The Cost of Culture:
Examining a Sociological Model of the Opioid Epidemic in the United States

Submitted by: Ariana Martine Mapua

Undergraduate Economics Program
Tepper School of Business
Carnegie Mellon University

In fulfillment of the requirement for the
Tepper School of Business Senior Honors Thesis in Economics

Advisor:

Laurence Ales
Associate Professor of Economics
Tepper School of Business
Carnegie Mellon University

August 2018

Abstract

Since the late 90s, the United States has been dealing with an increased abuse of opioids, with more overdose deaths resulting in what we now call "the opioid epidemic." Drawing inspiration from J.D. Vance's 2016 memoir, *Hillbilly Elegy*, this paper examines several factors that may have contributed to the overdose deaths, focusing on sociological factors such as association rates, crime rates, and physical and mental distress. I use 2016 cross-sectional data from 412 counties, regressing the log of drug overdose death rate on several factors to construct a model of the epidemic, aiming to find specific factors to target when considering how to tackle the crisis. I also use Akaike information criterion (AIC) to determine the best factors to include in my model. I find that while sociological factors play some role in overdose deaths, this role is smaller than I had hypothesized. The model suggests instead that economic factors such as unemployment rate and income inequality are more important. However this model does not account for possible multicollinearity and endogeneity; thus further tests need to be conducted to better understand how these factors impact opioid overdose deaths.

1 Introduction

In 2016, 115 people died from drug overdoses each day.

Their main drug of choice: opioids.

While most pharmaceutical drugs are created with good intentions, their attractive side effects can lead to widespread abuse. This situation has been worsening in most, if not all parts of the US with growing rates of opioid abuse and deaths from overdose.

While officials and physicians nationwide scramble to treat this growing epidemic, we need to give more attention to the root causes that are still plaguing Americans. Most literature seems to focus on economic factors of the crisis, but we must consider all factors involved if we want to find some real solutions. Economic factors like income and employment are important, but in reality these can also be driven by many other factors, such as sociological ones.

In their review about the *Determinants of Increased Opioid-Related Mortality in the United States and Canada* (King et al. 2014), researchers identified some economic factors such as income inequality, poverty, and urbanization that drive the epidemic, but also stressed other factors like sociodemographic characteristics, media coverage, and geography. Dr. Nora Volkow, the director of the National Institute on Drug Abuse (NIDA) also wrote recently about how "Addressing the Opioid Crisis Means Confronting Socioeconomic Disparities," suggesting that economic factors are not the sole basis for drug abuse. She presented other factors like "access and quality of health care," which can go overlooked.

These factors are also explored in more accessible literature and exemplified in prose. *Hillbilly Elegy: A Memoir of a Family and Culture in Crisis*, is one such example. Authored by J.D. Vance, a venture capitalist from the Rust Belt, it expresses some sociological factors that may be driving misfortunes in the region, including drug abuse. Growing up in the Appalachian region of the US, Vance experienced firsthand the effects of drug abuse as his mother was addicted to prescription opioids, as well as many others within his hometown. While it is very easy to pin some of the drivers of the opioid epidemic on economic factors, and the drivers of economic factors on outside forces like the government, Vance does not let us forget about the sociological issues driving both of these circumstances.

1.1 Question

Based on my motivations for this paper, I want to know what factors are driving the opioid crisis in the US. Specifically, I want to know whether these factors are more sociological or economic, since from there we can evaluate current and proposed policy recommendations to alleviate the crisis.

1.2 Stakes

The opioid epidemic has blown up so much in the past few decades that in 2017 President Donald Trump declared it a Public Health Emergency. While substance abuse has been around for some time now, it has unfortunately reached such levels at an alarming rate that maybe bigger, and more immediate actions are needed. Human lives are at stake – both the addicts and others surrounding them who are also affected by the crisis.

Take American families into consideration. Many news stories outline how many have died from overdoses, some of who are parents who have left children behind. Most of these children are then taken in by their grandparents, and this has become such a common occurrence that many politicians are writing bills to aid these grandparents in this crisis (Senator Susan Collins), gifting them financial and social support from the federal government. A press release by Senator Collins states that "approximately 2.6 million children are being raised by their grandparents," and this number is set to grow as the epidemic worsens.

Even if overdoses don't result in deaths, addicted parents still create many problems for their families. Pregnant women who continue to use opioids, especially heroin, risk neonatal abstinence syndrome (NAS) as the baby can also become dependent on the drugs (How does heroin use affect pregnant women?). Parents who are addicts become the burden, neglecting their children. NBC News is one of many media outlets that report about the crisis, telling stories of neglected children who grow up with addict parents (Medina et al). These stories continue to be published quite frequently, supporting the alarming increase in substance abuse and alerting the need to combat the crisis.

There are also the economic costs to consider in the opioid crisis. While social issues first come to mind (familial rifts, detriment to work and communities), there are many spillover effects financially. The National Public Radio (NPR) described in one of their podcasts how opioid addiction could tear families apart, outlining the story of Kathryn Sexton from Muncie, IN who struggled with addiction for years (Noguchi). Her parents spent their pension and savings sending her to rehab, and were still left with carrying her posthumous financial responsibilities – student loans and phone bills, to name a few. This financial burden is a problem many others face as the crisis grows along with a heavy economic cost. The White House Council of Economic Advisors estimated the 2015 economic cost of the opioid crisis to be $504 billion, amounting to 2.8% of the GDP that year (2017). This figure includes healthcare costs, foregone earnings from employment, criminal justice costs, lost productivity, and the value of a statistical life (VSL, adjusted by age). This is a massive overinvestment of resources – not in that addicts do not deserve treatment – but because addiction can be combated before it reaches emergency levels. Focusing on prevention rather than treatment may bring down these financial costs.

1.3 Outline

I gathered information about drug overdose deaths, employment status, age, gender, income, insurance coverage, and occupations to use in my model, all from the year 2016 at the county level. While the information about drug overdose deaths came from the Centers for Disease Control and Prevention (CDC), a public health institution in the US, the rest comes from the Integrated Public Use Microdata Series (IPUMS-USA), which uses information compiled from the American Community Survey (ACS). Using these factors, I constructed a multivariate model to explore the role of sociological factors in drug overdose deaths. While much of my research is driven by these sociological factors, I found that economic factors are the most statistically significant variables, supporting most research on the opioid epidemic.

Many of these factors are interlinked. For example, employment status is related to health insurance, which is related to physical or mental distress. While the connection between those factors seems reasonable in the narrative, it may create issues in creating my model. The statistical issue that immediately comes to mind is multicollinearity – that there are high inter-correlations among my variables, and may affect the reliability of my data. I address multicollinearity later on in the paper, using Akaike information criterion (AIC) to select the best variables for my model.

Current literature focuses more on the origins of the opioid epidemic, as well as variations in factors across states. Most research agrees that pharmaceutical companies played the biggest role in the start of the epidemic and greatly increased use of prescription opioids, creating and knowingly marketing highly addictive drugs. Other factors previously examined include prescription rates, unemployment rates, the role of pharmaceutical companies, and drug trafficking. This paper complements this current literature as it examines both economic and sociological factors, and building a multivariate model to suggest the most influential factors in the epidemic. Although most information released to the general public is shown at the state level, I conduct my research at the county level, and choose my factors based on univariate trends that most research has noted: drug abuse is more prevalent in low-income areas, women are more likely to abuse prescription opioids, and overdose deaths are highest among the middle-aged group, for example.

In section 2, I examine the background of the epidemic. In section 3 I talk about my motivating factors for this paper, explaining more about Vance's memoir. In section 4 I discuss how I obtained and manipulated my data, and show my results in section 5. Section 6 concludes this paper and discusses future work.

2 Background

2.1 Opioids

Opioids are pain-relieving substances created from the opium poppy, and are primarily produced for medicinal purposes as some form of pain medication like anesthesia and cough medicine. These substances include morphine and heroin, and popular prescription drugs like codeine, oxycodone (Percocet, OxyContin), hydrocodone (Vicodin), methadone (Dolophine), fentanyl (Duragesic, Subsys), and other pain relievers. Opioids come in natural, semi-synthetic, and synthetic forms, but for the purpose of this paper I will use the term "opioids" to refer to all of them. Usually opioids are prescribed for some chronic or cancer pain, but have been increasingly prescribed for non-cancer pain (e.g. back pain, migraines). Aside from pain relief, taking any of these opioids produces side effects like drowsiness, sedation, respiratory depression, and a feeling of euphoria – all of which are attractive to users. With the added addictive nature of opioids, people continue using them and even resort to crushing and snorting, or injecting them for a more intense experience (SAMHSA). Regular consumption of opioids also builds a tolerance within users, so larger doses are needed to produce these effects and dependence or addiction develops with harsh withdrawal symptoms (nausea, vomiting, constipation, etc.). These dangers, as well as the attractive side-effects, put users at risk for overdoses. Manifested by respiratory depression, these overdoses can lead to death. This situation is what constitutes the growing opioid epidemic across the US.

2.2 Current Description of the Epidemic

Some fast facts about the opioid epidemic in 2016 from the US Department of Health and Human Services:

- 42,249 people died from overdosing on opioids – that's 116 people a day, and five times the number in 1999.

- 11. M people misused prescription opioids, 2.1 M of them for the first time, and 17,087 of the overdose deaths were from commonly prescribed opioids.

Of these people who abused opioids:

- 948,000 of them used heroin, 170,000 of them for the first time, and 15,469 of them overdosed and died.

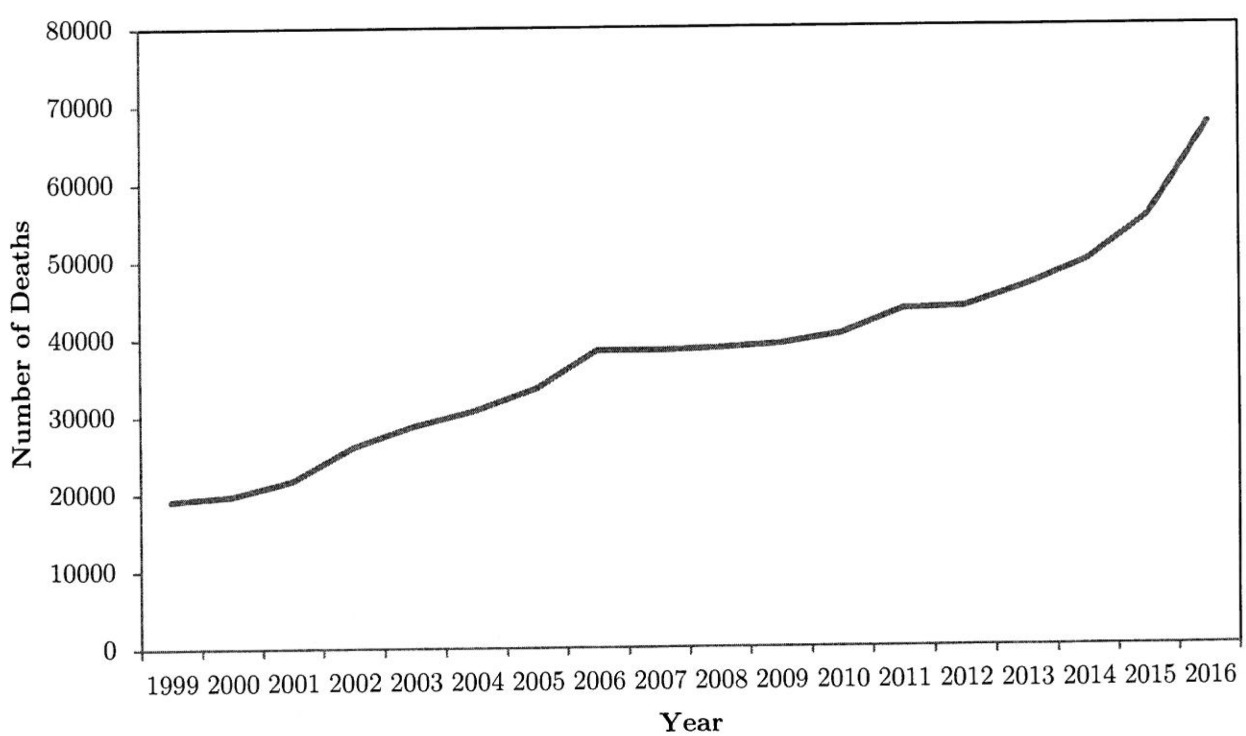
These numbers are daunting and are meant to shock us. The message here is that the opioid epidemic is reaching alarming levels and must be curbed to prevent further harm to people.

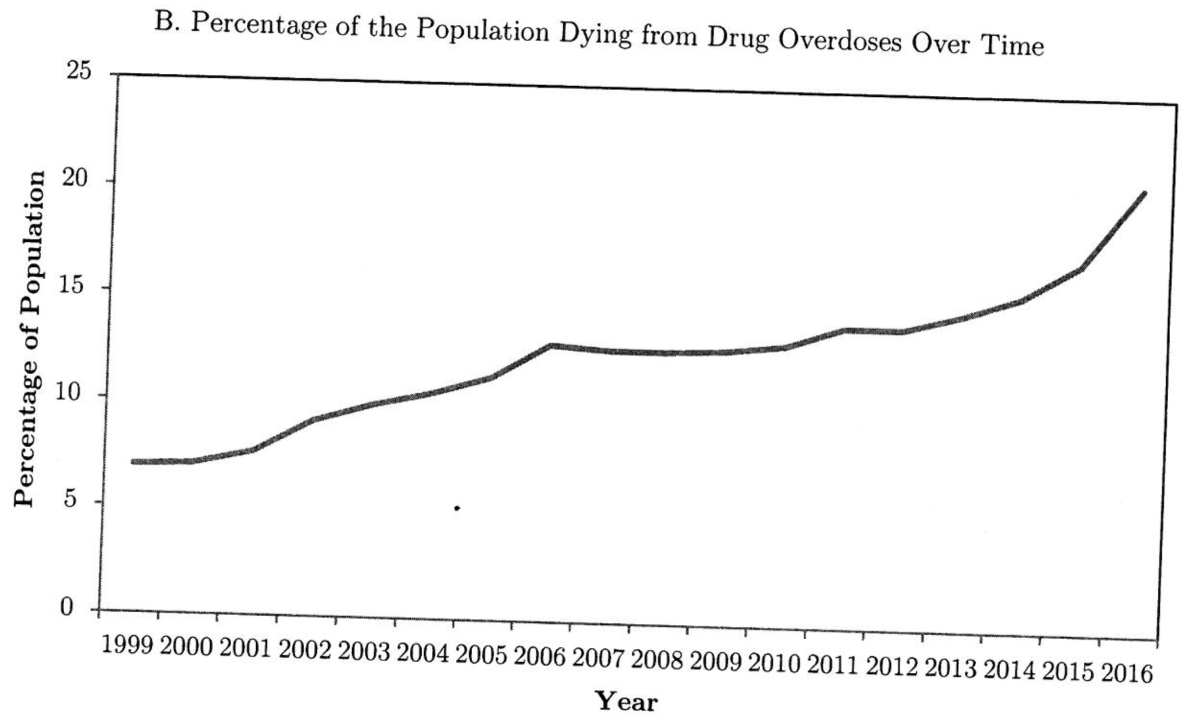
2.3 Trends and Rise of the Epidemic

The number (*and* the percentage) of people dying from drug overdoses in the US has been increasing every year. Opioids (in general) have risen to account for up to 60% of those deaths, about 40% of which are attributed to prescription drugs (HHS). These numbers make opioid overdoses one of the leading causes of deaths of Americans under 50. I provide some figures below, using data taken from the CDC to give a picture of the rise of drug overdose deaths in the US.

Figure 1: Deaths from Drug Overdoses in the US, 1999-2016

A: Number of Drug Overdose Deaths Over Time





Source: CDC WONDER

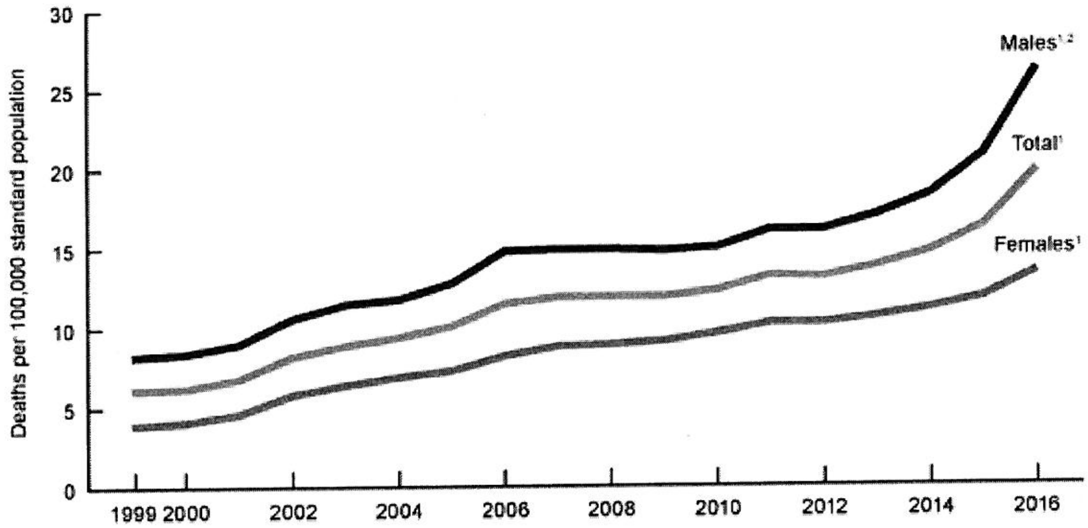
Observing Figure 1, there have been clear increases in the number of deaths from drug overdoses yearly, except between 2006-2009 when the number flattens out slightly. However the slope of the line seems to pick up after 2010, and the steepest part of the line is between 2015-2016, suggesting the biggest increase in the number of overdose deaths in that time. One big thing to note about this graph is that not once does the line fall – there have been no decreases in the yearly number of deaths from overdoses since the millennium started. Should this trend continue, we will be seeing more deaths in the coming years.

2.3.1 Demographic Trends

The figures below show some of the demographic breakdowns in overdose deaths – by gender and by age group. These graphs are taken from the CDC and I provide these to give an initial idea of the trends in overdose deaths across the years.

While overall deaths from overdose are rising, it is intersting to note the breakdown by gender, as seen in Figure 2 below:

Figure 2: Age-adjusted Drug Overdoses Death Rates: United States, 1999-2016

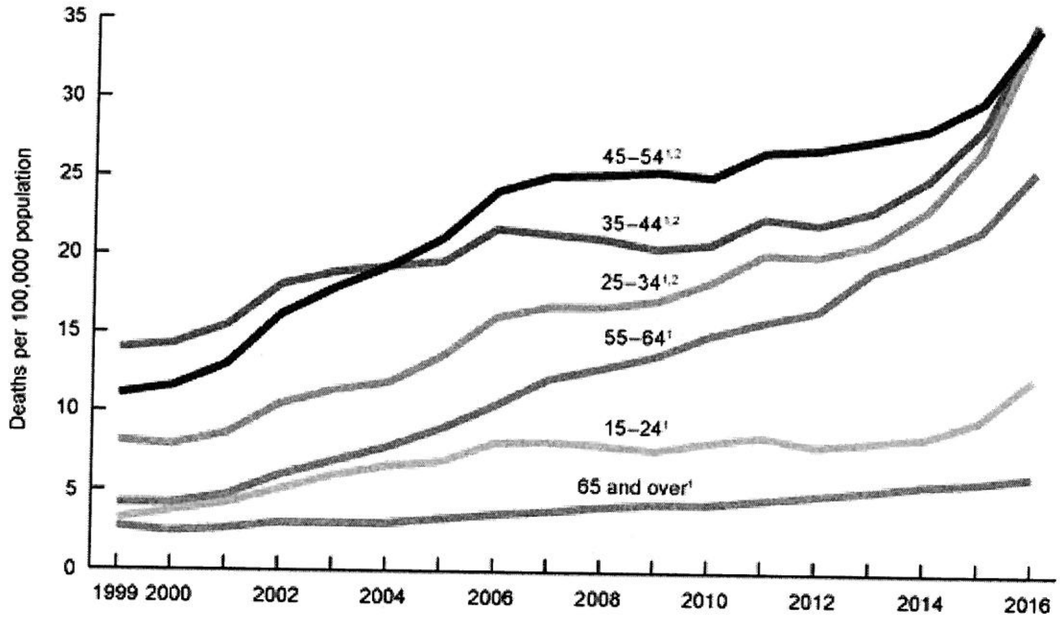


Source: CDC https://www.cdc.gov/nchs/products/databriefs/db294.htm

Throughout the entire time period, more males have overdosed than females. The CDC explains that "rates were significantly higher for males than females," especially in 2016 with $p < 0.001$. Also interesting to note is that both lines have very similar slopes – both changing at the same rate over time – so the gap between the gender differences stays pretty consistent.

Next I discuss drug overdoses by age group, as shown in Figure 3 taken from the CDC below:

Figure 3: Drug Overdose Death Rates, by Selected Age Group: United States, 1999-2016



Source: CDC https://www.cdc.gov/nchs/products/databriefs/db294.htm

There are several points to note in Figure 3. We see first that there were more deaths within the 35-44 age group from the late 90s till about 2004, then the 45-54 age group overtook it and has remained the most at-risk group ever since. With the latest data in 2016 however, we see the gap between the middle-aged (25-54) groups closing, with each group nearing 35 deaths per 100,000 population, a sign of the fast growth of the epidemic.

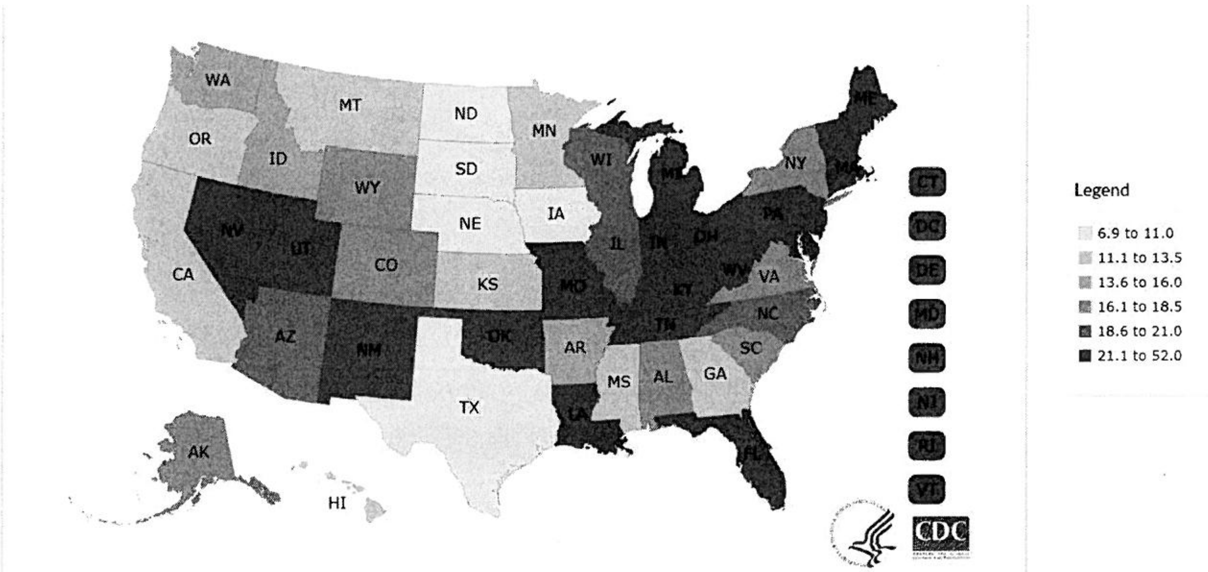
The lowest two age groups are the youngest and oldest ones: 15-24 and 65 and over. Age may not have a linear correlation with death rate then, but maybe an image that follows an inverse parabola, with the likelihood of drug abuse first increasing, and then decreasing with age. There may be something interesting to say about opioid abuse and age then.

In the mid-to-late 2000s, the 35-44 age group death rate declined slightly, while the 15-24 and 45-54 groups stagnated and all the other groups increased. Both the 55-64 and the 65 and over age groups are the only ones that have been increasing throughout decades, with deaths in the 55-64 group increasing at the fastest rate among all the age groups, and deaths in the 65 and over group increasing very slowly. The increasing use of opioids among older people may have started from health problems, then turned into abuse.

The CDC also notes that all the age groups show a "significant increasing trend from 1999-2016 with different rates of change over time, $p < 0.005$."

2.3.2 Geography

Figure 4: Number and Age-adjusted Rates of Drug Overdose Deaths by State, US 2016

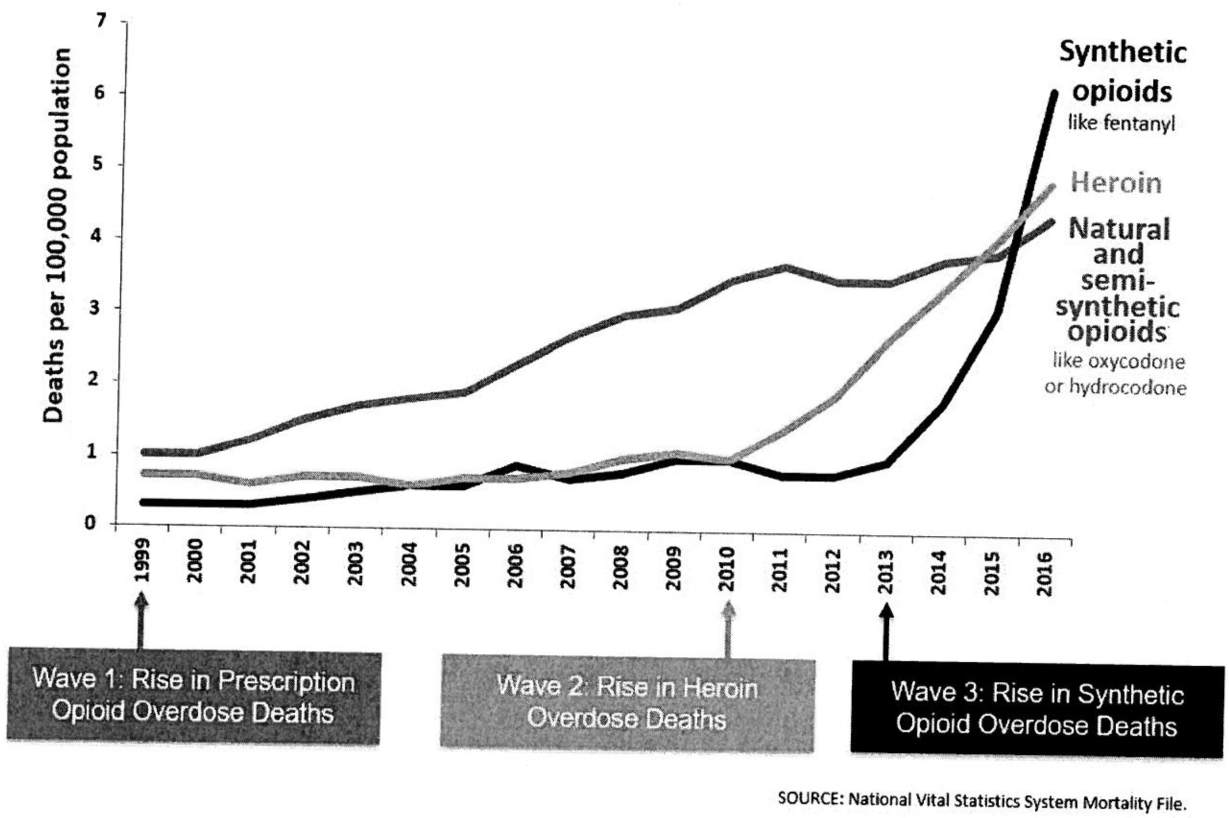


Source: CDC https://www.cdc.gov/drugoverdose/data/statedeaths.html

Consistent with the growing number of overdose deaths, there have been increases of deaths in almost every state in the US from 2015-16. The legend on Figure 4 indicates the age-adjusted rate of drug overdose deaths. It is interesting to note that some of the states with the highest death rates (those in the Midwest and South regions), are also some of the states with the highest numbers of opioid prescriptions per person (Figure 6). The CDC also keeps track of changes in overdose deaths from year-to-year. Despite the growing battle against the opioid epidemic, the US saw statistically significant increases in drug overdose death rates from 2014-15 in 19 states; from 2015-16 that number grew to 26, already present in the Northeast and expanding across the Midwest and the South regions.

2.3.3 Types of Opioids

Figure 5: 3 Waves of the Rise in Opioid Overdose Deaths



Source: CDC https://www.cdc.gov/drugoverdose/epidemic/index.html

The CDC outlines three distinct waves in the rise of opioid-related deaths, stemming first from the use of prescription opioids, then heroin, then synthetic opioids (e.g. fentanyl). While the deaths from prescription (natural and semi-synthetic) opioids remain at a steady increase, there is a huge spike in the deaths from heroin from 2010 onwards, and in synthetic opioids starting from 2013. These waves are just to note that people use different types of opioids and are focused on experiencing greater highs, and that more dangerous opioids are being developed for that purpose.

While prescription opioids are most commonly used, the graph shows that heroin is now on the rise. Most heroin addicts started their drug abuse with prescription opioids, switching to heroin because of its easier availability and lower price (SAMHSA). The lack of documentation involved in supplying and buying heroin (i.e. illegal markets) makes it easier to obtain, and its strong and immediate effects (especially when injected) makes it more attractive to users who want a greater high. Heroin is also highly addictive, with users easily building up a tolerance, making them highly susceptible to overdoses.

The spike in deaths from synthetic opioids, most notably fentanyl, may stem from their wider use because of their stronger effects – "up to 50 times more powerful than heroin" (Bebinger). With such a strong high, sales of prescription *and* illegally-made fentanyl are growing because of their popularity. Now some other drugs are being thrown in the mix (e.g. cocaine, heroin) to produce an even stronger high, and are driving up the overdose deaths from fentanyl.

2.3.4 Pharmaceutical Companies

According to NIDA, most research pinpoints the start of the opioid epidemic to the late 1990s. While opioids had already been around for some time then, it was only following Portenoy and Foley's paper *Chronic use of opioid analgesics in non-malignant pain* (1986) that more doctors began to prescribe them more widely, as the paper concluded that opioid use was safe. Aside from doctors overprescribing opioids to patients for (non-cancer) pain relief, research now shows that pharmaceutical companies also had a very strong hand in driving the crisis by both paying these doctors to prescribe their drugs, and heavily marketing their use.

The pharmaceutical companies Purdue Pharma and Insys Therapeutics were two of the biggest drivers of the widespread use of prescription opioids for chronic (non-cancer) related pain, heavily pushing the market to use their products: OxyContin and Subsys, respectively.

Throughout the late 90s, Purdue Pharma heavily promoted and marketed its newest product, OxyContin, to great success – their sales "grew from $48 million in 1996 to almost $1.1 billion in 2000" (Van Zee 2009). This process is analyzed in Art Van Zee's paper *The Promotion and Marketing of OxyContin: Commercial Triumph, Public Health Tragedy.* To increase sales, Purdue Pharma had to convince others that OxyContin had a low risk of addiction, and influence doctors to prescribe the drug more. The company conducted several national pain-management and speaker-training conferences across the county, training thousands of physicians, pharmacists, and nurses to become spokespeople for OxyContin. In an even bigger move, Purdue Pharma compiled data to identify "physicians with large numbers of chronic-pain patients," specifically targeting physicians who were "the most frequent prescribers of opioids" (Van Zee 2009). The company also targeted specific geographic areas in the US that had higher prescription rates, which were low-income areas in Maine, West Virginia, Kentucky, Virginia, and Alabama. Purdue Pharma also insisted that "the risk of addiction from OxyContin was extremely small," training its representatives with the message that risk was "less than one percent" (Van Zee 2009). It is now known that the company was aware of its users abusing OxyContin, crushing and snorting the drug, yet still promoted a low risk of addiction.

Similarly, Insys Therapeutics implemented a "speaker program" to promote its painkiller Subsys, to great sales success. This speaker program, which investigators now title a kickback scheme, paid doctors and other prescribers to promote the drug, and increased sales in one year by more than 1000% (Hughes).

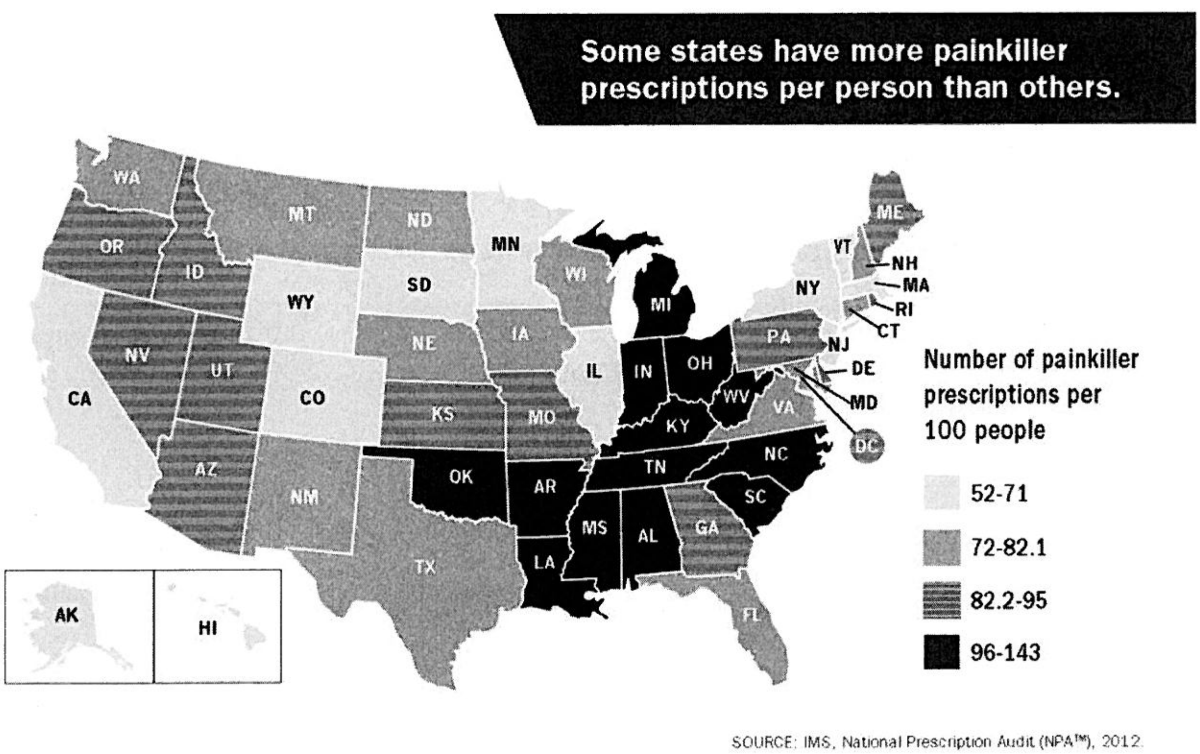
2.3.5 Prescription Rates

While opioids were invented and further modified for the benefit of pain relief, doctors overprescribing these opioids have definitely played a part in the increasing epidemic. Take for example, an emergency room doctor prescribing Ibuprofen, Vicodin, OxyContin, *and* Xanax for back pain. Although the Xanax was just prescribed as a sleeping aid, the number of drugs prescribed – and the combination of these drugs – increases the risk for addiction.

The CDC noted a great increase in the sales of prescription opioids from the late 90s till now – four times the amount prescribed in 1999 – accompanied by an increase in prescription opioid overdose deaths, but "no overall change in the amount of pain Americans report," warranting suspicion about the reasons for the increasing use of opioids in the US. The CDC also noted that there is "not enough evidence that opioids improve chronic pain, function, and quality of life," begging the question of why people turn to opioids and use them in the first place if they are not entirely effective anyway.

Opioid prescription rates vary across states, with the top states prescribing "almost 3 times as many opioid prescriptions per person" compared to low prescribing ones (CDC). The CDC notes that "health issues that cause people pain do not vary much from place to place," so there may be other factors involved the state variability of prescription rates. I provide this map taken from the CDC, using data from the IMS Health National Prescription Audit, to show the variability of prescription rates across states.

Figure 6: Map of Opioid Prescription Rates across the US



Source: CDC https://www.cdc.gov/drugoverdose/data/prescribing.html

As seen on the map, the states with the most prescriptions per person are concentrated in the Midwest and South regions. This observation is consistent with the information I provide in Figure 4, as the top prescribing states are among the states with the highest drug overdose death rates.

Most literature about the opioid epidemic focuses on those Midwest and Southern regions, but it is interesting to note that the next set of states with more painkiller prescriptions is concentrated in the West, sans California. Although the literature also points out that there is a big increase in overdose deaths in the Northeast, there seems to be relatively lower number of painkiller prescriptions in that area, suggesting other types of opioids being taken. However, the CDC also notes that "primary care providers account for about half of opioid pain relievers dispensed," raising concern about how others are obtaining these drugs. As I note just after this section, there are other (less legal) ways of obtaining prescription opioids, and it is easier to obtain prescriptions from some states (and counties) than it is from others.

This increase in prescription rates started in the 90s as pharmaceutical companies (see above) "reassured the medical community that patients would not become addicted to prescription opioid pain relievers" (NIDA Opioid Overdose Crisis). As previously discussed, these companies targeted areas with high prescription rates and drove them up, increasing drug sales and contributing to the growing spread of opioids.

The CDC outlines the most common sources of prescription opioids *for non-medical use* on their website, taken from the US National Survey on Drug Use and Health. The most common source is "given by a friend or relative for free," as possibly this is person's first introduction to opioids. The second most common source is "prescribed by ≥1 physicians," which is interesting since these prescription opioids are for non-medical users. This source might go back to how pharmaceutical companies provided incentives to doctors and prescribers to promote their drugs, increasing usage.

2.7 Reasons for substance abuse

The Substance Abuse and Mental Health Services Administration (SAMHSA) outlines the top reasons why adults misuse prescription drugs. Surveying adults (18 years and older) who misused prescription drugs in the past year, 11.5 million adults responded with the following reasons:

- To relieve physical pain
- To relax or relieve tension
- To experiment or see what the drug is like
- To feel good or get high
- To help with sleep
- To help with feelings or emotions
- To increase or decrease the effects of other drugs
- Because the respondent is "hooked' or has to have the drug
- For some other reason

Keeping these reasons in mind, we can understand a little more about why people turn to drugs in response to certain events. SAMHSA found that 63.4% of those surveyed misused prescription drugs to relieve physical pain, telling us first of all that these drugs are mainly being used for their intended purpose, but that is how addiction starts and can spiral downward from there. Once people start taking larger doses, or use the drug for longer than prescribed, drug abuse occurs.

2.8 Summary

The narrative of the opioid epidemic thus far can be summarized quite neatly. The rise began with research suggesting that opioids would be effective for treating acute, or chronic, non-cancer pain. Pharmaceutical companies then developed new prescription drugs, more potent than morphine, and heavily marketed them as low-risk painkillers. These companies implemented kickback schemes for more doctors to prescribe their drugs, leading to increased prescription rates and an increased risk of addiction. As these drugs relieve not only physical pain but also social stresses, people use them to cope with hardships. Thus events like economic downturns and violent crimes, or even general dissatisfaction or curiosity influence people to abuse drugs. The opioid narrative is common in many areas of the US, prompting officials to determine the main factors driving opioid abuse and thus alleviate the epidemic. This paper aims to do just that.

3 Motivating Factors

3.1 Economic Factors

I first discuss the motivating factors for the variables, building my model. As seen in Figure 4, the top five states with the highest death-from-overdose rates in 2016 were West Virginia, Ohio, New Hampshire, Pennsylvania, and Kentucky. Of these five states, three of them rank in the bottom half of US states in order of GDP per capita (New Hampshire 17th, Pennsylvania 20th, Ohio 25th, Kentucky 42nd, West Virginia 48th) (U.S. Bureau of Economic Analysis). Although New Hampshire and Pennsylvania are in the upper half of the rankings, I felt that GDP – and from there possibly income – would be a possible indicator of the usage of opioids in a given area. And where there is high usage of opioids, there may be high death rates from overdoses that follow. The regions with statistically significant increases in drug overdose deaths also include the Midwest and South, known for their declining economic state. With this observation in mind, I chose to look at economic variables such as average income and unemployment rate since there seems to be an economic correlation at hand.

While average income and unemployment are my main economic factors, other variables I decide to include stem from them, such as children in poverty and income inequality. I chose to include these other measures because these are stressors that in theory could lead one to substance abuse.

3.2 Sociological Factors

This paper comes at a time when the opioid crisis is truly relevant, with more news and research about the epidemic pouring in every day. In the search for factors driving the crisis however, I found that J.D. Vance's memoir, *Hillbilly Elegy*, provides much guidance and insight about the sociological causes of drug abuse. Vance gives suggestions as to what drives people to despair and start taking opioids – past actually being prescribed them. Vance's memoir, published very recently in 2016, tells a lot about the origins of the "hillbilly" stereotype, and the sociological/cultural and economic factors that sustain this image.

For many experiencing the economic downturn in the Midwest and Southern regions of the US, it is easy to blame the government for the decline in manufacturing jobs and overall employment. Opening up trade with other countries, especially China, has been accompanied by a decline in manufacturing since cheap labor and products are provided elsewhere. Vance argues however that trade is not solely to blame; he shows how culture has prompted this decline as well, telling us what he saw growing up in the Midwest.

The most common picture of the Midwest is a region of working-class whites living around the poverty line, with poor health and education levels. Vance describes a trip to Jackson, Kentucky, seeing "decrepit shacks rotting away," and "eyes, all looking at me...with an unsettling combination of fear and longing" (18). This scene, while a true image of Jackson, can be argued that it is merely an exaggerated stereotype. However Vance reminds his readers that people tend to "glorify the good and ignore the bad in ourselves" (20), making this truth difficult to accept. He refers to sociologists Carol A. Markstrom, Sheila K. Marshall, and Robin J. Tryon and their paper, *Resiliency, social support, and coping in rural low-income Appalachian adolescents from two racial groups* (2000), stating that these teens "learn from an early age to deal with uncomfortable truths by avoiding them, or by pretending better truths exist" (20). Although this culture builds some psychological resilience, it also prolongs problems in society.

Vance suggests this "hillbilly culture" is partly to blame for the raging opioid crisis, especially in the Appalachian region. The Appalachian region is known for its poor economic state, with high unemployment, low income, and negative economic growth. However, even though its residents face health, economic, and overall wellbeing issues, the Appalachian region refuses to see any progress made, instead being proud or ignorant of their lifestyle and shaming anyone who criticizes it, blaming it on widespread stereotypes. But these aren't mere stereotypes that Vance presents – these are the real stories of real people he has met along the way, and Vance attempts to explain them through the hillbilly culture these people live.

According to Vance, hillbilly culture includes the structure of their families (large, and tight-knit), religion (professed Catholics, though not necessarily practicing), and politics (Democrats-turned-Republicans). With such similar structures across the region, hillbillies cultivate "an intense sense of loyalty" and "fierce dedication to family and country" – they do not like outsiders or anyone they consider different from themselves. Also included in hillbilly culture is a toxic masculinity and "bizarre sexism" (3).

Most people would agree that better economic conditions would encourage these people to do better – that more jobs would lead to more happiness. However, despite this popular view, economic factors do not explain the whole story behind Appalachia's poor fortunes.

At least, this is what Vance tries to convince us. He tells the story of "Bob" who was always late, would skip a day of work weekly, and "often took three or four daily bathroom breaks, each over half an hour," yet complained when he was fired. He tells of his grandmother "Mamaw's" neighbor who "had no job and was proud of it" (19). He tells of a man who had "the time to make eight children but can't find the time to support them" (20-21).

What Vance describes is "a culture that increasingly encourages social decay instead of counteracting it" (7). Vance talks about the culture of "learned helplessness" and the strange sense of privilege that many people in his hometown carried, which leads people to stand proud of their poor states because they believe their actions and efforts have no real effects, thus extending the blame elsewhere.

Vance emphasized greatly how "the most important lesson in [his] life is not that society failed to provide [him] with opportunities," but that "society devoted too many resources too late in the game" (244). He also learned that "social class in America isn't just about money," and that improving people's lives meant "[extending] past their education and employment and into relationships they formed" (63). While education and employment can provide opportunities for income and social mobility, how one perceives education and employment and uses them leads to entirely different results. Companies provided work for people like "Bob," but those people, for whatever reason, did not view employment seriously. While Vance was able to escape being trapped in a declining town, many do not because of this learned helplessness, instead extending the blame elsewhere.

These aspects of hillbilly culture can thus discourage income and social mobility, contributing to economic stagnation – and even decline. Vance recalls the extreme sexism that plagued his community, citing how "as a child, [he] associated accomplishments in school with femininity. Manliness meant strength, courage, a willingness to fight, and, later, success with girls," resulting in "working-class boys...[doing] much worse in school because they view schoolwork as a feminine endeavor" (245-46). Those with less education face more economic challenges, which can result in substance abuse later on. Although quality and level of education can be linked to income, the effort and achievements people make in these areas are linked to gender and gender norms. This situation is one example of sociological factors behind the opioid epidemic.

This hillbilly culture also creates a potentially hostile environment. A child growing up in this culture may experience some traumatic events that could also drive them to opioids. Vance brings up the issue of Adverse Childhood Experiences (ACEs), which are "traumatic childhood events" that can impact one's adulthood (226). While some ACEs are physical e.g. "being pushed, grabbed, or having something thrown at you" (227), others are deeply psychological and thus have a more meaningful impact on what a child normalizes and expects later on in life. These mental/psychological ACEs include "feeling that your family didn't support each other," "having parents who were separated or divorced," and "living with an alcoholic or drug user" (227). According to Vance, researchers have noted a correlation between children with multiple ACEs and experiencing anxiety and depression, underperforming in school, and even instability in relationships – very mental and behavioral issues. While from past literature we see many correlations between economic factors and opioid abuse, Vance argues that culture is also partly to blame for the growing epidemic. Going back to ACEs, Vance cited Harvard researchers who found that '"significant stress in early childhood...result[s] in a hyperresponsive or chronically activated physiologic stress response, along with increased potential for fear and anxiety"' (Wood et al. 2011). With this heightened fear and anxiety, one growing up with regular traumatic events may start using drugs, such as opioids, to cope with the constant instability and stress. Additionally, if children normalize this everyday behavior around them – and *do not realize* that this behavior is uncommon elsewhere – they contribute to the culture and become part of the cycle.

Other researchers also suggest that sociological factors may be driving the opioid crisis. In their paper, *Bowling alone, dying together: The role of social capital in mitigating the drug overdose epidemic in the United States* (2017), Michael J. Zoorob and Jason L. Salemi explore the relationship between "social capital" and drug overdoses in the US. The two define social capital as "the extent and depth of social trust, norms, and networks." They found "a strong and statistically significant inverse association between county-level social capital and age-adjusted mortality due to drug overdose," suggesting that "social capital protects communities from drug overdose."

Dasgupta, Beletsky, and Ciccarone (2018)also discuss the social factors that drive people to abuse opioids, citing opioids as "a refuge from physical and psychological trauma, concentrated disadvantage, isolation, and hopelessness."

Dr. Nora Volkow of NIDA also supports this, questioning "how low social status might affect addiction risk." However, she also wonders whether this relationship might be bidirectional – "exclusion not only increases risk for using drugs but increased drug use can increase social isolation further, in a vicious cycle."

With these different factors in mind, I move on to detailing some economic, demographic, and sociological factors I have picked out to construct my model. From there

I conduct linear regression – ordinary least squares (OLS) – using R, and discuss my findings regarding which group of determinants significantly drives the opioid epidemic.

4 Method

My information about drug overdose deaths comes from the CDC, a public health institution in the US. Their Information and Communication Wide-ranging OnLine Data for Epidemiologic Research (WONDER) database provides public health information, including causes of death, so I used their data for deaths from drug overdoses. However, it must be noted that these are general drug overdose deaths, and do not necessarily indicate deaths from a specific drug. The CDC notes that "in approximately 1 out of 5 drug overdose deaths, no specific drug is listed on the death certificate," and that in many cases people overdose on multiple drugs; thus we cannot pinpoint a specific drug causing the death. I took information by county from the year 2016, with drug overdose as the induced and underlying cause of death. This data set served as my main independent variable for my model.

I obtained other key dependent variables from the US Census Bureau from the IPUMS-USA data base, containing US Census Data for social, economic, and health research. IPUMS-USA, under the University of Minnesota, takes information from the American Community Survey (ACS) and Current Population Survey (CPS) programs, in cooperation with the Bureau of Labor Statistics (BLS). The variables I was interested in were age, gender, health insurance coverage, income, (un)employment, educational attainment, and social workers and health psychologists as occupations. I also took the data from the year 2016. While the data set contained many other variables (group quarters status, detailed status of employment and educational attainment, other occupations) I only chose those I felt were relevant to this study and filtered the data using my criteria; I discuss my method of cleaning the data in the following section, and subsequently my rationale for my variable selection.

Other health indicators I needed came from the County Health Rankings & Roadmaps (CHR), a program designed to measure vital health factors and improve health and wellbeing. From their 2016 data set, also divided by county, I took information about social associations, single-parent households, and primary care physicians, poor physical and mental health days, education level, children in poverty, income inequality, violent crime rates, food insecurity, and race demographics.

After obtaining this data from the various databases, I compiled them into a master spreadsheet using Stata. I first loaded the data from the ACS (states, counties, population, occupation, employment status, health coverage, sex, average income) and started sifting through the information. To limit the data to just the 50 states, I dropped values that were under the different state groups/territories, which had certain STATE FIPS codes (ones over 56). I had to filter the data even more though under employment status and income. I dropped any values encoded under "not in labor force" and those who earned no income (code 999999). I assigned values of 0 and 1 for binary variables (employment status, health coverage, sex), for easier use in the model. I did the same under occupation, assigning 0 and 1 values for occupations that involved social work, and health psychologists as a subset of social workers.

After examining those initial variables, I added more from the County Health Rankings database, merging all the values using R and again dropping any counties with missing values.

Altogether I had 425 data points (out of 3007 counties), since so many values were dropped. From here I examine the univariate relationships between some variable and the log of death rate before moving on to constructing my multivariate model.

5 Results

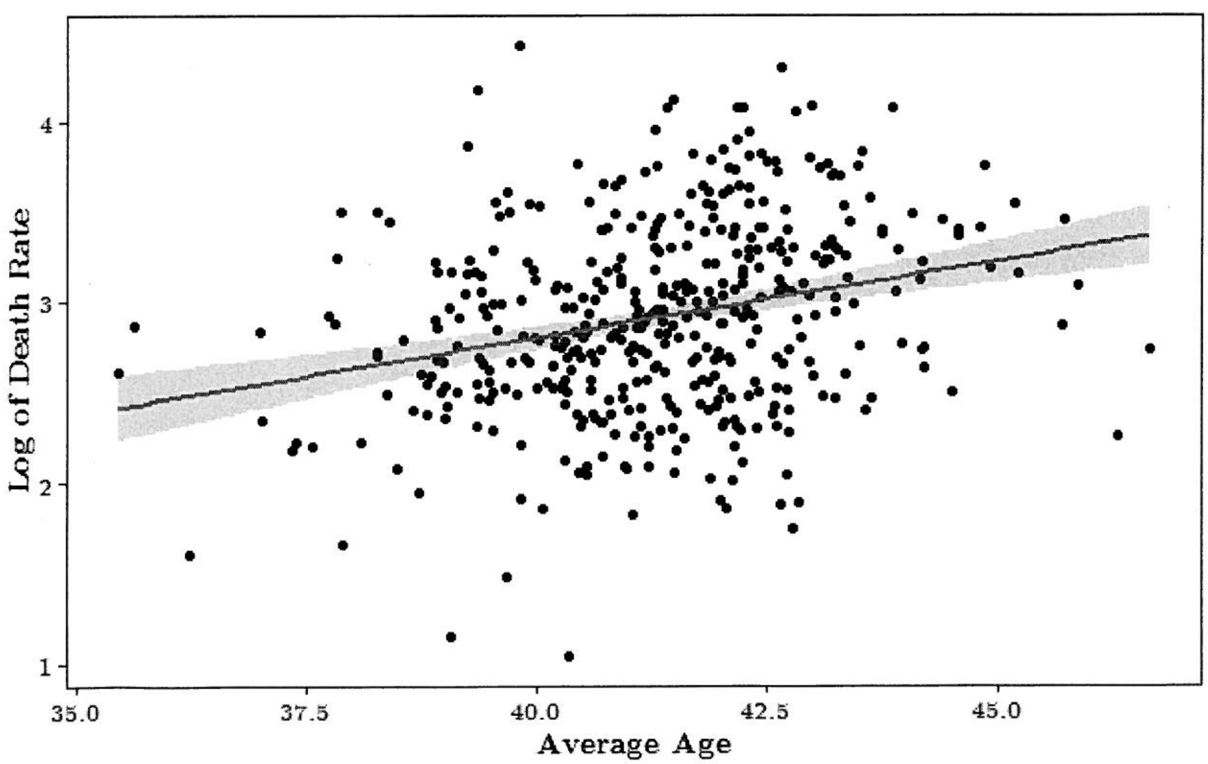
5.1 Descriptive Statistics: Univariate Relationships

For each of these variables, I explain my rationale behind choosing them, discuss some national statistics, then show and explain a scatter plot giving the perceived relationship between the discussed variable and (the log of) drug overdose death rate.

5.1.1 Age

Under demographic variables I looked at average age, the fraction of females, and the fraction of insured adults (under 65) within the county population. As I mentioned in the literature review, the top age group for deaths per 100,000 population was "45-54," although the groups "25-34" and "35-44" were among the top groups as well. Together these three age groups "had the highest rates of drug overdose deaths in 2016 at around 35 per 100,000" (CDC), and all the age groups shown in the graph demonstrate an upward trend in overdose deaths. Seeing the potential correlation from Figure 3, I became interested in including age as a factor.

Figure 7: Scatter Plot of Average Age vs. Log of Death Rate from Overdoses



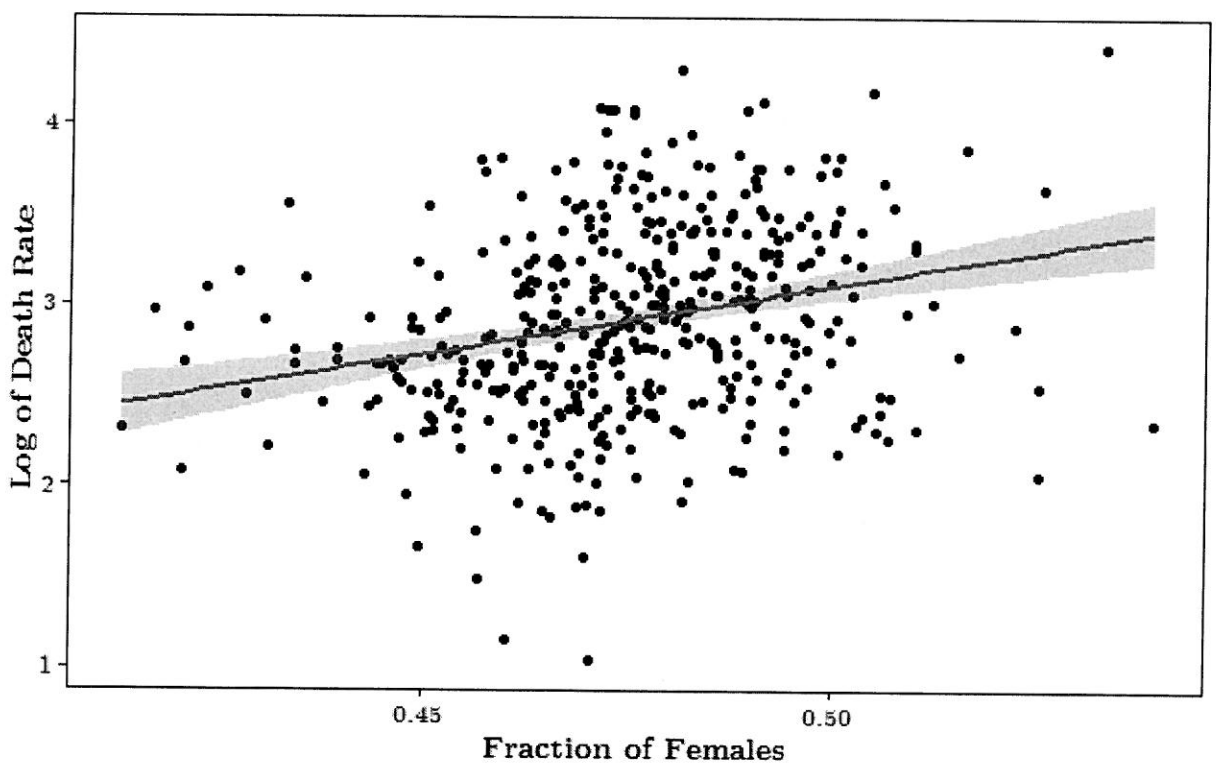
Source: CDC WONDER and IPUMS-USA

From the look of Figure 7, it looks like there is equal variance because the spread is about the same for all average ages tested. There is some change in variance as average age changes, but not significant enough to violate the assumption of equal variance. Although the values are clustered around the middle, there looks to be a moderate positive correlation between the two variables.

5.1.2 Gender

The other demographic variable I selected – fraction of females – was because the CDC also noted that the death rate for females was lower than for males. However, the CDC also found that "women are more likely to use prescription opioids more than men," making an interesting point about how gender might affect not just the usage of opioid, but the type of opioids taken as well.

Figure 8: Scatter Plot of Fraction of Females vs. Log of Death Rate from Overdoses



Source: CDC WONDER and IPUMS-USA

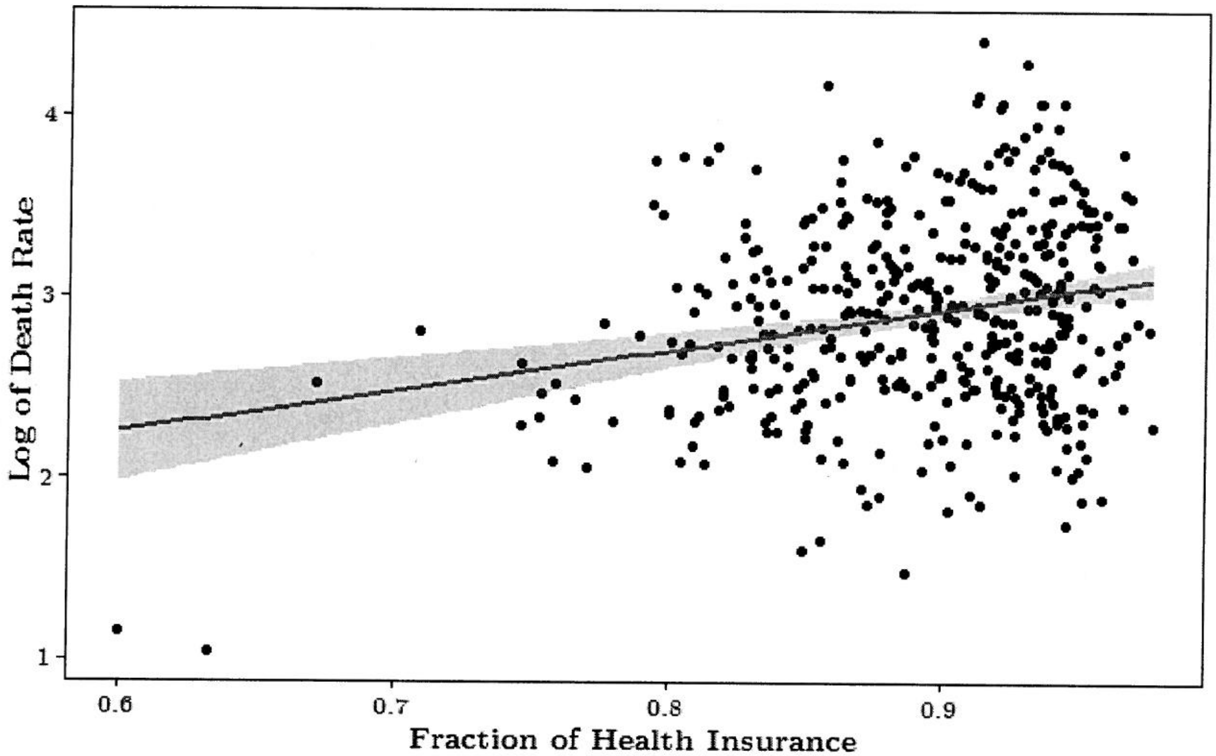
Looking at Figure 8, most of the values are clustered around the middle, and there seems to be a very slight upward trend. This trend could support what the CDC has mentioned about females being more likely to take prescription opioids.

5.1.3 Insurance

Having insurance may be a determinant in overdose deaths because it is easier for insured people to obtain opioids than others. Americans with insurance would be more likely to seek medical help than those without because it would be more affordable. Better affordability could also involve access to a wider range of drugs, increasing the risk of addiction and abuse. Health insurance comes with some moral hazard and overconsumption of health care, so the risk of using opioids for chronic non-cancer and even acute pain grows with being insured.

According to the US Census Bureau, 91.2% of people in the US had health insurance coverage in 2016. This number was an increase from the previous year (90.9%) and has been slowly increasing since the introduction of the Affordable Care Act in 2010 (CDC). Although this model mainly focuses on data from 2016, it would be interesting to note through time series data if increasing health insurance coverage would connect to the increasing over death rates as well.

Figure 9: Scatter Plot of Health Insurance vs. Log of Death Rate from Overdoses



Source: CDC WONDER and IPUMS-USA

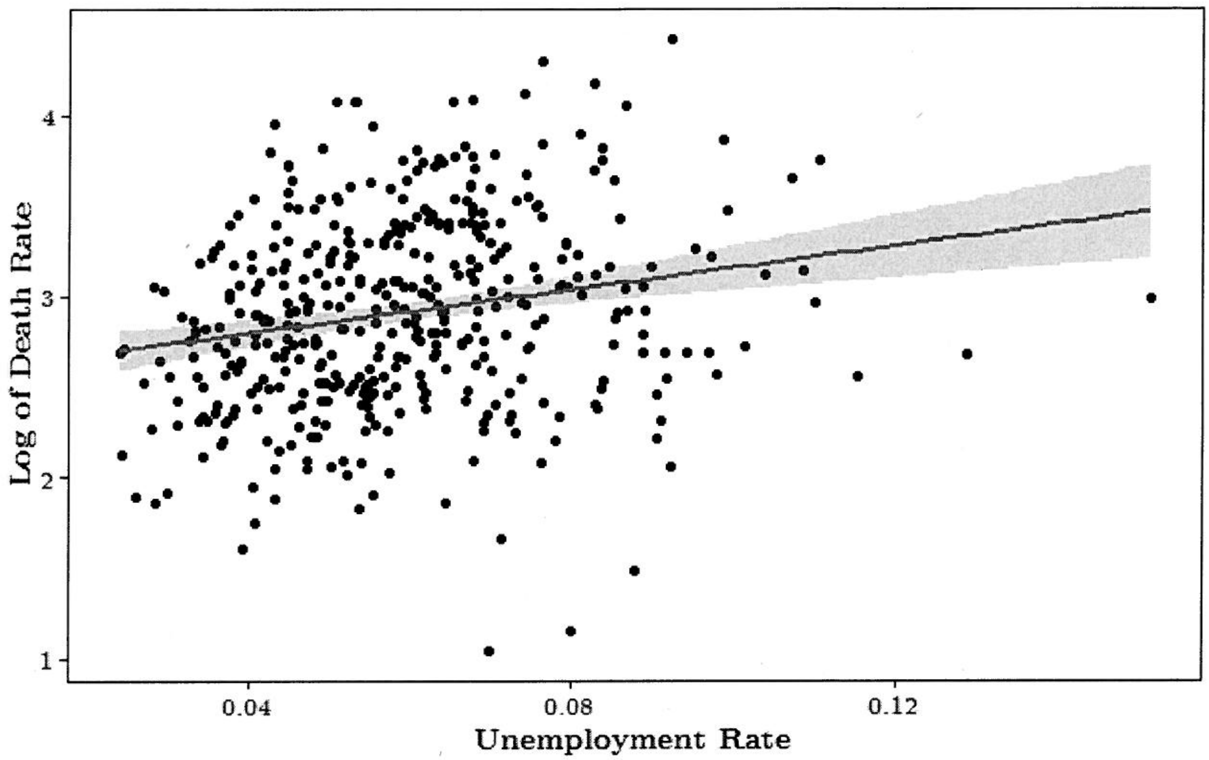
Based on Figure 9, there seems to be a positive relationship between the two variables. However, in this graph, equal variance is definitely violated because most of the values are clustered to the right and taper off to the left – the spread gets smaller as the fraction of health insurance decreases.

5.1.4 Unemployment

Unemployment is one of the main economic variables that come to mind when considering similarities between the Midwest and South regions of the US.

The average unemployment rate in the US in 2016 was 4.9%, according to the Bureau of Labor Statistics. Broken down by region, the average unemployment rate in the Northeast was 4.8%, in the Midwest was 4.7%, in the South was 4.9%, and in the West was 5.1%

Figure 10: Scatter Plot of Unemployment Rate vs. Log of Death Rate from Overdoses



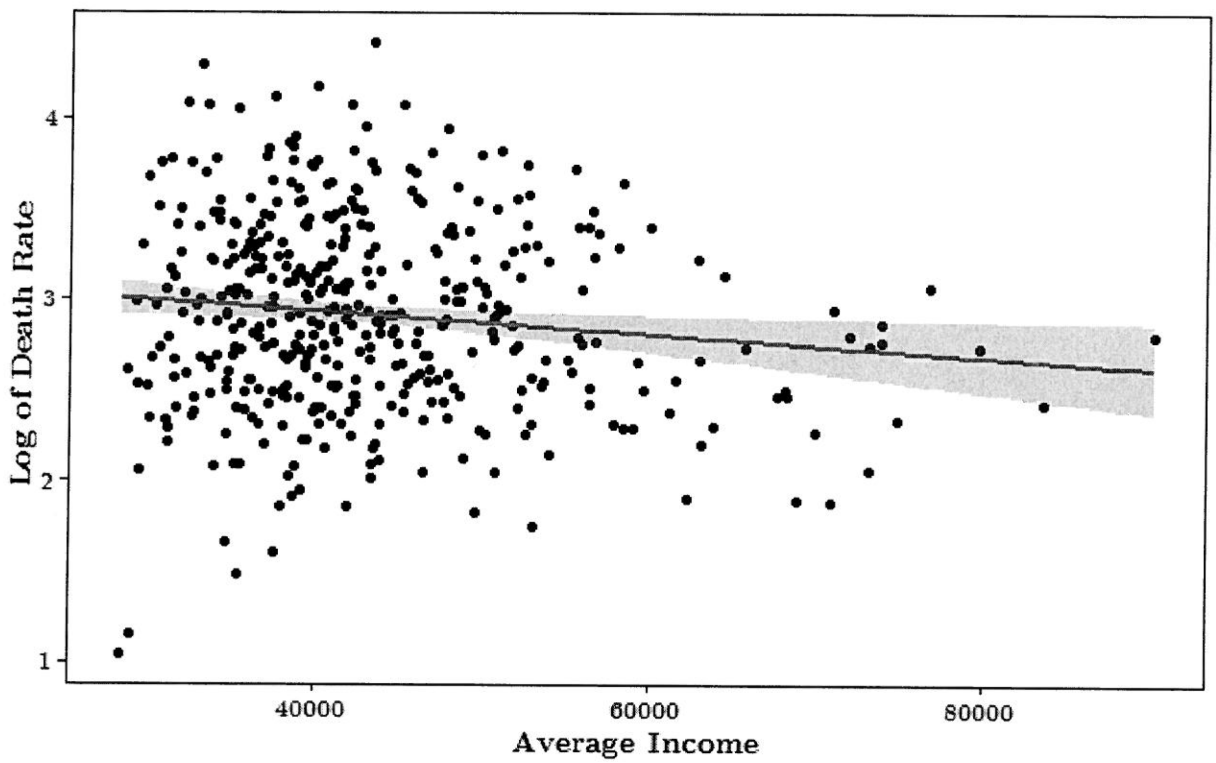
Source: CDC WONDER and IPUMS-USA

Figure 10 above shows a positive correlation between unemployment and log of death rate, which is supported in the literature. Researchers Alex Hollingsworth, Christopher J. Ruhm, and Kosali Simon examine in their paper *Macroeconomic Conditions and Opioid Abuse* (2017)how conditions like unemployment affect drug use. They find that with a one percentage point increase in unemployment, there is a 3.6% rise in opioid overdose deaths, and a 7% increase in opioid overdose emergency department visits. However, they examine unemployment data at the state rather than county-level, enlarging some of their estimated effects.

5.1.5 Average Income

Similar to unemployment, another measurable economic factor is average income.

Figure 11: Scatter Plot of Average Income vs. Log of Death Rate from Overdoses



Source: CDC WONDER and IPUMS-USA

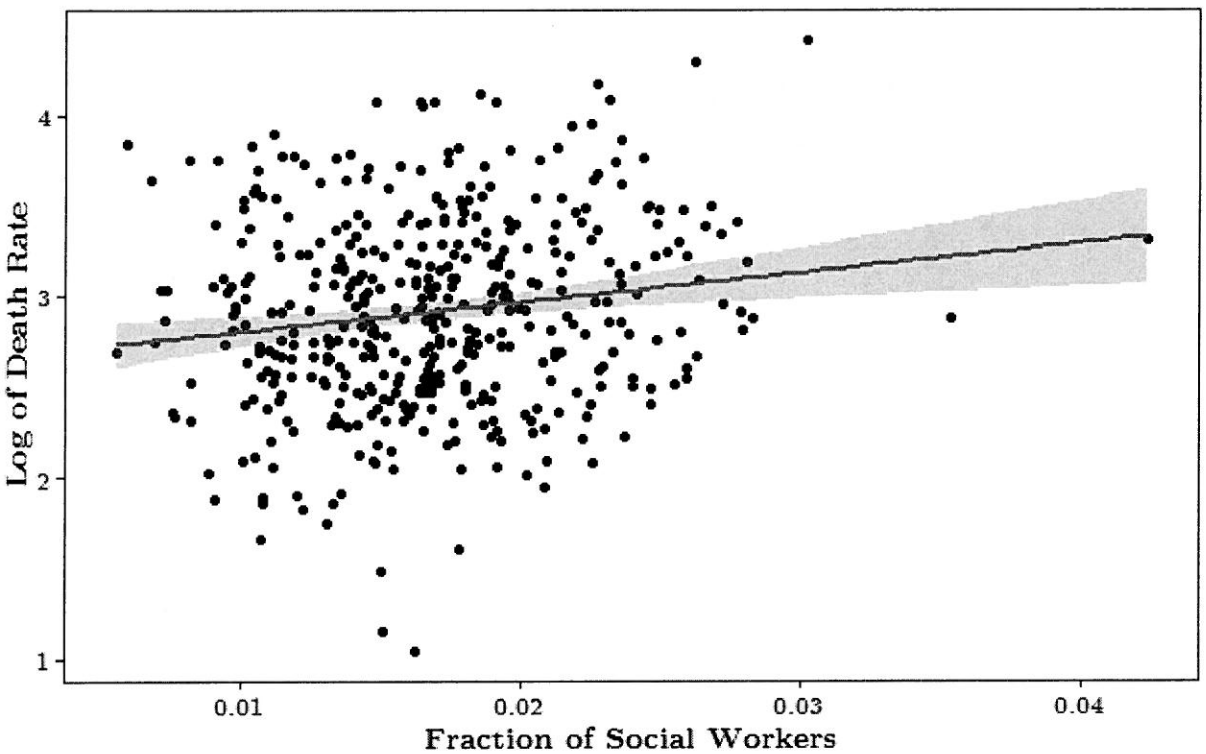
There looks like a negative correlation between average income and the log of death rate, which we can expect since having less income could create more stress in one's life, increasing risk for substance abuse. This information seems consistent with the maps from the CDC where we see more overdose deaths in states with lower income per capita.

5.1.6 Social Workers

Social workers are those who help people solve various problems in their lives, like issues with health, school, and welfare (BLS). I decided to include social workers as a factor in my model because of the question J.D. Vance raises in his memoir *Hillbilly Elegy* regarding the role social workers play in the opioid crisis. Within his hometown of Middleton, OH, there is a large population of social workers, but he also notes the large number of overdose deaths that occur around him.

According to the BLS, the number of social workers is projected to grow by 16% in the next year- much faster than the national average of 7%. As social workers can help curb the opioid epidemic – both in treatment and prevention – this growing number seems to be encouraging.

Figure 12: Scatter Plot of Fraction of Social Workers vs. Log of Death Rate from Overdoses



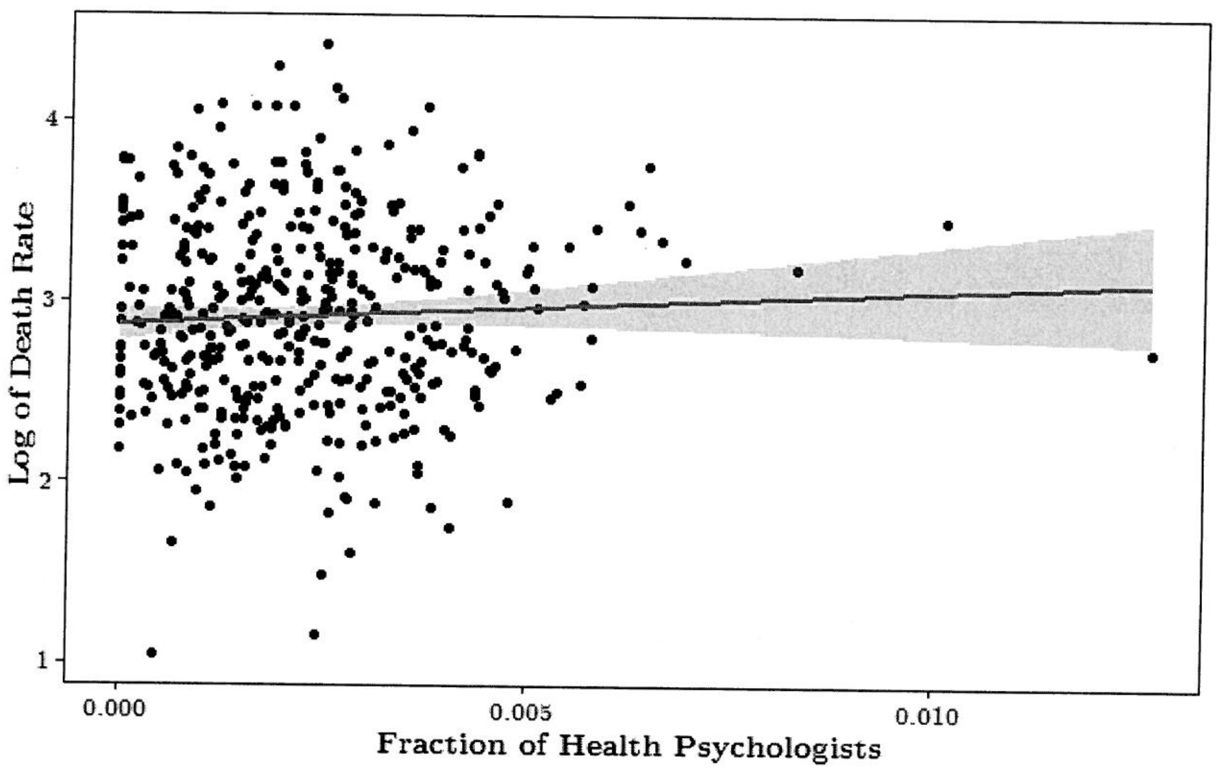
Source: CDC WONDER and IPUMS-USA

There is a positive correlation between the two variables. However, while it is easy to interpret this result as the presence of social workers driving up overdose deaths, it may actually be the reverse – that overdose death rates are increasing the demand for social workers in a county.

5.1.7 Health Psychologists

I chose to include health psychologists in my model because most of the literature agrees that stress increases risk of substance use and abuse. Similar to social workers, health psychologists would help prevent or treat opioid addiction as they examine how behavior impacts health and physical well-being. Better health can decrease risk of substance abuse because people will not need drugs if they are not in pain, and may feel more stable in their lives, thus decreasing the urge to turn to drugs to cope with stress.

Figure 13: Scatter Plot of Fraction of Health Psychologists vs. Log of Death Rate from Overdoses



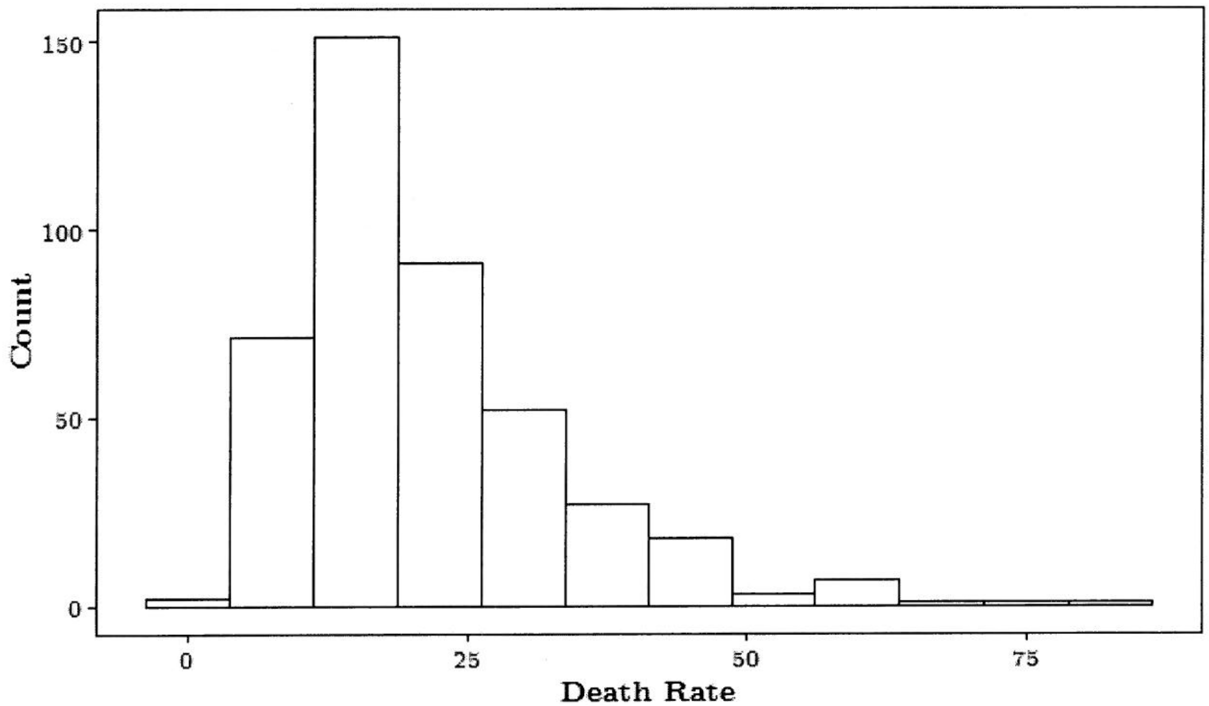
Source: CDC WONDER and IPUMS-USA

Similar to the fraction of social workers, there seems to be very little correlation between these two variables. This result may be because the fraction of health psychologists is a subset of the fraction of social workers. However, this result may change once I combine all these variables into my multivariate model.

5.2 Initial Data Work

Using R, I first constructed a histogram of the US death rates from overdoses in 2016.

Figure 14: Histogram of Death Rate from Overdoses

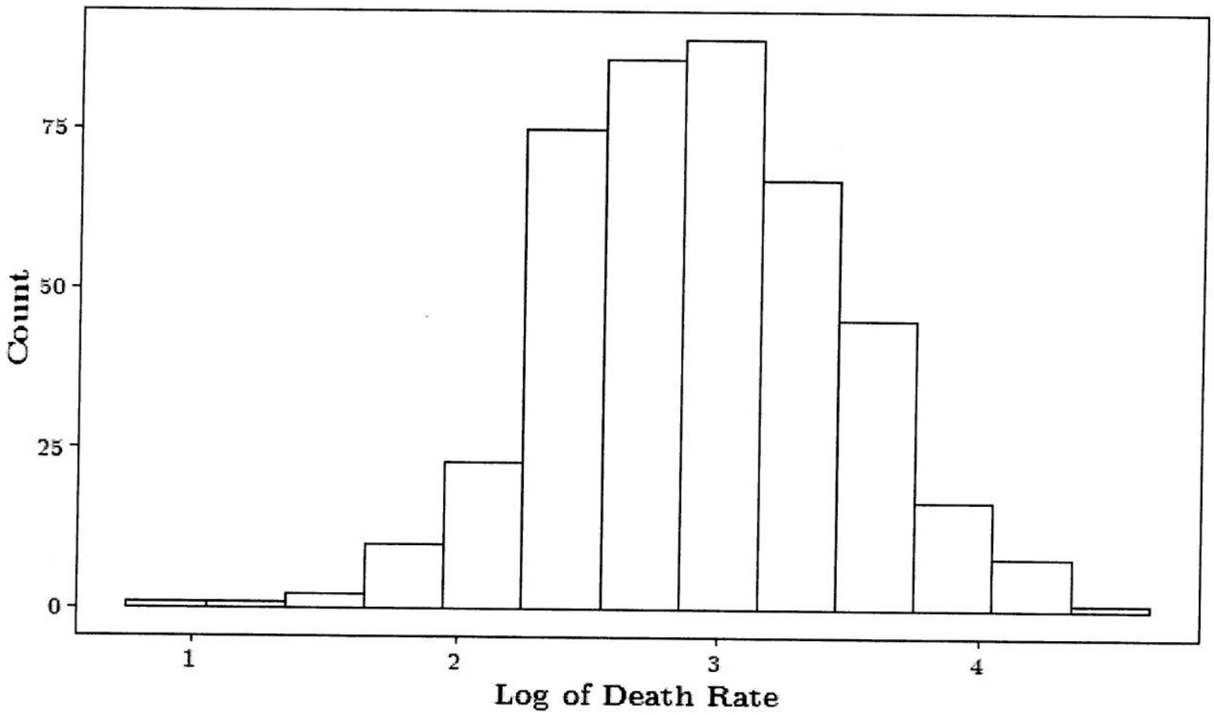


Source: CDC WONDER

This initial step was to see what kind of statistical distribution I was dealing with, and from there lead me to the types of interpretations I could make.

Based on the histogram, the data is skewed to the right; thus clearly not a normal distribution. This figure is concerning since it would be difficult to work with the data and conduct statistical tests. It would also be difficult to generalize my results if my data were not normally distributed. I then used a log transformation to correct the data, yielding a better result:

Figure 15: Histogram of Log of Death Rates from Overdoses



Source: CDC WONDER

This histogram now looks more normally distributed, and I am now ready to conduct regressions on my data.

My next step would be to conduct OLS between the log transformation on the dependent variable death rate and the independent variables that fell into the same categories (sociological, economic, demographic).

5.3 Constructing the Model:

The model I had first constructed when selecting these variables was this:[1]

$$\begin{aligned} \log(death\_rate) = {} & x_1 average\_age + x_2 frac\_female + x_3 frac\_health\_insurance \\ & + x_4 average\_income + x_5 unemp\_rate \\ & + x_6 frac\_social\_w + x_7 frac\_health\_psych + \varepsilon \end{aligned} \tag{1}$$

[1] frac_female = fraction of females in the county population; frac_health_insurance = fraction of the county population that has health insurance; unemp_rate = unemployment rate; frac_social_w = fraction of social workers in the county population; frac_health_psych = fraction of health psychologists under social workers

5.3.1 Model: Sociological Variables Only

I first wanted to see how the model would look with just the sociological variables, since I was concerned with the sociological side of the opioid epidemic. Using only the fraction of social workers and health psychologists in the population, my initial model looked like this:

$$\log(death\_rate) = x_1 frac\_social\_w + x_2 frac\_health\_psych + \varepsilon \tag{2}$$

Taking Equation 2, I used the regression function on R, which outputted the results shown below:

Table 1: Model with Occupational/Sociological Variables Only

	Dependent variable:
	log(death rate)
frac_social_w	15.795***
	(5.242)
frac_health_psych	9.495
	(16.449)
Constant	2.626***
	(0.093)
Observations	425
R^2	0.024
Adjusted R^2	0.019
Residual Std. Error	0.527 (df = 422)
F Statistic	5.190*** (df = 2; 422)
Note:	*p<0.1; **p<0.05; ***p<0.01

Based on the results in Table 1, if the fraction of social workers in a county population increases by 0.01, the log of death rate would increase by .15. The variable of social workers is significant at the 0.001 level, and its coefficient suggests that it largely impacts death rate. However, the variable of health psychologists is not a significant predictor. This regression may indicate then that sociological factors play a role, but as strong as I hypothesized them to be.

However, as I stated earlier, we must consider the possibility of endogeneity here, which may skew our results. The positive correlation between the presence of social workers and death rates may actually be the other way around – the increase in death rates from opioids may motivate more social workers to come to the county to aid in the issue.

5.3.2 Adding in Economic Variables

I wanted to see how the model would change once I added in the economic variables, now using this equation:

$$\log(death\_rate) = x_1 average\_income + x_2 unemp\_rate + x_3 frac\_social\_w + x_4 frac\_health\_psych + \varepsilon \tag{3}$$

Table 2: Model adding in Economic Variables

	Dependent variable:
	log(death rate)
average_income	-0.00000
	(0.00000)
unemp_rate	5.104***
	(1.543)
frac_social_w	11.888**
	(5.267)
frac_health_psych	22.472
	(17.359)
Constant	2.486***
	(0.204)
Observations	425
R^2	0.062
Adjusted R^2	0.053
Residual Std. Error	0.518 (df = 420)
F Statistic	6.982*** (df = 4; 420)
Note:	*p<0.1; **p<0.05; ***p<0.01

When I add in the economic variables of average income and unemployment rate into my model, the coefficients change. The effect of the fraction of social workers decreases, as it goes from 15.79458 to 11.89. It still remains a significant variable however, suggesting that sociological factors still have some role to play in the epidemic. If unemployment rate increases by 1 percentage point, the log of death rate increases by 0.05104, and this variable is significant at the 0.001 level.

5.3.3 Adding in Demographic Variables

At this point I now add the demographic variables, arriving at the model I constructed earlier in Equation 1. With the full set of variables, the regression output gives this summary:

Table 3: Model adding in Demographic Variables

	Dependent variable:
	log(death_rate)
average_age	0.106***
	(0.015)
frac_female	3.934***
	(1.304)
frac_health_insurance	2.836***
	(0.480)
average_income	-0.00001***
	(0.00000)
frac_social_w	6.572
	(5.216)
frac_health_psych	7.391
	(15.672)
unemp_rate	5.647***
	(1.388)
Constant	-5.713***
	(0.827)
Observations	425
R^2	0.266
Adjusted R^2	0.254
Residual Std. Error	0.460 (df = 417)
F Statistic	21.632*** (df = 7; 417)
Note:	*p<0.1; **p<0.05; ***p<0.01

As seen in Table 3, once I add the demographic variables, the sociological variables no longer contribute meaningfully to the strength of the model. The coefficient of the fraction of social workers drops once again, and in this model it is no longer a significant variable. The fraction of health psychologists remains an insignificant variable, suggesting that sociological variables do not contribute meaningfully to drug overdose deaths. If I remove them from my model, I obtain these results:

Table 4: Initial Model of Log of Drug Overdose Death Rate, US 2016

	Dependent variable:
	log(death rate)
average_income	-0.00001***
	(0.00000)
unemp_rate	5.752***
	(1.381)
average_age	0.101***
	(0.014)
frac_female	4.407***
	(1.222)
frac_health_insurance	2.957***
	(0.465)
Constant	-5.726***
	(0.804)
Observations	425
R^2	0.263
Adjusted R^2	0.255
Residual Std. Error	0.460 (df = 419)
F Statistic	29.980*** (df = 5; 419)
Note:	*p<0.1; **p<0.05; ***p<0.01

All the variables are significant at the 0.01 level, and what I have left now suggest that these are the variables that best fit my model – that maybe sociological variables are not the main drivers of the opioid epidemic. The most important factors for now, it seems, are unemployment rate and gender, both showing strong positive relationships with drug overdose death rate.

5.4 Evaluating the model through AIC

To test my variables again more formally, I then used Akaike information criterion (AIC) to give me the best variables for my model. AIC is a way of selecting the best from a wide range of variables, picking which ones best help a model. We want to minimize AIC so we can identify redundant predictors – ones that won't help our model – and subsequently remove them.

There are two ways (or directions, rather) to conduct AIC: forward and backward. Forward refers to adding variables from a selection, and backward refers to removing variables to yield the best model.

According to my AIC results, removing the variables fraction of social workers and fraction of health psychologists respectively would yield a lower AIC, thus the best variables to include in the model would be the ones under the economic and demographic domains, supporting the steps I did above. When I conducted forward AIC, it also gave the same result.

Thus the best model we have so far is this:

$$\begin{aligned} \log(death\_rate) &= x_1 average\_age + x_2 frac\_female + x_3 frac\_health\_insurance \\ &\quad + x_4 average\_age + x_5 unemp\_rate + \varepsilon \end{aligned} \tag{4}$$

Despite these results, I must be cautious about the relationships I found. For some variables like unemployment, we risk endogeneity – if there is some causality involved, we do not know in which direction it points. While the lack of a job may push some people to start using and abusing drugs, it is highly reasonable that using drugs would lead to unemployment. In the case of endogeneity, we would have to use an instrumental variable in place of unemployment to isolate causality and remove any correlation with the error term ε.

Although this model suggests that sociological variables play some role, it is not as large as I had initially hypothesized. The sociological variables I included in this initial model are insignificant, and thus seemingly play a very small role in drug overdose death rates. However, this result may change once I add more variables to the equation.

5.5 Constructing the final model

After my preliminary investigation of variables, I add more taken from the County Health Rankings to find which ones have the greatest impact and are most useful to my model. I grouped these variables again into economic, demographic, and sociological categories.

Other economic variables I decided to include were children in poverty, income inequality ratio, and health care costs. While I had included unemployment and average income earlier, I thought these other variables added a more thorough picture of the economic situation of a county, and would contribute to substance abuse.

Under demographic variables, I added primary care physician rate,[2] percent of single-parent households, race,[3] proficiency in English, percent of the population in a rural setting, and marital statuses.[4]

Under sociological variables, I added mentally unhealthy days, high school graduation rate, some college,[5] association rate,[6] violent crime rate,[7] percent frequent physical and mental distress, percent food insecure, percent insufficient sleep, and homicide rate.

In merging the data, I had to drop 13 data points because they were missing values under the new variables I had added. This data manipulation brings me now to only 412

---

[2] "Number of PCPs per 100,000 population"

[3] African American, American Indian/Alaskan Native, Asian, Native Hawaiian/other Pacific Islander, Hispanic, Non-Hispanic White

[4] Married, widowed, divorced, separated, never married

[5] "Percentage of adults age 25-44 with some post-secondary education"

[6] "Number of membership associations per 10,000 population"

[7] "Number of reported violent crime offenses per 100,000 population"

data points, which is a slight concern considering the generalities I want to discuss in the data set.

After adding all of these variables to my model, R yielded this regression output:

Table 5: Final Regression Output

	Dependent variable:
	log(death rate)
average_age	0.082*** (0.027)
frac_female	-2.859* (1.681)
frac_health_insurance	0.859 (0.713)
average_income	-0.00000 (0.00001)
unemp_rate	4.773*** (1.566)
mentally_unhealthy_days	0.190 (0.163)
pcp_rate	0.0002 (0.001)
high_school_graduation_rate	0.008** (0.004)
some_college	0.008 (0.005)
children_in_poverty	-0.007 (0.008)
income_inequality_ratio	-0.092 (0.063)
perc_single_parent_households	0.012 (0.008)
association_rate	-0.017* (0.009)
violent_crime_rate	0.0002 (0.0001)

perc_frequent_physical_distress	-0.072*
	(0.037)
perc_frequent_mental_distress	0.069
	0.052)
perc_food_insecure	-0.008
	(0.012)
perc_insufficient_sleep	0.036***
	(0.010)
health_care_costs	0.00003
	(0.00002)
residential_segregation_index_nonwhite_white	0.002
	(0.003)
homicide_rate	0.018**
	(0.009)
perc_african_american	-0.011
	(0.035)
perc_american_indian_alaskan_native	-0.008
	(0.037)
perc_asian	-0.015
	(0.037)
perc_native_hawaiian_other_pacific_islander	0.010
	(0.093)
perc_hispanic	-0.005
	(0.034)
perc_non_hispanic_white	0.012
	(0.035)
perc_not_proficient_in_english	0.016
	(0.016)
perc_ rural	-0.003
	(0.002)
married	-0.412
	(0.302)
widowed	-0.374
	(0.303)

divorced	-0.384
	(0.301)
separated	-0.288
	(0.305)
never_married	-0.383
	(0.302)
Constant	35.601
	(30.486)
Observations	412
R^2	0.548
Adjusted R^2	0.507
Residual Std. Error	0.377 (df = 377)
F Statistic	13.418*** (df = 34; 377)
Note:	*p<0.1; **p<0.05; ***p<0.01

As seen in the table, having many variables complicates the model, and most of them appear to be insignificant. With all these variables, only average age, fraction of females, unemployment rate, high school graduation rate, association rate, frequent physical distress, insufficient sleep, and homicide rate appear to be significant.

5.6 Reevaluating the model using AIC

Since I had so many variables, the best (and most formal) way of narrowing them down was, again, to use AIC.

5.6.1 Backward:

In conducting backward AIC, R yielded the following results:

Table 6: Model constructed using AIC, Backward direction

	Dependent variable:
	log(death_rate)
average_age	0.083***
	(0.020)
frac_health_insurance	1.351**
	(0.591)
unemp_rate	4.402***
	(1.470)
high_school_graduation_rate	0.007**
	(0.003)
income_inequality_ratio	-0.098**
	(0.048)
association_rate	-0.015*
	(0.008)
violent_crime_rate	0.0003*
	(0.0001)
perc_frequent_physical_distress	-0.068**
	(0.031)
perc_frequent_mental_distress	0.110***
	(0.034)
perc_insuffcient_sleep	0.037***
	(0.008)
health_care_costs	0.00003
	(0.00002)
homicide_rate	0.020**
	(0.008)
perc_african_american	-0.007**
	(0.003)
perc_non_hispanic white	0.018***
	(0.002)
perc_rural	-0.005***
	(0.002)
married	-0.155***

	(0.044)
widowed	-0.124**
	(0.055)
divorced	-0.125***
	(0.045)
never_married	-0.123***
	(0.044)
Constant	8.382*
	(4.296)
Observations	412
R^2	0.536
Adjusted R^2	0.513
Residual Std. Error	0.375 (df = 392)
F Statistic	23.819*** (df = 19; 392)
Note:	*p<0.1; **p<0.05; ***p<0.01

At this point nearly all the factors narrowed down by AIC are significant (health care costs were insignificant, but the AIC calculations still preferred to keep it in the model).

Most of the regression results looked consistent with what I had observed in the literature and hypothesized as well. Average age, having health insurance, and unemployment rate remain significant variables, all positively correlated with the log of death rate. Among those three variables – and all others included in the model – unemployment rate has the highest coefficient with a 1 percentage point increase in unemployment accompanied by a 0.04402 increase in the log of death rate. This coefficient is slightly lower than the result in Table 4, but that may be because of the addition of several new variables in my model.

Aside from the usual demographic and economic variables present, I found some of the variables AIC calculated to include to be related with the sociological factors Vance presents in his memoir. The criterion suggested including association rate, violent crime rate, homicide rate, and different marital statuses as part of the model. I previously discussed how different researchers examined social associations as a factor that may influence people to abuse opioids, and the negative coefficient from the regression seems to be consistent with their research: a lower association rate is accompanied by a higher log of death rate.

Vance and other researchers also assert that violent or traumatic events lead people to start taking opioids, and their notions are also supported in my model. Both violent crime and homicide rates are significant variables and show positive correlations, suggesting that people use opioids, possibly as a coping mechanism, in a greater presence (and higher frequency) of traumatic events. However, the model also shows a negative correlation between several non-married statuses and the log of death rate, which seems inconsistent with the story Vance presents of ACEs. A greater married population is correlated with a lower log of death rate, and this result seems reasonable because a more stable environment is present. As both the married and non-married statuses yield negative correlations, this result is confusing, and makes me worried about collinearity involved, especially between the factors of widowed, divorced, and never married.

One strange result is that that the percent of frequent physical distress people experience is negatively correlated with the drug overdose death rate. One might think that more physical distress would lead one to start using opioids for pain relief, but this result slightly supports the CDC's observation that there was no positive correlation between prescription opioid sales and the amount of pain Americans report, and maybe no relationship at all.

However, we see that there is a relatively strong positive correlation between the frequency of mental distress and drug overdose death rate. This result seems reasonable since it is common for people to cope with stress using drugs, and the sedative/euphoric effects of opioids makes them attractive for users.

5.6.2 Forward:

Table 7: Model constructed using AIC, Forward direction

	Dependent variable:
	log(death_rate)
widowed	0.003
	(0.026)
perc_hispanic	0.006**
	(0.003)
mentally_unhealthy_days	0.116
	(0.157)
frac_health_insurance	1.411**
	(0.607)
health_care_costs	0.00003
	(0.00002)
homicide_rate	0.018**
	(0.008)
perc_frequent_physical_distress	-0.078**
	(0.032)
perc_asian	0.001
	(0.006)
perc_insufficient_sleep	0.037***
	(0.009)
perc_non_hispanic_white	0.022***
	(0.004)
married	-0.030***
	(0.007)
average_age	0.079***
	(0.020)
unemp_rate	4.233***
	(1.489)
perc_rural	-0.005***
	(0.002)

separated	0.105**
	(0.044)
high_school_graduation_rate	0.008**
	(0.003)
association_rate	-0.016*
	(0.008)
violent_crime_rate	0.0003**
	(0.0001)
income_inequality_ratio	-0.086*
	(0.045)
perc_frequent_mental_distress	0.080
	(0.050)
Constant	-4.353***
	(0.998)
Observations	412
R^2	0.536
Adjusted R^2	0.512
Residual Std. Error	0.375 (df = 391)
F Statistic	22.546*** (df = 20; 391)
Note:	*p<0.1; **p<0.05; ***p<0.01

Compared to Table 6, Table 7 has 20 variables instead of 19. Most of them are the same, with minimal differences: Table 7 includes "mentally unhealthy days" while Table 6 does not, Table 6 includes the percentage of African-Americans while Table 7 includes the percentage of Asians and Hispanics, and Table 6 includes the percentage of divorcees while Table 7 includes separated.

While most of the variables between the two models are the same, some of their coefficients differ (though not by much); for example, the coefficient of average age in the backward model is 0.083, while in the forward model it is 0.079. These numbers are negligible, and as such, I will focus on the variables each model chooses to include (and exclude) and how those variables impact the opioid narrative.

Some variables supporting the sociological factors for drug abuse include the association rate, and the percent of separated people within a population. Again consistent with research expressing the relationship between social capital or connections and substance abuse, both models show a significant negative relationship between the two. This result tells us not to neglect social factors when considering the causes and extent of the epidemic.

However, both models also suggested excluding the percentage of primary care physicians in the county, seemingly ignoring the role physicians have played in worsening the epidemic. Opioid literature has repeatedly stressed this role, with physicians being careless or bought-off to prescribe more opioids, so it is very strange for both models to remove such a seemingly important factor.

Interestingly enough, both models suggested discarding gender and average income. While most of the literature shows a correlation between gender and overdose death rates, and even the use of prescription opioids, both models suggest that there is no real evidence for causation. Excluding average income seems contrary to the literature as well because there is some evidence of pharmaceutical companies targeting low-income areas. Both models indicate that income inequality is a significant variable, so that might be a better factor to consider instead of just income itself.

In any case, unemployment is the strongest factor for drug overdose deaths in both models. This result supports previous research (Hollingsworth, Ruhm, and Simon 2017) and what we observe in the data: Figure 4 shows a map of drug overdose deaths by state, and the states with the highest death rates are known to be states with poor unemployment rates.

Although the models show it is a significant factor, suggesting some causality, we still have reason to suspect some endogeneity involved – unemployment may lead someone to substance abuse, but substance abuse can most definitely also lead to unemployment. We would need to conduct more tests to identify reverse causality, if there is any.

6 Conclusion and Further Works

This paper assesses the opioid epidemic from a sociological standpoint using cross-sectional data. To start, I first chose common variables from the literature and observed data that are linked to substance abuse and addiction, such as gender, age, income, and unemployment. However I wanted to add other types of variables to get a bigger, more complete picture of the epidemic so it can better be addressed. Drawing inspiration from J.D. Vance's memoir *Hillbilly Elegy*, I chose to include sociological variables as Vance describes how culture has shaped behavior and driven most of the problems appearing in the Appalachian Region, a hotspot for drug abuse. I quantified these cultural findings, such as violence and household structure, and added them to my model.

Regressing all the variables I had chosen on the log of death rate, I found that it is really a combination of economic and sociological variables that are driving these drug overdose deaths. The most important economic variables are unemployment rate and income inequality ratio. Having health insurance, as well as the presence of physical and mental distress are positively correlated with drug overdose deaths because people with insurance have better access to prescription opioids, and use them when distressed. Surprisingly however, the rate of primary care physicians in an area does not help my model, and according to the calculations is better off excluded from the model. Lastly, association rate is a significant variable in the model, supporting previous research about how social ties affect drug abuse.

As this is just an initial analysis of the epidemic, focusing on data from the year 2016, future work should examine panel data since they might yield better results. Future work should also include additional variables that proxy for others to better address the issue of multicollinearity.

Additionally, other variables of interest might include shipments, the role of China, and manufacturing employment. Prescription data would especially be of great use here since most users start with prescription opioids. As far as I know, there is only data on Medicare members, who do not represent a large population of opioid addicts.

However this research does fall short of the true extent of the epidemic. As I stated earlier, the data I use are all drug overdose deaths, and are not drug-specific. While the CDC notes that in 2016 opioids accounted for about 60% of overdose deaths, this number might not be so accurate because they cannot always pinpoint the true cause of death. Many overdose deaths involve a cocktail of drugs, and autopsies cannot always tell them apart. The number from the CDC may also be understated as some deaths can go unreported. Officials have started to treat deaths as homicides to hold someone accountable, but treating overdoses as crimes may deter people from seeking help and reporting deaths. While both issues – underlying cause of death and overdose reports – are difficult to tackle, a more accurate count of opioid overdose deaths would be better for the model.

This data and research is also only focused on drug overdose *deaths*, which is the most extreme result of addiction. As we want to tackle the opioid epidemic before addicts reach the point of death (prevention vs. treatment), it might be better to examine other dependent variables, such as just overdoses, or even addiction itself, before the issue escalates. While the CDC publishes some vital statistics on overdoses, this data is taken from ER visits and is not readily available to the public.

Although I used county-level data to obtain an in-depth picture of the opioid epidemic, the way I chose to merge my data may have skewed and biased my results. In choosing my variables and merging my various data sets, I ended up dropping counties that had missing data in any of my variables. While the US officially has 3,007 counties (which would make for 3,007 data points), my final model only contained 412.

Lastly, while I attempted to model the "hillbilly" characteristics Vance outlined in his memoir, most went unquantified. I could not create quantitative measures for the culture of learned helplessness, immunity to hard work, feelings of privilege, and denying the truth.

Although Markstrom, Marshall, and Tryon were able to measure some qualitative variables, such as social support, wishful thinking, and resiliency, they used survey data and quantified those results for their study. This lack of data was a limitation of my study since I only had access to observational data, and limited access at that. Going out into these opioid-stricken regions and collecting survey and/or experimental data would possibly yield more detailed results and would be better additions to this sociological model.

Addiction is a serious issue. The opioid epidemic has been devastating to the US economy, and – more importantly – to people's lives. However we choose to tackle this crisis, we should consider how these solutions will affect people's lives and hopefully do more to study and treat addiction, or better yet, curb this risk altogether.

7 Appendix

Appendix Table 1: AIC of initial model, backward direction

	Df	Sum of Sq	RSS	AIC
<none>			88.561	-654.57
- frac_female	1	2.7498	91.311	-643.57
- unemp_rate	1	3.6658	92.227	-639.33
- average_income	1	5.0491	93.610	-633.01
- frac_health_insurance	1	8.5509	97.112	-617.40
- average_age	1	10.3877	98.948	-609.43

Step: AIC = -654.57
log(death_rate) ~ average_age + frac_female + frac_health_insurance
+ average_income + unemp_rate

Appendix Table 2: AIC of initial model, forward direction

	Df	Sum of Sq	RSS	AIC
<none>			88.561	-654.57
+ frac_social_w	1	0.301639	88.259	-654.02
+ frac_health_psych	1	0.012851	88.548	-652.63

Step: AIC=-654.57
log(death_rate) ~ frac_female + average_age + unemp_rate + frac_health_insurance
+ average_income

Appendix Table 3: AIC of final model, backward direction

	Df	Sum of Sq	RSS	AIC
<none>			54.987	-789.74
- health_care_costs	1	0.3652	55.352	-789.01
- association_rate	1	0.4958	55.483	-788.04
- violent_crime_rate	1	0.5283	55.515	-787.80
- income_inequality_ratio	1	0.5856	55.572	-787.37
- perc_african_american	1	0.6489	55.636	-786.91
- perc_frequent_physical_distress	1	0.6856	55.672	-786.63
- high_school_graduation_rate	1	0.7032	55.690	-786.50
- widowed	1	0.7034	55.690	-786.50
- frac_health_insurance	1	0.7323	55.719	-786.29
- homicide_rate	1	0.7871	55.774	-785.88
- never_married	1	1.1006	56.087	-783.57
- divorced	1	1.1074	56.094	-783.52
- perc_rural	1	1.1631	56.150	-783.12
- unemp_rate	1	1.2575	56.244	-782.42
- perc_frequent_mental_distress	1	1.4325	56.419	-781.14
- married	1	1.7308	56.718	-778.97
- average_age	1	2.4423	57.429	-773.83
- perc_insufficient_sleep	1	2.8481	57.835	-770.93
- perc_non_hispanic_white	1	8.3900	63.377	-733.23

Step: AIC=-789.74
log(death_rate) ~ average_age + frac_health_insurance + unemp_rate
+ high_school_graduation_rate + income_inequality_ratio + association_rate
+ violent_crime_rate + perc_frequent_physical_distress +
perc_frequent_mental_distress + perc_insufficient_sleep + health_care_costs +
homicide_rate

+ perc_african_american + perc_non_hispanic_white + perc_rural + married
+ widowed + divorced + never_married

Appendix Table 4: AIC of final model, forward direction

	Df	Sum of Sq	RSS	AIC
<none>			55.019	-787.50
+ frac_female	1	0.232061	54.787	-787.24
+ perc_african_american	1	0.162197	54.857	-786.72
+ perc_native_hawaiian_other_pacific_islander	1	0.123874	54.895	-786.43
+ some_college	1	0.110835	54.908	-786.33
+ percent_single_parent_households	1	0.073411	54.945	-786.05
+ perc_food_insecure	1	0.044484	54.974	-785.83
+ children_in_poverty	1	0.035462	54.983	-785.76
+ perc_american_indian_alaskan_native	1	0.024843	54.994	-785.69
+ residential_segregation_index_nonwhite_white	1	0.022840	54.996	-785.87
+ perc_not_proficient_in_english	1	0.016039	55.003	-785.62
+ average_income	1	0.013949	55.005	-785.60
+ pcp_rate	1	0.000199	55.019	-785.50
+ never_married	1	0.000119	55.019	-785.50
+ divorced	1	0.000100	55.019	-785.50

Step: AIC=-787.5
log(death_rate) ~ widowed + perc_hispanic + mentally_unhealthy_days
+ frac_health_insurance + health_care_costs + homicide_rate
+ perc_frequent_physical_distress + perc_asian + perc_insufficient_sleep
+ perc_non_hispanic_white + married + average_age + unemp_rate
+ perc_rural + separated + high_school_graduation_rate + association_rate
+ violent_crime_rate + income_inequality_ratio + perc_frequent_mental_distress

References

American Psychological Association. "Pursuing a Career in Health Psychology." http://www.apa.org/action/science/health/education-training.aspx.

Barnett, Jessica C., and Edward R. Berchick. "Health Insurance Coverage in the United States: 2016, Current Population Reports." U.S. Census Bureau. September 2017. https://www.census.gov/content/dam/Census/library/publications/2017/demo/p60-260.pdf

Centers for Disease Control and Prevention. "Drug Overdose Death Data." December 19, 2017. https://www.cdc.gov/drugoverdose/data/statedeaths.html.

——. "Prescribing Data." Centers for Disease Control and Prevention. August 30, 2017. https://www.cdc.gov/drugoverdose/data/prescribing.html.

——. "Understanding the Epidemic." August 30, 2017. https://www.cdc.gov/drugoverdose/epidemic/index.html.

Centers for Disease Control and Prevention, National Center for Health Statistics. Compressed Mortality File. 1999-2016 on CDC WONDER Online Database, released December 2017. Data are from the Compressed Mortality File 1999-2016 Series 20 No. 2V, 2017, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. Accessed at http://wonder.cdc.gov/cmf-icd10.html on Mar 25, 2018.

Council of Economic Advisers. 2017. "The Underestimated Cost of the Opioid Crisis." https://www.whitehouse.gov/sites/whitehouse.gov/files/images/The Underestimated Cost of the Opioid Crisis.pdf.

Dasgupta, Nabarun, Leo Beletsky, and Daniel Ciccarone. 2018. "Opioid Crisis: No Easy Fix to Its Social and Economic Determinants." *J Public Health* 108: 182–86. https://doi.org/10.2105/AJPH.2017.304187.

Hollingsworth, Alex, Christopher J. Ruhm, and Kosali Simon. 2017. "Macroeconomic Conditions and Opioid Abuse." *Journal of Health Economics* 56 (December): 222–33. https://doi.org/10.1016/j.jhealeco.2017.07.009.

Hughes, Evan. 2018. "The Opioid That Made a Fortune for Its Maker - and for Its Prescribers." The New York Times. The New York Times. May 2, 2018. https://www.nytimes.com/interactive/2018/05/02/magazine/money-issue-insys-opioids-kickbacks.html.

King, Nicholas B, Veronique Fraser, Constantina Boikos, Robin Richardson, and Sam Harper. 2014. "Determinants of Increased Opioid-Related Mortality in the United States and Canada, 1990-2013: A Systematic Review." *American Journal of Public Health* 104 (8): e32-42. https://doi.org/10.2105/AJPH.2014.301966.

Lipari, Rachel N., Ph.D., Matthew Williams, Ph.D., and Struther L. Van Horn, M.A. "Why Do Adults Misuse Prescription Drugs?" Substance Abuse and Mental Health Services Administration. July 27, 2017.

https://www.samhsa.gov/data/sites/default/files/report_3210/ShortReport-3210.html.

Markstrom, Carol A, Sheila K Marshall, and Robin J Tryon. 2000. "Resiliency, Social Support, and Coping in Rural Low-Income Appalachian Adolescents from Two Racial Groups." *Journal of Adolescence* 23: 693–703. https://doi.org/10.1006/jado.2000.0353.

Medina, Daniel A., Kate Snow, ML Flynn, and Eric Salzman. "A Generation at Risk: Children at Center of America's Opioid Crisis." NBC News. October 11, 2017. https://www.nbcnews.com/storyline/americas-heroin-epidemic/generation-risk-children-center-america-s-opioid-crisis-n806456..

National Institute of Drug Abuse. "How Does Heroin Use Affect Pregnant Women?" June 8, 2018. https://www.drugabuse.gov/publications/research-reports/heroin/how-does-heroin-abuse-affect-pregnant-women.

——. "Opioid Overdose Crisis." March 06, 2018. Accessed April 07, 2018. https://www.drugabuse.gov/drugs-abuse/opioids/opioid-overdose-crisis.

Noguchi, Yuki, writer. "Parents Lose Their Daughter And Their Life Savings To Opioids." Transcript. In *Morning Edition*. National Public Radio. April 19, 2018.

Portenoy, R.K., and K.M. Foley. 1986. "Chronic Use of Opioid Analgesics in Non-Malignant Pain: Report of 38 Cases." *Pain* 25 (2): 171–86. http://www.ncbi.nlm.nih.gov/pubmed/2873550.

Ruggles, Stevent, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. IPUMS USA: Version 8.0. Minneapolis, MN: IPUMS, 2018. https://doi.org/10.18128/D010.V8.0

Susan Collins United States Senator for Maine. "Bill to Help Grandparents Raising Grandchildren Due to Opioid Crisis Passes Senate." March 22, 2018. https://www.collins.senate.gov/newsroom/bill-help-grandparents-raising-grandchildren-due-opioid-crisis-passes-senate.

U.S. Bureau of Economic Analysis. "Per capital real GDP by state (chained 2009 dollars)." 2017. https://www.bea.gov/iTable/iTable.cfm?reqid=70&step=10&isuri=1&7003=200&7035=-1&7004=sic&7005=1&7006=xx&7036=-1&7001=1200&7002=1&7090=70&7007=-1&7093=levels#reqid=70&step=10&isuri=1&7003=1000&7035=-1&7004=naics&7005=1&7006=xx&7036=-1&7001=11000&7002=1&7090=70&7007=-1&7093=levels

U.S. Bureau of Labor Statistics. "Occupational Outlook Handbook: Social Workers." July 11, 2018. https://www.bls.gov/ooh/community-and-social-service/social-workers.htm.

——. "Regional and State Unemployment – 2016 Annual Averages." February 28, 2017. https://www.bls.gov/news.release/archives/srgune_02282017.pdf

U.S. Department of Health and Human Services. "What Is the U.S. Opioid Epidemic?" March 6, 2018. Accessed June 08, 2018. https://www.hhs.gov/opioids/about-the-epidemic/index.html.

Vance, J. D. *Hillbilly Elegy: A Memoir of a Family and Culture in Crisis.* New York, NY: Harper, 2016.

Volkow, Nora. "Addressing the Opioid Crisis Means Confronting Socioeconomic Disparities." National Institute of Drug Abuse. October 25, 2017. https://www.drugabuse.gov/about-nida/noras-blog/2017/10/addressing-opioid-crisis-means-confronting-socioeconomic-disparities.

Wood, David L, Mary I Dobbins, Marian F Earls, Andrew S Garner, Laura Mcguinn, John Pascoe, Behavioral Pediatrics, Benjamin S Siegel, and Jack P Shonkoff. 2011. "The Lifelong Effects of Early Childhood Adversity and Toxic Stress." *Pediatrics* 129 (1): 232–46. https://doi.org/10.1542/peds.2011-2663.

Zee, Art Van. 2009. "The Promotion and Marketing of OxyContin: Commercial Triumph, Public Health Tragedy." *American Journal of Public Health* 99 (2): 221–27. https://doi.org/10.2105/AJPH.2007.131714.

Zoorob, Michael J., and Jason L. Salemi. 2017. "Bowling Alone, Dying Together: The Role of Social Capital in Mitigating the Drug Overdose Epidemic in the United States." *Drug and Alcohol Dependence* 173 (April): 1–9. https://doi.org/10.1016/j.drugalcdep.2016.12.011.