, 2012).

Vertical transmission is rare and has primarily been recorded in HIV positive mothers whose vulnerable immune system had allowed for an abnormally extensive infection (Denniston et al, 2014).

Sexual transmission is also rare and has primarily been recorded spreading during male-to-male intercourse. The extent of the risk continues to be a topic of research, but general consensus seems to indicate that the average risk is low (Rooney and Gilson, 1998). Two factors are known to increase the risk, rough and/or violent sexual penetration and HIV infection (Denniston et al, 2014). These two factors also increase the risk of heterosexual intercourse where the predominant risk is for women who have sex with an infected partner. No significant evidence exists that there is a transmission risk during women-to-women sexual intercourse.

In summary, intravenous drug consumption is the dominant form of transmission, followed far after by unsafe healthcare settings and sharing of contaminated tools. For the purpose of this thesis both vertical transmission and sexual transmission will be disregarded due to their weak and still undeveloped linkages. That is not to say that they might not affect the data but the lack of understanding around their risks and their linkages to a biological phenomenon makes construction of further theory difficult and impractical.

2.3. The Importance of Social Factors for Hepatitis C

It is important to note that all of the relevant risk factors for HCV infection are caused by human actions. Contrary to certain other infectious diseases HCV can not spread through animals, food, the air, etc and must instead be delivered through invasive procedures. As such the disease is a prime example where changing individual behaviour and implementing social policy can have massive impact. Furthermore, the actions and risks detailed therein are not caused by random chance but can be claimed to be the result of social and economic conditions. This segment will detail how the three major risk factors, intravenous drug use, unsafe healthcare settings and sharing of contaminated tools can be indicated to result from socioeconomic processes. It will conclude with a summary of the most relevant aspects to study in this thesis. Primarily lessons regarding these concepts are drawn from research on the HIV/AIDS epidemic, which partially has similar modes of transmission, and from HCV specific research. All data and studies will be drawn from a North American context unless otherwise stated.

2.3.1. INTRAVENOUS DRUG USE[9]

Drug abuse has long been claimed to have association with socioeconomic conditions and been heavily attributed to lower economic classes and racial groups. Data suggest that on an average this claim might be a result of record and research bias, as well as communal visibility of sales and that communities of different socioeconomic classes have equal levels of drug consumption (Saxe et al, 2001).

Most importantly in relation to this thesis however is that poor individuals are dramatically more likely to experience negative side effects from their consumption, including HCV infection. The reasons for that are several; first of all, the stigmatization of being poor and a drug user is much higher than being richer and a consumer according to a study by Room (2009) and secondly poor people are more often exposed to the dangerous side-effects, such as reused or contaminated. A study by Magura et al (1989) provided early work on the rationalization regarding needle sharing and infectious disease. Studying HIV/AIDS the group found that intravenous drug users shared needles due to four primary reasons, 1) not owning personal injection equipment, 2) financial incentives, 3) a fatalistic attitude towards the associated risks and 4) peer pressure. The latter is a geographical concept and will be explored later, but the first three are mostly individual matters.

The reasons for not owning injection equipment, beyond the financial ones, were heavily associated with the stigma of being able to be discovered as a user and by the self-identification of being an addict. Combing Magura et al (1989) and Room (2009) it become evident that there is a high possibility that disenfranchised groups not owning their own equipment even if they could. Having the potential to occur greater stigmatization and loss of social connectivity to their community they would be more inclined to share needles than their non-disenfranchised counterpart.

[9] The reasons for why specific groups are more likely to commit risk health behaviour is debated topic and one that needs careful consideration. Characterizing such risks is important to understand how to tackle both the direct effect and the externalities. It is also important to remember that logic which attributes certain stigmatized behaviours to specific groups have been used to argue for racist and segregating policies in both the past and the present. Furthermore, spread of such ideas have potentially contributed into lowered bridging social capital between socioeconomic groups, creating both political and social tensions. As such it is vital to characterize the full causal linkages when attributing behaviours to specific groups and to be as mindful as possible about its potential misuse.

Financial incentives to share needles punish the poor directly. Making the choice between needle sharing and other purchases is difficult for people who are addicted and in poor economic situations. Purchasing injection equipment would potentially mean trading it off against other objects, potentially food, housing or consuming more drugs.

Fatalistic attitudes towards the risks associated with needle sharing encompasses two mindsets. The first is the notion that due to past behaviour one is already likely to have caught a disease from needle sharing and that future behaviour will not change it. The second is that the impact of the disease is not harmful enough for it to matter to them. The first of this is a circular problem, if the assumption and data that disenfranchised individuals are more likely to share hold true due to other reasons this created a loop where the risk increases. The second impacts people who have low self-value disproportionally due to a lower value placed on their lives. Saraceno and Barbui (1997) postulated that disenfranchisement has linkages to depression and other self-worth lowering mental conditions. In combination it could be assumed that disenfranchisement, lowing self-worth, would increase the risk of individuals possessing both forms of fatalistic attitudes.

Note that disenfranchisement can come in different forms and along both social and economic dimensions. While some logic is applicable only to poverty, a vast amount of it is applicable to other forms as well. Data from the United States support these assumptions, the poor are shown to be significantly more likely to share needles than the rich, and so are Hispanics and African Americans compared to Whites. This is troubling as there is no confirmed correlation between higher drug use due to poverty and both African Americans and Hispanics are less likely to use drugs than Whites (McCabe et al, 2008).

2.3.2. UNSAFE HEALTHCARE SETTINGS

Unsafe healthcare settings are not created due to random chance. They are the product of continued negligence by healthcare workers, healthcare systems and regulatory agencies. Maintaining contaminated healthcare spaces is not only ethically wrong but legally as well. But the problem is that even in places that are discovered to break regulation and shut down, the harshest consequence is still placed on the patients who were treated therein.

Denniston et al (2014) showed through a review of HCV epidemiology that annually 1-2 new outbreaks occur due to improper healthcare settings. The primary source of infections are haemodialysis centres and has been for decades (Tokars et al, 2004). Substantial policy measurements have been taken to mitigate that risk, with relative success (Denniston et al,

2014 and CDC, 2009). New data is shifting the focus towards non-haemodialytic settings, such as local clinics and alternative medical offices. These places have had lesser regulatory control in recent years than their haemodialytic counterparts. Thompson et al (2009) identified four reasons for why these cases were occurring, having studied 33 different outbreaks between 1998 and 2008, 1) Improper medical training, 2) Professional oversight, 3) Ease of licencing and 4) Improper public awareness.

Improper medical training among healthcare workers (HCWs) works along two dimensions. The first is the quality of their baseline education and the second is their continued education regarding novel discoveries, novel health burdens and simply reminders to adhere to safety practices. Professional oversight is the lack of control responsible actors have for the knowledge and practices their employees have. Licencing issues have been recorded amongst well established physician and nurses, due to the difficulty of removing medical licences and assigning direct responsibility for unsafe practices. Lastly, public awareness of the danger of hospital settings need to be improved in order to be able to put pressure on healthcare systems to adapt to best practice.

Furthermore, Thompson et al (2009) makes the claim that the decentralization of the United States healthcare system, with a rising amount of healthcare being provided in non-hospital settings is a determinant. This decentralization causes administrative issues with regulations as more places need to be constantly administered, leading to regulatory oversight. Petrosillo et al (2001) contributes into this field with evidence that understaffing in non-hospital settings contributes into HCV infection rates. With higher workloads HCWs are more likely to neglect proper sanitary regulations.

Additionally, Denniston et al (2014) recognized that the ability to bring hospitals who perform malpractice to justice is difficult for both institutions and individuals. The process is expensive, the evidence required is difficult to acquire and for individuals it often comes at personal costs regarding time. Both Thompson (2009) and Denniston et al (2014) recognized that the found cases of healthcare settings causing HCV is probably just the tip of the iceberg.

Lastly it is important to remember that one of the most unsafe healthcare settings is one where individuals are unable to access healthcare at all. In the United States with healthcare being run primarily by private actors and with one of the highest drug costs in the world there are many people who are denied this essential service. This both delays the diagnosis of HCV

and makes it so that certain diagnosed individuals are unable to access treatment (Henry, 2018).

2.3.3. CONTAMINATED TOOLS

Contaminated tools, such as tattooing and piercing equipment is rare or even possibly non-existent in professional parlours in the United States (Tohme and Holmberg, 2012). This risk seems to solely exist within homes and in prisons. Two causes exist for performing at home procedures, financial and educational ones. Individuals who are unaware of disease risk from such events and individuals who cannot pay for a professional parlour are more likely to expose themselves to such risks.

Incarceration is a major risk factor for HCV. Within prison environment contaminated tools are far more common than on the outside, with no official parlours operating within. Traditions of tattooing and other body modifications amongst incarcerated individuals have furthered this issue (Spauling and Thomas, 2012).

The origin point of HCV prevalence in incarceration environments potentially has to do with the socioeconomic composition of inmates. Many of the risk factors previously described for HCV are also risk factors for incarceration. Lower income individuals and ethnic minorities are incarcerated at much higher rates than privileged individuals (Carson, 2018). Such a pattern causes a loop in risk factors.

2.4. DOES GEOGRAPHY MATTER?

Geography matters when looking for patterns of infectious disease. Spread between people, it is important to understand how individuals operate in space to spread disease. This segment will provide two main contributions; first it will characterize how the factors described in the previous section cluster in space, based partially on theories regarding residential sorting, gentrification and resource access, secondly it will provide unique spatial characteristics which influence HCV rates and in conclusion it will link these segment together to show how important spatial characteristics are unevenly distributed over space and between socioeconomic groups.

2.4.1. CLUSTERING OF SOCIAL FACTORS

Social factors are not randomly distributed over space but are often concentrated along geography, due to social and historical processes. Where individuals can and choose to live

are dependant on social and economic conditions and the quality of public services is often dependant on the ability of local resources and institutions.

Economic classes cluster in space (Plucinski, et al, 2013 and Lichter et al, 2012) and their separation are part of a broader set of studies concerning what is referred to as residential sorting. Based in economic concepts of utility it postulates the argument that individuals sort themselves into different neighbourhoods depending on the utility they perceive to get out of that neighbourhood's characteristics. Combining that with resource limitations regarding commonly sough after characteristics and we get a basic explanation of why housing prices are spatially distributed. Commonly sought-after characteristics which are limited in their supply will ensure that certain neighbourhoods or areas occur price increases along relatively simple supply-demand logic. As costs of housing increases so does generally the average income of the residents in a particular space, due to affordability. Most obvious of these is the linkages between local quality resources, such as schools to housing prices. Kane (2006) found that an increase in standardized tests scores in schools were strongly correlated to increased housing costs in the same area in the United States. Similar concepts have been applied to other sought-after resources such as green spaces (McCord et al, 2014), quality public transportation, crime (Hellman and Naroff, 1979) and other often sought-after characteristics (Hopkins, 2017). Under the assumption that certain characteristics are sought after equally it becomes the upper economic classes which are able to access them under a housing market with liberal price mechanisms[10]. An additional overarching issue with this phenomenon is that location specific improvement in sought after characteristics do not apply their benefits to the local population in the long term. As the land which they live on is improved gentrification mechanism often ensure that individuals are unable to pay for rising rents on their housing. Wolch et al (2014) showed that direct attempts at providing lower-income areas in the United States with access to green spaces had the paradoxical issue of raising the market value of the properties located therein.

Even furthermore this system creates inherent feedback mechanisms which entrench this position. Wealth in many countries, including the United States, is a strong proxy value for both national and local political influence (Hacker and Pierson, 2014). While worrying for the long-term democratic stability of countries, it has direct effects upon the distribution of

[10] Even if housing prices are controlled similar patterns may emerge. Sought after characteristics may themselves cost, such as education and healthcare, and the only individuals who move to access them will be those that can afford to. If price control mechanisms are to be used in order for equal access those need to be holistically considered.

local resources. With increased local political influence higher income individuals have the capacity to influence decisions in their favour, promoting policies and agendas which allow improvement in spaces of their interests. As such they can petition for increased local school funding, the construction of healthy spaces and so forth to their own advantage. As this political power clusters in space the areas which become relevant to this group to improve become smaller. Wolch et al (2014) showed this type of pattern in the construction of healthy spaces where newly built, publicly funded, green spaces were predominantly located in upper-class neighbourhoods. This process of income segregation is increasing in the United States (Bischoff and Reardon in Logan, 2014).

Separating income segregation and modern-day ethnic segregation is at times difficult due to the overlap between ethnic minorities and economic poverty. It is rooted in a deep racist past with many characteristics of legal segregation still looming over the United States. In the modern day such segregation is a result of inequal economic opportunities and sets of unknown factors (Chang, 2006). It is known to continue however, with minority groups often creating enclaves or distinct geographic communities. With both of these factors correlating to access to healthcare, education (Van der Berg, 2007) and to increased incarcerating (Carson, 2018) risk it is natural that those factors cluster as well.

Different states in the United States have different capacity for social policy measurements and different levels of institutional quality (Melusky, 2016). As such the distribution of poor education, low-quality healthcare and affordable healthcare is likely to vary depending on state. U.S Census Bureau data (2018) confirms this association relative to education, between the lowest and highest rank in high school degree attainment there is over 10% difference. The difference for higher education is even more pronounced. Similarly, healthcare spending per capita on state level, healthcare cost for the individual and the number of individuals employed in healthcare per capita vastly differs between states. Similarly, special resources to target drug addiction such as special clinics or needle exchange programs differs between spaces (Des Jarleis et al, 2004).

2.4.2. HCV AND THE SPATIAL DIMENSION

HCV will always be a geographical concern due to its status as an infectious disease. It's inability to arise from nothing means that only people who are exposed to infected individuals will ever have a chance of getting it. Beyond that, for the risk factors described in this chapter there is research done (Jackson and Henriksen, 1997) which postulate that such

behaviours will spread in local spaces . The reason for this is the concept of social networks and how they cause changes in individuals' opinions and actions.

Magura et al (1998) found that one of the most significant risk factors for needle sharing is peer pressure and general acceptance from surrounding communities. Individuals who associated with people who shared injection needles were more likely than the average to also share, even if they possessed personal injection equipment. The risk for drug uses itself has been found by Williams and Latkin (2007) to be buffered by social contact with individuals who are not in that risk group. Lastly, Merzel and D'Afflitti, (2003) showed that community-based health intervention is a powerful tool for health policy. The reason for this is that communities form internal, and sometimes subconscious, opinions regarding health matters and behaviours, drawn from local key figures and values. Logically, the same aspects apply to healthcare behaviour if seen as a community, certain places have evolved to misplace the importance on sterilized environments and tools.

Jointly, combining these three ideas with risk factors, segregation and resource allocation allows us to understand how behaviour, arising perhaps due to random chance or a specific historical event become rooted in place. Spread through peer interactions and communities' specific ideas become accepted depending on historical circumstances. This also has the potential to make short term influencing factors spread, if an individual get exposed to a specific community norm they have the potential to spread that to other when their environment changes, for both good and bad.

2.5. CONCLUSION

HCV has clear social risk factors. In this chapter multiple ones were identified, and it was additionally argued that they would differ over space depending on local historical events. This research also contributes into a broader field of study regarding the social determinants of health and health equity. In summary seven were identified as vital to understand.

Economic vulnerability
Ethnic identity
Education
Healthcare standard
Healthcare access
Incarceration rates
Geography

The theory which will be explored in this paper is that all of these factors have a correlation with increased risk of HCV infection.

3. METHODOLOGY

This chapter will contain the methodological details of the empirical study, sources for data, the studies limitations and the hypothesis for the study. For the purpose of this thesis a quantative study of the discussed variables in chapter 2, section 4, and their relationship to HCV prevalence on a U.S state level was conducted using SPSS and ArcGIS. The overarching goal was to run a geographically weighted regression (GWR).

3.1. DATA COLLECTION

Seven variables were chosen to represent the factors which were identified in the conceptual framework in addition to HCV prevalence. Due to availability of data all data was taken for the year 2016.

- Incarceration rates were represented by the numbers of prisoners as a percentage of the total population. Data was collected from Carson (2018) who had collected inmate totals during 2016. Modification of the data was conduced by calculating the percentage using total population data from the 2012-2017 United States Census Bureau Estimates.
- Ethnic identity data was collected from the 2012-2017 United Sates Census Bureau Estimates. As the two largest ethnic minorities, and with supporting literature of structural disenfranchisement (Carson, 2018; Lichter et al, 2012; McGabe et al 2007), African American or Black and Hispanic or Latino[11] were chosen. Percentages rates were taken directly from the census data. The data for each group was used separately during the study to account for heterogeneity.
- Education was represented as the total percentage of the population who had completed high school education or higher. The data was taken directly from the United States Census Bureau Estimates 2012-2017.
- Healthcare standards were estimated from the percentage of the population employed in the healthcare sector. Other metrics were searched for but without success due to either using irrelevant calculations or not providing methodology. The total population employed in the healthcare sector was gathered from the United States

[11] The United States Census Bureau allows self -dentification of ethnic identity. With disagreeing opinions in the public regarding the correct term for ethnic identities African American and Black are one group, and Hispanic or Latino is one group.

Census Bureau Estimates 2012-2017 and the percentage was calculated from the total population from the same source.

- Healthcare accessibility was estimated from the percentage of people who lacked health insurance. With no state provided healthcare, outside of dialysis, access to healthcare is often dependant on having insurance. Important to note is also that both financial situation and ethnic identity is a limiting factor on healthcare. Ethnic identity due to discrimination in the healthcare system (Dula, 1994) and financial situation due to inability to pay for external healthcare costs which insurance does not cover. The data was provided by the Census Bureau Estimates 2012-2017.
- Financial vulnerability was measured by percentage of people who had been living in poverty during the last 12 months. The data was provided by the Census Bureau Estimates 2012-2017.
- HCV prevalence was collected from estimates in Rosenberg et al (2018). They drew from data from the National Health and Nutrition Examination Survey (NHANES) and applied a model to calculate current rates. The NHANES data itself was unrepresentative as a sample because both homeless and incarcerated individuals were excluded from the sample. Their methodology used HCV mortality, narcotic overdose rates and spatial demographics to estimate the rate for 2016. The NHANES census itself used presence of HCV RNA as the metric of infection. The data was taken in total number and converted into rates per 100000 individuals for the purpose of this thesis.
- The shapefile for the United States was taken from the United States Census Bureau for 2016. It was modified to remove Guam, American Samoa and the United States Virgin Islands.

3.2. RESEARCH DESIGN

The empirical study was conducted in four segments;

First, a correlation analysis was run in SPSS in order to calculate the Pearson Correlation for all of the independent variables to the dependant variable. All of the independent variables were then analysed in two categories, to establish their potential in a regression analysis. First their correlation coefficient was studied in order to establish a potential relationship, defined for this study as a value of greater than 0.3 or smaller than -0.3. The correlation coefficient range was between -1 and 1, with both -1 meaning total negative correlation and 1 meaning

total positive correlation. Secondly the significance, also known as p-value or alpha value, was studied (see 3.3 Methodological Theory for further details). All independent variables which did not fulfil the correlation coefficient or significance level were not continued with for the regression element of the thesis, these were; Hispanic population percentage and percentage of population not under health insurance.

Second, linear regression was performed with the remaining independent variables in SPSS and the results analysed. In this segment the significance of the model as a whole, significance of the individual regression coefficients and the collinearity diagnostics were evaluated (see 3.3 Methodological Theory for further details). All of the variables passed the collinearity diagnostics, but high school education level and individuals employed in healthcare percentage did not pass the significance level.

Third, a second linear regression was performed in SPSS with the independent variables not passing the former significance test removed. Significance and collinearity were once again evaluated, and all remaining variables passed. The formula for the model was calculated.

Fourth, a GWR was performed in ArcGIS. In order to check for the validity of the results two steps were taken prior. First a linear regression was performed once more, this time in ArcGIS to ensure the same results. Secondly the standardized residuals were analysed in a Moran's I test to ensure that the residuals were randomly spatially distributed. Both of these tests passed and the GWR was run.

All the relevant data was collected and formatted into either graphs, tables or maps.

3.3. METHODOLOGICAL THEORY

The GWR is an adaptation of Ordinary Least Square (OLS) with the ability to break away from the assumed global relationship between the studied variables. An OLS based regression will use the following model:

$$y = B_0 + B_1x_1 + B_2x_2 + B_3x_3 \ldots + B_px_p + \varepsilon$$

B_0 is the intercept of the model, the value of y when all other values are equal to 0

y = The dependant variable.

$B_{1,2,3\ldots p}$ = The regression coefficient / slope of an independent variable

$x_{1,2,3\ldots p}$ = The value of the independent variables

ε = The value of error term, all remaining factors not included as independent variables in the model

The OLS assumes that the relationship the model produces is constant over the dependant variable's sample and works well when dealing with natural phenomenon where one is trying to calculate natural laws. For social phenomenon however each part of the sample may vary in how influential a specific independent variable is. For geographical research the GWR allows the researcher to take such variance into account over spatial dimensions. Through analysis of its result the researcher can draw conclusions regarding the relative importance of a specific independent variable to a specific place. Moreover, the method allows for testing of variations in the validity and strength of the model over space.

In order to establish underlying rationalization for running a regression on a set of variables a relationship between the dependant variable, in this case HCV prevalence, and the independent variables, in this case education, poverty, ethnicity, education, healthcare access, healthcare standards and incarceration, an investigation into their correlation needs to be performed. The Pearson's Correlation, also known as Pearson's R, which was used for this thesis evaluates the potential linear relationship between two variables. It ranged from -1 to 1, where -1 represent perfect negative correlation, 1 represents perfect linear correlation and 0 no linear correlation. For this thesis values under -0.3 and over 0.3 were chosen as representing potential relationships and therefore continued with into the regression analysis.

The second evaluated factor in the correlation analysis and in all of the regressions was the significance of the evaluated relationships. Statistical significance evaluates the chance of an erroneous rejection of the null hypothesis, the null hypothesis stating that there is no relationship between two values. The significance level for this thesis was set at 0.05, meaning that a 5% chance of an erroneous rejection was acceptable. The 0.05 was chosen per academic tradition (Lindenmayer and Burgman, 2005).

When evaluating the strength of a regression the R^2, also known as coefficient of determination, is used to evaluate the percentage of the dependent variable is explained by the dependent variables. For a simple linear regression, the R^2 is simply the correlation coefficient squared. For a multivariable regression the R^2 is the sum of the regression coefficients squared, assuming that there are no correlations between the independent variables. If that is the case, the R^2 calculation takes that into account and produces a R^2 lower than the sum of its parts. During the GWR local R^2 values are generated which explain the predictive strength of the model depending on location. The predictor's coefficients ($B_{1,2,3...p}$) are the values which indicate the importance of the individual independent variables

for the complete model. They can take the form of any integer or fraction and may be positive or negative. The coefficient means that for each 1 increase in its related independent variable the dependant variable changes at the value of the coefficient.

Correlation between independent variables, known as collinearity for two variables or multicollinearity for more than two, is a problem in statistics as it makes the potential evaluation of the individual coefficients difficult to evaluate. For this purpose, collinearity diagnostics were run and the Variance inflation factor (VIF) and Tolerance was evaluated. Accepted value for VIF was sub 5 and for Tolerance over 0.3 (Hair et al, 1995). Note that collinearity does not affect the R^2 due to already taking it into account, it only affects the reliability of the individual predictor's coefficients.

Before the GWR could be run, the Moran's I needed to be run. The Moran's I test evaluated the residuals in the regression for spatial autocorrelation. The residuals are the difference between the observed value of the dependant variable and the expected value of the dependant variable according to the regression model. The smaller the residuals the better the model fits. Spatial autocorrelation is performed in order to evaluate the distribution of the local residuals and if their distribution is random. The value of Moran's I ranged between -1 and 1, with 1 meaning clustered distribution of similar values and -1 meaning clustering of dismissive values (Lin and Wen, 2011), clustering would imply a missing key explanatory variable which accounted for it, for example.

3.3.1. LIMITATIONS

This study carries four major limitations; potential missing independent variables, unknown transmission factors, the small sample of the data and the reliability of the secondary data gathered. Furthermore, time lag and migration are not taken into account during the study.

It is probable that missing explanatory variables were excluded from the model as a whole during the analysis. The essential one was for injectable drug use, which was sought for but not found, this variable would have allowed a deeper analysis on the dangers of injectable drug use contrary to injectable drug use in combination with other factors. But also, other variables, not discovered during the literary review may have been found and will therefore account to the residual in the model. It is also possible that the transmission of HCV is not fully understood. While there is strong and reliable data on risk factors, as Thompson et al (2009) points out there is still a large amount of HCV cases for which the cause was unknown. Lastly it is possible that the variables chosen in this study does not represent the

overall factor which this thesis claims, due to error in theory. As such, dismissal of potential relationships should not be done without supporting theory or additional data.

The data sample was relatively small with n=51 for all variables except for incarceration rate where n=50, due to the lack of prisons in the District of Columbia. This number of occurrences to not technically fulfil the recommended data size for a GWR, where the lower end is 200 (Andersson, 2017)[12]. This needs to be taken into consideration during the interpretation of the results. While an argument could be made that the sample was complete due to encompassing all U.S states a similar study could be performed on smaller scale with similar conceptual and methodological backing.

This thesis relies on the validity of the secondary data gathered from the United States Census Bureau and Rosenberg et al (2018). Sampling errors or other statistical errors in those sources would corrode the result of this study. Analysis of their methodology found no obvious fault and the sources were deemed reliable. The only potential for fault is in the HCV prevalence count in Rosenberg et al (2018) where diagnosed cases of HCV was used as a baseline for the estimates. According to the CDC (2018) a significant number of HCV cases may be undiagnosed, however without additional data on how many those would be there was no way to account for this phenomenon.

Lastly time and migration may play a role in this study. As the data only indicate prevalence, not infection rates there could be a significant difference in the environment where people contracted HCV and where they are living now. Transmission data may have been more reliable but was not found in a reliable source. As there would also be time lag between infection and diagnosis data that data was not more reliable than prevalence.

Lastly it is also important to note that the OLS and the GWR both only calculate linear relationships. If the relationship has another form the method will not produce relevant results.

3.3.2. HYPOTHESIS

[12]Andersson only makes the recommendation for large sample sizes. The 200 recommendation comes internally within ArcGIS.

H_0: Education, poverty, healthcare access, healthcare standards, African American or Black identity, Hispanic or Latino identity and incarceration rates have no relationship with HCV prevalence rates.

H_{A1}: Higher education levels have a negative relationship with HCV prevalence rates.

H_{A2}: Lower healthcare access have a positive relation ship with HCV prevalence rates.

H_{A3}: Lower healthcare standards have a positive relationship with HCV prevalence rates.

H_{A4}: Higher rates of African American or Black population has a positive relationship with HCV prevalence rates.

H_{A5}: Higher rates of Hispanic or Latino population has a positive relationship with HCV prevalence rates.

H_{A6}: Incarceration rates has a positive relationship with HCV prevalence rates.

Positive relationship means that as the values of the independent variable increases so does the dependant variable, negative means the opposite.

4. RESULTS

This chapter will provide all relevant tables, graphs and maps produced during the study. Each segment will be accompanied with a description of the results.

4.1. CORRELATION ANALYSIS

Hispanics or Latino and Healthcare access neither passed the correlation coefficient or the significance standards necessary for being used in the regression model. (See pages 34-35 for scatterplots of all variables). The remaining variables all had a coefficient which indicated a moderate correlation.

		HCV Prevalence	Poverty	Hispanics or Latino	Education
HCV Prevalence	Pearson Correlation	1	.587**	0.199	-.382**
	Sig. (2-tailed)		0.000	0.161	0.006
	N	51	51	51	51
		Incraceration	African American or Black	Healthcare Access	Healthcare Standard
	Pearson Correlation	.443**	.341*	0.174	-.437**
	Sig. (2-tailed)	0.001	0.014	0.221	0.001
	N	50	51	51	51

Table 1: Correlation analysis for HCV prevalence and chosen variables

4.2. REGRESSION ANALYSIS 1

Coefficients[a]							
	Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
Constant	2507.678	1467.820		1.708	0.095		
Poverty	33.927	15.949	0.364	2.127	0.039	0.397	2.517
Prisoners	454.435	216.671	0.291	2.097	0.042	0.603	1.658
African American or Black	-9.732	3.474	-0.391	-2.801	0.008	0.597	1.676
Education	-27.464	15.210	-0.347	-1.806	0.078	0.314	3.183
Healthcare Standards	27.346	59.865	0.056	0.457	0.650	0.763	1.311

Table 3: Coefficients for regression one

Model Summary				
R		R Square	Adjusted R Square	Std. Error of the Estimate
	.700[a]	0.489	0.431	179.507361865767000

Table 2: Model summary for regression one

Regression one had a R^2 of 0.489 which a implies a 48.9% description of HCV prevalence from the used variables. However, education and healthcare standard were above the significance level for the variables and as such had to large risk of erroneously rejecting the null hypothesis.

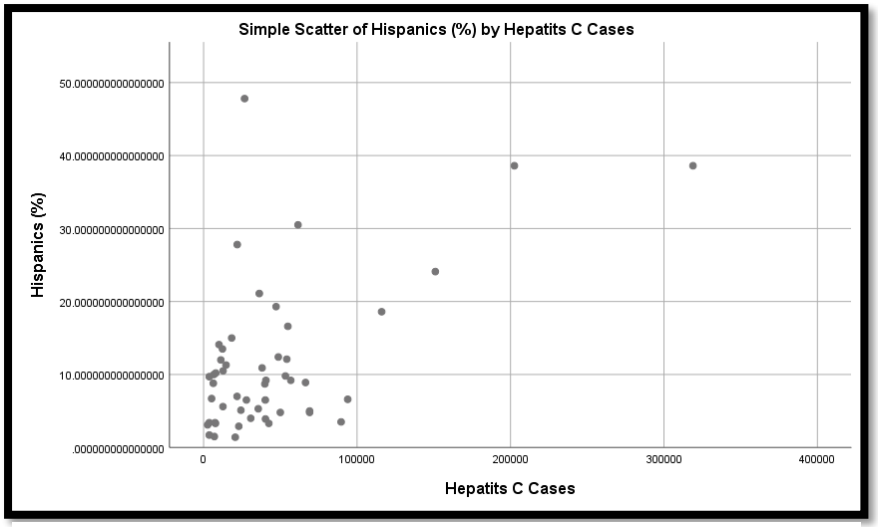


Figure 2: Scatterplot of correlation between HCV prevalence and Hispanics or Latinos

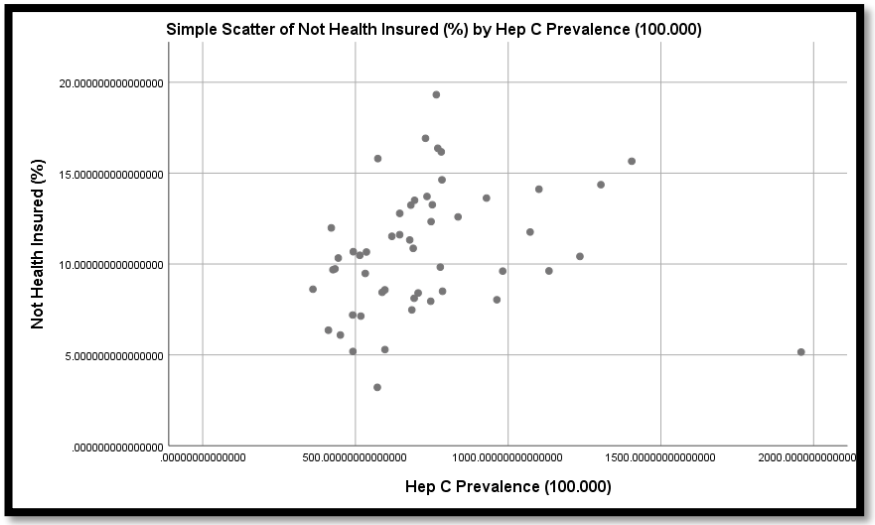


Figure 3: Scatterplot of the correlation between HCV prevalence and Healthcare insurance rates

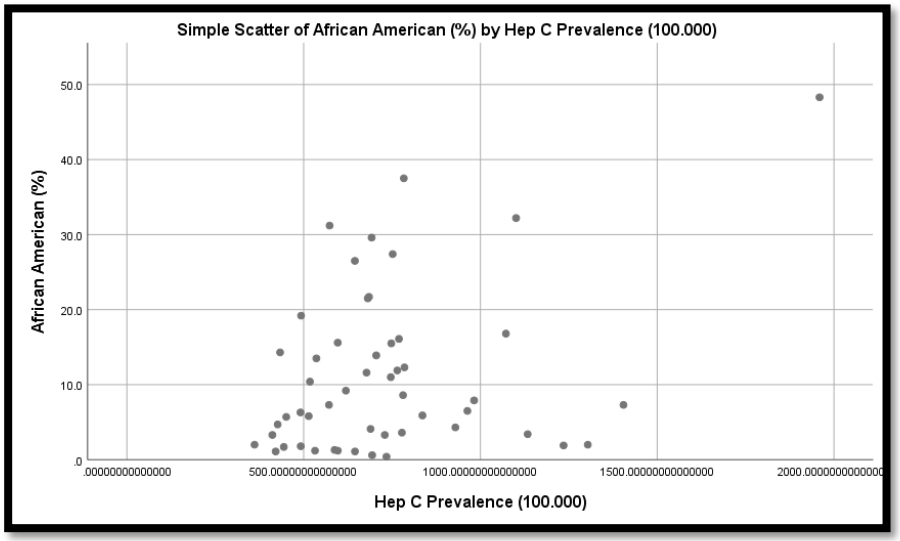


Figure 5: Scatterplot of correlation between HCV prevalence and African American or Black

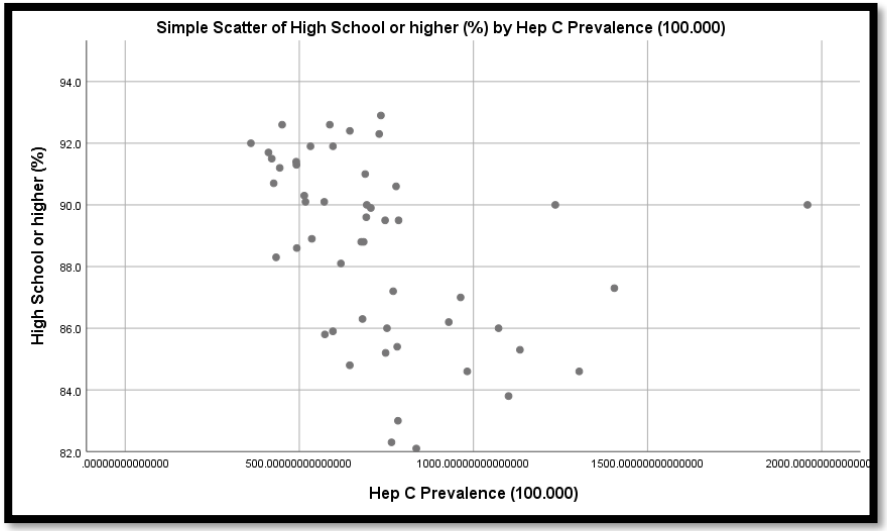


Figure 4: Scatterplot of correlation between HCV prevalence and education rates

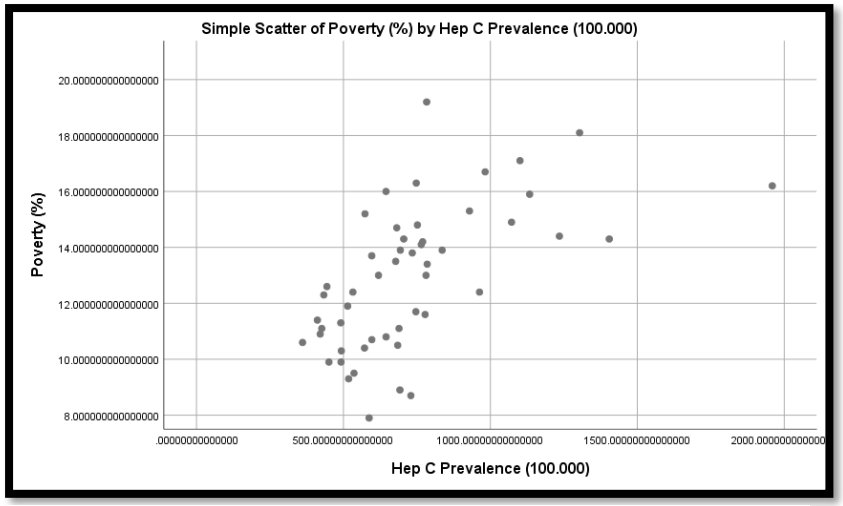


Figure 7: Scatterplot of correlation between poverty and HCV prevalence

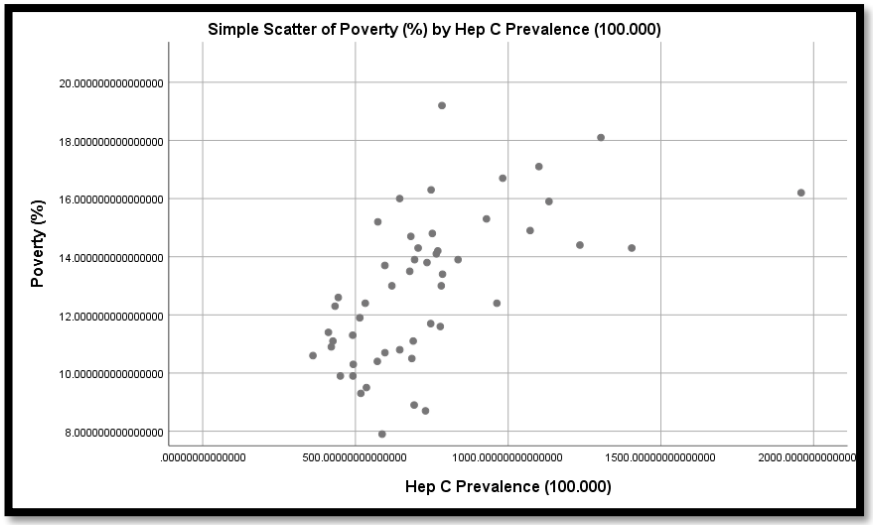


Figure 6: Scatterplot between HCV prevalence and poverty

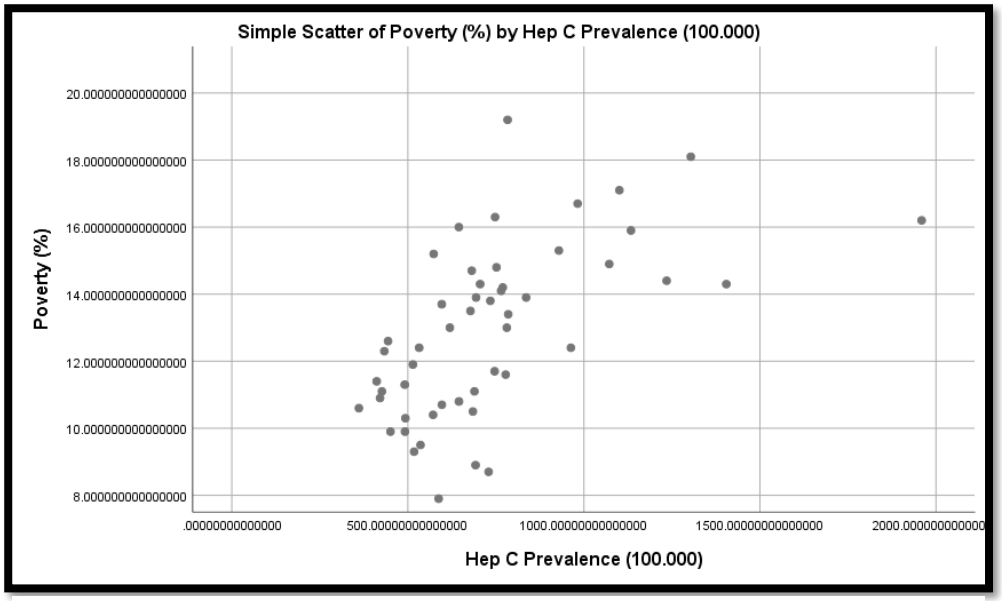


Figure 8: Scatterplot between HCV prevalence and incarceration rates

4.3. REGRESSION ANALYSIS 2

Coefficients[a]							
	Unstandardized Coefficients		Standardized Coeffi	t	Sig.	nearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
Constant	-85.866	136.224		-0.630	0.532		
Poverty	52.885	11.990	0.567	4.411	0.000	0.722	1.385
Prisoners	452.755	209.636	0.290	2.160	0.036	0.662	1.511
African American or	-6.952	3.132	-0.279	-2.219	0.031	0.754	1.326

Table 5: Coefficients for regression two

Model Summary				
R	R Square	Adjusted R Square	Std. Error of the Estimate	
.672[a]	0.452	0.416	181.950140016702000	

Table 4: Model Summary for regression two

The R^2 value for regression two implied a 45.2% explanation of HCV prevalence from the independent variables. All variables passed significance and collinearity tests.

4.4. MORAN'S I

The Moran's I indicated no spatial autocorrelation, which means the residuals were distributed randomly

4.5. GEOGRAPHICALLY WEIGHTED REGRESSION

The GWR indicated low spatial difference in the R^2 over space, range 0,0003, meaning a 0,03% difference in total effect. The local coefficients also had small range of difference over space. Both of these results indicate small spatial variation of the risk factors for HCV. (See maps in the following section for detailed results).

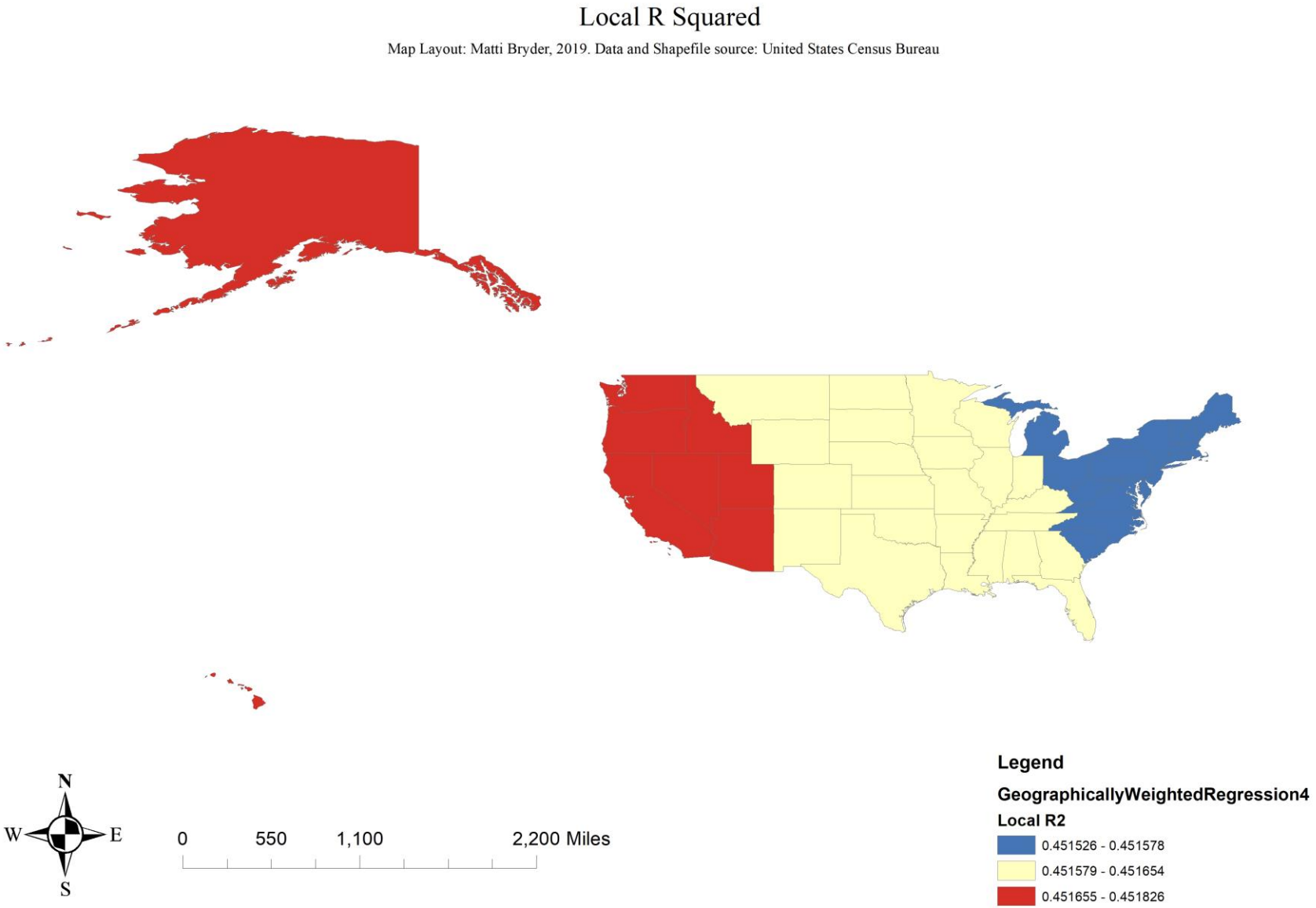


Figure 9: Map of localized R^2

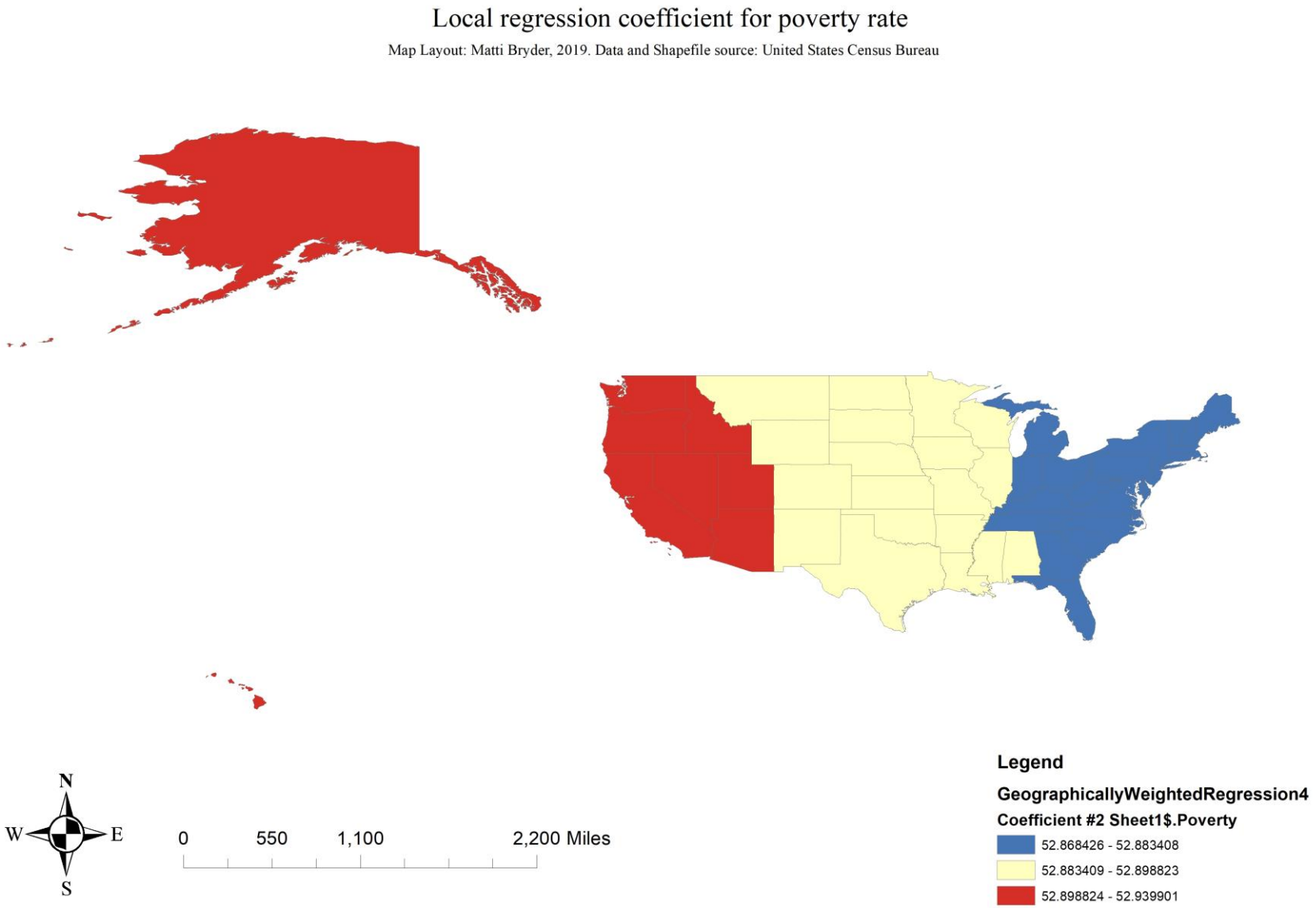


Figure 10: Local regression coefficient for poverty rate

4.6.

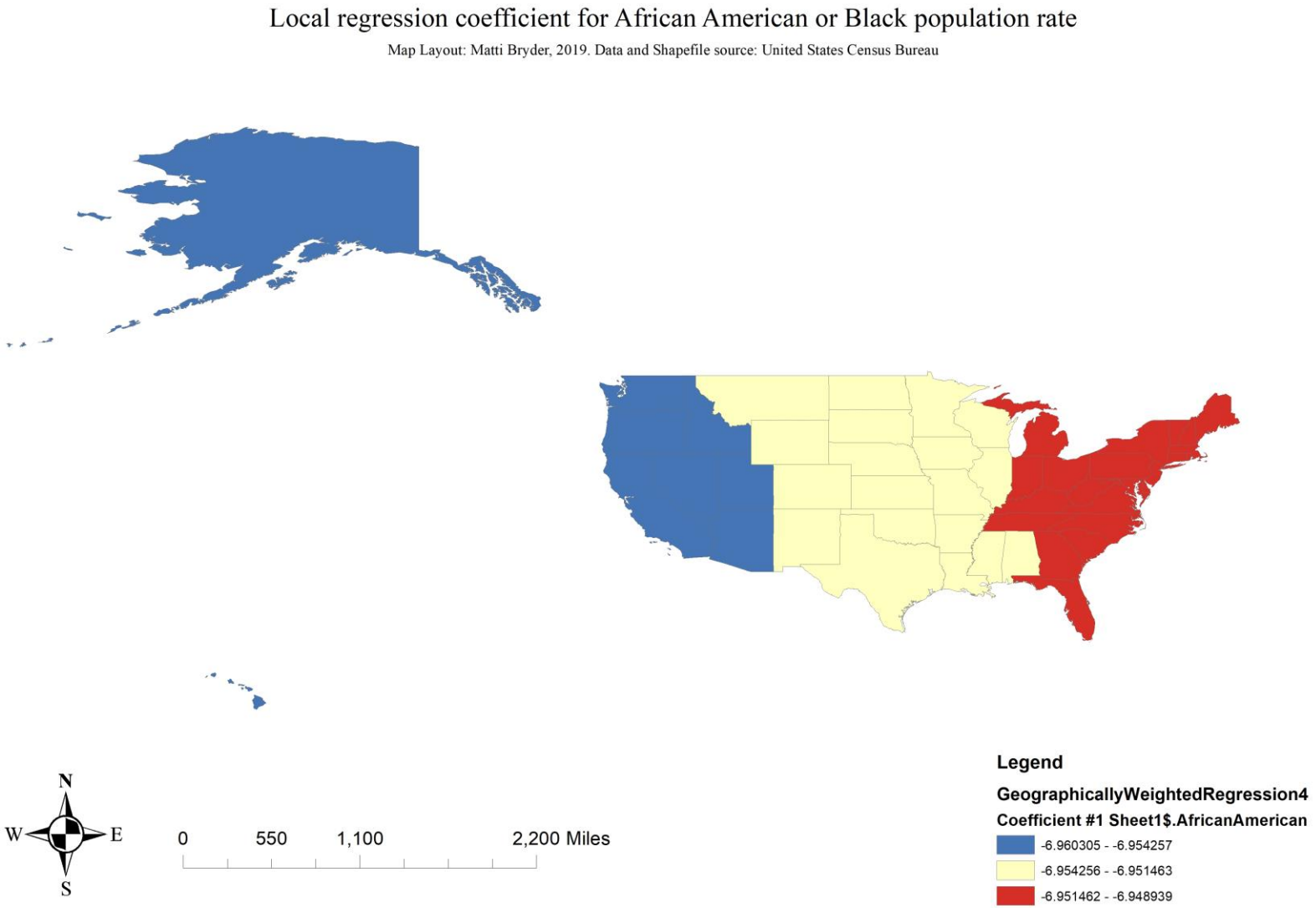


Figure 11: Local regression coefficient for African American or Black population rate

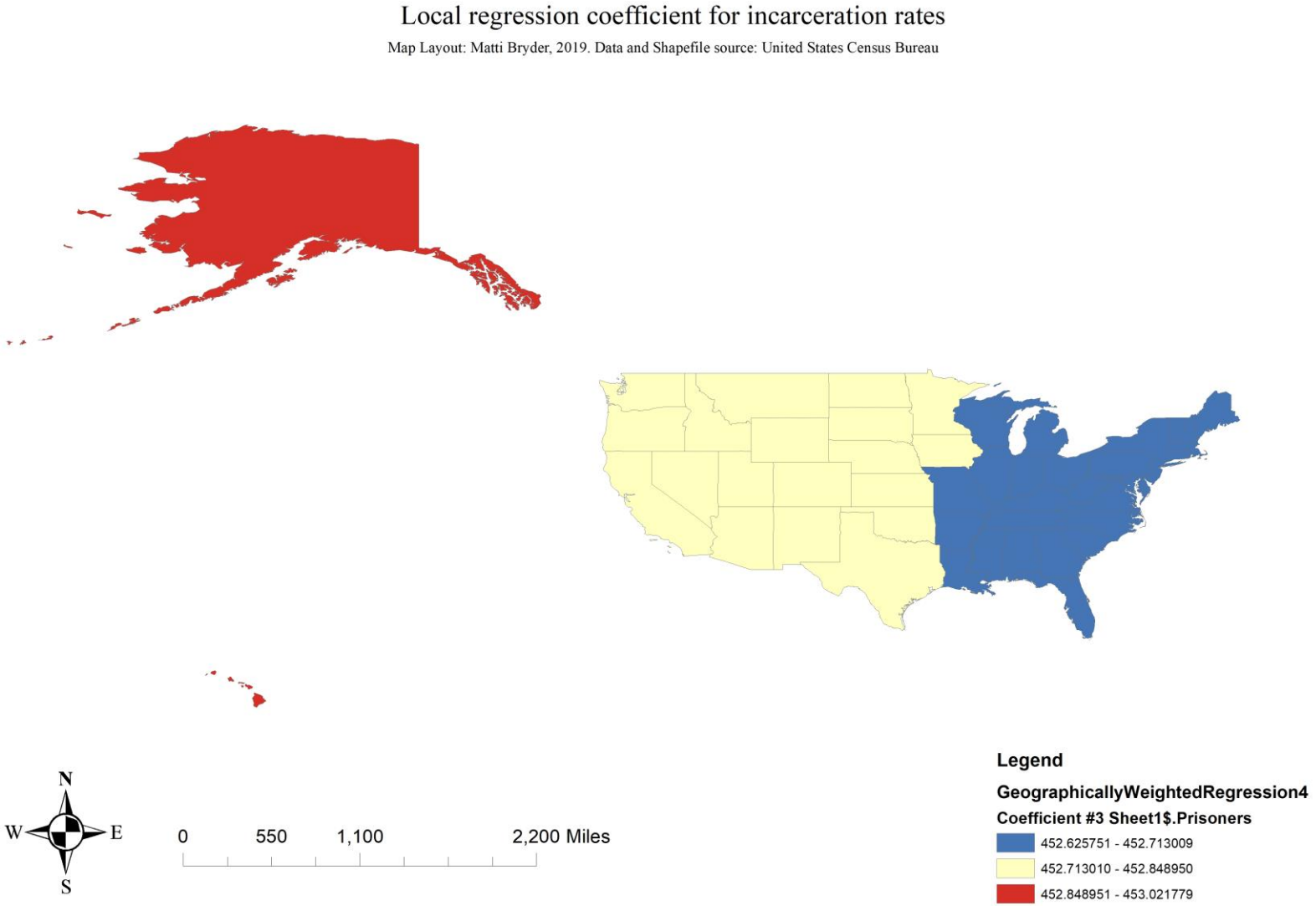


Figure 12: Local regression coefficients for incarceration rate

5. DISCUSSION

This chapter will present the findings, set those findings in relationship to the conceptual framework and lastly provide concepts for alternative research.

5.1. FINDINGS

The final regression, with an explanatory power of 45.2% of HCV prevalence included only three of the initial seven variables. Both Hispanics or Latino identity and health insurance levels had no correlation with HCV prevalence rates. Additionally, high school graduation rate and percentage of population in the healthcare sector did not pass significance levels for their coefficients during regression analysis. The reason for why their significance in passed in correlation was not deeply analysed but it was assumed to have to do with potential levels of multicollinearity which increased the sample size need for these valuables to be significant. Their removal however caused a 4% decrease in the explanatory power of the model. The final model formula from the second regression was produced as following:

$$HCV\ Prev = -85.866 + 52.885(Poverty) + 452.755(Incarceration\ rate) - 6.952(African\ American\ or\ Black\ population\ rate)$$

Of interesting note is that the coefficient variable for African American or Black population rate indicated a positive relationship, in the regression analysis it turned into a negative relationship. The implications of this is that after adjusting for the fact that African Americans or Blacks have a higher rate of imprisonment and a higher rate and higher rates of poverty, their population rate is actually negatively correlated to HCV prevalence. The other independent variables were positive. In total the regression estimated that for each 1% of poverty there would be an increase of 52.885 cases of HCV per 100000 people, for each 1% of the population incarcerated there would be an increase of 452,755 cases of HCV per 100000 people and for each percentage of African American or Blacks population percentage there would be a decrease of 6.952 cases of HCV per 100000 people.

As such, this thesis rejects one of the common assumptions regarding HCV prevalence in the United States, that there is an explicit racial bias in the disease's distribution. While the disease is guaranteed more prevalent amongst African Americans or Blacks, as both the correlation coefficient indicates and previous research (Whitt & Fleckenstein, 2008) shows, according to the regression model this is due to poverty and incarceration, not ethnic identity itself. While poverty is a question of ethnic identity to a large degree in the United States this

conclusion serves to reinforce the counterargument to the ideas that African Americans or Blacks have on an average, due to culture or community, a riskier culture or communities regarding health-related matters. They simply have a higher chance of being poor or incarcerated, and that is what is causing HCV risk factors. Poverty and incarceration played expected roles in the regression and were positively correlated to HCV prevalence. Additional data by Blankenship et al (2005) claimed that the prevalence of HIV in African American or Black populations and communities was due to their higher incarcerations. Upon release they were more likely to transmit it to members of their own ethnic group due to higher contact on multiple fronts, similar logic may be applied to HCV infection.

There are multiple reasons to question the validity of the model, the change in correlation direction for Africa Americans or Blacks could happen due to a variety of errors. Additionally, the constant produced is negative which is impossible for HCV prevalence which can only range down to 0. Lastly, it is important to note that the residuals for the regression are quite large indicating that the model is non descriptive of reality (see appendix 1: Table of residuals).

The GWR provided minimal insight into a potential spatial variation in how the independent variables could influence HCV prevalence. The results indicated that space had little influence on both the strength of the model as a whole or on the individual variables. Assuming the model would be valid, this would mean that the independent variables measured were almost equally important in every state. However, for the same reasons as the global linear regression with the additional fact that the scale of study was large, with each state encompassing a highly heterogenous population, there is significant room to critique and be skeptical of the study results.

The perhaps most useful analysis conducted was the correlation analysis which provided multiple factors which correlated with HCV prevalence. While these factors might themselves be correlated it provides a basis for further research (see 5.4. Further Research).

In conclusion the quantative analysis performed poorly. The limitations, choice of data and scale of study produced results of little to no direct use. Broadly the data implies certain relationships but not in a quantifiable way.

The major finding in the study is the relationships between the independent variables and their consequences (see Appendix 2: Full correlation table between all variables). With many of the variables being correlated to each other it raises the question on how to attribute risk to

them. Supported by theory from chapter 2, this thesis believes that the most likely answer is that there is a combination of impact, however those factors need to be viewed holistically and on smaller scale than was approached in this study.

Additional theory would also confirm the relationship between the dependent variables, for example Van der Berg (2007) showed the relationship between poverty and education within the United States. With such relationships, running regressions like the one performed does little to improve knowledge regarding the issue. The multitude of data which exists regarding the connection between health situation and social conditions may need review. Current literature points out causalities which may not be there, but simply drawn from correlations.

5.2. ALTERNATIVE RESEARCH

The research performed in this study could have been improved primarily by using smaller scale data in order to portray the heterogeneity of the states themselves. This would have allowed for a higher detail regarding the study. Secondly variables with higher independence from each other, or already adjusted for each other, should have been used.

The issue with smaller scale study is the availability of data for disease prevalence in the Untied States. The data gathered for this thesis was already an estimation with potential error and there is surprisingly low detailed data regarding HCV for such a rising health related phenomenon. While this author understands the ethnical discussion of personal data gathering regarding disease and the publication of such data, even on geographical scales, there are difficulties in understanding the social dimensions of the disease due to this issue. More troubling is the apparent large degree of undiagnosed HCV cases, which could screw any research done. As such a major suggestion is the gathering of higher amount of diagnosis data on HCV in the country. Assuming the validity of the study, the primary further research which should be conducted is to understand what exactly causes both poor people and incarcerated individuals to be so highly exposed to HCV risks and how to mitigate that phenomenon. Even if seen as a personal issue, the disease is placing an economic and social burden on society and will need to be tackled in the foreseeable future.

Lastly, with the social risk factors for HCV infection still not being fully understood it might be a good approach to go bottom-up, instead of the top-down approach that was conducted in this thesis. Instead of trying to estimate the social predictors based on literature more studies on individuals who have an HCV infection and how their social circumstances might have influenced their exposure to risk factors could be desired. After such research has been

produced the common social circumstances could be mapped out and used for policy measurements.

6. CONCLUSION

According to the Declaration of Human Rights *"Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family, including food, clothing, housing and medical care and necessary social services"*. Yet in todays world a multitude of social factors influence our everyday health situation. This thesis aimed to contribute into that discussion and expand the western focus of non-communicable diseases to communicable ones.

The socially disadvantaged and the economically disenfranchised have a higher risk of being exposed to HCV, the reasons are not fully understood. But just the base correlations from this study indicate that HCV has a higher presence in the states where these people live. By just simple higher exposure people who commit certain actions put themselves at higher risk of contracting the disease, something that in today's world should be prevented.

This thesis did not produce relevant quantifiable data, but it provided a baseline for how to approach the issue for the future. For disease smaller scale studies with good supporting data is most likely needed for the production of relevant results. Effort should be made to identify those who have contracted HCV, for their own sake and ability to be treated, and for the ability to understand the social pattern of disease. Understanding of HCV is affected by social phenomenon could contribute into the understanding for other diseases with similar transmission pathways.

More work needs to be done, both on the research front and policy front to address the health issues of social inequality. If nothing else this thesis would make the final claim that the inequalities presented within are not only objects of study, but a small representation of a larger set of social injustices.

Acknowledgments

I would like to thank my ever-patient supervisor for her assistance. Additionally, I thanks my partner for her invaluable assistance providing me space, proof-reading and snacks while I worked. Lastly, I would like to thank my family for their never-ending support of my goals, no matter how little they understand them.

7. REFERENCES

Andersson, J. (2017). Using Geographically Weighted Regression (GWR) to explore spatial variations in the relationship between public transport accessibility and car use. *Master Thesis, Department of Physical Geography and Ecosystem Science, Lund University*.

Assari, S., & Lankarani, M. (2016). Education and Alcohol Consumption among Older Americans; Black–White Differences. *Frontiers In Public Health*, *4*. doi: 10.3389/fpubh.2016.00067

Bakalar, N. (2012). Hepatitis C Deaths Creep Past AIDS Study Finds. Retrieved from https://www.nytimes.com/2012/02/28/health/research/hepatitis-c-deaths-creep-past-aids-study-finds.html?_r=0

Bischoff K and Reardon SF (2013) *Residential segregation by income, 1970-2009*. In Logan, JR (ed.), Diversity and disparities, New York, NY: Russel Sage Foundation.

Boschma, J. (2016). Where Schools Are Separate and Still Unequal. Retrieved from https://www.theatlantic.com/education/archive/2016/03/separate-still-unequal/471720/

Braveman, P., & Gottlieb, L. (2014). The Social Determinants of Health: It's Time to Consider the Causes of the Causes. *Public Health Reports*, *129*(1_suppl2), 19-31. doi: 10.1177/00333549141291s206

Carlson, A., & Frazão, E. (2014). Food costs, diet quality and energy balance in the United States. *Physiology & Behavior*, *134*, 20-31. doi: 10.1016/j.physbeh.2014.03.001

Carson, A. (2018). *Prisoners in 2016*. U.S. Department of Justice.

CDC. (2019). Hepatitis C Information | Division of Viral Hepatitis | Retrieved from https://www.cdc.gov/hepatitis/hcv/index.htm

Chang, V. (2006). Racial residential segregation and weight status among US adults. *Social Science & Medicine*, *63*(5), 1289-1303. doi: 10.1016/j.socscimed.2006.03.049

Chen, J., Vargas-Bustamante, A., Mortensen, K., & Ortega, A. N. (2016). Racial and ethnic disparities in health care access and utilization under the Affordable Care Act. *Medical care*, *54*(2), 140.

Chetty, R., Hendren, N., Kline, P., Saez, E., & Turner, N. (2014). Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility. *American Economic Review*, *104*(5), 141-147. doi: 10.1257/aer.104.5.141

Consumer Reports.org (2017), *"Are you really covered? Why 4 in 10 Americans can't depend on their health insurance". Archived from the original on September 22, 2008. Retrieved April 10,2012.*

Denniston MM, Jiles RB, Drobeniuc J, Klevens RM, Ward JW, McQuillan GM, et al (2014). Chronic Hepatitis C Virus Infection in the United States, National Health and Nutrition Examination Survey 2003 to 2010. *Ann Intern Med.* ;160:293–300. doi: 10.7326/M13-1133

Des Jarlais, D. C., McKnight, C., & Milliken, J. (2004). Public funding of US syringe exchange programs. *Journal of urban health : bulletin of the New York Academy of Medicine*, *81*(1), 118-21.

Dimsdale, J. (2008). Psychological Stress and Cardiovascular Disease. *Journal Of The American College Of Cardiology*, *51*(13), 1237-1246. doi: 10.1016/j.jacc.2007.12.024

Dula, A. (1994). African American Suspicion of the Healthcare System Is Justified: What Do We Do about It?. *Cambridge Quarterly Of Healthcare Ethics*, *3*(03), 347. doi: 10.1017/s0963180100005168

Engels, F. (1845). *The condition of the working class in England*. London: Electric Book Co.

Escobedo, L., & Peddicord, J. (1996). Smoking prevalence in US birth cohorts: the influence of gender and education. *American Journal Of Public Health*, *86*(2), 231-236. doi: 10.2105/ajph.86.2.231

Goodman, R., Bunnell, R., & Posner, S. (2014). What is "community health"? Examining the meaning of an evolving field in public health. *Preventive Medicine, 67,* S58-S61. doi: 10.1016/j.ypmed.2014.07.028

Hacker, J., & Pierson, P. (2010). Winner-Take-All Politics: Public Policy, Political Organization, and the Precipitous Rise of Top Incomes in the United States. *Politics & Society*, *38*(2), 152-204. doi: 10.1177/0032329210365042

Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. & Black, W. C. (1995). Multivariate Data Analysis (3rd ed). New York: Macmillan.

Harriss, K., & Salway, S. (2009). Long-term ill-health, poverty and ethnicity. *Ethnicity And Inequalities In Health And Social Care*, *2*(3), 39-48. doi: 10.1108/17570980200900022

Healthline. (2018). The Effects of Hepatitis C on Your Body. Retrieved from https://www.healthline.com/health/hepatitis-c/effects-on-the-body#1

Hellman, D., & Naroff, J. (1979). The Impact of Crime on Urban Residential Property Values. *Urban Studies*, *16*(1), 105-112. doi: 10.1080/00420987920080081

Henry B. (2018). Drug Pricing & Challenges to Hepatitis C Treatment Access. *Journal of health & biomedical law*, *14*, 265-283.

HHS.gov. (2018) Hepatitis C. Retrieved from https://www.hhs.gov/opa/reproductive-health/fact-sheets/sexually-transmitted-diseases/hepatitis-c/index.html

Hiza, H., Casavale, K., Guenther, P., & Davis, C. (2013). Diet Quality of Americans Differs by Age, Sex, Race/Ethnicity, Income, and Education Level. *Journal Of The Academy Of Nutrition And Dietetics*, *113*(2), 297-306. doi: 10.1016/j.jand.2012.08.011

Hopkins, E. (2017). The influence of public transportation on housing values. *International Journal Of Sustainable Development & World Ecology*, *25*(3), 206-215. doi: 10.1080/13504509.2017.1382398

Jackson, C., & Henriksen, L. (1997). Do as I say: Parent smoking, antismoking socialization, and smoking onset among children. *Addictive Behaviors*, *22*(1), 107-114. doi: 10.1016/0306-4603(95)00108-5

Kane, T. (2006). School Quality, Neighborhoods, and Housing Prices. *American Law And Economics Review*, *8*(2), 183-212. doi: 10.1093/aler/ahl007

Kim M. Blankenship, Amy B Smoyer, Sarah Bray, Kirstin Mattocks. (2005). Black-white disparities in HIV/AIDS: the role of drug policy and the corrections system. *Journal of health care for the poor and underserved*, *16*(4 Suppl B), 140-56.

Kristie L. Ebi, David M. Mills, Joel B. Smith, & Anne Grambsch. (2006). Climate Change and Human Health Impacts in the United States: An Update on the Results of the U.S. National Assessment. *Environmental Health Perspectives, 114*(9), 1318.

Kullgren, J. T., McLaughlin, C. G., Mitra, N., & Armstrong, K. (2012). Nonfinancial barriers and access to care for U.S. adults. *Health services research, 47*(1 Pt 2), 462-85.

Lee, A., & Maheswaran, R. (2010). The health benefits of urban green spaces: a review of the evidence. *Journal Of Public Health*, *33*(2), 212-222. doi: 10.1093/pubmed/fdq068

Leon, D. (1998). Social Stress and the Russian Mortality Crisis. *JAMA*, *279*(10), 790. doi: 10.1001/jama.279.10.790

Lichter, Daniel & Parisi, Domenico & C. Taquino, Michael. (2012). The Geography of Exclusion: Race, Segregation, and Concentrated Poverty. Social Problems. 59. 364-388. 10.1525/sp.2012.59.3.364.

Lin, C. H., & Wen, T. H. (2011). Using geographically weighted regression (GWR) to explore spatial varying relationships of immature mosquitoes and human densities with the incidence of dengue. *International journal of environmental research and public health*, *8*(7), 2798-815.

Lindenmayer, David; Burgman, Mark A. (2005). "Monitoring, assessment and indicators". *Practical Conservation Biology* (PAP/CDR ed.). Collingwood, Victoria, Australia: CSIRO Publishing. pp. 401–424. ISBN 0-643-09089-4.

Magura, S., Grossman, J. I., Lipton, D. S., Siddiqi, Q., Shapiro, J., Marion, I., & Amann, K. R. (1989). Determinants of needle sharing among intravenous drug users. *American journal of public health*, *79*(4), 459-62.

Mays, V., Cochran, S., & Barnes, N. (2007). Race, Race-Based Discrimination, and Health Outcomes Among African Americans. *Annual Review Of Psychology*, *58*(1), 201-225. doi: 10.1146/annurev.psych.57.102904.190212

McCabe, S. E., Morales, M., Cranford, J. A., Delva, J., McPherson, M. D., & Boyd, C. J. (2007). Race/ethnicity and gender differences in drug use and abuse among college students. *Journal of ethnicity in substance abuse*, *6*(2), 75-95.

McCord, McCord, Michael, McCluskey, William, Davis, Peadar, McIlhatton, Haran and Martin. (2014). Effect of public green space on residential property values in Belfast metropolitan area. Journal of Financial Management of Property and Construction. 19. 10.1108/JFMPC-04-2013-0008.

McFadden, D. L. (2016). Health and academic success: A look at the challenges of first-generation community college students. *Journal of the American Association of Nurse Practitioners*, *28*(4), 227-232.

McKeown, T., & Brown, R. (1955). Medical Evidence Related to English Population Changes in the Eighteenth Century. *Population Studies*, *9*(2), 119. doi: 10.2307/2172162

Melusky, Benjamin (2016) Institutional Capacity in the American States, Doctoral Dissertation, University of Pittsburgh.

Merzel, C., & D'Afflitti, J. (2003). Reconsidering community-based health promotion: promise, performance, and potential. *American journal of public health*, *93*(4), 557-74.

Monsanto, D., Kasprzyk, D., & Rimer, B. (2008). *Theory of reasoned action, theory of planned behavior, and the integrated behavior model*. Jossey-Bass, Wiley Imprint.

Nakano T, Lau GM, Lau GM, Sugiyama M, Mizokami M (December 2011).An updated analysis of hepatitis C virus genotypes and subtypes based on the complete coding region". *Liver Int.*. doi:10.1111/j.1478-3231.2011.02684.x. PMID 22142261.

Nelson, P., Mathers, B., Cowie, B., Hagan, H., Des Jarlais, D., Horyniak, D., & Degenhardt, L. (2011). Global epidemiology of hepatitis B and hepatitis C in people who inject drugs: results of systematic reviews. *The Lancet*, *378*(9791), 571-583. doi: 10.1016/s0140-6736(11)61097-0

Olshansky, S., Antonucci, T., Berkman, L., Binstock, R., Boersch-Supan, A., & Cacioppo, J. et al. (2012). Differences in Life Expectancy Due to Race and Educational Differences Are Widening, And Many May Not Catch Up. *Health Affairs*, *31*(8), 1803-1813. doi: 10.1377/hlthaff.2011.0746

Petrosillo, N., Gilli, P., Serraino, D., Dentico, P., Mele, A., Ragni, P., … Collaborative Grp. (2001). Prevalence of infected patients and understaffing have a role in hepatitis C virus transmission in dialysis. *AMERICAN JOURNAL OF KIDNEY DISEASES*, *37*(5), 1004–1010.

Plucinski, M., Ngonghala, C., Getz, W., & Bonds, M. (2012). Clusters of poverty and disease emerge from feedbacks on an epidemiological network. *Journal Of The Royal Society Interface*, *10*(80), 20120656-20120656. doi: 10.1098/rsif.2012.0656

Room, R. (2005). Stigma, social inequality and alcohol and drug use. *Drug And Alcohol Review*, *24*(2), 143-155. doi: 10.1080/09595230500102434

Rooney, G., & Gilson, R. J. (1998). Sexual transmission of hepatitis C virus infection. *Sexually transmitted infections*, *74*(6), 399-404.

Rosenberg ES, Rosenthal EM, Hall EW (2018), et al. Prevalence of Hepatitis C Virus Infection in US States and the District of Columbia, 2013 to 2016. *JAMA Netw Open*

Rosenberg, M. (2013). Health geography I. *Progress In Human Geography*, *38*(3), 466-475. doi: 10.1177/0309132513498339

Ross J, Clarke, Elizabeth A. Clarke, Dominique Roe-Sepowitz & Richard Fey (2012) Age at Entry into Prostitution: Relationship to Drug Use, Race, Suicide, Education Level, Childhood Abuse, and Family Experiences, Journal of Human Behavior in the Social Environment, 22:3, 270-289, DOI: 10.1080/10911359.2012.655583

Ross, C., & Wu, C. (1995). The Links Between Education and Health. *American Sociological Review*, *60*(5), 719. doi: 10.2307/2096319

Saraceno, B., & Barbui, C. (1997). Poverty and Mental Illness. *The Canadian Journal Of Psychiatry*, *42*(3), 285-290. doi: 10.1177/070674379704200306

Saxe, L., Kadushin, C., Beveridge, A., Livert, D., Tighe, E., Rindskopf, D., Ford, J. Brodsky, A. (2001). The visibility of illicit drugs: implications for community-based drug control strategies. *American journal of public health*, *91*(12), 1987-94.

Schrager, L., Friedland, G., Feiner, C., & Kahl, P. (1991). Demographic characteristics, drug use, and sexual behavior of i.v. drug user with AIDS in Bronx, New York. *Public health reports (Washington, D.C. : 1974)*, *106*(1), 78-84.

Schultz, S. (1990). From isolation to commitment the role of the community in values education. *New Directions For Student Services*, *1990*(50), 91-100. doi: 10.1002/ss.37119905009

Shepard, C. W., Finelli, L., & Alter, M. J. (2005). Global epidemiology of hepatitis C virus infection. *The Lancet. Infectious Diseases*, *5*(9), 558–567.

Spaulding, A. C., & Thomas, D. L. (2012). Screening for HCV infection in jails. *JAMA*, *307*(12), 1259-60.

Thoits, P. (2010). Stress and Health: Major Findings and Policy Implications. *Journal Of Health And Social Behavior*, *51*(1_suppl), S41-S53. doi: 10.1177/0022146510383499

Thompson, N. (2009). Nonhospital Health Care–Associated Hepatitis B and C Virus Transmission: United States, 1998–2008. *Annals Of Internal Medicine*, *150*(1), 33. doi: 10.7326/0003-4819-150-1-200901060-00007

Tohme, R. A., & Holmberg, S. D. (2012). Transmission of hepatitis C virus infection through tattooing and piercing: a critical review. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*, *54*(8), 1167-78.

Tokars, J., Finelli, L., Alter, M., & Arduino, M. (2004). Special Article: National Surveillance of Dialysis-Associated Diseases in the United States, 2001. *Seminars In Dialysis*, *17*(4), 310-319. doi: 10.1111/j.0894-0959.2004.17339.x

United States Census Bureau (n/d) "2013-2017 American Community Survey 5-Year Estimates". *factfinder.census.gov*. Retrieved 2019-03-07.

Van der Berg, S. (2008). Poverty and education. *Education policy series*, *10*, 28.

Virchow, R., & Rather, L. (1985). *Collected essays on public health and epidemiology*. New Delhi: Amerind Pub. Co.

Wen, M., Zhang, X., Harris, C., Holt, J., & Croft, J. (2013). Spatial Disparities in the Distribution of Parks and Green Spaces in the USA. *Annals Of Behavioral Medicine*, *45*(S1), 18-27. doi: 10.1007/s12160-012-9426-x

Whitt, K., & Fleckenstein, J. (2008). Hepatitis C in African Americans. *Current Hepatitis Reports*, *7*(4), 139-144. doi: 10.1007/s11901-008-0028-9

WHO 2 | The determinants of health. (2018). Retrieved from https://www.who.int/hia/evidence/doh/en/

WHO. (2018). Hepatitis C. Retrieved from https://www.who.int/news-room/fact-sheets/detail/hepatitis-c

Wilkins, T., Malcoms, J., Raina, D., & Scade, R. (2010). Hepatitis C: Diagnosis and Treatment. *American Family Physician*, *81*(11), 1351-1358.

Williams, C. T., & Latkin, C. A. (2007). Neighborhood socioeconomic status, personal network attributes, and use of heroin and cocaine. *American journal of preventive medicine*, *32*(6 Suppl), S203-10.

Williams, C. T., & Latkin, C. A. (2007). Neighborhood socioeconomic status, personal network attributes, and use of heroin and cocaine. *American journal of preventive medicine*, *32*(6), S203-S210.

Wolch, J., Byrne, J., & Newell, J. (2014). Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'. *Landscape And Urban Planning*, *125*, 234-244. doi: 10.1016/j.landurbplan.2014.01.017

Woolf, S., Johnson, R., Phillips, R., & Philipsen, M. (2007). Giving Everyone the Health of the Educated: An Examination of Whether Social Change Would Save More Lives Than Medical Advances. *American Journal Of Public Health*, *97*(4), 679-683. doi: 10.2105/ajph.2005.084848

Woolhandler, S., & Himmelstein, D. (2007). Competition in a publicly funded healthcare system. *BMJ*, *335*(7630), 1126-1129. doi: 10.1136/bmj.39400.549502.94

Yu, C. I., & Chiang, B. L. (2010). A new insight into hepatitis C vaccine development. *Journal of biomedicine & biotechnology*, *2010*, 548280.

Zein N. N. (2000). Clinical significance of hepatitis C virus genotypes. *Clinical microbiology reviews*, *13*(2), 223-35.

APPENDIX 1: TABLE OF RESIDUALS

Predicted	Observed	Residual
836.625925	1234.986482	398.360557
909.777714	1133.639599	223.861886
983.230628	982.12373	-1.106897
660.018057	962.990355	302.972297
762.958386	835.867124	72.908738
744.231136	785.00232	40.771185
763.958	780.994795	17.036795
626.657522	777.867461	151.209939
633.65906	746.446747	112.787687
812.490395	734.166314	-78.32408
876.466876	693.621285	-182.845591
333.456477	692.254862	358.798385
644.349053	689.082017	44.732965
642.546171	684.583187	42.037016
665.015286	645.118545	-19.896741
715.876136	619.390052	-96.486084
598.333172	596.566358	-1.766815
648.255027	596.291182	-51.963845
420.279878	587.360759	167.080882
477.315181	571.597886	94.282705
424.458303	535.882621	111.424318
644.305241	532.075406	-112.229836
525.264082	517.659985	-7.604098
661.304867	514.201832	-147.103035
536.829877	492.777658	-44.052219
611.181787	491.716413	-119.465373
654.199538	491.033707	-163.165831
486.980205	450.619936	-36.360269
776.70747	443.818567	-332.888903
621.183065	433.247594	-187.935471
598.016644	426.379621	-171.637023
578.657992	420.867068	-157.790924
627.559078	411.367588	-216.191489
573.354709	360.954697	-212.400012
794.58794	677.755851	-116.832089
982.845384	928.970325	-53.875059
956.445786	1404.548048	448.102262
707.91456	681.328151	-26.58641
783.419596	1072.049649	288.630053
856.980057	764.76079	-92.219267
1013.471252	1303.651738	290.180486
850.695775	644.783106	-205.912669
965.992783	783.204002	-182.788781
746.145254	573.318157	-172.827098
705.734917	751.645703	45.910786
940.73501	747.752797	-192.982213
950.247847	1100.761022	150.513175
783.842644	769.57549	-14.267154
763.756521	705.21833	-58.538191
632.740182	729.349588	96.609407

APPENDIX 2: FULL CORRELATION TABLE BETWEEN ALL VARIABLES

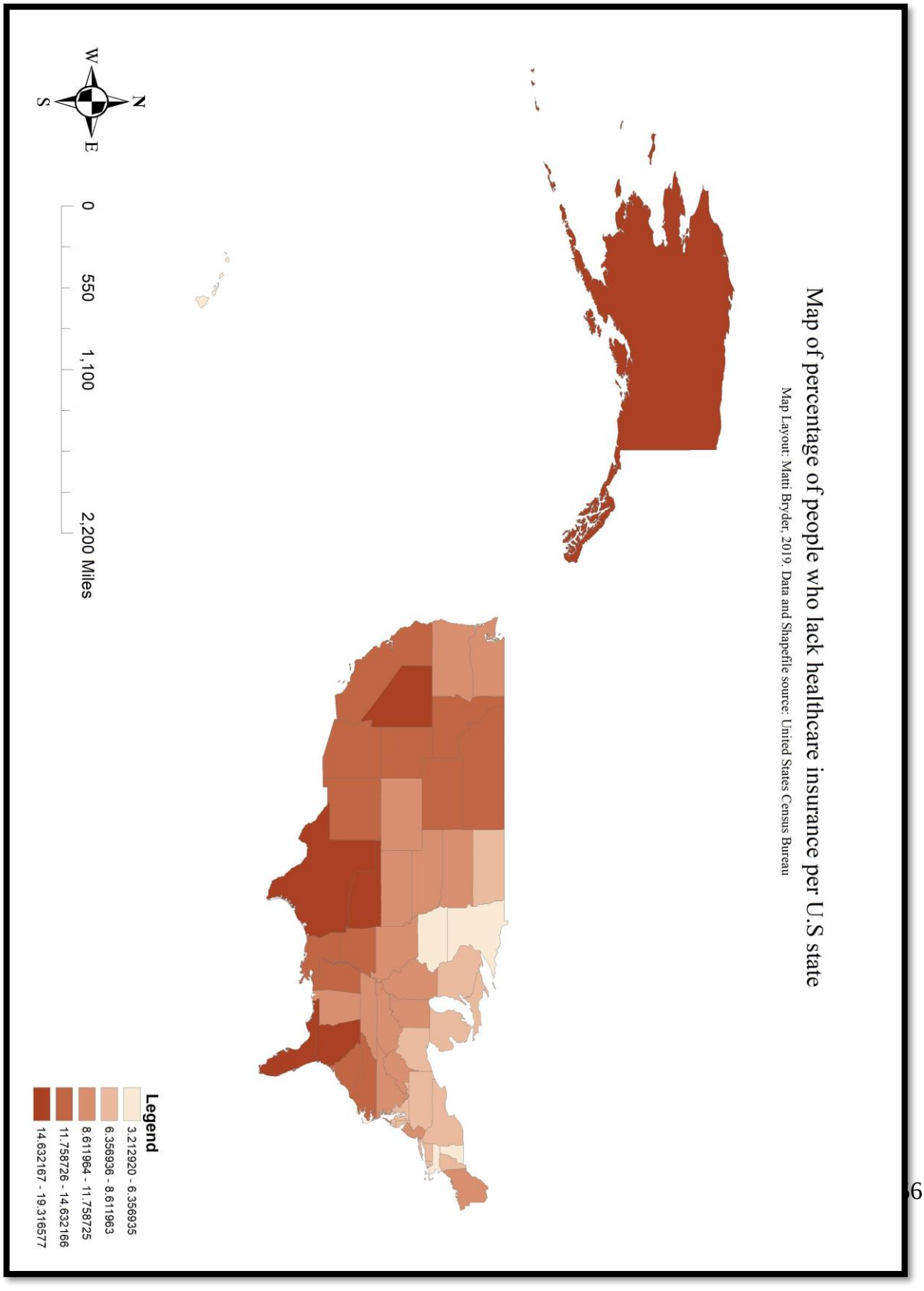
Correlations

		Prevalence	Poverty (%)	(%)	(%)	Healthcare per (%)	Prisoners (%)	American	Insured
Hep C Prevalence (100.000)	Pear	1	.587**	0.199	-.382**	-.437**	.443**	.341*	0.174
	Sig.		0.000	0.161	0.006	0.001	0.001	0.014	0.221
	N	51	51	51	51	51	50	51	51
Poverty (%)	Pear	.587**	1	0.148	-.731**	-.361**	.500**	.416**	.438**
	Sig.	0.000		0.300	0.000	0.009	0.000	0.002	0.001
	N	51	51	51	51	51	50	51	51
Hispanics (%)	Pear	0.199	0.148	1	-.432**	-.446**	0.016	-0.121	.396**
	Sig.	0.161	0.300		0.002	0.001	0.912	0.397	0.004
	N	51	51	51	51	51	50	51	51
High School or higher (%)	Pear	-.382**	-.731**	-.432**	1	.324*	-.502**	-.455**	-.510**
	Sig.	0.006	0.000	0.002		0.020	0.000	0.001	0.000
	N	51	51	51	51	51	50	51	51
Employed in Healthcare per (%)	Pear	-.437**	-.361**	-.446**	.324*	1	-.385**	-.287*	-.495**
	Sig.	0.001	0.009	0.001	0.020		0.006	0.041	0.000
	N	51	51	51	51	51	50	51	51
Prisoners (%)	Pear	.443**	.500**	0.016	-.502**	-.385**	1	.465**	.575**
	Sig.	0.001	0.000	0.912	0.000	0.006		0.001	0.000
	N	50	50	50	50	50	50	50	50
African American (%)	Pear	.341*	.416**	-0.121	-.455**	-.287*	.465**	1	0.089
	Sig.	0.014	0.002	0.397	0.001	0.041	0.001		0.536
	N	51	51	51	51	51	50	51	51
Not Health Insured (%)	Pear	0.174	.438**	.396**	-.510**	-.495**	.575**	0.089	1
	Sig.	0.221	0.001	0.004	0.000	0.000	0.000	0.536	
	N	51	51	51	51	51	50	51	51

**. Correlation is significant at the 0.01 level (2-tailed).

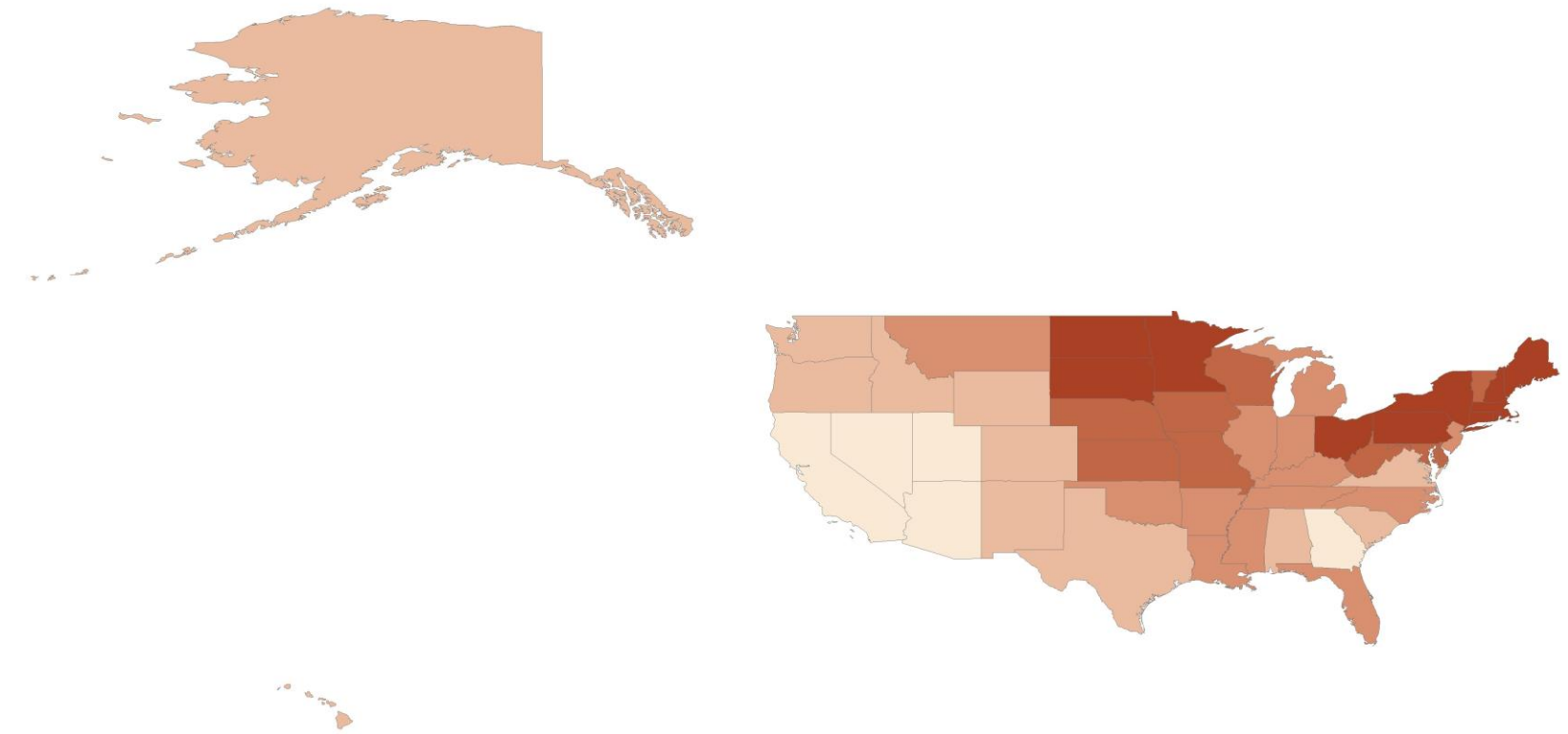
*. Correlation is significant at the 0.05 level (2-tailed).

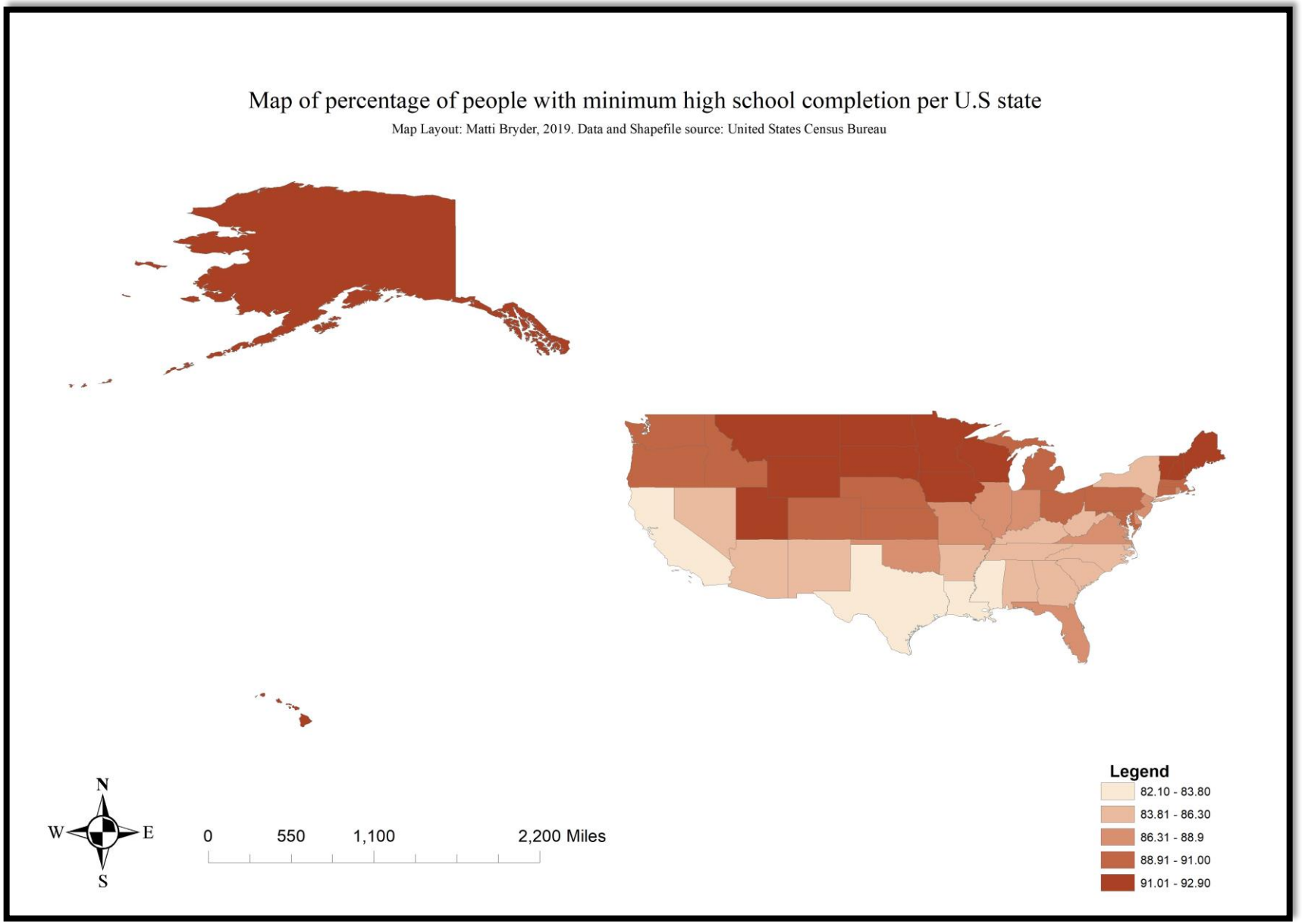
Appendix 3: Maps of the distribution of all the independent variables

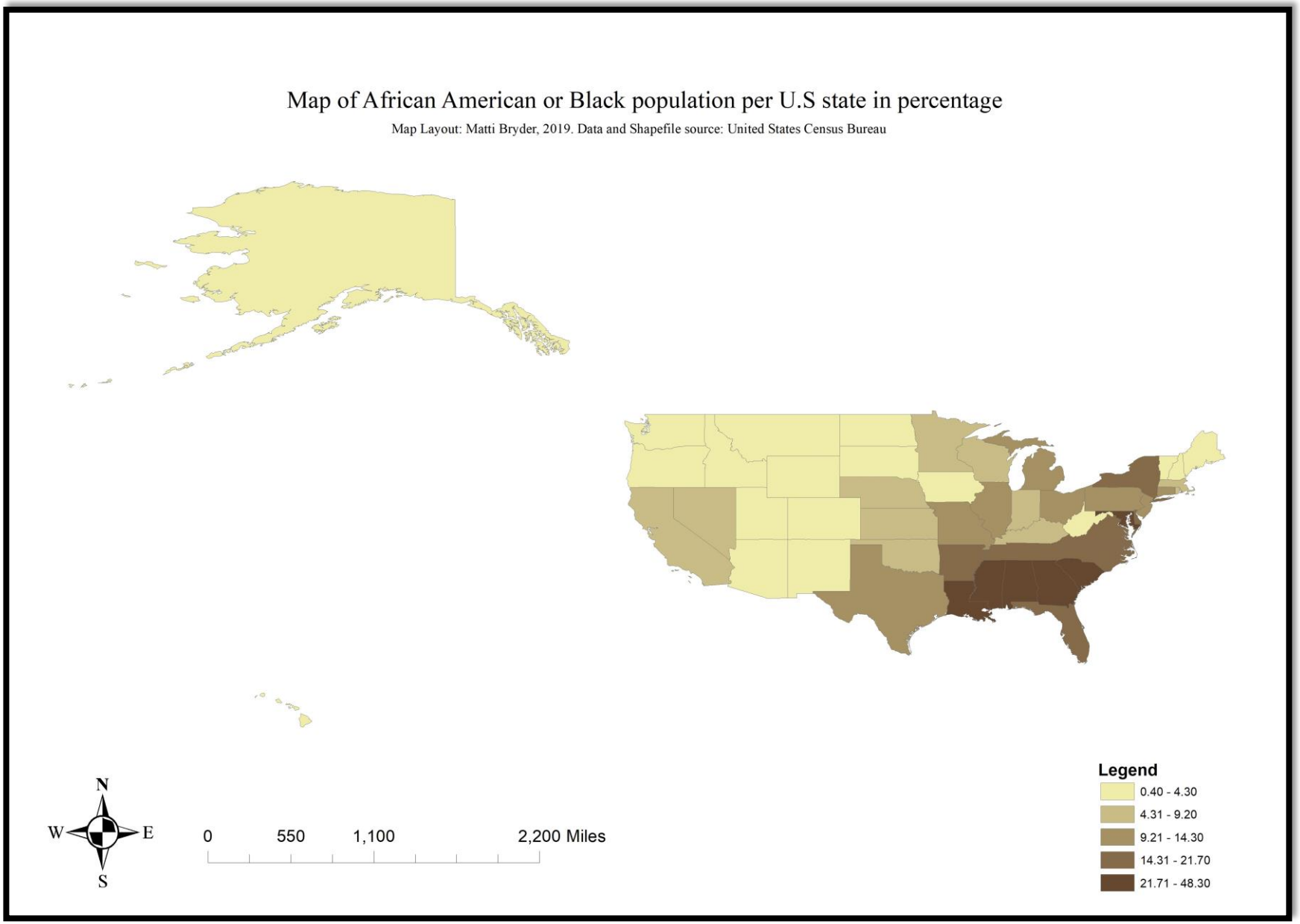


Map of percentage of people who work in the healthcare sector per U.S state

Map Layout: Matti Bryder, 2019. Data and Shapefile source: United States Census Bureau

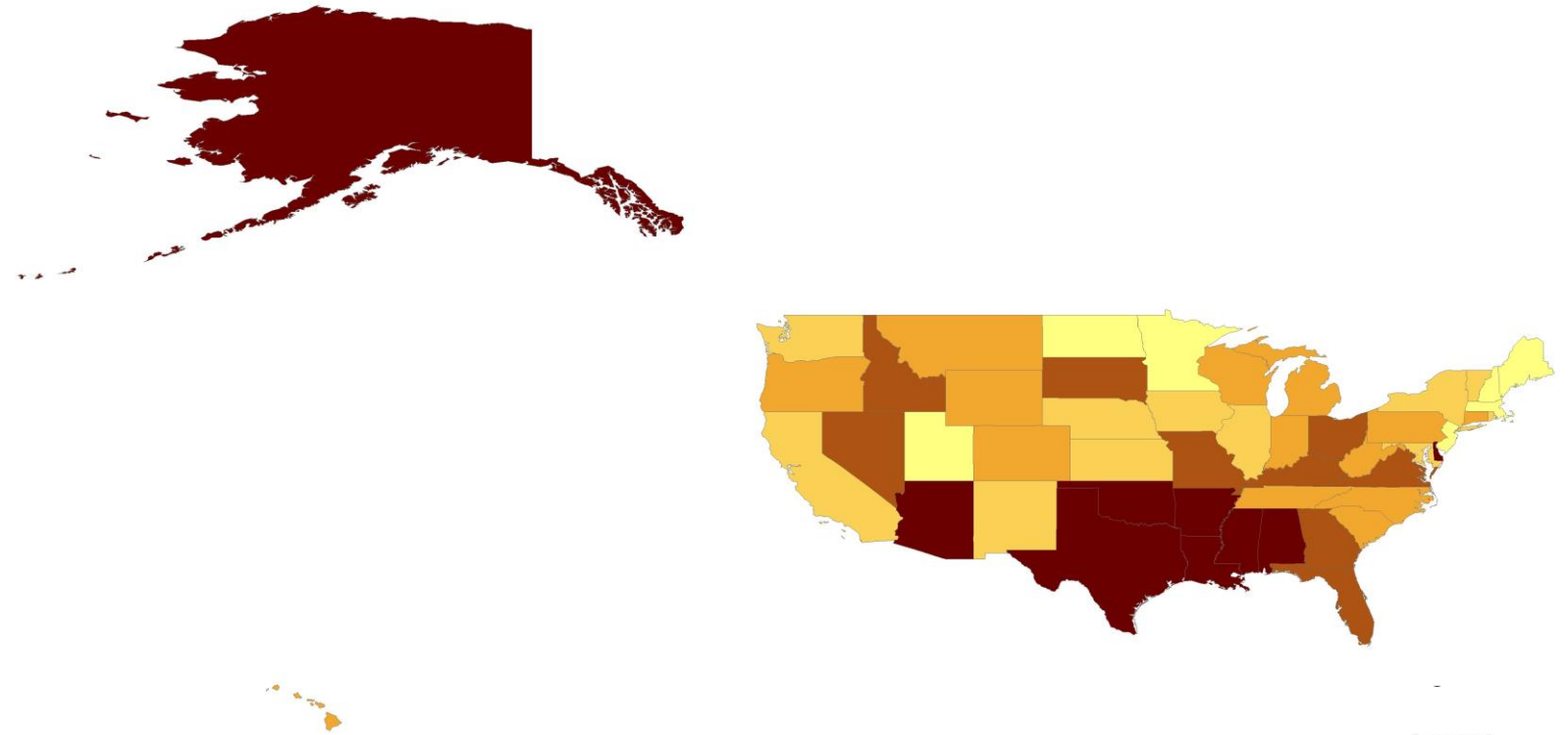



Map of percentage of people with minimum high school completion per U.S state
Map Layout: Matti Bryder, 2019. Data and Shapefile source: United States Census Bureau
N
W
E
S
0
550
1,100
2,200 Miles
Legend
82.10 - 83.80
83.81 - 86.30
86.31 - 88.9
88.91 - 91.00
91.01 - 92.90


Map of African American or Black population per U.S state in percentage
Map Layout: Matti Bryder, 2019. Data and Shapefile source: United States Census Bureau
N
W
E
S
0
550
1,100
2,200 Miles
Legend
0.40 - 4.30
4.31 - 9.20
9.21 - 14.30
14.31 - 21.70
21.71 - 48.30

Map of population incarcerated in percentage by U.S state

Map Layout: Matti Bryder, 2019. Data and Shapefile source: United States Census Bureau



Map of poverty percentage by U.S state

Map Layout: Matti Bryder, 2019. Data and Shapefile source: United States Census Bureau