The Healthy, the Wealthy, and the Wise: The Study of Geographic and Socioeconomic Variation in Health Outcomes

Alison Margolis
Faculty Advisor: Professor Jyoti Khanna
Colgate University, Department of Economics
Hamilton, NY 13346

April 2013

Abstract:

I examine whether multiple health-related quality of life (HRQoL) measures are stratified by socioeconomic status (SES) and geographic location in the United States, and whether these disparities come from the demand side of the patient or the supply side of physicians. Data are from the 2005/2006 National Health Measurement Study (NHMS), a telephone survey of a nationally representative sample of U.S. adults. Regression analyses test whether education, income, and geographic location each have independent associations with three different HRQoL measures. The results suggest that SRH is determined primarily from the demand side variables, specifically, education and income. However, geographic location does play a role in the choices that individuals make, which in turn affect their health outcomes.

I. Introduction

Access to health care is one of the most important issues in health care today, as it is the most relevant element of the Affordable Care Act enacted in 2009 and plays an integral role in the health status of individuals. The issue of health insurance coverage is a persistent question driven by the decisions of policy-makers and others that monitor health insurance programs such as Medicare, Medicaid and other medical assistance programs. More important is the study of the health care outcomes of these individuals and the gap that exists between those who do have health insurance coverage and those who do not have coverage. However, access to care is not the only determinant of individual health outcomes. Health also depends on the area in which the individual resides, that is, an urban or rural area.

During the past few decades, Americans have continued to experience improvements in health, as seen by increased life expectancy and declining rates of morbidity. Life expectancy, for instance, has increased from 70 years in 1980 to 80 years in 2008 (U.S. Census Bureau). While some of these improvements are driven by technological innovations, others result from a growing awareness of health-related issues that arise from behavioral modifications. A case in point is the prevalence of smoking in the United States, in which the percent of individuals who smoke decreased from 25.3% in 1990 to 21% in 2009 (U.S. Census Bureau). However, these improvements in health have been uneven across both geographical areas and socioeconomic status (SES) levels. The health of individuals who live in rural areas, for instance, has improved less than the health of those who live in urban areas (Eberhardt and Pamuk, 2004). In 2001, the death rate among persons who lived in the most rural

counties was 31% higher than those who lived in the most urban counties (Eberhardt and others, 2001). One explanation for the poorer health outcomes in rural areas could be the lack of access to medical care. However, rural areas are also characterized by higher rates of poverty, lack of health insurance coverage, and limited access to primary health care providers (Healthreform.gov).

Many studies have shown that there is an uneven distribution of physicians around the United States. Physicians are less likely to locate in rural areas than in urban areas, and in addition, the more highly specialized the physician, the less likely he or she will settle in a rural area. In 2011, while 19.2% of the U.S population lived in rural areas, only 11.4% of physicians practiced in rural locations (Rosenblatt, Chen, Lishner, 2010). Furthermore, the Bureau of Health Professions' Office of Shortage Designation reports that in 2011, 65% of the primary health care professional shortages were in rural America. As a result of the shortage of access to care in rural areas, individuals must travel to obtain needed care, which may be inconvenient and costly for them. Jeremy Mattson (2010) demonstrated in a study that the average distance traveled was 17 miles for routine health checkups, 42 miles for chronic diseases, and 17 miles for emergency care. Because of this, residents of rural areas are unable to seek the medical care they need, and therefore, face worse medical conditions than those who live in urban areas.

On the other hand, health outcomes may not be due to the supply of physicians, but rather due to socioeconomic status and lifestyle choices of the individual. Those who have a lower SES tend to live in worse housing conditions and are exposed to a more disadvantaged surrounding environment. In addition, the exchange of ideas through news, media, and neighbors provides access to information on how to maintain a good

health status. In urban areas, there is constant communication and contact with others, exposure to media, access to healthier foods and other resources that improve one's health, while those who live in rural areas are usually isolated from these resources. Therefore, it is not surprising to see that those who live in rural areas fare worse when it comes to health than those who live in urban areas. Furthermore, the more socioeconomic disadvantaged areas may diminish the resources necessary to maintain organizations such as churches, schools and other clubs that act as social networks and provide a flow of information between residents.

Policy makers have recognized that communities at different urbanization levels differ in their demographic, environmental, economic, and social characteristics and understand that these characteristics greatly influence the magnitude and types of health care problems that several individuals face. However, the question still remains whether it is access to medical care or socioeconomic status that affects health outcomes. What causes this gap in health care outcomes between individuals? Are health outcomes better in areas with a larger supply of physicians than in those with fewer physicians? Do demographic and economic characteristics directly affect individual health? Or is geographic location an issue, in which those who have better health choose to live in areas with healthier resources? Many economists believe that socioeconomic status is the explanation for the gap in health care outcomes between individuals, while others, such as Culter and Sheiner (1999) find that demographic factors explain less than 5% of health variation. I will try to answer these questions by integrating concepts from the Health Behavior Model (HBM), which was developed by Ronald Andersen in 1975. This model considers three determinants of health care use: predisposing factors (demographic

characteristics), enabling factors (health insurance and affordability), and need for care (health status and disability). The HBM considered transportation and geography as factors that could enable or impede health care utilization. Using this model as a framework to understanding the determinants of health care use and using survey data from 2005 and 2006, this paper will explore the differences in health measures among rural and urban residents and determine the factors that contribute to health outcomes.

II. Literature Review

A number of papers have studied the effects of socioeconomic status and location on individual health, noting that education and financial resources are the driving factors behind poor health. A number of other papers have also studied the geographical preference of physicians and hospitals and the effect of the supply of physicians on individual health. Few, however, have examined the health care gap between urban and rural areas looking at it from both the supply side and the demand side. It is important to determine if the health outcome disparities between rural and urban areas come from the number of physicians available or if the differences come from the demand for health care from the patient. I address this gap by investigating whether neighborhood socioeconomic and geographic levels are associated with health outcomes, net of individual characteristics. By doing this, I will try to determine if the health care disparities in rural and urban areas are a result from supply side or demand side issues.

Because the U.S population is highly segregated by income (Massey 1996), there may be lower levels of health care utilized in impoverished communities because the residents of these areas are unable to afford the care they need. For example, Andersen et al. (2002) found that those living in areas with high rates of poverty and unemployment

were less likely to have seen a doctor in the previous year than those living in other areas. This study suggests that there is an association between community level characteristics and health care utilization. However, in this study, health care utilization is an indicator of access to care. Therefore, it does not capture the healthy individuals with high incomes and generous insurance plans who could have greater access to care but choose not to use the care because they do not need it. Therefore, Kirby and Kaneda (2005) expand on this study and examine how neighborhood socioeconomic disadvantage is associated with access to health care. This study examines the effects of both community- and individual-level factors on access to care simultaneously. It measures access to health care explicitly, uses data based on a nationally representative sample of individuals, and uses individual and community-level data in the same year. Kirby and Kaneda use the behavioral model of care utilization, in which the determinants of health care use are grouped into three categories: demographic characteristics, health insurance and affordability, and health status and disability. Using the Medical Expenditure Panel Survey, this study found that living in disadvantaged neighborhoods reduces the likelihood of having a usual source of care and of obtaining recommended preventive services, while it increases the likelihood of having unmet medical need. However, they also find that these associations are not explained by the supply of health care providers. Expanding on this study, I will use the behavioral model of care utilization, as a framework to understand the determinants of health care utilization in urban and rural areas.

More specifically suited to looking at health outcomes in rural areas and spatial behavior, Arcury et al. (2005) determines the importance of geography and spatial

behavior as predisposing and enabling factors in rural health care utilization. This study analyzes the health services utilization of the number of health care visits in the past 12 months for regular check-up care, chronic care, and acute care across different sociodemographic variables. This study focuses on transportation in rural areas as being an important enabling factor in preventing rural residents from obtaining needed care. When there is a poor public transportation system, poorly maintained sidewalks and roads, and less access to personal vehicles, it makes it more difficult for people to find the means to travel to care. Living in rural areas exacerbates this issue and prevents residents from obtaining the care they need, contributing to their worse health outcomes. Looking at several other enabling factors, Arcury et al. found that having a driver's license, use of provided rides, and distance for regular care were significantly related to health care utilization for regular check-up and chronic care. In addition, geographic measures, such as predisposing and enabling factors, were related to regular check-up and chronic care, but not to acute care. Their results show the importance of geographic and spatial behavior factors in rural health care utilization. They also indicate continuing inequity in rural health care utilization that must be addressed in public policy. This study provides motivation behind the objective in my study, as I am trying to show that geographic variation will have an effect on health outcomes for individuals.

Moving away from the effect of geographic issues on health care utilization, other papers have solely looked at socioeconomic status levels and its impact on health. One of these studies, done in 1998 by James P. Smith, examined the effects of socioeconomic status on individual health. Their model included the complex dimensionality of health status, which produces heterogeneity in health outcomes, the two-way interaction

between health and economic status, and the separation of anticipated from unanticipated health or economic shocks. Using the Health and Retirement Survey (HRS) of 7,702 households, Smith found that there exists a strong positive association between levels of household income, or wealth, and self-reported health status. For example, he finds that HRS respondents in excellent health have 2.5 times as much household income and 5 times as much household wealth as respondents in poor health. Smith's paper demonstrates the demand side issue of health outcomes, in that higher incomes will lead to better health outcomes. In addition to income being a possible factor in determining health disparities, we can also look to particular geographical regions. Brett O'Hara and Carole Poppoff (2002) examine the effect of geographical region on the probability of individuals being uninsured. This paper uses a multilevel technique, or a hybrid of synthetic estimation and hierarchical modeling, to increase the reliability of their estimates. They find that state level data combined with demographic data, including age, race, sex, and ethnicity, improves health insurance estimates. Although both of these papers were useful in helping me determine some of the factors that would affect health outcomes, neither of them looked at the supply side issue. Therefore, I looked to several other papers that studied how health varies with the availability of physicians.

Paul Newacheck and others (2003) assess the presence of income gradients using four income groups and outcome variables that include health status, health insurance coverage, access to and satisfaction with care, utilization, and unmet health needs. This study uses survey data from the 1999 and 2000 National Health Interview Series, and focuses on adolescents aged 10 to 18 years old. Although they focus on adolescent health and health care, they find that higher family income is associated with a higher likelihood

of having a usual source of care, obtaining needed care, receiving any care from a doctor, and being seen in a doctor's office. More specifically, adolescents in poor and near-poor families have a five to six times higher likelihood of being uninsured as their counterparts in middle- and higher-income families. Although this study focuses on adolescents, it addresses many of the issues that I examine in my paper and demonstrates the importance of socioeconomic status on health outcomes.

Several other papers have not only looked at socioeconomic status as a factor in health status, but also but also patterns in lifestyle choices in urban and rural areas. Eberhardt and Pamuk (2004) examine differences in health measures among rural, suburban, and urban residents and the factors that contribute to these differences, including mortality, chronic diseases, and risk-factors. Cigarette smoking is one such example. In 1997, the prevalence of cigarette smoking among U.S adolescents and adults was 19% and 32% higher for those who lived the most rural counties relative to their urban counterparts. Eberhardt and Pamuk and many others have also looked at how individuals spend their time during leisure. They found that leisure-time physical inactivity was 50% higher among adults who lived in the most rural counties than among their urban counterparts. Therefore, risk factors such as obesity and smoking appear to be more common among rural residents and are related to higher mortality rates and prevalence of chronic health conditions in rural areas. However, it is important to note that there is a selection issue in these outcomes. There may be patterns of higher rates of obesity and higher mortality rates in rural areas, but individuals with these issues may be choosing to live in rural areas over urban areas. Nonetheless, these articles demonstrate

that the demand for health care is dependent on income, geographical region, and lifestyle choices at the individual level.

On the other hand, many studies have tried to examine the effects of the number of physicians in an area on the health of individuals. In 1997, Kevin Grumbach, Karen Vranizan, and Andrew Bindman examined whether patient's reports of access to care are associated with physician supply or some other population characteristics. They find that residents of rural areas with lower levels of physician supply are more likely to be uninsured, poor, and members of minority groups. These factors explain much, and for some access measures virtually all, of the differences in access to care across physician supply groups. Therefore, this study shows that health outcomes are explained by the characteristics of the population rather than by the supply of the physicians in the area. Taking this study one step further, Barbara Starfield, Leiyu Shi, Atul Grover, and James Macinko (2005) examine the effects of specialist supply on population's health, looking at the relationship between specialist physician supply and death rates. Their findings show a negative relationship between the state-level supply of primary care and death from stroke, infant mortality and low birth weight, and all-cause mortality. However, they found no evidence for a positive effect of specialist physician supply on the population's health. This finding was particularly interesting because it reflects the issue in America's health care system today. The United States performs well when it comes to cancer survival rates, but has very high rates of potentially preventative deaths. This shows that the health care system in the U.S. is dominated by specialist care, but the specialists do not provide the preventive care that general practitioners provide. Both the Grumbach and the Starfield studies show that although a higher supply of physicians may

lead to better health outcomes, it is not the most powerful explanatory variable when looking at why health outcomes vary between areas. Therefore, it must be coming from the demand side as to why individuals in rural, remote areas fare worse when it comes to health than those in urban areas.

Although some of these studies explore how the demographic characteristics of the population determine individual health while others explore how the supply of physicians affect health outcomes, few combine the two sides to see if health outcomes result from the supply or the demand side. Using the aforementioned articles, I will expand on the studies by combining both the supply side of health care and the demand side for health care into one study to see why the gap in health outcomes exists. Rather than only focusing on the demand side, I will also separate the study into urban and rural areas. Looking at health issues from the demand, as well as urban and rural characteristics, I will analyze in my paper why health outcomes differ between individuals and whether it is due to the resources they have available to them or their SES levels.

III. Data

Health is a complex concept to define and measure. Although we can look to see if individuals have specific diseases or disabilities, there is no unique, direct way of characterizing individual general health. Therefore, health-related quality of life (HRQoL) measures were established to quantify overall health by combining observations of various determinants of health (pain, chronic diseases, mental health, and social abilities). In my study, the HRQoL measures used are general self-related health (SRH), body mass index (BMI) levels, and current smoking levels. The data are from the

National Health Measurement Study (NHMS), which is a random-digit-dial telephone survey of a nationally representative sample of 3,844 U.S. adults in the 48 contiguous states, aged 35-89 years old. This survey was conducted between June 2005 and August 2006, with a response rate of 46%. Sampling weights in this survey were computed based on the sample selection algorithm and adjusted by poststratification so that the final weighted case distribution was as close as possible to that of the target population. SRH was asked early in the survey, followed by randomly ordered questionnaire sets used to compute scores for different HRQoL measures. Summary statistics of the population sample is shown in the table below.

Table 1. Sample Characteristics for All Variables

	Mean (Std. Deviation)
Urban	60% (0.231)
<High School	12.7% (0.333)
High School	30.2% (0.459)
Some College	22.3% (0.416)
4-year college or more	34.8% (0.489)
<$20,000	20.3% (0.402)
$20,000-$34,999	18.2% (0.386)
$35,000-$74,999	30.6% (0.461)
>$75,000	30.9% (0.465)
Private Insurance	46.6% (0.500)
Married	47.4% (0.500)
White	66.6% (0.472)
Male	42.7% (0.495)
Chronic Disease	46.3% (0.479)
Number Kids	2.19 (1.50)
Age	60.2 (14.0)

Note: Total number of observations = 3,844; urban = population > 50,000

Measures

The SRH measure is used as an ordinal scale ranging from 1 (poor health) to 5 (excellent health) in which respondents were asked to rank their health as poor, fair, good, very good, or excellent. This measure has been used in several large national surveys, and although it seems like a simple question, it has been a usually robust predictor of later health outcomes, including disability and mortality (Idler & Benyamini, 1997).

The BMI is available from the self-reported height and weight of the individual. The BMI is a measure for human body shape and is calculated by the individual's body mass (in pounds) divided by the square of their height (in inches). The BMI level serves as a proxy for the body fat percentage among ratios of height and weight. In addition, the BMI can be used as a means of classifying sedentary individuals, or populations, with an average body composition. A BMI of 18.5 to 25 may indicate optimal weight, a BMI lower than 18.5 suggests the individual may be underweight, while a number over 30 may indicate the individual is obese.

The current smoking status of the individual is measured on an ordinal scale of 1 to 3, with 1 representing that the individual never smokes and 3 representing that the individual smokes everyday. This smoking status variable is considered as a health condition and risk factor to the individual. The NHMS uses about 11 common health conditions, with follow up questions to differentiate subclasses of the disease and past and current treatment.

Income reflects household income over the past year and is broken up into four categories: less than $20,000, $20,000-$34,999, $35,000-$74,999, and $75,000 and more.

Education was broken up into four categories as well: less than high school graduate, high school graduate, some post-high school education, and a 4-year college degree or more. The location variable was coded as living in an urban or rural area, with the urban variable defined as having a total population of more than 50,000 people. As for the HRQoL measures, I looked at overall self-related health, the BMI level, and the current smoking status level.

Model

In order to examine health variations within socioeconomic status levels and geographic regions, I conducted multivariate ordinal probit regressions for the SRH and smoking measures, and a Weighted Least Squares regression for the BMI levels. All of the regressions controlled for age, gender, race, marital status, and number of kids. There were missing data in some of the variables, with the following percent missing: education (0.5%), income (8.3%), urban (2.7%), and SRH (0.5%). I first examine whether the HRQoL health measures are associated with the geographical region (urban or rural) of the individual alone, and then look to examine if this relationship stays significant when controlling for socioeconomic status levels, specifically, education and income. If health outcomes are primarily dependent on the demand side, rather than the supply side, then adding these SES variables to the model should eliminate the significance of the geographic location on the health measure. Controlling for SES will signal how much location plays a role in the production of individual health outcomes. Analyses of my model were performed using the STATA system that incorporated the survey weights. The final model used in my paper is as follows:

[1] SRH = β_0 + $\delta_1$urban + $\delta_2$less than high school + $\delta_3$highschool + $\delta_4$somecollege + δ_5<\$20,000 + δ_6\$20,000-\$34,999 + δ_7\$35,000-\$74,999 + β_1Insurance + β_2Married + β_3Race + β_4Gender + β_5Age + β_6# kids + ε

[2] Smoke Status = β_0 + $\delta_1$urban + $\delta_2$less than high school + $\delta_3$highschool + $\delta_4$somecollege + δ_5<\$20,000 + δ_6\$20,000-\$34,999 + δ_7\$35,000-\$74,999 + β_1Insurance + β_2Married + β_3Race + β_4Gender + β_5Age + β_6# kids + ε

[3] BMI = β_0 + $\delta_1$urban + $\delta_2$less than high school + $\delta_3$highschool + $\delta_4$somecollege + δ_5<\$20,000 + δ_6\$20,000-\$34,999 + δ_7\$35,000-\$74,999 + β_1Insurance + β_2Married + β_3Race + β_4Gender + β_5Age + β_6# kids + ε

IV. Results

In these models, I examine whether multiple levels of SES (income and education) and geographic location are separately associated with multiple HRQoL measures, including SRH, smoking status, and BMI levels. Three models were estimated for each measure to determine the effects of living in an urban area on individual health outcomes. The first model estimates the effect of living in an urban area on the HRQoL measure without taking into account SES factors; the second model controls for education, while the third model controls for both education and income. The first regression estimates the general SRH score, the second regression estimates current smoking status, and the third regression estimates BMI levels. Tables 2, 3, and 4 summarize the results of these regressions, respectively. Each model controls for marital status, type of insurance, gender, age, and number of kids. For all HRQoL measures, individuals at lower levels of education and income reported lower SRH scores, higher

frequencies of smoking, and higher BMI levels compared to those in the highest education and income levels.

An ordered probit model was used to estimate the general SRH score. The dependent variable was measured on 1-5 scale, ranging from poor health to excellent health, where 1=poor, 2=fair, 3=good, 4=very good, and 5=excellent. The estimated odds ratios of these results are presented in Table 2. Summarizing the results from the final model in this table, we see that lower levels of both education and income have statistically significant negative impacts on the general SRH measure, while living in an urban area is insignificant in determining SRH. Having a lower education level is associated with a decrease in the likelihood of having a better SRH score, relative to having a college degree. Similarly, having a lower income level is associated with a decrease in the likelihood of having a higher SRH score, relative to those with income levels greater than $75,000. For example, on average and estimated at the margin, those with less than a high school degree have a 4.6% (odds ratio = 0.432) decrease in the odds of having a higher general SRH score relative to those with a college degree, and those with an income level less than $20,000 have a 6.7% (odds ratio = 0.631) decrease in the odds of having a higher general SRH score relative to those with an income level greater than $75,000. These coefficients are significant at the 1% significance level. The magnitude of these associations, however, becomes weaker as the income and education level increases. Hence, after some level of income and education is reached, the benefits of these variables plateaus. In addition, the urban/income interaction term demonstrates that higher income levels weaken the effect of living in an urban area on the individual's SRH score. As for the individual characteristics, having private insurance and being

white are positively associated with SRH scores compared to their counterparts at the 1% significance level, while being a male and being older are negatively associated with their SRH score, relative to their counterparts. However, marital status, living in an urban area,

Table 2. Odds Ratios from Ordered Probit Regression on Self-Related Health (SRH)

	SRH (1)	SRH (2)	SRH (3)
Urban	0.168 (0.073)***	0.147 (0.073)**	0.522 (0.312)
Education			
<High School		-0.673 (0.064)***	-0.432 (0.075)***
High School Grad.		-0.372 (0.043)***	-0.247 (0.048)***
Some college		-0.282 (0.045)***	-0.206 (0.048)***
Ref. Group: 4 yr coll			
Income			
<$20,000			-0.631 (0.076)***
$20,000-$34,999			-0.320 (0.063)***
$35,000-$74,999			-0.190 (0.048)**
Ref. Group: >$75,000			
Private Insurance	0.402 (0.039)***	0.329 (0.039)***	0.209 (0.043)***
Married	0.132 (0.037)***	0.110 (0.037)***	0.015 (0.042)
White	0.326 (0.037)***	0.287 (0.038)***	0.217 (0.040)***
Gender	-0.016 (0.036)	-0.029 (0.036)	-0.065 (0.038)*
Age	-0.010 (0.0015)***	-0.009 (0.002)***	-0.009 (0.002)***
Number Kids	-0.006 (0.016)	0.003 (0.016)	0.010 (0.018)
Urban*Income	No	No	-0.074 (0.041)*
# Observations	3,844	3,844	3,844

Robust Standard Errors in parentheses
*Significant at 10%
**Significant at 5%
***Significant at 1%

and the number of kids show no statistical significance in their impact on the general SRH of the individual in the final model. These results demonstrate that once income and education are controlled for, the urban/rural divide in health becomes insignificant, which may suggest that health outcomes are primarily dependent on socioeconomic status, rather than on the location in which the individual resides. This is consistent with previous research, showing that health variation between individuals is dependent on socioeconomic status (demand side), rather than on the availability of physicians around them.

An ordered probit model was also used to estimate the odds of the smoking status of the individual. The dependent variable was measured on a 1-3 scale, ranging from never smoking to smoking everyday, where 1=never, 2=sometimes, and 3=everyday. The estimated odds ratios of how much the individual smokes are presented in Table 3. Although the results shown in Table 2 indicate that living in an urban area has an insignificant effect on general health, the results in the final model in this regression show that the geographic factor has a significant impact on how much an individual smokes. Furthermore, having private insurance, being married, age, and the number of kids also all have a significant, negative impact on how much the individual smokes. The final model indicates that living in an urban area is significantly associated with a decrease in the likelihood that the individual will smoke more frequently, compared to their rural counterparts. This coefficient is significant at the 1% significance level. On the other hand, having a lower level of education and income positively impacts how much the individual smokes, relative to having the highest level of education and income. For example, those with less than a high school degree are significantly more likely to smoke everyday than those with a college degree, and those with income less than $20,000 are significantly more likely to smoke everyday than those with an income of more than $75,000. Both of the odds ratios are significant at the 1% significance level. However, the magnitude of these ratios becomes weaker as the education and income level increase. Compared with the general SRH score, these results demonstrate that both geographic location and SES levels play a significant role in the likelihood of an individual smoking more frequently. As expected, those with lower SES levels will have a higher likelihood of smoking more, relative to those with higher SES levels, while those

who live in an urban area have a lower likelihood of smoking more, relative to their rural counterparts.

Table 3. Odds Ratios from Ordered Probit Regression on Current Smoking Status

	Smoker (1)	Smoker (2)	Smoker (3)
Urban	-0.247 (0.080)***	-0.240 (0.079)***	-0.239 (0.082)***
Education			
<High School		0.462 (0.067)***	0.349 (0.071)***
High School Grad.		0.300 (0.047)***	0.224 (0.049)***
Some college		0.234 (0.053)***	0.185 (0.054)***
Ref. Group: 4 yr. coll.			
Income			
<$20,000			0.369 (0.077)***
$20,000-$34,999			0.204 (0.062)***
$35,000-$74,999			0.127 (0.047)***
Ref. Group: >$75,000			
Private Insurance	-0.256 (0.043)***	-0.201 (0.043)***	-0.104 (0.047)**
Married	-0.232 (0.043)***	-0.211 (0.042)***	-0.141 (0.044)***
Race	-0.044 (0.051)	-0.025 (0.051)	-0.008 (0.051)
Gender	-0.035 (0.039)	-0.022 (0.039)	-0.016 (0.040)
Age	-0.024 (0.019)***	-0.023 (0.002)***	-0.024 (0.002)***
Number Kids	-0.036 (0.015)**	-0.033 (0.015)**	-0.031 (0.015)**
# Observations	2,033	2,033	2,033

Robust Standard errors in parentheses
*Significant at 10%
**Significant at 5%
***Significant at 1%

The last model is a weighted least squares regression of individual BMI levels. The BMI level is calculated by the individual's body mass (in pounds) divided by the square of their height (in inches) and is a continuous variable. BMI can be thought of as a proxy for body fat percentage, or how "healthy" the individual is in terms of exercise and weight control. The results of the final model, presented in Table 4, indicate that all of the independent variables significantly impact the individual BMI level. Living in an urban area, having private insurance, being white, age, and the number of kids are all associated with a decrease in the BMI level at the 1% significance level. For instance, on average, living in an urban area is associated with a 0.731 decrease in the BMI level, relative to

those living in a rural area. Conversely, having lower SES levels is associated with an increase in the BMI level, compared to those with higher SES levels. In this model,

Table 4. Estimates from Weighted Least Squares Regression on BMI Level

	BMI (1)	BMI (2)	BMI (3)
Urban	-0.940 (0.367)***	-0.862 (0.363)**	-0.731 (0.387)*
Education			
<High School		3.37 (0.361)***	2.85 (0.378)***
High School Grad.		1.80 (0.231)***	1.47 (0.242)***
Some college		2.16 (0.247)***	1.97 (0.251)***
Ref. Group: 4 yr. coll.			
Income			
<$20,000			1.70 (0.395)***
$20,000-$34,999			1.13 (0.315)***
$35,000-$74,999			0.535 (0.224)**
Ref. Group: >$75,000			
Private Insurance	-1.15 (0.226)***	-0.708 (0.228)***	-0.363 (0.258)
Married	0.145 (0.218)	0.372 (0.215)*	0.719 (0.242)***
Gender	0.962 (0.192)***	1.15 (0.189)***	1.16 (0.189)***
White	-1.09 (0.248)***	-0.966 (0.245)***	-0.838 (0.244)***
Number Kids	-0.321 (0.084)***	-0.301 (0.083)***	-0.341 (0.086)***
Age	-0.038 (0.009)***	-0.039 (0.009)***	-0.047 (0.009)***
Chronic Disease	2.12 (0.200)***	1.87 (0.199)***	1.79 (0.200)***
R^2	0.0533	0.0810	0.0845
# Observations	3,783	3,783	3,783

*Significant at 10%
**Significant at 5%
***Significant at 1%

having less than a high school degree is associated with a 2.85 higher BMI level, compared to having a college degree, and having income of less than $20,000 is associated with a 1.70 increase in the BMI level, compared to having an income of more than $75,000. Both of these coefficients are significant at the 1% significance level. However, similar to the results in Tables 2 and 3, the effect of education and income on the BMI level becomes weaker as education and income level increases.

All three of these regressions demonstrate that SES levels play an important and significant role in predicting individual health outcomes. As seen in the results presented in Tables 2, 3, and 4, having lower education and income levels is associated with worse

health outcomes, relative to having the highest level of education and income. Although some of the other independent variables become insignificant in the final models, the SES variables consistently stay significant at the 1% significance level. We can separate these health measures into two sides: health status and behavioral choices, where the SRH score represents the individual's health status while the smoking status and BMI level indicate individual behavioral choices that affect their health. We see that when controlling for SES levels, the urban variable becomes insignificant in determining the individual's SRH, but stays significant in predicting the individual's smoking status and BMI level. Therefore, although SES levels may have an affect on general health, the location in which the individual resides, specifically an urban or rural area, influences the choices that they make, that in turn affect their health. This suggests that health status may be determined primarily from demand side variables, specifically education and income, while choices which affect health outcomes are determined primarily from the geographic location variable.

V. Discussion/Conclusion

Although one of the goals of my research was to demonstrate that HRQoL measures were stratified across SES levels and geographic location in the United States, the urban factor showed surprising results. While the education and income levels were consistently the strongest predictors of health in all models, the geographic variable fluctuated throughout. Living in an urban area allows for easier transportation and access to health care, provides awareness about how to maintain a healthy lifestyle, and offers a variety of resources that can improve individual health. Therefore, it is expected that those who live in isolated, rural areas may not be as aware as their urban counterparts of

the consequences of smoking, eating healthy, and getting the right amount of exercise. The results in this study, however, do not directly support this idea, but rather imply a different story. The results of this study are mixed; for individual smoking status and BMI levels, location plays a significant role. For SRH status, however, the urban variable is not significant. This may indicate that once we control for education and income, the geography variable does not pick up any other additional information. This may also suggest a selection bias in the data implying that, on average, individuals with higher income and higher educational levels or those with better health choose to live in urban areas. Once we account for these SES variables, the urban variable fails to have a significant effect. Smoking and BMI levels, on the other hand, which are to some extent behavioral outcomes, are affected by the location variable.

Thus, we could interpret the results to suggest that individuals who live in urban areas have better health, not because of the location in which they reside and the resources or culture around them, but because they are naturally healthier people and hence, opt to live in areas with healthier cultures and environments. However, living in an urban area does impact smoking and BMI levels; but again, we need to exercise some caution in that non-smokers and relatively low BMI individuals may choose to live in urban areas. To correct the inherent selection bias, we would need data that followed individuals who changed location which would ideally pick up any changes in health status that may result from such a change in location. However, that information was unavailable in this dataset. Another issue to consider when thinking about the effects of living in an urban area on BMI and smoking levels is the influence of the culture of the location. It is expected that people will alter their behaviors based on the choices and

actions by people around them. For example, an individual who is surrounded by people who live sedentary lifestyles, who are overweight, or who smoke a lot may assume that these choices are the norm and, therefore allow themselves to take on these unhealthy habits. One way to account for the "culture" of an area could be to analyze the effect of relative BMI levels on the individual BMI levels. This would explain if population weight-related choices affect individual weight-related choices and reveal how much individual choices are influenced by the culture of the area around them. However, there were no variables that measured average BMI levels, so I was unable to account for this in my study.

It is also important to note the subjectivity of the SRH measure. Because the SRH measure is subjective to the individual, the insignificance of the geographic location in the SRH regression may suggest that the individual's SRH is more dependent on their socioeconomic status than on their geographic location. Therefore, having higher (lower) SES may make one "feel" healthier (sicker) than they actually are. This idea illustrates that although SES levels may enable health care utilization and, therefore lead to better health outcomes, it may have a psychological effect on the individual's well-being as well. Despite some limitations to my study, my results are consistent with other studies, which show that health outcomes are dependent on SES variables.

In conclusion, similar to other studies performed on variations in health outcomes across socioeconomic status levels, my results show that higher levels of education and income are associated with better HRQoL measures, but the magnitude of these associations becomes weaker as education and income level increase. In addition, higher income levels weaken the effect of living in an urban area on individual SRH, as

demonstrated by the interaction term. The results from my models may suggest that individual health status is determined primarily from the demand side variables, specifically income and education, but the choices that individuals make, which in turn affect health outcomes, are influenced by geographic location. These results are indicative of the impact of the socioeconomic status that, to some degree, may be related to the residence of the individual. Nonetheless, the results imply that socioeconomic status levels explain much of the health variation that exists in the United States, but geographic location does play a role in influencing individual behavioral choices. Because of the health disparities that exist between education and income levels, we cannot expect overall population health to significantly improve in the United States unless it improves for those in the lowest socioeconomic levels first. Until the gap between the richest and the poorest, and the most educated and the least educated diminishes, health variation across socioeconomic status levels and geographic locations will continue to exist in society.

References

Arcury, Thomas A., Wilbert M. Gesler, John S. Preisser, Jill Sherman, John Spencer, and Jamie Perin. "The Effects of Geography and Spatial Behavior on Health Care Utilization among the Residents of a Rural Region." *Health Services Research* 40:1 (2005)

Eberhardt, Mark S., and Elsie R. Pamuk. "The Importance of Place of Residence: Examining Health in Rural and Nonrural Areas." *American Journal of Public Health* 94:10 (2004.

Eberhardt, Mark S., DD Ingram, and DM Makuc. "Urban and Rural Health Chartbook." *National Center for Health Statistics* (2001)

Grumbach, Kevin, Karen Vranizan, and Andrew B. Bindman. "Physician Supply And Access To Care In Urban Communities." *Health Affairs* 1:16 (1997): 71-86.

Kirby, James B., and Toshiko Kaneda. "Neighborhood Socioeconomic Disadvantage and Access to Health Care." *Journal of Health and Social Behavior* 46.15 (2005)

Newacheck, Paul W., Yun Yi Hung, Jane M. Park, Claire D. Brindis, and Charles E. Irwin, Jr. "Disparities in Adolescent Health and Health Care: Does Socioeconomic Status Matter?" *Health Services Research* 38.5 (2003): 1235-252.

O'Hara, Brett, and Carole Popoff. "Do Regions Have an Impact on a Person's Health Insurance Status? A Hierarchical Modeling Approach for Small Area Estimates." *Journal of Economic and Social Measurement* (2002): 143-59.

Rosenblatt, Roger A., Frederick M. Chen, Denise M. Lishner, and Mark P. Doescher. "The Future of Family Medicine and Implications for Rural Primary Care Physician Supply." *WWAMI Rural Health Research Center* (2010).

Smith, James P. "Socioeconomic Status and Health." *AEA Papers and Proceedings* (1998): 192-96.

Starfield, Barbara, Leiyu Shi, Atul Grover, and James Macinko. "The Effects Of Specialist Supply On Populations' Health: Assessing The Evidence." *Health Affairs* (2005): 97-107

"Transportation, Distance, and Health Care Utilization for Older Adults in Rural and Small Urban Areas." *Small Urban & Rural Transit Center* (2010)

United States. The Congress of the United States. Congressional Budget Office. By David Auerbach and Chapin White.