Assessing the Impacts of an Aging Population on Rising Healthcare and Pharmaceutical Expenditures within the United States

Rahul Sharma

Professor Gillian D. Saunders-Schmidler, Faculty Advisor Professor Grace Kim, Economic & Seminar Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.

Duke University Durham, North Carolina 2018

Acknowledgements

First and foremost, I would like to thank my advisors Dr. Saunders-Schmidler and Dr. Kim for their time and patience throughout the process of writing this thesis. Their support, kindness, and knowledge have been crucial and granted me the opportunity to explore the topic presented here.

Additionally, I would like to thank my classmates from my Honors Thesis Seminars. Your constructive criticism and helpful suggestions were instrumental in the construction of this final product.

Abstract

This paper studies the impact of aging on rising healthcare and pharmaceutical expenditures in the United States with the goal of contextualizing the future burden of public health insurance on the government. Precedent literature has focused on international panels of multiple countries and hasn't identified significant correlation between age and healthcare expenditures. This paper presents a novel approach of identifying this correlation by using a US sample population to determine if age impacts an individual's consumption of healthcare services and goods. Results suggest that age has a significant impact on healthcare and pharmaceutical expenditures across private and public insurance.

JEL classification: H51; H53; I12; I13; I18; I38

Keywords: Health Economics; Insurance; Welfare: Medicare, Medicaid; Healthcare Expenditures; Government; Public Health

1. Introduction

While the link between aging and increasing healthcare and pharmaceutical expenditures seems relatively self-evident, economists have struggled to identify a clear, causal link between population aging and increasing healthcare expenditures. This is likely due to a variety of different factors that can contribute to an individual's necessity to spend on healthcare including demographic factors (race, gender, geographic location, etc.) as well as genetic susceptibility, presence of certain chronic diseases, social and cultural influences, and many more. Understanding the impact of aging on healthcare expenditures has important policy implications, especially in the United States, where total healthcare expenditures and pharmaceutical expenditures have experienced rapid, continuous growth for over 40 years and the population is projected to face significant aging in the near future. If age is identified to be a key driver of healthcare expenditures, then there are multiple policy options that must be identified such as increasing the social security tax rate or improving preventative care options for the elderly.

Thus, this study aims to determine the link between aging and increasing healthcare and pharmaceutical expenditures in the United States in an effort to bring awareness to politicians and policymakers while encouraging others to conduct further research in this area. On net, the results of this study will likely help address a highly partisan issue and determine the validity of pervasive concerns surrounding the ability of public insurance programs to sustain an aging US population.

Since 2000, prescription drug expenditures in the United States have nearly tripled to 360 billion dollars per year ("Prescription Drug Expenditure U.S. 1960-2018", 2017) with the US outpacing every other country in the world in per capita drug spend by nearly \$1,000. Some speculate that surges in healthcare expenditures are "largely driven by brand-name drug prices that have been increasing in recent years at rates far beyond the consumer price index."

(Kesselheim, Avorn, & Sarpatwari, 2016) Individuals in the US pay 50% to 60% more for brand name pharmaceuticals compared to individuals in European countries such as France or Germany. (Paris, 2014) This is due to the US having laxer regulation on pharmaceutical pricing and insurers, both public and private, having little to no say in the pricing mechanisms of the pharmaceutical industry.

This leads to an unfair burden on American citizens as Sheldon Lubar and Timothy Smunt (2017) from the University of Wisconsin have found, "Americans, in fact, are heavily subsidizing the healthcare costs of the rest of the world because of non-competitive and unfair drug pricing." (Lubar & Smunt, 2017) However, competition and deep market penetration of offpatent generic pharmaceuticals have caused generics to compose greater than 80% of the market, but only 28% of the pharmaceutical spend. (Paris, 2014) As a result, individuals in the United States are being overcharged for patented, brand-name drugs and therapies, especially those that exist without competition. What is equally concerning is that rapidly increasing pharmaceutical expenditures will place greater financial burdens on already cash-strapped public insurance programs as shown in Figure 1¹ where Medicare expenditures on pharmaceuticals have increased by nearly 100% from 2006 to 2016. With current tax legislation set to cut nearly \$1.5 trillion from Medicare and Medicaid over the next decade, public insurance plans are very likely to go bankrupt if current trends continue, leading to high levels of uninsured individuals across the United States. (Morse, 2017)

¹ Data taken from Center for Medicare and Medicaid Service (CMS) National Health Estimates ("National Health Expenditure Accounts," 2018)



Figure 1: United States Pharmaceutical Expenditures by Source of Funds (2000-2016)

Additionally, as the costs of pharmaceuticals continues to increase, there is a strong likelihood that medical non-adherence will continue to rise (McGuire & Iuga, 2014) and thus will lead to increased numbers of last-resort emergency room visits and overall poorer health, both of which are extremely economically detrimental. Figure 2² shows how an estimated 15 million individuals (left axis) in 2015 did not fill out a prescription or follow a doctor-recommended treatment plan, corresponding to approximately 5% of the US population over the age of 2 (right axis). (Kennedy & Wood, 2016) This has particularly problematic economic consequences as non-adherence may lead to \$100-\$300 billion in avoidable medical costs per year which accounts for 3-10% of total US healthcare expenditures. (McGuire & Iuga, 2014) Furthermore, non-adherence is a continuous cycle in which individuals do not adhere to proper treatment regimes, as a result get sicker, thus requiring more expensive treatments. These costs get pushed onto the patient, which further encourages more non-adherence.

² Graph drawn from Kennedy & Wood, 2016

The problem of medical non-adherence is likely to be amplified in the elderly population as a result of lower employment and less disposable income with which to purchase the necessary treatment options. It is important to note that medical non-adherence is not limited to just pharmaceuticals. Rather it is defined as the process of a patient not following a prescribed treatment regime as recommended by a doctor. (Hugtenburg, Vervloet, Dijk, Timmers, & Elders, 2013) This can include medical treatments, visits to specialists, and other forms of healthcare service utilization recommended by a primary care physician. Medical non-adherence also has long term negative economic impacts on the healthcare system as individuals who do not adhere to treatment regimens are more likely to visit the Emergency Room. (Blanchard, Madden, Ross-Degnan, Gresenz, & Soumerai, 2013) This is problematic because non-adherent individuals are likely unable to pay for the hospital services after treatment and thus either go into debt or the charges are written off by the hospital. While this study does not directly evaluate the impact on health outcomes in the context of increased healthcare expenditures and aging, economic indicators of medical non-adherence might potentially be observed and are important to note.



Figure 2: Rates of Medical Non-adherence in the United States (1999-2015)

However, rising expenditures in healthcare are not simply limited to pharmaceuticals. In fact, US healthcare expenditures, on a whole, have experienced a similar rate of growth to pharmaceutical expenditures with an average 4.8% annual increase from 2000 to 2016 as shown in Figure 3³. ('National Health Expenditure Projections', 2016) Additionally, the growth rate is expected to increase to 5.6% annually by 2025 and account for just under 20% of the US GDP. ("National Health Expenditure Accounts," 2018) Joseph Dieleman (2017), researcher at the University of Washington, identified the three driving factors behind the continual increase in healthcare expenditures.

First, he finds that the costs of services have steadily been increasing, stating "Price and the variety and complexity of services is the largest driver of health care spending increases." (Dieleman et. al, 2017) In other words, the increasing cost of purchasing specialized equipment and necessary investment into training specialist means that the costs of effective treatments are steadily increasing. Second, the trend in medicine is towards the idea of personalized medicine, whereby treatments are tailored to the patient as opposed to generic treatment regimens. Highly differentiated goods and services lead to increased costs as companies need to maintain acceptable profit margins on goods that will have low volume of utilization. Third, individuals are spending more on specific conditions. Along with the idea of more personalized options for treatment, people are spending less on treating diseases with common remedies and are spending more on treatments that are patient-specific such as chronic diseases like arthritis, which require highly specialized, expensive treatments. (Dieleman et. al, 2017)

³ Data taken from Center for Medicare and Medicaid Service (CMS) National Health Estimates, ("National Health Expenditure Accounts," 2018)



Figure 3: United States Total Healthcare Expenditures by Source of Funds (2000-2016)

It seems clear that healthcare expenditures are increasing at an alarming rate. Politicians and research alike must pay attention to these trends and attempt to identify the reasons behind this growth and identify ways in which policy can be leveraged to prepare for the inevitable critical point at which expenditures will overwhelm the current system, especially of public insurance programs. Additionally, it is important to note which services the elderly are using disproportionately in order to effectively target spending on certain services. As shown in Figure 4⁴, the elderly are responsible for a majority of hospital and nursing care expenditures in the United States. While these results are not entirely surprising and this study will not evaluate spending on different healthcare services, they showcase a potential for future study on what specific services the elderly prefer to spend on and how targeted policy mechanisms can lower prices on insurers and patients as well.

⁴ Data taken from Center for Medicare and Medicaid Service (CMS) National Health Estimates, ("National Health Expenditure Accounts," 2018)



Figure 4: Total US Healthcare Expenditure Breakdown by Service/Good by Age Group in 2012

Potentially correlated with increasing expenditures is likely the aging population in the United States. The Census Bureau has identified the amount of individuals who will be older than 65 years is set to nearly double from 2012 to 2050, largely due to the baby boomer generation aging. (Ortman, Velkoff, & Hogan, 2014) More specifically, the Census has identified that the percent of the population that is over 65 has grown approximately 2.5% from 2012 to 2017 alone. (Ortman et. al, 2014) (Table 1⁵) Additionally, the NIH has identified that the global life expectancy is projected to increase approximately 13 years over the span of 2012 and 2050. (Senthilingam, 2017) Both these trends are particularly informative as the elderly are the most likely subset of the population to heavily utilize healthcare goods and services and increased life expectancy would prolong use of these goods and services. Furthermore, the Census has identified that key drivers of mortality such as smoking and other medically harmful habits have significantly decreased over the past decade. (Ortman et. al, 2014) Currently, the NIH has determined that a typical 65+ year old individual faces an average of \$12,000 annually

⁵ Data taken from Center for Medicare and Medicaid Service (CMS) National Health Estimates, ("National Health Expenditure Accounts," 2018)

in uncovered (out-of-pocket) prescription drug costs. (Knickman & Snell, 2002) The Congressional Budget Office conservatively estimates that this number could increase roughly 2.6% per year. (Cook, 2013) This is likely to have strong impacts on public and private insurers alike. In fact, in 2016 the CBO found that 64.3% of health expenditures were covered by the public sector. (Himmelstein & Woolhandler, 2016) Thus the increasing population of elderly, if positively correlated with aging, can have severe implications in regards to the national budget. Especially in the current political climate where social services and healthcare are at the forefront of policy debate, gaining a better understanding of the future monetary needs of our public healthcare beneficiaries is extremely important.

2014-2060				
	Year	Under 18	18-64	65+
	2014	23%	62%	15%
	2020	22%	61%	17%
	2030	21%	58%	21%
	2040	21%	58%	22%
	2050	20%	58%	22%
	2060	20%	57%	24%

Table 1: Percent of Population by Age Group from 2014-2060

Percent of Population by Age Group from

In this study, I will attempt to utilize a novel approach of identifying the impact of aging on increasing health expenditures by focusing on a subset of the US population as opposed to an international panel of multiple countries. The United States represents a unique case study due to the structure of the insurance market and the freedom of choice of healthcare services and goods. Section 2 provides a review relevant literature that has been conducted on this topic, specifically identify shortcomings of current research and identify the specific niche I hope to explore. Section 3 provides a theoretical framework adapted from researchers to help guide the construction of this study. Section 4 evaluates the use of the 2015 Medical Expenditure Panel

Survey (MEPS) as the data source for this studies as well as its benefits and limitations and the construction of the dataset for this study. Section 5 provides the empirical specification being utilized in this study. Section 6 presents the results of the study and a discussion of the relationships identified. Section 7 provides insight into the impact of these results and provide a direction to move forward.

2. Literature Review

2.1 An Aging United States Population

The United States Administration on Aging suggests that by 2030, greater than 1/5th of the US population will be over the age of 65. ("*Fueled by Aging Baby Boomers, Nation's Older Population to Nearly Double in the Next 20 Years, Census Bureau Reports*", 2014) This aging is not localized to any one part of the country, rather the median age is growing almost universally across all of the United States, both states and territories. ("2010 Census Shows Nation's Population Is Aging," 2011) The reasoning for this population aging is pretty well-understood by most researchers. This decade marks the transition of the baby-boomers into this segment of the population as well as the fact that individuals are having fewer kids, thus decreasing the number of younger individuals and reducing the population growth rate. Thus, not only is the number of individuals over the age of 65 increasing, the percent of the population that is above this age threshold is also increasing. This can lead to a multitude of problems for the United States such as stunted economic growth and increased dependencies on social services.

The RAND Corporation (2016) found that while older individuals are less likely to be employed, there is an increasing trend whereby older individuals are retiring later. (Maestas, Mullen, & Powell, 2016) This aging workforce is leading to less per-work efficiency and can be directly correlated to slower economic growth. (Maestas et. al, 2016) This is extremely problematic because elderly workers could be crowding out younger, more efficient workers and layoffs tend to simply increase dependencies on stipends and other government-supported programs including unemployment benefits. Researchers from the NIH (2004) highlight another critical problem: caring for the aging baby boomers. They found that "Even at current expenditure levels, there is a significant amount of unmet need for long-term care among the frail [and elderly], and no foreseeable end in upward pressure on per diem service costs." (Harrington & Estes, 2008) This study highlights that even in a situation where healthcare expenditures are growing, needs are still not being met. This is compounded by the crowding-out of young labor mentioned above, leaving a smaller population to support the elderly.

Finally, this increasing burden of the elderly is reducing the revenue that is flowing into the government. In a study aimed at determining the impact of aging on tax revenues, the Kansas Federal Bank (2011) determined that with the current projections state tax revenue alone will likely be 1.1% less or show a loss of \$8.1 million per state by 2030. (Felix & Watkins, 2011) This highlights the cyclical nature of the problem. Insofar as the number of individuals that are relying on the system increases and the number of individuals that are paying into the system is decreasing, increasing healthcare expenditures are disproportionately burdening governmentbacked, public insurance systems which can have catastrophic impacts if left unchecked.

2.2 Insurance Systems in the United States

The United States is very unique in how the insurance market is setup. Many countries around the world operate under a universal healthcare system where the government is the major (or sole) provider of health insurance for all of the citizens. The United States has a multi-payer system that consists of multiple private insurers and the government as a provider of public insurance. In addition, insurance in the EU tends to be a lot less restrictive in terms of options for care. Specifically, individuals are able to walk into any health practice and receive service while most private insurers in the United States restrict the specific practices that can be used, specifically by practice with which the insurer has relationships with. (Gold, 2011) This also extends to pharmacies as well. Furthermore, the presence of a marketplace of insurers in the United States allows for a general level of competition to help decrease the costs between insurers as opposed to a single offer. That being said, the price of European healthcare tends to be less costly on net due to the ability of the government to subsidize healthcare and place direct

regulations on pharmaceutical companies. Additionally, US regulations on the pharmaceutical market differ vastly from European regulations in their subsidies for research and development (R&D) as well as allowing pharmaceutical companies to actively market their brand-name drugs directly to the public, a practice that is not allowed in the EU. On a whole, "[European] government systems also are the only large drug buyers in most of these countries, giving them substantial negotiating power." (Whalen, 2015)

The U.S. market, by contrast, is highly fragmented, with bill payers ranging from employers to insurance companies to federal and state governments." (Whalen, 2015) However, there is a possibility that the US can adopt the best practices of many "natural" policy experiments that have occurred all throughout Europe to help place constraints on pharmaceutical prices and lead to an overall increase in health. Many critics of the universal healthcare system and European practices claim that the inherent competition that is caused by a privatized insurance and pharmaceutical system with minimal government interference will lead to competitive pricing as natural competition with a capitalist system should do. However, this ultimately has failed to prove true as reports from the Human Development Records Office show that European countries consistently have a much higher Health Index than the United States. (*Health Index*, 2013) This is uniquely important in the realm of economics due to the fact that recent reports have found that healthier areas tend to have faster job and income growth. (Frenk, 2004) It is in the best interests of the United States to place a strong emphasis on improving overall health to ensure future economic growth.

Currently, US insurance providers can be broadly classified into public and private. Public insurers are usually a form of government insurance, funded by either state or federal governments, whereas private insurers are private companies from which different insurance plans can be purchased by an individual. There are various debates about how differences in public vs. private insurance can impact health outcomes and how it impacts healthcare providers (doctors, hospitals, etc.) however this is outside the scope of this study. In this study, private insurance as well as the 3 biggest public insurers: Medicare, Medicaid/CHIP, and Tricare/CHAMPVA, will be included in the regression analysis. Understanding the structure and segment of the population that each public insurer serves will help identify if aging does indeed correlate with higher expenditures within each of these insurance schemes.

While public insurance programs are often cheaper than private insurance plans, this does not necessarily mean that individuals are more likely to spend under public insurance plans as opposed to private ones. In fact, research supports the opposite conclusion. The Center for Medicare and Medicaid Services (2015) identified that both Medicare and Medicaid rate of increase in expenditures was lower than that of private insurance and was predicted to remain this way through 2023⁶. (Altman, 2015) The study determined that from 2007 to 2023, private insurance costs would increase by 104% while Medicare and Medicaid costs would only increase by 63%. (Altman, 2015)

However individual research on Medicare and Medicaid produced mixed results. A study using data from the 2015 Consumer Expenditure survey found that households covered under Medicare spent almost 2x more of their household spending on health-related expenses as compared to not Medicare households. (Cubanski, Orgera, Damico, & Neuman, 2018) While this is not a direct comparison between public and private insurance plans, it still represents an interesting trend, especially as Medicare already consumes a substantial portion of the national budget. There are likely two reasons for this trend. First, since only individuals with disabilities and individuals over the age of 65 are eligible for Medicare, they likely need to spend more on

⁶ Data after 2023 was not available

healthcare than individuals who aren't eligible for Medicare. Second, individuals who are eligible for Medicare are likely to have a lower income (disabilities & old age) and thus spend a greater percent of their income on health without necessarily spending more. The study goes on to find that older households covered under Medicare spend more than younger households covered under Medicare, indicating a potential correlation between age and increased expenditures. (Cubanski et. al. 2018)

Conversely, research using Medical Expenditure Panel Survey (MEPS) data from 2005 has found that families covered under Medicaid had 26-37% less expenditures than families in comparable socioeconomic situations covered under private insurance. (Clemens-Cope, Holahan, & Garfield, 2016) These results were further corroborated by a two-way study in which researchers found that households with Medicaid spent more when exposed to private plans and vice versa. (Hadley & Holahan, 2003) The results of these studies are particularly curious as private insurance is associated with lower co-pays and premiums for services and thus was expected to experience higher utilization than private insurance. However, it is possible to explain these results with two hypotheses. The first is that acceptance of Medicaid by service providers is considerably limited compared to private plans and thus as individuals have less access to affordable care, they are less likely to spend. The second is that Medicaid specifically covers low-income individuals and thus are, on net, have a lower propensity to spend.

No research has exclusively examined the impact of Tricare coverage on increased healthcare expenditures, likely due to the fact that it is the smallest of the three public insurance programs. These mixed results in studies discussed above represent a majority of the minimal literature that evaluates differentiated spending by age and insurance provider. This gap in the literature can be addressed by the results of this study and thus better understand how different insurance plans impact spending behaviors on healthcare by individuals.

Medicare is a government insurance plan which, as of 2015, has 55 million beneficiaries. (Kaiser Family Foundation, 2017) Medicare is available to American Citizens (of at least 5 years) who are over the age of 65 and/or qualify for disability benefits from the Social Security Administration. Medicare is the primary government health provider for the elderly, and is the biggest source of government healthcare expenditures. This is because individuals over the age of 65 are unlikely to receive healthcare benefits from their employers, as they have likely retired, and private insurance is likely too expensive to be a feasible option. Additionally, Medicare funding comes, for the most part, from payroll and income taxes that are classified as Social Security Tax. Since everyone is required to pay these taxes, individuals are likely to want to reap the benefits of the program when they are eligible. (AARP, 2011)

Medicare is split into four parts (A,B,C,D) which encompass all the critical parts of insurance including hospital insurance, medical insurance, and prescription insurance. (AARP, 2011) It is important to note that while Medicare is a government healthcare insurer, it is not completely free. Beneficiaries still must pay copays, fees, and out-of-pocket expenditures, however by participating in this plan, they have access to subsidized care where the government pays for a substantial part of the service or good being utilized. Within the context of this study, it is important to note that a majority of the individuals covered under Medicare are already over the age of 65 and thus are in the oldest age group that is utilized in this study. As a result, if coverage under Medicare correlates strongly with healthcare expenditures, it is likely due to the fact that the average age of an individual covered under Medicare is high. However, with the exception of work done by Cubanksi et. al (2018), limited literature on the correlation between aging Medicare beneficiaries and increased expenditures exist.

Medicaid is a joint federal-state program designed to provide health insurance to low income individuals, specifically women & children. CHIP is an offshoot of the program designed

specifically for children from low income families who do not qualify for Medicaid coverage. Within Medicaid, states are responsible for the actual administration of service, while the federal government provides at least half the funding and sets the standards of eligibility (Rudowitz, 2016). Medicaid and CHIP combine to have approximately 67 million beneficiaries, a very big problem since Medicaid funding was one of the first budgets to be slashed by both state and federal governments during the recent recession and has yet to recover to post-recession levels. (Rudowitz, 2016) This is further complicated by the current political climate in which Congress seems likely to continue the trend of increasing cuts to social service programs such as Medicaid. (Matthews, 2017) While Medicaid provides insurance to those who are in need of it, multiple cost restrictions within the program and limitations on acceptance by healthcare providers thus not providing adequate coverage for a segment of the population at the highest risk for diseases and illness.

As a result, many individuals attempt to access healthcare through an employer or more recently through the insurance marketplace setup in the Affordable Care Act, leading to few individuals who simply have only Medicaid. Medicaid presents an interesting use case in terms of tying it to healthcare expenditures because of multiple factors that could compound this direct effect. More specifically, Medicaid has more limited coverage as compared to Medicare or private plans and simultaneously is accepted by fewer healthcare providers, all of which could potentially drive down total expenditures by Medicaid beneficiaries and decrease quality of care. Within the context of this regression, it will be interesting to compare Medicaid to other public insurance plans as well as private insurance to determine which program is more strongly associated with increased healthcare expenditures, something which hasn't been done in previous literature.

Tricare is a federal healthcare program that is run by the government, specifically for individuals who are a part of the armed forces and their families and allows them access to services both in the VA system and the civilian hospital system. ("Health Care Benefits for Dependents (CHAMPVA)," n.d.) CHAMPVA is a healthcare system that is run specifically through the VA hospital system and provides insurance to disabled veterans and their families. ("CHAMPVA", 2014) Both these programs were conceived given that many veterans returning from active duty, as well as their family, required access to healthcare which they could not afford on a government salary. As with most government employees, the government insurance program was intended to substantially reduce the cost of necessary treatments and other expenditures, both for individuals who had access to a VA and those who only could access a civilian hospital. Tricare covers substantially more than Medicaid and Medicare due to the smaller number of individuals who are beneficiaries as well as the fact that they are employed by the government and as a result it is a form of employer healthcare.

CHAMPVA is run specifically through the VA system and was offered as an addendum to Tricare coverage as there are many limitations on how veterans can access their benefits. Its specific goal was to target disabled veterans and family members who needed access to higher levels of care which would be too expensive at a civilian hospital. Thus, Tricare & CHAMPVA were designed as a full coverage insurance plan and a specialized coverage plan for individuals who needed access to extra healthcare services without making costs unsustainable, respectively. This study will be one of the first to evaluate the impact of Tricare insurance coverage on increasing healthcare expenditure. Minimal literature on the subject currently exists and thus this study will seek to provide novel insights into how Tricare beneficiaries spend compared to other public insurance plans as well as private plans.

This study is unique in its specific examination of different insurance plans and understanding the impact of each plan on healthcare expenditures and, more specifically, the impact of aging within each of these insurance plans of healthcare expenditures. By providing insight into how beneficiaries of different health insurance plans spend more or less relative to each other and identifying the impact of aging within each plan, this study can begin to identify specific causes of healthcare expenditure increases and help insurers construct a system which minimizes expenditures without sacrificing quality of care.

2.3 Impact of Aging on Increased Health Expenditures

The impact of changing demographic factors in the United States and countries around the world is hotly debated due to conflicting results and a variety of confounding factors that can influence increasing healthcare expenditures. Most researchers have concluded that aging has minimal impact on increasing healthcare expenditures. Researchers postulate multiple other reasons for increases in healthcare expenditures such as increasing population size, inflation, increasing GNP, technological advancements, unhealthier lifestyles, and unnecessary expenditures. According to the Center for Economic Policy and Research (2007), studies conducted in Germany & the Netherlands have found that aging, which nearly mimics the aging found in the United States, only explains 0.5-0.7% of annual health expenditure growth. (Steinmann, Telser, & Zweifel, 2007) Other researchers have used the United Kingdom and Canada as proxies for the United States and have identified aging as only responsible for 2-14% of the increase in healthcare expenditures. (Meijer, Wouterse, Polder, & Koopmanschap, 2013) Researchers identified aging in Sweden to account for just 13% of the total increase in healthcare expenditures. (Gerdtham, 1993) Comparatively, researchers identify increases in price of services and number of services rendered as responsible for more than 50% of increases in

services. (Dieleman et. al, 2017) Others identify GNP and population growth as responsible for as much as 65% of the increase in expenditures. (Getzen, 1992)

However, other literature that evaluates aging and healthcare expenditures within the United States had identified that aging is responsible for, on average, 23% of the increase in healthcare expenditures. (Gregersen, 2013) Additionally, research that evaluated an entire panel of OECD countries identifies a red-herring effect in how many of these studies are conducted whereby not accounting for aging caused for an underestimation of health expenditures in all models of greater than 30%, however when it is included in regressions, aging can only be accurately identified as having on average a 1.5% on health expenditures on a whole. (Getzen, 1992) However, even these minimized impacts are refuted by other researchers. Researchers from the University of Adelaide (2002), using data from the United Kingdom, found that it is nearly impossible to isolate the impact of aging on healthcare expenditures and that the rapid increase in inflation and GNP of countries obscures any minimal impact from socio-demographic factors. (Seshamani & Gray, 2002) Of particular note is the limited studies that specifically identify the impact of aging on healthcare expenditures within the United States. Gregersen (2013) and Schenider & Guralnik (1990) remain the only research into aging in the United States and both papers conclude that aging retains a statistically significant impact on increasing healthcare expenditures. However, papers that focus on countries in Europe or use a global panel of countries are unlikely to corroborate those findings.

Many of the studies that attempt to identify the impact of aging on increasing health expenditures do so from a global perspective and conduct a comparative analysis of countries around the world. However, in doing so these studies are not able to capture the impact of differing insurers and sociodemographic factors, both of which can have impacts on health expenditures. Studies that have used country-specific data have been predominantly focused on

European countries, specifically the UK, Sweden, Germany and the Netherlands. These studies are not necessarily applicable to the United States because, based on OECD data in 2002, the United States is an extreme outlier compared in international comparisons of healthcare expenditures, whereby they spend more than 50% more per capita on health-related expenses than the next highest country. (Anderson, Reinhardt, Hussey, & Petrosyan, 2003) Additionally, as discussed earlier, the US multi-payer health insurance system is vastly different from other European country single-payer systems. Examining US national data, specifically responses of households in a random sample around the United States, removes the confounding factors of GNP, growing populations, and increasing prices which are constant amongst an entire population. If we were to assume that these factors are unlikely to change and are endemic to the growth of any country, a direct observation of the population will give policy makers a greater insight into understanding aging specifically as it impacts increases in US healthcare expenditures and within the specific insurance market that is unique to the US.

This study aims to address the lack of research on understanding the impacts of aging on healthcare expenditures specifically within the United States. We hope to do this in three unique ways. First, this study analyzes healthcare expenditures as a consumption of a good/service governed by the behavior of individuals as a function of their specific circumstances. This decision making is likely based on a variety of factors including age, education, income, race, gender, and marital status among many others. (Ubel, 2010) Thus in this study, I hope to take a new look at understanding the impact of an aging population on increasing health expenditures in the United States in a manner that has been seldom done in the literature before. By conducting an analysis on a dataset that represents only the US population and integrates the factors that are likely to influence their decision making on healthcare expenditures, this study is able to specifically determine how the aging in the US population will affect US healthcare expenditures

in the near future. While spending on healthcare often carries a greater consequence than spending on other goods such as entertainment, healthcare spending is not mandated and thus is likely to be governed by similar cost-benefit analyses that individuals use to determine their consumption of other goods.

Second, this study attempts to categorize aging in a different manner. While most research currently examines changes in median age, this is unlikely to yield significant results as the median age is likely to change very slow year-over-year. Thus, this study aims to identify the impact on an individual's expenditures as they age and move between census groups⁷. This use of census groups was identified by Di Matteo and is addressed in Section 3. (Di Matteo, 2005) Since current population estimates (Table 1) showcase population aging as the percent of the population greater than 65 increasing, this method of measurement is likely to yield higher degrees of explanatory power on the impact of aging on healthcare expenditures. This is further enhanced by the use of a recent data source, the 2015 Medical Expenditure Panel Survey, which provides a more accurate base from which projections can be determined and is more likely to accurately represent the current population. Of particular note is studies that have used this manner of conducting an analysis have found a "strong correlation between population's age and per capita health expenditures." (Lopreite & Mauro, 2017) Stefan Felder, researcher at the University of Sydney (2017), clarifies this in his findings that "the share of the population older than 65 turns out to be a significant and positive determinant of health expenditure as a share of GDP." (Piabuo & Tieguhong, 2017)

Third, this research specifically attempts to analyze the impact of the United States' unique insurance network and identify the impact of aging on beneficiaries of both public and

⁷ Explained in Data Section: (10-18, 19-44, 45-64, 65+)

private insurance. Specifically in regards to US policy studies agree that "the increasing elderly share explains most of the growth of the public sector size, particularly because this age group increases functions of public expenditure mainly benefiting their group: social welfare and health." (Sanz & Velázquez, 2007) In other words, government and public insurers are considerably more susceptible to the impacts of aging on expenditures. With minimal previous literature differentiating healthcare expenditures by insurance type, this research aims to provide new insights into understanding spending patterns between different insurance groups and whether aging-related impacts on healthcare expenditures are magnified with public beneficiaries as opposed to private beneficiaries.

3. Theoretical Model

Specifically observing the United States healthcare system as the focus of this study establishes a simplistic consumption function for understanding how individuals make choices related to healthcare expenditures.

$$h = f(Y, A, z)$$

Equation 1: Model Drawn from Di Matteo

In this model developed by Di Matteo (2005), healthcare expenditures (h) are a function of income (Y), age (A), and a multitude of sociodemographic variables (z). (Di Matteo, 2005) In doing so this model establishes a set group of factors that impact an individual's behavior of consumption of healthcare goods and services. In general, a consumption function is defined as:

$$C = \beta * Y + \alpha$$

Equation 2: General Consumption Function

where α is autonomous consumption (consumption when income is at 0), β is the marginal propensity to consume (MPC), and Y is income. In this specific model, we assert that consumption of healthcare services and goods is predicated on multiple factors thus we can modify the consumption function in the following manner.

$$C = \beta * Y + \varphi * A + \delta * z$$

Equation 3: Updated Consumption Function including Age & Sociodemographic Factors

In this model, β , ϕ , and δ are the marginal propensities to consume associated with income, age, and other sociodemographic factors respectively. Given that individuals with no income are unable to spend on anything (including healthcare) we assume the value of α to be 0. This linear consumption model serves as the basis of the OLS regression model used in this study (See Empirical Specification). The coefficients are indicative of the trends of how an individual will spend on healthcare services and goods as the variables increases or decreases. For example, ϕ represents the amount by which healthcare expenditures will increase (or decrease) as an individual moves between census age groups. This can also be loosely interpreted as the percent of the total population within a given census age group when applying this equation to an entire population. Mathematically, the value of ϕ represents:

$$\varphi = \frac{\Delta C}{\Delta A}$$

Equation 4: Mathematical Representation of the MPC of the Age term

where Y, A, and z are completely independent of one another. Thus, if this model is able to identify the value of ϕ in a sample population, it can be extrapolated to the entire US population. However, this method of analysis assumes that e remains uniform across a population which is not realistic in practice (explained below). The value of e can be impacted by a variety of factors such as perceived health status, success rate of a specific drug or procedure, etc. all of which cannot be included or measured and included in this model. The assumption of independence of Y, A, and z does not necessarily hold true in practice and might reduce the explanatory power of the model, however inclusion of a number of sociodemographic factors is likely to improve robustness and fit and generally conclusively signal a trend (positive or negative).

The utilization of this model is novel insofar as it analyzes healthcare expenditures at the individual level, however doing so requires a key assumption. The model assumes that healthcare expenditures follow a linear consumption model and thus are linearly correlated to all of the variables that are being studied. However, Okunade and Sraratdecha (2000) note that "non-linear functional form may be more appropriate for modelling OECD health data." (Okunade & Sraratdecha, 2009) As most literature in this field have constructed linear regressions, this study will as well. However, this does highlight potential further research and experimentation to identify what models best capture the relationship between aging and healthcare expenditures.

Studies that have attempted to isolate the impacts of age on health expenditures have yielded conflicting results, specifically due to the fact that there are "changing health expectations and demands across population cohorts, the effect of new techniques and technologies, demographic uncertainty and even age related changes on the overall cost of health services." (Di Matteo, 2005) This model addresses the problem by assuming that factors such as increased costs, technological advancement, demands for services, and GDP/GNP is uniform across the population and impacts everyone identically. Obviously, this does not hold true in practice. For example, increased prices will likely impact wealthier individuals less than lower-income individuals. Since these factors and their impacts have been addressed heavily in previous literature, they have been excluded from the current model. This is both beneficial and harmful as it allows for the impact of other individual-specific factors to be analyzed but not including these factors potentially reduces the robustness of the model.

This model categorizes an individual's healthcare expenditures as any other expenditure an individual would make; that is to say that individuals make choices to spend money on pharmaceuticals, hospital services, etc. based on the conditions they find themselves in and the value that they place on each condition. While the model is constructed assuming healthcare expenditures trends can be explained by a general consumption function, the limited variable set that includes income, age, and the z vector of sociodemographic factors in unlikely to allencompassing in identifying all the key factors driving expenditure as well as the specific utility placed on each factor. Since it is impossible to construct a specific, weighted utility function that accurately represents the entirety of the US population, this study will identify certain demographic factors that are likely to have high levels of impact on most individual's decision to spend on healthcare services and goods while assigning each factor an equal weight on an individual's consumption function.

In this model, income and age are highlighted as two key factors that would influence an individual's desire to spend on healthcare goods and services. Again this brings up a key shortcoming of the model. While this model was constructed to understand an individual's desire to spend, it does not directly capture an individual's ability to pay for specific goods or services. While income might serve as proxy for ability to spend, this model is unable to differentiate between individuals who prefer not to spend and individuals who want to spend but cannot (medical non-adherence). In this way, the results of this study will be somewhat ambiguous as a variety of different situations could explain the results. However, Di Matteo identified income as an important variable impacting healthcare expenditures and thus can likely serve as a proxy for ability. Given that individuals with greater means to spend (higher income) will be more likely to spend on their health, it is likely that income is positively correlated with health expenditures. Additionally, aging individuals are also likely to be positively correlated with health expenditures due to the fact that elderly individuals are likely going to be in need of more regular care and are more likely to get sick and require care. However, income and age are also highly correlated with a host of other factors that might influence healthcare expenditures which are included in the z term.

The z term in this utility function serves as a representation of a host of demographic variables that could impact an individual decision making on healthcare expenditures. Di Matteo only utilizes geographic and access to care variables, however this study includes factors such as race, marital status, gender, diagnosis of chronic diseases, education level, and employment among a whole host of other variables. While this study will aim to isolate the factors that are most likely to impact the decision of a large majority of the population, it is very likely that our model will not include all of the potential variables that any individual will consider. The factors that were chosen were based on examining regressions of previous studies (See Section 2) and

utilizing economic theory to postulate on specific variables that would likely have an impact on healthcare expenditures (such as geographic region or employment).

However, variables that would be included in the model of this study were limited by the variables that were present within the MEPS dataset. Furthermore, the demographic factors that were chosen attempted to be as objective as possible. Since the MEPS survey is subject to high degrees of error due to the survey-style gathering of information and self-reporting of individuals, factors that could be answered definitively were preferred to be included in the model. For example, an individual's marital status or gender is extremely unambiguous and thus serves as an effective control variable within this study as opposed to variables such as perceived healthcare status, which is completely subjective from individual to individual. This method removes a lot of potential error and subjectivity from the model but also lessens its explanatory power as some variables will not be included in the model. These include values such as access to healthcare options, cultural and societal differences in perception of healthcare, quality of healthcare options, access to employer insurance, presence of an abusive relationship and many more. Not only would the addition of these variables add additional complexity to the regression, there is no clear way to standardize responses across the entire sample size. As mentioned earlier, the use of variables in the construction of the model for this study are limited to the results of the MEPS survey and the questions that it uses.

Finally, this model and the choice of specific demographic values to include in the z term serve as a good model that can be utilized in making policy decisions. By choosing variables that have trends which can be tracked analyzed, such as age, marital status, gender distributions, race distributions, individuals with chronic diseases etc. this model can be utilized to make predictions about how health expenditures will be impacted by demographic changes in the population. This is particularly salient when looking at the two key variables of the regression,

income and age. Average household income and median age are two variables that are extensively tracked and are available to researchers and policy makers alike. Thus if a significant correlation to healthcare expenditures is identified by either or both variables, then policy makers will be able to more accurately predict trends and thus foresee problems. This is important especially for public insurance programs which are funded by the federal, state, and local governments. Understanding the demographic of individuals who are using public insurance and how changes in those demographics will impact expenditures, and by extension use of insurance, will be helpful in making more informed policy decisions in the future.

4. Data

Data from this study was taken from the 2015 Medical Expenditure Panel Survey (MEPS) Household Component. The MEPS is an annual survey that began in 1996 that contacts families, individuals, insurers, and medical providers to collect data on the cost, breadth, and scope of health insurance that individuals in the United States are offered as well as information on how individuals utilize healthcare services as well as the cost and coverage of these services. The survey is split up into two components, the Household and Insurance component. This study will rely predominantly on the Household survey which focuses on responses regarding service utilization and expenditure for individuals and households for the year. This component includes "demographic characteristics, health conditions, health status, use of medical services, charges and source of payments, access to care, satisfaction with care, health insurance coverage, income, and employment." (Smith, 2009) The breadth of data covered in this survey will allow for the control of multiple variables that could have significant impacts on healthcare expenditures.

The 2015 survey was the most recent data source available at the time this study was conducted and thus was the most likely to serve as an accurate sample of the current population. Due to time constraints⁸ and sampling overlap year-over-year, a multi-year, cross section analysis could not be conducted over a multiple-year time frame, however it remains a future goal to do so. Additionally, a panel dataset of multiple years would be utilized in order to identify how the impact of different factors on healthcare expenditures have changed over time. While this would be an interesting extension of the current study, it exists outside of the scope of

⁸ Time constraints were due to the inability to get past "pay-walled" data sources plus confusion with researchers on right to use data sources for my original thesis which required me to pivot to a new topic and restart thesis work in late Dec. 2017.

the current goal. Given the lack of consensus among researchers about the impact of age on health expenditures, this study aims to use a single time point (2015) to analyze if age impacted healthcare expenditure decisions. If a statistically significant correlation is observed, then future research can focus on determining how the strength of this correlation has changed over time.

This survey data contains self-reported information from 35,427 individuals across the contiguous United States and is the most recent year for which data is currently available for public use. The survey is conducted by the Agency for Healthcare Research and Quality, a subsidiary of the US Department for Health and Human Services, and specifically is known for being one of the biggest sources of publicly-available health expenditure data. More importantly, this dataset is the only publicly-available source for information about insurance coverage and explicit healthcare expenditures by individuals. While other pay-walled data sources exist that have information about more explicit insurance coverage and exact values of medical expenditures, these data sources are often detached from the necessary demographic factors, including age. Such data sources include IQVIA (formerly IMS Health) and country-specific agencies⁹ which aggregate health expenditures but are often heavily protected and require federal government grant funding to access.

The MEPS dataset uses the best practices for utilizing a randomized sampling of individuals to survey. However, there is likely to be some bias in regards to the type of individual who would be willing to participate in this survey and be able to accurately report their expenditure values. This will be important to contextualize the results of this study. However, MEPS has taken clear measures to improve the validity of their data as well as improve the population sampling, including "oversampling" minorities, allowing for a more

⁹ Includes groups such as the Norwegian Pharmacy Association

robust and clear diversity of socioeconomic individuals. (Ezzati-Rice, Rohde & Greenblatt,

2008)

Data Cleaning Process with Cumulative Sample Size after each Cleaning Step			
Data Cleaning Process	# of people passing this filter and previous filters		
Total Raw Sample of MEPS-HC 2015	35,427		
Age >= 10 and <= 85	29,913		
Total Yearly Income is Known & $\geq $7,500^{10}$	20,246		
Education Level Known	20,008		
Cancer, Diabetes, & High Blood Pressure	19,975		
Diagnosis Known			
Employment Status Known at Each Round	19,853		
Total Healthcare Expenditure Known	16,161		

Table 2: Data Cleaning Process with Cumulative Sample Size after each Cleaning Step

In order to ensure that the survey contained sufficient data to conduct a robust analysis for this study, the data was cleaned, the specifics of which are shown in Table 1. The first step of pruning was to identify an age range of individuals who would be included in the survey. Most census data categorize age groups in the following manner: (under 18, 18-44, 44-65, 65-85, 85+). In the context of the 2015 survey, the oldest individual who was surveyed 85 while the youngest was 4. While the survey questions were likely answered by a guardian on behalf of minors (minor is anyone under the age of 18), the data set will not include individuals under the age of 10. This is important because especially in infant and younger years' healthcare and pharmaceutical expenditures tend to vary heavily dependent on the extremely variable health of the child. Additionally, the model of this study qualifies healthcare expenditures as a choice

¹⁰ Total wage includes items such as social security, pension, trust, veteran incomes, and other sources of income than just wages. \$7,500 is the amount of money an individual making minimum wage working 20 hours a week would make in a year. The extenuating circumstances for individuals beneath this line could have drastic, un-quantifiable impacts on pharmaceutical expenditure and is thus removed from the study.

made by an individual as a result of specific factors that impact the utility they will gain from making specific expenditure choices. However, individuals under the age of 10 are likely to have decisions made exclusively by parents or guardians and thus it is important to exclude these individuals from the sample as their expenditures are not a reflection of their decisions. By the age of 10, individuals likely retain some degree of autonomy on healthcare expenditure choices although they are still likely heavily influenced and paid for by parents or guardians. While the degree to which the expenditures of these individuals is a direct expression of their behavior and choices is ambiguous, they were included in the dataset as a baseline age group of which to compare. Furthermore, this group only accounts for a 176 individuals in the final dataset and thus would not lead to extreme bias and error. Access to healthcare along with a variety of cultural and sociodemographic factors play an extremely important role in expenditures that is not the same as an individual gets older. Thus, the age groups that will be used in this study will be the same as mentioned above and is present in other census survey datasets, however the under-18 category will be amended to only include individuals between the ages of 10 and 18, while the over 85 category will be removed as no one within the dataset will fit into this category. This method of age classification likely led to some introduced bias which should be considered when analyzing results.

Additionally, an important variable that will be included as a control within the regression is the total income that an individual has access to. The United States defines that an individual working part-time, making minimum wage would have a yearly salary of approximately \$7,500. While this is substantially under the federal poverty line for families of two or greater, it represents access to some semblance of monetary stability that will allow them to have access to necessary medical care and pharmaceuticals as needed. Individuals who report an income of less than \$7.5k are likely to be highly impacted by a variety of other factors related

to their socioeconomic status and cannot be controlled for within the parameters of this study. Additionally, it is very likely that the utility that these individuals have from healthcare expenditures is radically difference than the consumption utility of most of the population. Finally, individuals under this income are likely to underreport or inaccurately report expenditures due to a variety of factors including self-consciousness, and lack of knowledge as most of their support is likely coming from public sources. It is also important to note that income was self-reported value as opposed to choosing an income bracket, further adding to the error that could be experienced with this term. Thus, in order to ensure that a substantial set of data was utilized while also accounting for observations that are likely to skew data, this pruning method was utilized in this experiment.

The final steps of the data cleaning process were necessary including ensuring that education level, employment status, as well as health status are known for the observations that are used in the study. As with income, these are important control variables that will be utilized within the regression and as a result, their value for the observations used should be known and reported. These specific variables are likely to have strong correlations with healthcare and pharmaceutical expenditures and thus their value should be known and the lack of data on these states can potentially have a severe skewing effect on the regression results. Additionally, it was important that only the individuals who had reported their total healthcare expenditures be included in the dataset. Since this is the critical dependent variable that is being evaluated in this study, any observation that did not report any healthcare expenditures would skew the results by understating any correlation that would be identified. Additionally, the initial dataset coded a value of 0 for any individual who chose not to report their expenditures and thus these individuals were excluded from the dataset as well.
The MEPS dataset is a particularly valuable dataset primarily due to the robustness of variables that are included as well as the random population sample that it collects data from, thus minimizing bias and allows for clear identification of covariance and other sources of errors in regressions. However, it is limited to the contiguous United States and thus cannot necessarily be expanded to an international context as ballooning expenditures on pharmaceuticals and healthcare are a problem endemic to the international community. Furthermore, as mentioned above, the survey is conducted periodically throughout the year and done so over the phone. This necessarily disenfranchises a number of individuals from participating in the study and thus can introduce selective bias against these groups.

Finally, the overarching issue with a dataset that is predicated on individual responses to questions is that there is the potential for incorrect reporting of information, especially on information relating to exact expenditures which has the potential to skew results, especially when per capita expenditures that are estimated can have significant impact on error of results. In fact, it has been determined that individuals are often likely to underreport health expenditures due to the fact that a lot of expenditures could be hidden or be forgotten by the individual. (Hill, Zuvekas, & Zodet, 2011) However, many researchers have been using the data from this survey since its inception in 1996 to understand trends in medical expenditures and insurance and for use in policy determination and evaluation. On a whole, researchers agree that the MEPS is integral due to the "broad array of relevant and timely data that are available to researchers and policymakers to address the issues posed by the US healthcare system both today and in the future." (Smith) MEPS has been used by researchers and political analysts alike in studies and retains a lot of explanatory power in understanding decision making by patients regarding healthcare expenditures. A subset of literature that has used the MEPS dataset can be found in Table 1 of the Appendix.

Isolating the impact of age on pharmaceutical expenditures requires identifying the impact of other key demographic and health variables that are available in the MEPS data. The variables that were chosen to be included in the regression can be viewed in Table 2 of the Appendix with the coding values and predicted directionality of the coefficients of each variable provided in Table 3 of the Appendix. Table 4 of the Appendix provides the summary statistics of the key variables of the final dataset being utilized in this study. A visual inspection of the summary statistics confirms the hypothesis that the sample included in not extremely skewed by any of the sociodemographic factors. The only point of potential concern is that individuals who identify as white consist of 70% of the sample population. However, white individuals represented 76.9% of the population as of the 2016 Census, thus the population remains an accurate sample population for the entire United States. ("QuickFacts - United States," n.d.) Furthermore, the average income of the study sample was approximately \$42,000/year whereas the national mean income per capita in 2016 was \$46,550, confirming that the study sample was indeed representative. (United States Census Bureau, n.d.)

As discussed above, these variables were chosen due to their lack of subjectivity and ability to be normalized across every individual in the sample size. Furthermore, most of the literature that has been conducted in this field (See Section 2) have used at least some of the variables that will be included in this regression. This study was also limited by the variables that were included in the MEPS dataset. While the included variables likely are some of the most salient factors that influence an individual's healthcare expenditure decisions, there are definitely some variables that are missing which should be included in future research. This study is limited insofar as the only chronic diseases that are included are diabetes, cancer, and high blood pressure (proxy for heart disease). However, there are a variety of other diseases, that have high costs associated with them such as arthritis or AIDS which are not included in this regression.

Additionally, this study doesn't include any perceptions of health status, healthcare options, or access to healthcare all of which are likely to play a significant role in determining how individuals chose to spend on healthcare services and goods. Additionally, understanding the familial and community status of individuals is an important factor that might impact healthcare decisions which are not included in this study. This can include values such as size of family, number of children or grandparents in household, religion, and type of schooling (public, private, religious) amongst others. Within the context of this, it was important to not include too many variables as this also increases the number of potential interactions, which would further complicate the regression. Additionally, previous use of MEPS in research by Smith (2009), Hill et. al, (2011) & Ubel (2010) further conclude that the variables identified in this study are the ones that likely have the greatest impact on medical decision making. Finally, the goal of this study is to isolate the impact of aging on increasing health expenditures. However, it is unlikely that aging will ever be completely isolated from other sociodemographic factors such as income and education. Given that most of the previous research in this field has retained an international perspective and has not used a dataset like MEPS to isolate the impact of aging at a national level and predicated on sociodemographic factors, further research and future studies will help determine and narrow the scope of important variables to consider. In the context of this study, it is expected that the lack of the inclusion of the variables mentioned above are likely to introduce omitted variable bias (OVB) and could potentially cause results to be over or understated and reduce the overall fit of the model.

The model will seek to isolate the percent change in total healthcare expenditures and pharmaceutical expenditures between the different age groups. The log-linear model allows for the isolation of the effect of a unit change of the independent variable (movement between age groups) on the percent change of the dependent variable (total healthcare & pharmaceutical

expenditures). As mentioned above, most of the other variables included will be dummy variables that will be used as controls and ensure that the model specifically identifies and isolates the impact of aging, thus it is likely that these relationships will remain linear. It is also important to note that the variables mentioned above are likely to have some degree of covariance and the lack of an interaction term introduces error/bias into the analysis. As a result, interaction terms were integrated into the regression to capture correlation effects between different socioeconomic factors. (See Results) This model accounted for some of the identified correlations, however the exclusion of any interaction terms introduces covariance or collinearity into the model which reduces the significance of the results.

5. Empirical Specifications

On an international level, there are two macro-factors that can influence growing health expenditures in developed nations, the design of the major healthcare system and demographic factors. Specifically, we identify the following equation to determine health expenditures:

$$HCE = \alpha_0 + \alpha_1 LEV + \alpha_2 LEV^2 + \alpha_3 GDP + \alpha_4 RELP + \alpha_5 PHY + \alpha_6 POP65 + \alpha_7 POP15$$
$$+ \psi_1 GK + \psi_2 FFS + \psi_3 CA + \psi_4 PI + \psi_5 PR + \varepsilon$$

Equation 5: Healthcare Expenditures as a function of International Indicators

Specifically, within this model¹¹, healthcare expenditures (HCE) are predicated on the level of healthcare services that are present within a given country (LEV) as well as the GDP of the country, the relative price of healthcare (RELP), the doctor density (PHY), population demographics: POP65 (Population greater than 65), POP15 (Population younger than 15) and a host of dummy variables defining the specific design of the health care system including gate-keeping (GK), fee-for-service (FFS), capitated primary care (CA), public integration (PI), and public reimbursement (PR) and an error term ε .

Gatekeeping (GK) is a system in which any individual must go through their primary care physician (PCP) to have access to more specialized care such as specialists, surgeons, laboratories, or even going to the hospital. This method of insurance prevents excessive costs because it forces the patient to consult with a medical professional prior to accessing high cost services. Fee-for-service (FFS) insurance systems mimic those of the US private healthcare system and usually have a large number of privately owned insurers who charge specific fees and have certain coverage based on the service that is provided or good that is being purchased. Capitated primary care (CA) insurance schemes involve the payment of a set sum (capitation) to

¹¹ Drawn Okunade, Karakus, & Okeke, 2004)

healthcare providers based on the number of individuals that enroll with them for a predefined time period. Public integration (PI) is an insurance system in which a governmental entity controls the funding of care options with doctors, hospitals, and other service providers existing as private entities outside the insurance system. Public reimbursement (PR) is an insurance scheme in which individuals are responsible for the allocation of their healthcare expenditures and price of services are agreed upon by the insurers and the providers which gives individuals more choice on services, however limits providers to those who are in contract with the insurer. Within the context of the United States, the relevant healthcare systems are FFS/PR (Private Insurers) and PI/PR (Medicare, Medicaid, & Tricare).

All of these variables can have an impact on total healthcare expenditure of a given country. Of specific importance to this study is the correlations to the population age demographics as well as the type of insurance service that is being provided. Most models, including the United States, have an integration of PI and PR built into their system structure. This is particularly interesting because they have conflicting impacts on healthcare expenditures as explained by taking the first order conditions (FOC) of the relationships.

$$\frac{\partial HCE}{\partial PR} > 0$$

as public reimbursement systems tend to raise expenditures with all else equal Equation 6: FOC of Public Reimbursement Insurance on Healthcare Expenditures

$$\frac{\partial HCE}{\partial PI} > 0$$

as public integrated systems tend to restrain expenditure growth with all else equal Equation 7: FOC of Publicly Integrated Insurance on Healthcare Expenditures This model developed by Okunade et. al (2004) was constructed to analyze health expenditures at an international level to compare country health expenditures as a result of national GDP, median age, level of care available, and types of insurance systems available. However, as mentioned earlier, such an international construction is unlikely to capture the specific factors that encourage individuals to spend on healthcare. Thus combining this model with the individual consumption function framework adapted by Di Matteo (2005) results in an OLS regression model that replaces the national indicators (GDP etc.) with individual-specific indicators of healthcare consumption (age, race, gender etc.) as shown in Equations 8 and 9 where Z is a vector of the sociodemographic factors that will be utilized in this study.

Such a domestic look is important to specifically identify the impact of aging in the US insurance and healthcare system. Given the multi-payer insurance system and highly differentiated nature of healthcare services in the United States compared to other countries in Europe and across the world, international studies such as that by Okunade retain minimal explanatory power over identifying the specific circumstances that are causing US healthcare expenditures to grow faster than any other developed country. By using the models set forth in this study, we can identify the specific sociodemographic factors, including age, which are impacting healthcare expenditures in the United States and understand how trends in these factors can be used to predict future growth in healthcare expenditures within the United States.

 $ln(HCE) = \alpha_1 AGE + \beta_1 Medicare + \beta_3 Medicare + \beta_3 Tricare + \beta_4 Private + \psi_1 Cancer$ $+ \psi_2 Diabetes + \psi_3 HighBloodPressure + \phi Z + \varepsilon$

Equation 8: Log-Linear OLS Regression evaluating impact of Age, Insurance Coverage, Health Indicators, and Sociodemographic Variables on Total Healthcare Expenditures

 $\ln(PE) = \alpha_1 AGE + \beta_1 Medicare + \beta_3 Medicard + \beta_3 Tricare + \beta_4 Private$

+ β_5 PrivatePublic + ψ_1 Cancer + ψ_2 Diabetes + ψ_3 HighBloodPressure + $\phi Z + \varepsilon$

Equation 9: Log-Linear OLS Regression evaluating impact of Age, Insurance Coverage, Health Indicators, and Sociodemographic Variables on Total Pharmaceutical Expenditures Equation 8 observes the impact of age, here shown as the transition from one census age group as defined in the Section 4, on total health expenditures. The regression also includes critical variables including type of insurance coverage that an individual has as well as the presence of any critical illnesses (Cancer, High Blood Pressure, & Diabetes), as well as a multitude of sociodemographic factors that are listed in the empirical specifications and are included in the regression with the Z term. The specific illnesses were chosen as they represent the most common chronic diseases that affect households with high blood pressure serving as a proxy for most cardiovascular diseases. An error term is also included in the regression as it is impossible to take into account every factor that could potentially impact an individual's ability and necessity to spend on healthcare. Equation 9 examines the impact of the same variables, but specifically isolate the impact of aging on pharmaceutical expenditures (PE). This will allow a comparison of the impact of aging on both pharmaceutical and total expenditures to determine if aging is impacting pharmaceutical expenditures to a greater degree as compared to overall health expenditures or vice versa which has important policy ramifications.

$\ln(HCE_{public}) = \alpha_1 AGE + \psi_1 Cancer + \psi_2 Diabetes + \psi_3 HighBloodPressure + \phi Z + \varepsilon$

Equation 10: Log-Linear OLS Regression evaluating impact of Age, Health Indicators, and Sociodemographic Variables on Total Public Healthcare Expenditures

Equation 10 examines the specific impact of these variables on total public health expenditures which is defined as the sum of expenditures from Medicare, Medicaid, and Tricare combined. This will specifically identify how the sociodemographic and health variables will impact expenditures under public insurers and serve as a proxy for understanding how government expenditure on public health insurance will change in the future. Of important note is the lack of insurance terms in this regression. This was important to remove high degrees of collinearity as the dependent variable is directly correlated to insurance coverage.

This adapted model is likely to confirm the results of Gregersen (2013) and Schenider & Guralnik (1990), that healthcare expenditures in the United States are positively correlated with aging, however the significance and degree of impact may or may vary from current literature. While the manner in which aging has been captured varies by researcher and the degree of significance is impacted by the robustness of the study, most researchers have come to a consensus that in fact aging does increase health expenditures, which aligns with common sense understanding. However, this regression, by looking at the impact of aging between predefined age groups, and isolating the impact of pharmaceutical expenditures away from health expenditures can help capture how aging impacts one of the biggest portions of health expenditures, pharmaceutical expenditures. Furthermore, the log-linear model is utilized as it identifies the percent increase in expenditures for an individual as they move from one age group to the other. In turn, this value can be used to proxy population dynamics thus allowing for the prediction of increases in expenditures as a higher percent of the population will move into the over 65 age group. Furthermore, this model is isolated from any confounding factors that might focus on patient outcomes as opposed to just expenditures. As our model aims to identify sources of increasing healthcare expenditures and not the impact of aging on health outcomes, by focusing on factors that will impact behaviors and decisions regarding expenditures this model is able to remove the uncertainty surrounding outcomes.

The final goal of this study was to identify the impact of aging within each of the different insurance plans. This model is shown in Equation 11.

 $ln(HCE_{insurer}) = \alpha_1 AGE * Insurance + \psi_1 Cancer + \psi_2 Diabetes$ $+ \psi_3 HighBloodPressure + \phi Z + \varepsilon$

Equation 11: Log-Linear OLS Regression evaluating impact of Age, Health Indicators, and Sociodemographic Variables on Healthcare Expenditures by Insurer This model has an independent variable which was constructed as the interaction term between age and the insurance indicators from Equation 8 & 9 thus to specifically identify the impact of aging under each insurance plan. Equation 11 was utilized to look at aging under Medicare, Medicaid, Tricare, and Private insurance.

Future research should also attempt to identify the impact of aging on other subsections of expenditures such as ambulatory care or outpatient procedures, but that is outside the scope of this study. Finally, it's important to note that the regression including differences in coverage is something that is not commonly included in other statistical regressions. By including and differentiating on these factors, this study will be able to make novel contributions to understanding the impact of certain policy mechanisms on expenditures by public insurance providers and can provide detailed insight into planning for a currently aging US population.

6. Results

As previously stated, many of the control variables, along with the independent variable of age, are likely correlated with one another. For example, age and employment or race and income could be correlated which could potentially skew relationship between each individual variable and total healthcare expenditures. Thus, in order to understand the strength of the correlation between variables utilized in the model, a correlation matrix was constructed. (Table 5 in Appendix) The correlation matrix was utilized to determine which of the correlations between variables were strong enough to be included within the regression. There is no consensus amongst researchers about a set cut-off for correlation coefficients that determine significance, thus the absolute value of 0.250 (25%) was arbitrarily chosen. This value was chosen through visual inspection of the correlation matrix as well as application of economic theory to determine which correlations were, logically, more likely to be true and relevant to the model (income and education) and which were likely but not relevant to the regression (income and geographic regions).

However, while certain interactions were not included in the regression, understanding the implication of the correlation between the variables might help contextualize some of the results. In the correlation matrix, the values that are highlighted and bolded in red met the 0.250 threshold, but were not included in the final regression as initial tests determined that these interactions did not have a statistically significant impact on healthcare expenditures and thus did not improve the model's explanatory power or fit. These interactions included relationships that would have been expected, for example that income and Medicaid enrollment are negatively correlated as Medicaid was an insurance program specifically for individuals that have a low income. The values highlighted and bolded in green are correlations that were included in the regression as interaction terms because they represented a high degree of correlation with the

independent variable that is being tested in the study, age. The only exception to this are the correlations between health insurance plans and age. Identification of the impact of aging within each of these insurance schemes when observing all of the insurance plans in a single regression will likely result in statistically insignificant results while impacting the variance and coefficients of other variables. Thus investigating the impact of aging within each insurance plan was conducted as a separate regression

6.1 Primary Correlational Analysis Results: Chi-squared Test Results

In order to understand at a basic level if any correlation between age and healthcare expenditures existed, a chi-squared test was conducted between total healthcare expenditure quartiles and age group. (Table 3)

		Age Group		
Expenditure by Quartile	10-18 (Group 0)	19-44 (Group 1)	45-64 (Group 2)	65+ (Group 3)
\$0-\$149	15.45%	13.43%	7.27%	1.61%
\$150-\$1,999	4.55%	4.77%	2.62%	0.62%
\$2,000-\$349,000	80%	81.79%	90.06%	97.76%
\$350,000+	0%	0.02%	0.05%	0%
	Pearson $Chi2(9) = 60390$.	<i>19</i> $Pr = 0.00$	0	

Table 3: Chi-squared analysis results, Age on Total Healthcare Expenditures

The results of this chi-square test suggest that there is indeed a statistically significant relationship between age and expenditures. The percent of each age group in the expenditure group "\$2,000 to \$349,999" increased as the age group got older from 80% of age group 0 to 97.7% of age group 3. These initial results serve as a confirmatory test of my hypothesis that indeed age was indeed positively correlated with healthcare expenditures. Additionally, chi-squared test results were run on the significant interactions that were identified in the correlation matrix. The results of these chi-squared tests are shown in Tables 6-13 in the Appendix. While some of these interactions are not directly included in the regression, they might enhance the explanatory power of the results received from the OLS regression and help determine the reason for unexpected results.

In order to determine which of the identified variables (Table 2 in Appendix) would be included in the final regressions, it was important to note the impact of collinearity between variables that would skew the coefficient results. This was particularly important to consider in the case of multiple dummy variables that were indicative of a single sociodemographic factor. These included race, geographic regions, and education. While 3 different variables were constructed to signify race (White, Black, OtherRace), only white was included in the final regression. As is shown in the correlation matrix, these 3 variables indicated a high degree of collinearity and thus are likely to have extremely high variance and skew the correlations significantly. Additionally, since all three variables represent the entire population in a single category (race), at least 1 of the variables had to be removed. This same phenomenon was present with the dummy variables for education and geographic region as well and thus only *northeast* and *ColDeg* were included in the final regression. It is important to note that some collinearity was evidenced between the different insurance schemas. However, as the study aims to identify the impact of all the insurance schemes that are available to individuals in the United States, all four of the insurance variables remained in the regression. This benchmarked the results of the insurance variables against individuals with no healthcare insurance.

6. 2 Log-Linear OLS Regression Results: Total Healthcare and Pharmaceutical Expenditures

The log-linear OLS regression specified in Equations 8 and 9 were ran and the results are in Table 14 in the Appendix (Models 1 & 2).

Looking at the results of the regression as a whole, we see that both regressions have low R^2 values. This was to be expected due to the fact that the data was collected from a self-reported survey. As discussed previously, this method of data collection is prone to many errors and introduces a lot of noise that is likely to impact the fit of a specific model to the data. Furthermore, as discussed prior, there are likely to be a large number of variables that would

impact increasing healthcare expenditures, some of which are likely not included in this regression. Thus, low R^2 values are logical given the subset of variables that are being utilized. However, given the data source presented, an R^2 value of 0.24 and 0.27 respectively are indicative of statistically significant results. Furthermore, the p-value of the F-statistic in the case of both regression was 0 which indicates that the overall regression was statistically significant at the 1% level. Thus, the null hypothesis was rejected that none of the variables chosen in the regression had any impact on total healthcare expenditures or pharmaceutical expenditures.

The absolute measure of fit, the Root Mean Squared Error (RMSE) was 0.240 for total healthcare expenditures and 0.270 for total pharmaceutical expenditures. Since the range of values for total healthcare expenditures and total pharmaceutical expenditures are 0-13.4 and 0-12.94 respectively, it can be concluded that the RMSE values are within an acceptable threshold. It is important to note that there is no accepted cutoff for significance of RMSE values as they are conditional upon the values of the dependent variable. As the RMSE values are approximately 2% of the total range of both dependent variables, it can be concluded that the results of these regressions are statistically significant and individual coefficient p-values must be observed to determine if individual variables retained a statistically significant impact on healthcare expenditures.

The results of these regressions show that age has a statistically significant (at the 5% level) impact on both total healthcare and pharmaceutical expenditures. Movement between age groups was correlated with a 10% increase in total healthcare expenditures and a 13% increase in total pharmaceutical expenditures. While the impact seems minor, these results are more significant and show stronger correlation than the results of Steinmann & Zweifel (2007), Meijer et. al (2013), and Getzen (1992). The results of this study directly contradicts the findings of Seshamani & Gray (2002) as well as Dieleman et. al (2016) on a panel of OECD countries while

supporting the results of Gerdtman's (1993) findings in Sweden as well as the findings of Gregersen (2013) and Schenider & Guralnik (1990). Furthermore, while these values seem small, in the context of total healthcare expenditures of the United States they can have large impacts. Total healthcare expenditures in the US reached \$3.3 trillion in 2016. ("National Health Expenditure Accounts," 2018) Thus, if 1% of the population transitioned from the "44-65" age group to the "65+" age group, then total healthcare expenditures would increase by \$3.3 billion dollars. It is also important to note that the impact of aging was also mitigated by the presence of multiple interaction terms which included age as well. Generally speaking, the data confirms the hypothesis that aging does have an impact on healthcare expenditures and thus concerns over ballooning healthcare and pharmaceutical expenditures are extremely valid.

The regression results also bring out a number of interesting observations which, while not directly correlated to the impact of aging on expenditures, are cause for interest. The interaction terms for diabetes, cancer, and high blood pressure with age all have negative correlations with both total healthcare and pharmaceutical expenditures. The interactions with diabetes and cancer were significant at the 1% level in both models. This was initially confusing as one would assume that as you get older, you are likely to need to spend more money on treatment options for chronic diseases. The results directly contradict prior logic and the positive correlation with the isolated terms of aging, diabetes, cancer, and high blood pressure, all of which were strongly positive and significant at the 5% and 1% levels respectively. Additionally, the chi-squared tests in the Appendix (Tables 7-9) showed an extremely significant, positive correlation between aging and the contraction of the aforementioned chronic illnesses. However, the strong positive correlation between expenditures and the interaction between employment and age shows that individuals who retain the means to be able to pay for services and goods (i.e have money for being employed) are spending more as they get older.

However, as seen in the chi-squared test in Tables 10-13 in the Appendix, age is negatively correlated with employment at a statistically significant level as is having diabetes, cancer, or high blood pressure. Thus, aging individuals with chronic illnesses represent a portion of the population that are least likely to be employed and consequently are unlikely to have the means to pay for the necessary medical treatment, leading to non-adherence. There are other potential rationales for why older individuals with chronic illnesses are paying less such as coverage under a family member's plan or more general doctor visits reducing the necessity for additional external costs or even underreporting in the population are all potential other reasons for the negative correlation. However, given the strength of the coefficient of the interaction terms in the regression, we are comfortable concluding that the segment of the population that includes elderly individuals with chronic illnesses are most likely to experience high levels of medical non-adherence, although further inspection and research is required.

Of additional interest was the strong negative correlation (-0.92 on Total Healthcare Expenditures and -1.52 on Pharmaceutical Expenditures), significant at the 1% level, between employment and total healthcare and pharmaceutical expenditures (Table 14 in Appendix). As explained above, it follows logically that an individual who is employed is more likely to have access to more funds which affords them the ability to spend more on necessary healthcare services. However, the results of the regressions show this not necessarily to be the case. However, the strong positive correlation between the interaction term of aging and employment are suggestive of the fact that in general, individuals who are employed receive access to insurance through their employer and as such are removed from the expenditures incurred when using this insurance. In doing, individuals who are employed are likely to severely underreport many of their health expenditures. Additionally, employer health plans often cover spouses, children, and immediate family members, meaning a large portion of the population is likely

covered under an employer insurance plan. If individuals are not including these expenditures in their reporting, then the strong negative correlation is to be expected.

Most of the other sociodemographic factors retained the expected sign of the coefficient (Table 3 in Appendix). Of particular interest were the insurance variables. While the results of these regressions are not comparative between the different insurance types, the value of the coefficient is particularly informative. All the coefficients were significant at the 1% level across both models¹². Medicare was associated with the highest increase in expenditures relative to individuals without Medicare (86% and 92.2% higher respectively). Tricare was associated with the lowest increase in expenditures relative to individuals without Tricare (52.7% and 26.3% respectively) followed closely by Medicaid (60.1% and 60.5% respectively).

6.3 Log-Linear OLS Regression Results Limited to Public Expenditures

In order to determine if the trends identified in these two regressions would hold true specifically for expenditures that were covered by public insurance, an additional log-linear regression was run using Equation 10. The results are present in Table 14 in the Appendix (Model 3).

These regression results mirror many of the same trends that were observed in the previous regressions that were conducted above. The regression results indicate that the movement from one age group to the next was associated with a 12.4% increase in total healthcare expenditures statistically significant at the 5% level. This value was higher than Model 1 which found only a 10% impact. While these results might signify that impacts of aging on healthcare expenditures are likely to be magnified by beneficiaries of public insurance programs, the sample size of this regression is substantially smaller than the first two models

¹² Exception of Tricare when observing Pharmaceutical Expenditures (Significant at 5%)

(6,861 individuals). This subset of the population was utilized as many individuals did not report the breakdown of their total healthcare expenditures by insurance types. Even with the small sample size, the regression on a whole is statistically significant with similar R^2 values (0.232) and Root MSE values (1.8003) as Models 1 and 2, and many of the same trends are observed within this regression as were observed in Models 1 and 2.

The same correlations identified in Models 1 and 2 were present in Model 3, specifically the interaction terms between age and chronic illnesses retaining statistically significant negative values while the individual terms retain a statistically significant positive value. This suggest that even under a public insurance plan, services and goods are often too expensive and likely will lead to increased levels of medical non-adherence. The confirmation of these trends across all the regressions is extremely important as it validates the findings across multiple different dependent variables and thus can be broadly generalized beyond just the sample used in this study.

One unique trend identified in this model which was not present in Models 1 and 2 was the statistically significant (1%) negative correlation with income. While the model indicates that a \$1000 increase in income yields a 0.38% increase in healthcare expenditures, this directly contradicts the findings of Model 1 and 2. However, the most likely rationale for this is that individuals who are covered under only public insurance plans are likely to be elderly, disabled, or have low income. Therefore, the income variation in this population is likely to be minimal and individuals with higher income are likely exposed to better conditions and thus have less necessity to spend on healthcare goods and services.

Overall, this model is one of the first that has attempted to specifically identify how aging impacts specifically health expenditures by subsidiaries of public insurance plans. The statistically significant results of this model indicate that as the US population continues to age,

there is likely to be an increasing burden on public healthcare systems as healthcare expenditures increase as well.

6.4 Log-Linear OLS Regression Results, Assessing Aging within each Insurance Plan

The final part of this study was to determine the impact of aging within each of the insurance schemes (Medicare, Medicaid, Tricare, and Private). The regressions results can be viewed in Table 15 of the Appendix. All four models displayed similar overall fit with an R² value of 0.212 and Root MSE values of 1.5588-1.5603. Additionally, all four models had an F-stat p-value of 0, thus it could be assumed that the models were statistically significant.

These regressions identified a statistically significant (at the 1% level), positive correlation between aging and healthcare expenditures within each of the insurance plans, with the exception of Tricare. These regression results support the findings of Cubanski et. al (2018)¹³, and confirm the hypothesis that aging does lead to increased healthcare expenditures, regardless of insurance coverage. Additionally the impacts of aging were considerably stronger in Private Insurance and Medicaid (10.3% and 11% respectively) compared to Medicare (7%). These results are particularly interesting and suggest that Medicaid expenditures are likely to rise quicker than Medicare expenditures in response to aging. While the results of the Tricare regression were not significant, this was likely due to the fact that the sample population only accounted for 305 individuals who were beneficiaries of Tricare. Thus, potential outliers and a small sample size could lead to low statistical significance. These models additionally exhibit the same directionality and significance of Models 1-3, thus serving as an additional confirmatory test of significance.

¹³ Cubanski's results were specific to Medicare

The results of these models are extremely important and represent a novel approach at observing the impact of aging. Minimal research has attempted to evaluate impacts of aging within each insurance plan. By identifying Medicaid as the insurance plan with the strongest effect on increasing health expenditures as a result of aging, this study can begin to provide policy makers and researchers a benchmark from which to conduct further research and identify method of controlling expenditure growth within Medicaid.

7. Conclusion

The purpose of this study was to determine if aging retained a statistically significant impact on increased healthcare and pharmaceutical expenditures within the United States. Previous literature had failed to identify any significant impact of aging on healthcare expenditures. Previous research relied on the construction of an international model to identify the impact of aging which was not be able to capture the specific impacts of aging on the US population nor did it take into account a variety of factors that influence an individual's consumption of healthcare goods and services. The model used in this study was designed to identify the specific impact of aging in the United States when compared to a host of other sociodemographic factors that were likely to significantly impact an individual's desire to spend on healthcare goods and services. The results of the regression models supported the hypothesis that aging is correlated with increasing healthcare and pharmaceutical expenditures in the United States and this correlation holds true under all insurance plans, both private and public. Additionally, the results demonstrated high levels of medical non-adherence in elderly populations with chronic diseases such as diabetes, cancer, or cardiovascular disease. The pervasiveness of these correlations signify that specifically within the United States, an aging population is a signifier of increased healthcare expenditures and potentially worse health outcomes and that continual population aging is likely to continue this trend in the future.

These results should be particularly alarming to policy makers as the United States is already at a point where it is projected to have healthcare expenditures grow 1% faster than GDP from 2017-2026. (NHE Fact Sheet) This is particularly problematic as a report by the US Department of Health and Human Services found that increasing healthcare expenditures can lead to increased unemployment, increased cuts and limitations of government (public) insurance plans, and higher uninsurance rates. (*The Effect of Health Care Cost Growth on the U.S.*

Economy, 2012) This is because as healthcare expenditures increase, the insurance burden on employers increase and thus companies are likely to have to reducing hiring or switch to more part-time and contractor work where provision of healthcare is not mandatory.

Additionally, increased expenditures in healthcare necessarily mean less expenditures on other necessary programs such as education and thus the government is increasing limitations on individuals who have access to public insurance and reducing reimbursement rates which will likely lead to decreased acceptance of public insurance plans at healthcare service providers. On net, this will likely cause more individuals to go uninsured and increase medical non-adherence which, as stated earlier, has a cyclical effect on increasing hospitalization and overall healthcare expenditures. Finally, as healthcare expenditures increase, insurance companies are likely to increase premiums and costs of coverage plans which will disincentive or prevent individuals from accessing insurance and also lead to uninsurance rates. Overall, this is concerning because all of the reactionary measures that have been implemented as a result of rising healthcare expenditures are likely to exacerbate the problem in the near future and a more sustainable solution needs to be identified.

Further research is necessary in order to determine if the results of this study hold true more broadly and further identify key areas that can targeted to help slow down healthcare expenditures and prevent public insurance plans from going bankrupt while simultaneously preventing higher levels of medical non-adherence. Additional studies should be done with the intention of identifying which specific healthcare services and goods are most likely to be increasing a result of an aging population. Thus, insurance allocations can favor lower deductibles and higher coverage on select services which are more likely to be utilized and reduce coverage on other goods and services. Furthermore, more robust studies which take into account additional variables, specifically subjective variables such as access to healthcare services and perceived health status, should be done to determine if the significance of aging is present in larger populations and larger variable sets.

Finally, future studies should look at panel data to determine how the impact of aging on healthcare expenditures has changed over time. If the impact of aging on healthcare expenditures has increased overtime, this would signal a greater cause for alarm as this will likely compound the problem of overburdening insurance plans. Overall, policy makers must begin looking at austerity measures that will curb total spending while simultaneously preventing adverse effects such as high rates of uninsurance and medical non-adherence. It is our recommendation that more emphasis and money be placed on preventative care and education as well as nutrition and lifestyle coaching as this will overall improve population health and decrease the necessity of healthcare expenditures, however further research and testing is necessary to determine if this holds true in practice.

References

- 2010 Census Shows Nation's Population Is Aging. (2011, May 26). Retrieved from www.census.gov/newsroom/releases/archives/2010_census/cb11-cn147.html
- AARP. (2011, January 01). Understanding Medicare Part A, Part B, Part C and Part D -. Retrieved April 17, 2018, from http://www.aarp.org/health/medicare-insurance/info-01-2011/understanding_medicare_the_plans.html
- Altman, D. (2015, April 16). Public vs. Private Health Insurance on Controlling Spending. Retrieved April 17, 2018, from https://blogs.wsj.com/washwire/2015/04/16/public-vsprivate-health-insurance-on-controlling-spending/
- Anderson, G. F., Reinhardt, U. E., Hussey, P. S., & Petrosyan, V. (2003). It's The Prices, Stupid:
 Why The United States Is So Different From Other Countries. *Health Affairs*, 22(3), 89-105. doi:10.1377/hlthaff.22.3.89
- Bell, J. F., Zimmerman, F. J., Arterburn, D. E., & Maciejewski, M. L. (2012). Health-Care
 Expenditures of Overweight and Obese Males and Females in the Medical Expenditures
 Panel Survey by Age Cohort. *Obesity*, 19(1), 228-232. doi:10.1038/oby.2010.104
- Bhattacharyya, N. (2011). Incremental Health Care Utilization and Expenditures for Chronic Rhinosinusitis in the United States. *Annals of Otology, Rhinology & Laryngology, 120*(7), 423-427. doi:10.1177/000348941112000701
- Blanchard, J., Madden, J. M., Ross-Degnan, D., Gresenz, C. R., & Soumerai, S. B. (2013). The Relationship Between Emergency Department Use and Cost-Related Medication Nonadherence Among Medicare Beneficiaries. *Annals of Emergency Medicine*, 62(5), 475-485. doi:10.1016/j.annemergmed.2013.04.013
- Census Bureau, Public Information Office. (2014). Fueled by Aging Baby Boomers, Nation's Older Population to Nearly Double in the Next 20 Years, Census Bureau Reports.

Centers for Medicare and Medicaid Services. (2016). *Projections of National Health Expenditures: Methodology and model specification: 2015-2025.*

CHAMPVA. (2014, April 10). Retrieved from

www.va.gov/COMMUNITYCARE/programs/dependents/champva/index.asp

- Clemens-Cope, L., Holahan, J., & Garfield, R. (2016). Medicaid Spending Growth Compared to Other Payers: A Look at the Evidence. *The Kaiser Commission on Medicaid and the Uninsured*.
- Clift, J. B., & Kirby, J. (2012). Health Care Access and Perceptions of Provider Care Among Individuals in Same-Sex Couples: Findings from the Medical Expenditure Panel Survey (MEPS). *Journal of Homosexuality*, 59(6), 839-850. doi:10.1080/00918369.2012.694766
- Cohen, J. W., Cohen, S. B., & Banthin, J. S. (2009). The Medical Expenditure Panel Survey. *Medical Care*, 47. doi:10.1097/mlr.0b013e3181a23e3a
- Cook, A. (2013). Costs under Medicare's Prescription Drug Benefit and a Comparison with the Cost of Drugs under Medicaid Fee-for-Service (United States, Congressional Budget Office, Health, Retirement, and Long-term Analysis Division). Washington, DC: CBO.
- Cubanski, J., Orgera, K., Damico, A., & Neuman, T. (2018). The Financial Burden of Health Care Spending: Larger for Medicare Households than for Non-Medicare Households. *Kaiser Commission on Medicare*.
- Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation. (2012). *The Effect of Health Care Cost Growth on the U.S. Economy* (pp. 3-57).
- Di Matteo, L. (2005). The macro determinants of health expenditure in the United States and
 Canada: Assessing the impact of income, age distribution and time. *Health Policy*, 71(1),
 23-42. doi:10.1016/j.healthpol.2004.05.007

Dieleman, J. L., Squires, E., Bui, A. L., Campbell, M., Chapin, A., Hamavid, H., . . . Murray, C.
J. (2017). Factors Associated With Increases in US Health Care Spending, 1996-2013. *Journal of the American Medical Association*, *318*(17), 1668.
doi:10.1001/jama.2017.15927

- Erixon, F., & Van der Marel, E. (2011). WHAT IS DRIVING THE RISE IN HEALTH CARE EXPENDITURES? An Inquiry into the Nature and Causes of the Cost Disease. *ECIPE*.
- Ezzati-Rice, TM, Rohde, F, Greenblatt, J, Sample Design of the Medical Expenditure Panel Survey Household Component, 1998–2007. Methodology Report No. 22. March 2008. Agency for Healthcare Research and Quality, Rockville, MD.

http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr22/mr22.shtml

- Felix, A., & Watkins, K. (2011). The Impact of an Aging U.S. Population on State Tax Revenues (Kansas City Federal Bank). KS.
- Frenk, J. (2004). Health and the Economy: A Vital Relationship. Retrieved April 17, 2018, from http://oecdobserver.org/news/archivestory.php/aid/1241/Health_and_the_economy:_A_vi tal_relationship_.html
- Gerdtham, U. (1993). The impact of aging on health care expenditure in Sweden. *Health Policy*, 24(1), 1-8. doi:10.1016/0168-8510(93)90083-2
- Getzen, T. E. (1992). Population Aging and the Growth of Health Expenditures. *Journal of Gerontology*, 47(3). doi:10.1093/geronj/47.3.s98

Gold, S. (2011, May 11). How European nations run national health services: Belgium, France,
Germany and Sweden. Retrieved April 17, 2018, from
https://www.theguardian.com/healthcare-network/2011/may/11/european-healthcare-services-belgium-france-germany-sweden

Gregersen, F. A. (2013). The impact of ageing on health care expenditures: A study of steepening. *The European Journal of Health Economics*, *15*(9), 979-989. doi:10.1007/s10198-013-0541-9

- Hadley, J., & Holahan, J. (2003). Is Health Care Spending Higher under Medicaid or Private Insurance? *INQUIRY: The Journal of Health Care Organization, Provision, and Financing, 40*(4), 323-342. doi:10.5034/inquiryjrnl_40.4.323
- Harrington, C., & Estes, C. L. (2008). Health policy: Crisis and reform in the U.S. health care delivery system. Sudbury, Mass.: Jones and Bartlett.
- Health Care Benefits for Dependents (CHAMPVA). (n.d.). Retrieved from https://www.benefits.gov/benefits/benefit-details/318
- Hill, S. C., Zuvekas, S. H., & Zodet, M. W. (2011). Implications of the Accuracy of MEPS
 Prescription Drug Data for Health Services Research. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing, 48*(3), 242-259.
 doi:10.5034/inquiryjrnl_48.03.04
- Himmelstein, D. U., & Woolhandler, S. (2016). The Current and Projected Taxpayer Shares of US Health Costs. *American Journal of Public Health*, *106*(3), 449-452. doi:10.2105/ajph.2015.302997
- Hugtenburg, J., Vervloet, M., Dijk, L. V., Timmers, L., & Elders, P. J. (2013). Definitions, variants, and causes of nonadherence with medication: A challenge for tailored interventions. *Patient Preference and Adherence*, 7, 675-682. doi:10.2147/ppa.s29549
- Kaiser Family Foundation. (2017, November 22). An Overview of Medicare. Retrieved April 17, 2018, from http://www.kff.org/medicare/issue-brief/an-overview-of-medicare/.

- Kaplan, R. M., Fang, Z., & Kirby, J. (2017). Educational attainment and health outcomes: Data from the Medical Expenditures Panel Survey. *Health Psychology*, *36*(6), 598-608. doi:10.1037/hea0000431
- Kawatkar, A. A., Jacobsen, S. J., Levy, G. D., Medhekar, S. S., Venkatasubramaniam, K. V., & Herrinton, L. J. (2012). Direct medical expenditure associated with rheumatoid arthritis in a nationally representative sample from the medical expenditure panel survey. *Arthritis Care & Research*, 64(11), 1649-1656. doi:10.1002/acr.21755
- Kennedy, J., & Wood, E. G. (2016). Medication Costs and Adherence of Treatment Before and After the Affordable Care Act: 1999â□"2015. American Journal of Public Health, 106(10), 1804-1807. doi:10.2105/ajph.2016.303269
- Kesselheim, A. S., Avorn, J., & Sarpatwari, A. (2016). The High Cost of Prescription Drugs in the United States. *Journal of the American Medical Association*, *316*(8), 858. doi:10.1001/jama.2016.11237
- Knickman, J. R., & Snell, E. K. (2002). The 2030 Problem: Caring for Aging Baby Boomers. *Health Services Research*, *37*(4), 849-884. doi:10.1034/j.1600-0560.2002.56.x
- Lopreite, M., & Mauro, M. (2017). The effects of population ageing on health care expenditure: A Bayesian VAR analysis using data from Italy. *Health Policy*, *121*(6), 663-674. doi:10.1016/j.healthpol.2017.03.015

Lubar, S. B., & Smunt, T. L. (2017, October 27). Drug costs undercut U.S. healthcare system. Retrieved April 17, 2018, from http://www.jsonline.com/story/opinion/contributors/2017/10/27/drug-costs-undercuthealthcare-system/800566001/

- Machlin, S. R., & Soni, A. (2013). Health Care Expenditures for Adults With Multiple Treated Chronic Conditions: Estimates From the Medical Expenditure Panel Survey, 2009.
 Preventing Chronic Disease, 10. doi:10.5888/pcd10.120172
- Maestas, N., Mullen, K., & Powell, D. (2016). The Effect of Population Aging on Economic Growth, the Labor Force and Productivity. *National Bureau of Economic Research*. doi:10.7249/wr1063-1
- Matthews, D. (2017, October 26). Congress quietly passed a budget outline with \$1.8 trillion in health care cuts. Retrieved April 17, 2018, from http://www.vox.com/policy-andpolitics/2017/10/26/16526458/2018-senate-budget-explained
- Mcguire, M., & Iuga, A. (2014). Adherence and health care costs. *Risk Management and Healthcare Policy*, 7, 35-44. doi:10.2147/rmhp.s19801
- Meijer, C. D., Wouterse, B., Polder, J., & Koopmanschap, M. (2013). The effect of population aging on health expenditure growth: A critical review. *European Journal of Ageing*, 10(4), 353-361. doi:10.1007/s10433-013-0280-x
- Morse, S. (2017, October 23). Billions in proposed budget cuts to Medicare, Medicaid threaten provider compensation. Retrieved April 16, 2018, from http://www.healthcarefinancenews.com/news/billions-proposed-budget-cuts-medicaremedicaid-threaten-provider-compensation
- National Health Expenditure Accounts. (2018, January 08). Retrieved from https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html
- Okunade, A. A., Karakus, M. C., & Okeke, C. (2004). Determinants of Health Expenditure Growth of the OECD Countries: Jackknife Resampling Plan Estimates. *Health Care Management Science*, 7(3), 173-183. doi:10.1023/b:hcms.0000039380.42784.20

- Okunade, A. A., Suraratdecha, C., & Benson, D. A. (2009). Determinants of Thailand household healthcare expenditure: The relevance of permanent resources and other correlates. *Health Economics*, 19(3), 365-376. doi:10.1002/hec.1471
- Okunrintemi, V., Somlo, D., Spatz, E., Khera, R., Valero-Elizondo, J., & Salami, J. (2011).
 Racial Disparities in the Utilization of Cardiovascular Disease Preventive Care Services in the United States: Insight From Medical Expenditure Panel Survey 2014. *Risk Management Healthcare Policy*, 4, 1-7. doi:10.2147/RMHP.S15777
- Ortman, J. M., Velkoff, V. A., & Hogan, H. (2014). An Aging Nation: The Older Population in the United States (United States, Department of Commerce, Economics and Statistics Administration). Washington, DC: US Census.
- Paris, V. (2014, February 07). Why do Americans spend so much on pharmaceuticals? Retrieved April 17, 2018, from http://www.pbs.org/newshour/health/americans-spend-muchpharmaceuticals
- Piabuo, S. M., & Tieguhong, J. C. (2017). Health expenditure and economic growth a review of the literature and an analysis between the economic community for central African states (CEMAC) and selected African countries. *Health Economics Review*, 7(1). doi:10.1186/s13561-017-0159-1
- Prescription drug expenditure U.S. 1960-2018. (2017). Retrieved April 17, 2018, from http://www.statista.com/statistics/184914/prescription-drug-expenditures-in-the-us-since-1960/.
- QuickFacts United States. (n.d.). Retrieved from https://www.census.gov/quickfacts/fact/table/US/PST045216
- Rudowitz, R. (2016). Medicaid Financing: The Basics. *The Kaiser Commission on Medicaid and the Uninsured*.

- Sanz, I., & Velázquez, F. J. (2007). The role of ageing in the growth of government and social welfare spending in the OECD. *European Journal of Political Economy*, 23(4), 917-931. doi:10.1016/j.ejpoleco.2007.01.003
- Schneider, E. L. (1990). The Aging of America. *Journal of the American Medical Association*, 263(17), 2335. doi:10.1001/jama.1990.03440170057036
- Senthilingam, M. (2017, February 23). South Korea will take lead in life expectancy by 2030. Retrieved April 17, 2018, from http://www.cnn.com/2017/02/21/health/life-expectancyincrease-globally-by-2030/index.html
- Seshamani, M., & Gray, A. (2002). The impact of ageing on expenditures in the National Health Service. *British Geriatrics Society*, *31*(4), 287-294. doi:10.1093/ageing/31.4.287
- Shao, H., Mohammed, M. U., Thomas, N., Babazadeh, S., Yang, S., Shi, Q., & Shi, L. (2017).
 Evaluating Excessive Burden of Depression on Health Status and Health Care Utilization
 Among Patients With Hypertension in a Nationally Representative Sample From the
 Medial Expenditure Panel Survey (MEPS 2012). *The Journal of Nervous and Mental Disease*, 205(5), 397-404. doi:10.1097/nmd.000000000000618
- Smith, D. L. (2009). Disparities in patient-physician communication for persons with a disability from the 2006 Medical Expenditure Panel Survey (MEPS). *Disability and Health Journal*, 2(4), 206-215. doi:10.1016/j.dhjo.2009.06.002
- Steinmann, L., Telser, H., & Zweifel, P. S. (2007). Aging and Future Healthcare Expenditure: A Consistent Approach. Forum for Health Economics & Policy, 10(2). doi:10.2202/1558-9544.1041
- Ubel, P. A. (2010). Beyond Costs and Benefits: Understanding How Patients Make Health Care Decisions. *The Oncologist*, *15*(Supplement 1), 5-10. doi:10.1634/theoncologist.2010-s1-5

- UNDP (United Nations Development Programme), Human Development Report. (2013). *Health Index*. New York.
- United States Census Bureau. (n.d.). Selected Characteristics of People 15 Years and Over, by Total Money Income, Work Experience, Race, Hispanic Origin, and Sex. Retrieved from https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pinc/pinc-01.html
- Whalen, J. (2015, December 02). Why the U.S. Pays More Than Other Countries for Drugs. Retrieved April 17, 2018, from https://www.wsj.com/articles/why-the-u-s-pays-morethan-other-countries-for-drugs-1448939481

Appendix

Title	Author	Year Published	Year of MEPS	Findings
Health Care Expenditures for Adults With Multiple Treated Chronic Conditions: Estimates From the Medical Expenditure Panel Survey, 2009	Steven R. Machlin and Anita Soni	2013	2009	Individuals with chronic conditions have substantially larger health expenditures than those with fewer or no chronic conditions
Direct medical expenditure associated with rheumatoid arthritis in a nationally representative sample from the medical expenditure panel survey	Aniket A. Kawatkar et. Al	2012	2008	Individuals with Rheumatoid Arthiritis (RA) are likely to have statistically significant higher healthcare expenditures and pharmaceutical expenditures than those without RA.
Health-Care Expenditures of Overweight and Obese Males and Females in the Medical Expenditures Panel Survey by Age Cohort	Janice F. Bell et. Al	2012	2000- 2005	Observe and overweight individuals had significantly higher expenditures than the control group, mostly due to higher utilization of ambulatory care and prescription drugs. Overweight healthcare expenditures are more substantial and emerge at a younger age in females compared to males.
Incremental Health Care Utilization and Expenditures for Chronic Rhinosinusitis in the United States	Neil Bhattacharyya	2011	2007	Individuals with Chronic Rhinosinusities (CR) are likely to have statistically significant higher healthcare expenditures than the control group. This is due to increases in healthcare service utilization as well as increased pharmaceutical expenditures.
Health Care Access and Perceptions of Provider Care Among Individuals in Same-Sex Couples: Findings from the Medical Expenditure Panel Survey (MEPS)	Joseph B. Clift and James Kirby	2012	1996- 2007	Same-sex couples differ in their healthcare consumption of goods and services from different-sex married couples. These differences include healthcare access, general health, and patient-provider interactions. Differences in other sociodemographic factors such as income, sex, age, or education did not explain these differences.
Racial Disparities in the Utilization of Cardiovascular Disease Preventive Care Services in the United States: Insight From Medical Expenditure Panel Survey 2014	Victor Okunrintemi et. Al	2017	2014	The results of this study found statistically significant racial disparities in the utilization of cardiovascular disease (CVD) preventative services in the United States which highlight significant, necessary growth in improving health equity

Table 1: Previous uses of MEPS Dataset in Economic Literature

Evaluating Excessive Burden of Depression on Health Status and Health Care Utilization Among Patients With Hypertension in a Nationally Representative Sample From the Medial Expenditure Panel Survey (MEPS 2012)	Hui Shao et. Al	2017	2012	Across a spectrum of health services and goods including outpatient, inpatient, and prescription medicine, individuals with depressive disorders are likely to have substantially higher healthcare expenditures while still suffering lower mental and physical health scores.
Educational Attainment and Health Outcomes: Data From the Medical Expenditures Panel Survey	Robert M. Kaplan, Zhengyi Fang and James Kirby	2017	2012	Individuals with higher levels of education attainment were more likely to self-report having better physical health status. This trend cannot be explained by a host of sociodemographic factors that were analyzed in this experiment.

Variable Name Variable Type Definition Age Continuous Variable that categorizes an individual based on what age group they fall into (0: youngest to 3: oldest) Continuous Variable that characterizes an individual's total income in the calendar Income year of 2015 by \$1000 (Total Income/1000) *Lnpharmtotexp* Continuous Variable that is the natural log of the total pharmaceutical expenditures of an individual in the calendar year of 2015. Variable that is the natural log of the healthcare expenditures covered Lnpublicexp Continuous by public insurance of an individual in the calendar year of 2015. Continuous Variable that is the natural log of the total healthcare expenditures of **Lntotexp** an individual in the calendar year of 2015. Continuous -Variable that is used to categorize the different age groups used in the Agegroup String (Summary variable "Age" for the purpose of conducting a summary statistics Variable) analysis. Coverage Continuous -Variable that is used to categorize the different coverage types that an String (Summary individual could have for the purpose of conducting a summary Variable) statistics analysis. Continuous -Variable that is used to categorize different spending groups by total TotspendGroup String (Summary pharmaceutical expenditures for the purpose of conducting a summary Variable) statistics analysis. Black Dummy Variable that characterizes an individual if they identify as Black Dummy Variable that characterizes an individual if they had received or knew Cancer of a previous diagnosis of cancer in the calendar year of 2015. Variable that characterizes an individual if they have at least received ColDeg Dummy their bachelors or associates degree, but also includes individuals with professional degrees. Diabetes Dummy Variable that characterizes an individual if they had received or knew of a previous diagnosis of diabetes in the calendar year of 2015. Employ Dummy Variable that characterizes an individual who was employed at 2 of 3 time points at which the survey was conducted throughout the calendar year of Female Dummy Variable that characterizes an individual who identifies as female. GED Dummy Variable that characterizes an individual if they have at least received their GED but has not received a college degree. Variable that characterizes an individual if they had received or knew Hibp Dummy of a previous diagnosis of high blood pressure in the calendar year of 2015. someHS Dummy Variable that characterizes an individual if they have completed any level of education below receiving their GED. Dummy Variable that characterizes an individual if they were married at the Marital end of 2015. Medicaid Dummy Variable that characterizes an individual if they were covered by Medicaid/CHIP for at least 6 months throughout the calendar year 2015 and no private insurance. (may have other public insurance)

Table 2: Variables with Type and Definition

Medicare	Dummy	Variable that characterizes an individual if they were covered by Medicare for at least 6 months throughout the calendar year 2015 and no private insurance. <i>(may have other public insurance)</i>
Midwest	Dummy	Variable that characterizes an individual who lives in the Midwest of the contiguous United States at the end of 2015.
Northeast	Dummy	Variable that characterizes an individual who lives in the Northeast of the contiguous United States at the end of 2015.
OtherRace	Dummy	Variable that characterizes an individual if they identify as another minority group that is not Black
Private	Dummy	Variable that characterizes an individual if they were covered by Private Insurance for at least 6 months throughout the calendar year 2015 and no public insurance.
PrivatePublic	Dummy	Variable that characterizes an individual if they were covered by Private Insurance and some form of Public Insurance (Medicare, Medicaid, Tricare) for at least 6 months throughout the calendar year 2015.
South	Dummy	Variable that characterizes an individual who lives in the South of the contiguous United States at the end of 2015.
Tricare	Dummy	Variable that characterizes an individual if they were covered by Tricare/CHAMPVA for at least 6 months throughout the calendar year 2015 and no private insurance (may have other public insurance).
West	Dummy	Variable that characterizes an individual who lives in the West of the contiguous United States at the end of 2015.
White	Dummy	Variable that characterizes an individual if they identify as White
Variable		Predicted Sign of
---------------	---	-------------------
Name	Coded Values	Coefficient
Age	0: 10-18	Positive
	1: 19-44	
	2: 45-64	
	3: 65+	
White	1: White (Binary)	Positive
Black	1: Black (Binary)	Negative
OtherRace	1: Other Race (Binary)	Negative
someHS	1: Some High School Completed (Binary)	Negative
GED	1: GED Obtained (Binary)	Negative
ColDeg	1: College/Professional Degree Obtained (Binary)	Positive
Marital	1: Married	Negative
Female	1: Female	Positive
Employ	1: Employed 2/3 of the Year	Positive
Income	Continuous Variable (Table 4)	Positive
Cancer	1: Diagnosed with Cancer	Positive
Diabetes	1: Diagnosed with Diabetes	Positive
Hibp	1: Diagnosed with High Blood Pressure	Positive
Midwest	1: From Midwest	Negative
South	1: From South	Negative
West	1: From West	Positive
Northeast	1: From Northeast	Positive
Medicaid	1: Beneficiary of Medicaid (No Private)	Positive
Medicare	1: Beneficiary of Medicare (No Private)	Positive
Tricare	1: Beneficiary of Tricare (No Private)	Positive
Private	1: Beneficiary of Private Insurance (No Public)	Positive
PrivatePublic	1: Beneficiary of both Private and Public Insurance	Positive
Employ_age	0: Unemployed or Below Age 18	Positive
	1: Employed and Age 19-44	
	2: Employed and Age 45-64	
	3: Employed and Age 65+	
Diabetes_age	0: Not diagnosed with diabetes or Below Age 18	Positive
	1: Diagnosed with diabetes and Age 19-44	
	2. Diagnosed with diabetes and Age 65+	
Cancer age	0: Not diagnosed with cancer or Below Age 18	Positive
cuncer_uge	1: Diagnosed with cancer and Age 19-44	1 Ostive
	2: Diagnosed with cancer and Age 45-64	
	3: Diagnosed with cancer and Age 65+	
Hibp_age	0: Not diagnosed with high blood pressure or Below Age 18	Positive
_	1: Diagnosed with high blood pressure and Age 19-44	
	2: Diagnosed with high blood pressure and Age 45-64	
	3: Diagnosed with high blood pressure and Age 65+	

Table 3: Variables with Predicted Directionality

Table 4: Descriptive Summary Statistics for Dataset									
	(1)	(2)	(3)	(4)	(5)				
VARIABLES	Ν	mean	sd	min	max				
age	16,161	1.804	0.780	0	3				
Inpharmtotexp	12,533	5.817	2.131	0	12.94				
Intotexp	16,161	7.430	1.755	0	13.40				
Inpublicexp	6,861	7.327	2.052	0	12.95				
white	16,161	0.701	0.458	0	1				
black	16,161	0.187	0.390	0	1				
otherrace	16,161	0.0819	0.274	0	1				
income	16,161	41.46	37.67	7.500	409.1				
someHS	16,161	0.151	0.358	0	1				
GED	16,161	0.554	0.497	0	1				
ColDeg	16,161	0.294	0.456	0	1				
marital	16,161	0.513	0.500	0	1				
female	16,161	0.542	0.498	0	1				
northeast	16,161	0.162	0.369	0	1				
midwest	16,161	0.209	0.406	0	1				
south	16,161	0.363	0.481	0	1				
west	16,161	0.266	0.442	0	1				
diabetes	16,161	0.129	0.336	0	1				
cancer	16,161	0.112	0.315	0	1				
hibp	16,161	0.395	0.489	0	1				
tricare	16,161	0.0189	0.136	0	1				
medicare	16,161	0.157	0.364	0	1				
medicaid	16,161	0.145	0.352	0	1				
private	16,161	0.519	0.500	0	1				
PrivatePublic	16,161	0.109	0.312	0	1				
employ_age	16,161	1.067	0.872	0	3				
income_age	16,161	74.33	78.37	0	1,227				
diabetes_age	16,161	0.298	0.810	0	3				
hibp_age	16,161	0.877	1.175	0	3				

Table 5: Correlation Matrix of Independent Variables in Dataset

⊳ e																						
ivate 🔻 a																					-1	-0.3923
are v pr																				-	-0.1442	0.0746
icatio + into																			1	-0.0353	-0.4287	-0.0524
are v meo																		1	0.1926	0.1428	-0.4481	0.5112
▼ medi																	1	0.2822	0.0601	0.0368	-0.2326	0.4331
 Nibp 																1	0.1419	0.1878	0.0215	0.046	0.1589	0.2948
× cancer															-1	0.058	.3025	.1888	- 10801	.0061	- 1664 -	.2481
 diabete 														1	2106	2315	3097 0	5319 0	2282 0	0941 0	5426 -0	5087 0
▼ employ													1	358	467 -0.	162 -0.	717 -0.	026 -0.	224 -0.	174 -0.	192 0.	318 -0.
▼ West												1	1 8	303 0.0	505 -0.0	0.0- 140	-0.0	148	542 0.0	258 0.0	224 0.0	016 -0.0
▼ south											1	11	91 -0.4	25 -0.0	79 0.0	35 0.0	04	04 0.0	47 -0.0	19 0.0	48 -0.0	11 0.0
▼ midwest										1	1	4 -0.38	5 -0.30	2 0.00	3 -0.01	7 0.01	8.	3 -0.03	5 -0.03	3 -0.01	2 0.03	4 -0.00
northe-t											-0.226	-0.332	-0.26	-0.006	-0.003	-0.000	-0.01	0.006	60.0	-0.041	-0.032	0.018
female										0.001	-0.0074	0.0189	-0.0146	-0.061	-0.0223	0.0184	-0.0448	0.0246	0.0916	-0.0277	-0.0387	0.0113
vantel								1	-0.1453	-0.0309	0.0232	-0.0181	0.0242	0.1186	-0.001	0.0054	-0.0136	-0.1101	-0.234	0.0015	0.1927	0.0749
olDeg ▼							1	0.16	-0.0138	0.016	-00.09	-0.0379	0.0362	0.1364	-0.0985	0.0065	-0.1148	-0.134	-0.2055	-0.0207	0.2424	-0.0274
ED →						1	-0.7204	-0.0737	0.02	-0.0197	0.0488	0.0248	-0.0555	-0.0004	0.0274	0.005	0.0435	0.0024	0.0145	0.0356	-0.0511	-0.0161
omeHS v G					-	-0.4708	-0.2727	-0.1013	-0.0101	0.007	-0.0562	0.0138	0.0308	-0.173	0.0873	-0000	0.0856	0.1671	0.2414	-0.023	-0.2374	0.0572
ome v s				-1	-0.2261	-0.213	0.4101	0.2033	-0.1443	0.017	-0.0049	-0.0637	0.0597	0.2867	-0.0944	-0.0177	-0.0845	-0.1941	-0.2656	-0.0122	0.3299	-0.0153
ierre 🔻 in				0.083	-0.0354	-0.1046	0.1419	0.0812	-0.0154	-0.0133	-0.0717	-0.0931	0.1784	0.041	-0.0176	-0.0615	-0.0413	-0.0126	-0.0181	-0.0165	0.0605	-0.0186
× 01		-	-0.143	-0.1068	0.0079	0.0885	-0.1027	-0.166	0.065	-0.0002	-0.0396	0.2089	-0.1908	-0.0452	0.0747	-0.0595	0.1242	0.0351	0.1184	0.0001	-0.0591	-0.0089
ite v bl6	1	-0.7341	-0.4577	0.0469	0.0162	-0.0162	0.0049	0.111	-0.0524	0.0123	0.0799	-0.119	0.0459	0.0126	-0.0494	0.0903	-0.0775	-0.0204	-0.1012	0.004	0.0156	0.0354
▼ Wit	e.	ž	errace	me	teHS		Deg	ital	ale	theast	west	t,		yold	betes	cer		dicare	dicaid	are	ate	
	whit	blac	othe	inco	som	B	ColL	mar	fem.	nort	mid	Sout	wes	emp	diat	Can	hibp	med	med	trica	priv	age

			Insurance C	overage					
Expenditure by	Medicaid	Medicare	None	Private	Private &	Tricare			
Quartile	Only	Only		Only	Public	only			
\$0-\$149	11.19	2.05	15.33	8.6	1.87	6.72			
\$150-\$1,999	3.5	0.87	3.99	3.6	0.79	1.49			
\$2,000-\$349,000	85.31	97.08	80.64	87.76	97.34	91.79			
\$350,000+	0	0	0.04	0.04	0	0			
Pearson Chi2(15) = 453.8807 Pr = 0.000									

Table 6: Chi-squared Results: Insurance Coverage and Total Healthcare Expenditures

Table 7: Chi-squared Results: Age Group and High Blood Pressure

	Age Group							
High Blood Pressure	0	1	2	3				
0	100	83.34	53.49	29.32				
1	0	16.66	46.51	70.68				
		Pearson Chi2(3) = 3.0e+03	Pr = 0.000					

Table 8: Chi-squared Results: Age Group and Cancer

	Age Group						
Cancer	0	1	2	3			
0	100	97.27	89.45	71.87			
1	0	2.73	10.55	28.13			
		Pearson Chi2(3) = 1.5e+03	Pr = 0.000				

Table 9: Chi-squared Results: Age Group and Diabetes

	Age Group								
Diabetes	0	0 1 2 3							
0	100	96.33	83.83	75.16					
1	0	3.67	16.17	24.84					
		Pearson Chi2(3) = 1.0e+03	Pr = 0.000						

Table 10: Chi-squared Results: Age Group and Employment

	Age Group						
Employed	0	1	2	3			
0	88.18	10.5	21.88	80.96			
1	11.82	89.5	78.12	19.04			
	F	Pearson Chi2(3) =	5.8e+03 Pr = 0.000	D			

Table 11: Chi-squared Results: High Blood Pressure and Employment

High Blood Pressure							
Employed	0	1					
0	19.13	48.35					
1	80.87	51.65					
	Pearson Chi2(1) = 1.5e+03	Pr=0.000					

	Cancer	
Employed	0	1
0	26.88	60.76
1	73.12	39.24
	Pearson Chi2(1) = 866.3108	Pr=0.000

Table 12: Chi-squared Results: Cancer and Employment

Table 13: Chi-squared Results: Diabetes and Employment

	High Blood Pressure						
Employed	0	1					
0	19.13	48.35					
1	80.87	51.65					
	Pearson Chi2(1) = 1.5e+03	Pr=0.000					

	(1)	(2)	(3)
VARIABLES	Total Healthcare Expenditures	Pharmaceutical Expenditures	Public Expenditures
white	0.239***	0.233***	0.143***
	(0.0273)	(0.0364)	(0.0482)
income	0.00217***	0.000774	-0.00375***
	(0.000379)	(0.000530)	(0.000877)
ColDeg	0.220***	0.238***	-0.0917
0	(0.0288)	(0.0401)	(0.0607)
marital	-0.0438*	-0.0774**	-0.360***
	(0.0258)	(0.0351)	(0.0477)
female	0.367***	0.173***	0.0416
	(0.0251)	(0.0344)	(0.0461)
northeast	0.232***	0.273***	0.235***
	(0.0330)	(0.0444)	(0.0565)
employ	-0.920***	-1.516***	-1.800***
	(0.0906)	(0.128)	(0.138)
diabetes	1.304***	2.469***	1.200***
	(0.128)	(0.163)	(0.233)
cancer	1.053***	0.915***	0.383
	(0.151)	(0.190)	(0.302)
hibp	0.473***	0.690***	0.438***
	(0.0783)	(0.105)	(0.157)
tricare	0.527***	0.263**	
	(0.0883)	(0.105)	
medicare	0.860***	0.922***	
	(0.0592)	(0.0739)	
medicaid	0.601***	0.605***	
	(0.0450)	(0.0584)	
private	0.755***	0.690***	
	(0.0400)	(0.0541)	
PrivatePublic	1.150***	1.001***	
	(0.0610)	(0.0776)	
age	0.0999**	0.130**	0.124**
	(0.0408)	(0.0562)	(0.0498)
employ_age	0.168***	0.398***	0.309***
	(0.0400)	(0.0548)	(0.0604)
diabetes_age	-0.250***	-0.499***	-0.180**
	(0.0525)	(0.0664)	(0.0866)
cancer_age	-0.222***	-0.290***	0.0347
	(0.0588)	(0.0731)	(0.107)
hibp_age	-0.0280	-0.0556	-0.0318
	(0.0369)	(0.0489)	(0.0619)
Constant	6.108***	4.514***	7.216***
	(0.0943)	(0.134)	(0.124)
Observations	16,161	12,533	6,861
R-squared	0.240	0.270	0.232
Root MSE	1.5312	1.8221	1.8003
F-Stat P-Value	0.0000	0.0000	0.0000

Table 14: Regression Results, by Expenditure Type

Robust standard errors in parentheses

^{***} p<0.01, ** p<0.05, * p<0.1

Table 15: Regression Results, by Insurance Plan				
	(1)	(2)	(3)	(4)
VARIABLES	Medicare	Medicaid	Tricare	Private
white	0.230***	0.239***	0.229***	0.230***
	(0.0276)	(0.0276)	(0.0276)	(0.0276)
income	0.00318***	0.00334***	0.00310***	0.00277***
	(0.000383)	(0.000385)	(0.000381)	(0.000383)
ColDeg	0.278***	0.284***	0.272***	0.258***
-	(0.0290)	(0.0291)	(0.0290)	(0.0290)
marital	-0.0241	-0.0101	-0.0300	-0.0502*
	(0.0257)	(0.0259)	(0.0257)	(0.0259)
female	0.390***	0.384***	0.389***	0.382***
	(0.0254)	(0.0254)	(0.0254)	(0.0254)
northeast	0.244***	0.228***	0.246***	0.247***
	(0.0335)	(0.0336)	(0.0336)	(0.0335)
employ	-1.293***	-1.314***	-1.347***	-1.348***
	(0.0607)	(0.0596)	(0.0594)	(0.0594)
diabetes	1.196***	1.156***	1.158***	1.140***
	(0.129)	(0.128)	(0.129)	(0.129)
cancer	0.912***	0.853***	0.875***	0.804***
	(0.153)	(0.152)	(0.152)	(0.153)
hibp	0.251***	0.181**	0.188**	0.140*
	(0.0757)	(0.0746)	(0.0746)	(0.0753)
employ age	0.287***	0.288***	0.292***	0.251***
eb.o/8e	(0.0266)	(0.0266)	(0.0266)	(0.0275)
diabetes age	-0.191***	-0.177***	-0.171***	-0.162***
	(0.0527)	(0.0526)	(0.0527)	(0.0527)
cancer age	-0 142**	-0 110*	-0 123**	-0.0912
cancer_age	(0.0591)	(0.0588)	(0.0588)	(0.0591)
hibn age	0 113***	0 150***	0 150***	0 175***
	(0.0343)	(0.0335)	(0.0335)	(0.0340)
medicare age	0.0700***	(0.0000)	(0.0000)	(0.00 10)
	(0.0148)			
medicaid age	(0.0110)	0.110***		
		(0.0198)		
tricare age		(0.0130)	0.0526	
theare_age			(0.0325)	
private age			(0.0000)	0.103***
h				(0.0181)
Constant	7.018***	7.011***	7.077***	7.069***
	(0.0490)	(0.0483)	(0.0468)	(0.0467)
Observations	16,161	16,161	16,161	16,161
R-squared	0.212	0.212	0.211	0.212
Root MSE	1.5594	1.5589	1.5603	1.5588
F-stat P-value	0.0000	0.0000	0.0000	0.0000

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1