FACTORS AFFECTING THE PRACTICE LOCATION DECISIONS OF GENERAL PRACTICE DOCTORS IN THE PHILADELPHIA AREA

LUKE STUBY
SPRING 2014

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Finance and Economics
with honors in Economics

Reviewed and approved* by the following:

Edward Green
Professor of Economics
Thesis Supervisor

David Shapiro
Professor of Economics
Honors Adviser

* Signatures are on file in the Schreyer Honors College.
ABSTRACT

The Affordable Care Act will provide insurance for much of the previously uninsured population. The increase in the amount of insured patients sparks the question of access to care and whether there will be equality in care. Doctors face the decision of where to locate in the beginning and during their careers and this study will focus on factors that affect the practice location decisions of general practice doctors. The study will focus on the Philadelphia metropolitan and suburban area. A Poisson regression was employed to predict the location decisions of doctors and determine whether factors such as household income or demographics affected doctor location decisions. Policy implications based on the results are discussed and should help to provide more insight into policies directed toward redistributing doctors.
# TABLE OF CONTENTS

List of Figures ......................................................................................................................... iii
List of Tables .............................................................................................................................. iv
Acknowledgements ................................................................................................................ v
Chapter 1 Introduction .............................................................................................................. 1
Chapter 2 Literature Review .................................................................................................... 3
  Discussion of Previous Work .................................................................................................. 3
  Discussion of Inadequate Areas ............................................................................................ 8
  Literature Review Conclusion ............................................................................................. 10
Chapter 3 Data and Research Method ...................................................................................... 11
  Research and Data Collection Methodology ...................................................................... 11
  Discussion of Model Used .................................................................................................. 12
Chapter 4 Findings and Implications ....................................................................................... 14
  Data Results ........................................................................................................................ 14
  Policy Implications ............................................................................................................ 18
  Unresolved Problems and Issues ....................................................................................... 19
Appendix A Zip Code Maps .................................................................................................... 21
Appendix B Additional Regression Tables ............................................................................. 23
Works Cited .............................................................................................................................. 26
LIST OF FIGURES

Figure 1: Final Poisson Regression for Number of Doctors ........................................15
LIST OF TABLES

Table 1: Average Change from 1980 to 2008 ................................................................. 14

Table 2: Sensitivity Analysis at the 25th and 75th Quartiles .......................................... 17
ACKNOWLEDGEMENTS

I want to thank my parents for guiding me throughout my four years of college and always believing in my decisions.

My warmest acknowledgement goes to my Nana who has made life at college so much easier and never once asked for anything in return.

I’d like to thank Professors Ed Green, Jim Tybout and David Shapiro for guidance and support before and during the thesis writing process.
Chapter 1

Introduction

In the world of health care economics, the “Obamacare” legislation that became law on March 23, 2010 and is currently being enforced is at the forefront of debate. The Affordable Care Act (ACA) changes the landscape for health care and for doctors, who will be stretched ever thinner as all Americans are required to become insured. The Association of American Medical Colleges states that there will be 45,000 primary care physicians needed by 2020 to meet demand (AMAA 2014) and this shortage is attributable to both supply and demand issues. Due to the Affordable Care Act, it is estimated that 32 million more patients will be entering the health care system in 2014 alone (AMAA 2014). This growth in demand greatly outpaces the growth in supply since it takes 7 years for new medical students to practice (AMAA 2014). The lag between the demand and supply shift will inevitably cause problems with service and how service is delivered. To understand the possible changes in service, scrutiny should be given to doctors’ practice locations and the factors that influence the doctors’ decision on where to practice.

This paper will address the questions: “Is the general practice doctor distribution within urban areas skewed based on income inequality or other factors and has this distribution changed in the past twenty-seven years from 1982 to 2009?” This question is very relevant to health care under the ACA because general practice doctors are on the frontline of diagnosing and treating patients and therefore crucial to being equally accessible for patients. As discussed by Laia Palencia (2013), a project researcher for INEQ-CITIES, and colleagues, “An equitable service is one that matches access to need, regardless of ability to pay.” This study will focus on the question of whether truly equitable service, based on Palencia’s quote, exists within urban areas.
by reviewing the distribution of doctors in tangency with social and economic factors. The Affordable Care Act requires everyone to obtain insurance, which matches Palencia’s definition of equitable service, however other costs could be inflicted upon patients. Patients might find unequal service in the form of wasted time, waiting lists or travel costs. A shortage of doctors might lead to wait times that cause patients to travel outside the city limits or cause patients’ conditions to worsen if not treated in a timely manner. This study, by determining the distribution of doctors, will attempt to see if such costs are present.

This study will also bring insight into the distribution of general practice doctors which is not as well-known as the distribution of specialist doctors. Specialist doctors are usually assumed to be located in high wealth or high access areas such as near hospitals. This study intends to illuminate whether general practice doctors vary more in their location and the reasons for such variation. Also, the second question of the change in doctor distribution over the past 27 years will help to answer the question of an uneven distribution of doctors and the driving factors behind their decisions.

A study by Bindman et al. yielded results about doctors’ locations pertinent to this study. They stated, “Communities with high proportions of Black and Hispanic residents were four times as likely as others to have a shortage of physicians, regardless of community income” (1305). This paper will answer the question as to whether race, income or other unobserved variables are affecting how doctors choose their practice locations.

Policy implications based on the data will also be discussed, with emphasis on how to correct doctors’ decisions on where to practice if the doctor distribution is skewed based on income inequality or other demographics. Doctors could reflect on where their services would be needed the most and choose to act based on that information. Policymakers could aid in the doctors’ efforts by creating new policies to encourage doctors to move to underserved areas and even the distribution out.
Chapter 2

Literature Review

Discussion of Previous Work

Previous research on doctor distributions is largely discussed in the context of the entire doctor population or within a rural versus urban setting. Numerous papers cover these topics and provide an analog to researching and discussing the doctor distribution within urban areas only.

Dussault and Franceschini (2006) try to measure doctor geographical imbalances and the many factors that affect doctors’ decisions on where to locate. The definition of an imbalance is more doctors located in one location compared to another based on certain measures. Their focus is primarily to mitigate the difference in doctors to population in rural areas, as urban areas are already well-staffed. In trying to reason why such doctor imbalances exist, two measures are discussed: normative and economic (Dussault and Franceschini 2006). Normative measures focus on the location of doctors based on social norms and economic measures focus on the equilibrium of doctors in the market and classic economic theory of an increase in real wages leading to a larger supply of doctors. After a discussion of measurements, Dussault and Franceschini (2006) explore five factors that highlight the majority of doctor geographic imbalances. The five factors relate to the doctor, how their place of work is organized, current issues within the healthcare system, the structure of public institutions such as education, and the policy environment.

Individual factors relating to the doctor mainly deal with the background of the doctor and current circumstances such as a spouse that affects them. The culture and career advancement are largely related to the second factor of how their place of work is organized. The third factor is general and directly relatable to what the current healthcare system is and how its operation would affect
the doctor. The fourth factor is how national institutions affect the healthcare market and system as a whole. Lastly, the policy environment, relating to cultural norms such as gender inequality, affects doctor location preferences (Dussault and Franceschini 2006).

Although no actual model or data is presented in this paper, Dussault and Franceschini address bridging the gap left by other researchers and suggest further policy implications that could be made in order to solve the doctor geographical imbalance. My proposed thesis directly relates to Dussault and Franceschini’s paper, however focusing more on identifying if a discrepancy in doctor distribution is seen in urban areas and then addressing the implications and factors affecting such a distribution. Dussault and Franceschini give many factors to input as a function of doctor distribution such as medical school locations, workloads, individual factors and demographics of the work area. Many of the factors can be quantified and used to try and create a definitive model of doctor geographical imbalances.

Meredith Rosenthal, Joseph Newhouse and Alan Zaslavsky (2005) examined the change in the geographic distribution of all types of physicians from 1979 to 1999 and measured the access to physicians that patients have within rural areas in the United States. The increase in the number of physicians within that timespan spurred the researchers to better the understanding of the unequal distribution of doctors by area (Rosenthal et al 2005). Rosenthal et al (2005) collected relevant data through several sources: physician data through the American Medical Association, location data from previous studies and research, and population data through the U.S. Census. This data was then used to conduct three types of measures: physicians per population at the county level, how close a patient is to the nearest physician, and the caseloads of the physicians (Rosenthal et al 2005). These three measures were chosen because each offsets an imprecise feature of an analysis. The analysis of physicians per population at the county level allows for comparison by county but does not account for patients being able to move between counties to see a given physician. Therefore, Rosenthal et al (2005) also use the measure of patient distance
to a physician. The final measurement to eliminate imprecision is analysis of caseloads per physician. In total, these three measures together provide a more accurate picture of patient geographical access to physicians. The results of the study indicated that although physicians increased in numbers, there was still a disparity in geographic access between rural and urban areas. Another interesting result was that rural areas closer to urban areas have smaller population to doctor ratios than rural areas away from cities. Lastly, Rosenthal et al (2005) state that such measures used overestimate the disparity in the distribution of doctors which should be taken into account. Rosenthal et al. has extremely relevant research to my study on doctor distributions across time. The study provides deep insight into the geographic distribution of physicians and the access that patients have to physicians. In context with my thesis, the same measures could be performed on a zip code basis to provide a more thorough picture of access and a better way to rank those geographical areas. The methods Rosenthal et al. use give insight on the imprecision of certain studies with a similar analysis and will allow a study with less error to be completed urban only doctor imbalances. This study also helps to provide additional considerations to where I get my data from and its limitations.

In “Spatial accessibility of primary care: concepts, methods and challenges,” Mark Guagliardo (2004) focuses on the spatial accessibility of primary care, how to correctly measure it and next steps to better understanding accessibility and its effect on patients, primarily centered on urban areas. Spatial Accessibility is the combination of supply and the distance of primary care doctors and is much more useful in urban settings. Guagliardo (2005) concentrates on urban areas and explains how some spatial measurements do not matter in an urban setting such as the measurement of distance to nearest physician. In addition to distance to nearest physician, Guagliardo discusses three other measures of spatial accessibility: average distance to a set of providers, physician to population and gravitational models of provider influence. Guagliardo (2005) postulates that other newer methods of measurements, relating to advances in geographic
information systems, will better help researchers determine not only physician accessibility but also help researchers determine how geographic distribution affects the healthcare provided.

There are also many barriers to research of this kind. Not all providers of care are doctors; some are nurse practitioners and physician assistants (Guagliardo 2005). Also, there is no complete list of physicians, as commonly used sources such as the American Medical Association do not represent all of the physician population. In conclusion, Guagliardo (2005) expands outside the realm of just finding skewness in doctor geographical distribution and states that the next area of research for even greater impact is to see how this skewness impacts the patients receiving care.

This research directly relates to measurements I plan on using within my thesis such as the average distance to a set of physicians. He focuses on how to measure the primary care service within urban areas which is a topic where relatively little research has been done. Guagliardo sheds light on some new and different studies that could possibly give a more accurate picture of the geographic distribution of doctors. He also explains some of the restrictions for a study looking at the distribution of doctors, although most limitations cannot be avoided and the limitations will also be discussed in my thesis.

In “Do Physicians Locate as Spatial Competition Models Predict? Evidence from Alberta”, Dr. Malcolm Brown (1993) tests the basic spatial competition model with actual data to see how the model predicts actual doctor counts in specific regions in Alberta. First, Dr. Brown (1993) describes some basic characteristics of a spatial model: “physicians locate where they expect to maximize income, each person in the population represents the same demand for a physician’s services, and each physician has the same production capacity.” The model is another way to predict where doctors will locate based on income and other factors such as population size. The model is flexible enough to bend some of the characteristics such as a doctors are income maximizers and allows for other utilities to be accounted in the model. Dr. Brown (1993) produced the model and discovered that urban areas had too many doctors while rural areas had a
shortage of almost the surplus of doctors in the urban areas. Overall, the hypothesis was correct that a physician is more likely to be located in an area that has a higher population density, however, the model also found significance with other factors such as high-quality restaurant prevalence and proximity to an urban center (Brown 1993). Dr. Brown discusses the policy implications of his research and suggests a different approach than most researchers. Dr. Brown believes that policy costs of giving financial incentives to doctors in order to relocate to rural areas would be too high and mostly ineffective. Instead, he proposes that the problem is immaterial and that with today’s transportation advances, physical distance is a non-factor and not a constraint to receiving quality care (Brown 1993).

Dr. Brown’s research provides interesting dialogue to the topic of doctor distributions and how to predict the doctors in a certain area. In relation to my thesis, I could use a model to replicate his idea of trying to predict the doctor market in Philadelphia and then compare it with the actual data results to test the accuracy of the model. He also mentions the relevance of such a test, and although distance might not be the largest factor, I do believe it affects the caseloads per doctor and can interact with the quality of patient care.

In “Differential Effects of Economic Factors on Specialist and Family Physician Distribution in Illinois: A County-Level Analysis”, Martin Mistretta (2007) addresses unequal doctor distributions in Illinois by using the three-stage least squares estimate to predict reasons why doctors were distributed a certain way. What differs from previous research is Mistretta’s use of the three-stage least squares model which allows the researcher to look at two or more endogenous variables working at the same time (Mistretta 2007). The physician data was gathered from the American Medical Association and the American Osteopathic Association. The model was run with three types of doctors: primary care, family and specialist (Mistretta 2007) using four categories of variables which were supportive facilities, socioeconomic and
demographic characteristics of an area, sociocultural considerations, and need for medical service. The findings for primary care doctors yielded the result that four variables were highly significant: specialist to population, percent system hospitals, public school expenditures and hospital beds per population (Mistretta 2007). For the other relevant group, family doctors, four variables were highly significant: per capita income, percent system hospitals, specialists to total population and total population. From these findings, Mistretta (2007) concludes that economic factors influence the decision of where to locate for family doctors such as per capita income and total population while specialists care more about quality of life factors. Therefore, family doctors will want to stay in high patient volume areas such as urban counties and away from the less populated rural areas. Mistretta (2007) suggests the typical policy approach of providing subsidies for family doctors to encourage movement to these rural areas.

In relation to my research question, this study provides another way to calculate the distribution of doctors in specific locations and is even more predictive in nature. The three-stage least squares model allows for many variables, all of which are not related to household income, as is my thesis. However, as Mistretta discovered, family doctors were most affected by economic factors and the focus of my study is general practice doctors, a bulk of which is family doctors. Mistretta, like many other researchers discussed, gives new variables to consider in my research and pushes the boundaries on what to consider as a whole within my study.

Discussion of Inadequate Areas

The question of a geographical imbalance of doctors has been asked and answered for many years with varying results of success. The typical focus within this realm has been the difference in number of doctors within rural versus urban areas. Not much thought or research has been given to the topic of geographical imbalances of doctors in just urban areas, as urban areas already have a higher population of doctors. This has left a large area to consider when forming policies and determining how care is delivered in the country. Another problem with the existing
literature is the inaccuracy of the measures used and the number of “good” variables to predict where doctors will locate. Many studies have concluded that income is related to where a doctor will locate but the same studies vary on the significance of variables such as association with hospitals, different demographic variables and other non-economic factors. Therefore, every paper has a slightly different set of variables that predict where doctors locate and how accessible care is in those markets. One reason for this is a lack of accuracy due to the wide range of models and analyses used. Researchers have used models that were as simple as the distance between a physician and a group of patients to geographic information systems (GIS) that can help to more accurately measure the distances and distributions of doctors. Researchers also have used models incorporating ten to twenty variables with simultaneity included such as the three least squares model. The point is that there is no uniform or best approved way of researching the discussed topic and this leads to uncertainty about the accuracy of results. Another idea that is contributing to the possible inaccuracy of results is the commonly used data source called the AMA Masterfile. The American Medical Association is the leading authority on gathering data from doctors practicing in the United States. However, this data source sometimes does not provide all the data needed and might not include all doctors in the United State. Therefore, some of the research studies could be inaccurate, although the inaccuracy would be the same throughout all studies.

Lastly, there is inadequate research on how the distribution of doctors affects patient care. Most studies are focused on explaining the mal-distribution of doctors and how to remedy that situation while forgetting to mention what the measurable impact might be on patients and the care they receive in the area. This area of study could lead to more insight on whether the doctor mal-distribution is hurting patient care and what policy implications could be useful.


**Literature Review Conclusion**

Through research of previous studies pertaining to the geographical distribution of doctors, I have been able to see the evolution of studies and new areas that could potentially be helpful to perfecting the research method I will use in my thesis. Using geographic information systems and the different models described in the described research could give a more accurate picture and allow for comparison of the models. Much consideration in the research was put on not only the distribution of doctors but the variables that would cause such a distribution. Variables not related to income comprised a large part of what many researchers used and is something to take into more consideration because determining exactly what causes doctors to settle in certain areas will be crucial to the success of policymakers in trying to get doctors to locate to areas that they may not prefer.
Chapter 3

Data and Research Method

Research and Data Collection Methodology

In order to collect data to analyze the Philadelphia region’s doctor distribution, a set area displayed in Appendix A was defined as the ideal area to focus for the study. The area focuses on Philadelphia and the surrounding suburban areas which consist of 80 zip codes. The study left out any New Jersey or Delaware zip codes on the assumption that residents in the Philadelphia area would stay within state for medical treatment. Also an area of zip codes within Philadelphia, also shown in Appendix A, was removed from the study in order to focus on primarily general practice doctors or doctors not associated with a large teaching hospital. The goal of the selected area is to provide a generally large area in and around a metropolis with a large enough doctor sample to accurately estimate doctor decisions’ on where to locate. Throughout this paper, the area known as Philadelphia County will be referred to as Philadelphia while all other zip codes will be summarized as the Non-Philadelphia area.

Two main groups of data were collected for the study: data on the number of doctors and their zip code location and characteristics of the populations in every zip code in and surrounding Philadelphia. The data was collected in 1982 and 2008 for a comparison across time.

All doctor data was acquired from the American Medical Association and their compilation of doctors. The American Medical Association’s 41st Directory of Physicians in the United States was used for 2008 data and the predecessor, the American Medical Directory, was used for 1982 doctor data. Doctors were chosen by specific practice type, focusing on general practice services. This resulted in 15 acceptable practice types, although most were defined as
internal medicine, pediatrics and general practice. The American Medical Association allows doctors to give a primary and secondary specialty designation and doctors were included whether the indication of general practice knowledge was a primary or secondary specialty. The doctors were also chosen based on position designation. Second year residents and full-time doctors were chosen to best represent the total population of qualified and available doctors.

The second main data group, the population characteristics within the zip codes, is broken into six data categories: mean household income, median household income, percentage of population below the poverty line, population amount, percentage of population that is African American, and the occupancy rate of dwellings. All of the data was acquired from the United States Census Bureau, directly through the website and through the site Social Explorer, which aggregates Census data. Again the data was collected for the years 1980 and 2010 by zip code which are the closest census reports to the data on doctor locations.

**Discussion of Model Used**

The type of model used can greatly influence the outcome of the study and much deliberation was used to find a model that dealt with panel data across time. Due to the nature of the data collected, being panel and count data, the model used is a population-averaged Poisson model. One of the major assumptions in regard to the Poisson model is that the conditional variance equals the mean or \( \text{Var}(y \mid x) = \text{E}(y \mid x) \) (Woolridge 2002). The model’s independent variables produce coefficients that are expressed in log count therefore a one unit change in the independent variable would change the dependent variable by the log of that independent variable’s coefficient. In order to interpret the estimated model, the independent variables’ coefficients have to be exponentiated. This will give the true prediction for the dependent variable \( y \). The Poisson model used was a GEE (Generalized Estimating Equation), meaning that the equation is estimating the average response over the population (Woolridge 2002). This can be
seen small range in the number of predicted doctors for each zip-code, given the regression coefficients.
Chapter 4

Findings and Implications

Data Results

It is important to first talk about some summary statistics for the population of doctors and the characteristics of the zip-code areas used as it greatly influenced the modeling strategy. Table 1 shows the average changes in the Philadelphia and Non-Philadelphia areas from 1980 to 2008 for the variables used in the study.

Table 1: Average Change from 1982 to 2009

<table>
<thead>
<tr>
<th>Variable</th>
<th>Philadelphia</th>
<th>Non-Philadelphia</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Doctors</td>
<td>-28.3%</td>
<td>6.5%</td>
</tr>
<tr>
<td>Mean HH* Income</td>
<td>4.3%</td>
<td>26.5%</td>
</tr>
<tr>
<td>Median HH* Income</td>
<td>-2.9%</td>
<td>11.5%</td>
</tr>
<tr>
<td>% Below Poverty Level</td>
<td>48.9%</td>
<td>56.8%</td>
</tr>
<tr>
<td>Total Population</td>
<td>-14.5%</td>
<td>23.9%</td>
</tr>
<tr>
<td>Race: % White</td>
<td>-7.7%</td>
<td>-17.1%</td>
</tr>
<tr>
<td>Race: % Black</td>
<td>907.6%</td>
<td>560.1%</td>
</tr>
<tr>
<td># of Housing Units</td>
<td>-6.2%</td>
<td>37.7%</td>
</tr>
<tr>
<td>% Occupancy</td>
<td>-1.0%</td>
<td>-1.8%</td>
</tr>
</tbody>
</table>

The number of doctors in Philadelphia dropped significantly over the time period along with a decrease in the general population of Philadelphia. Non-Philadelphia experienced a large increase in population with a more modest increase in the number of doctors. The next test to see if doctors were migrating from the city to the suburbs was to calculate the ratios of doctors to patients across time for both Philadelphia and Non-Philadelphia and see if they were close to equal. The easiest way to do this is to take the change in doctors divided by the change in population for both areas. This yielded the result of a 0.839 doctor to patient ratio for Philadelphia and a 0.860 doctor to patient ratio for non-Philadelphia through time. The ratios illustrate that
although there were different population shifts in the areas, the access to care as a whole stayed the same. Those results shifted the target of the study from solving why doctors would pick suburbia over the city to an individual look at each zip-code and focus on why doctors make their location decisions. The primary focus of how doctors base their decision on which zip-code to locate to will be the wealth of a zip-code, measured by household mean, median income and poverty level and the race make-up of the zip-code, although other variables are accounted for in the regression as well.

The Poisson regression was run in STATA and the results on the main pooled regression is presented in Figure 1.

**Figure 1: Final Poisson Regression for Number of Doctors**

<table>
<thead>
<tr>
<th>GEE population-averaged model</th>
<th>Number of obs</th>
<th>= 160</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group variable:</td>
<td>zipcode</td>
<td>Number of groups</td>
</tr>
<tr>
<td>Link:</td>
<td>log</td>
<td>Obs per group: min</td>
</tr>
<tr>
<td>Family:</td>
<td>Poisson</td>
<td>avg</td>
</tr>
<tr>
<td>Correlation:</td>
<td>exchangeable</td>
<td>max</td>
</tr>
<tr>
<td>Scale parameter:</td>
<td>1</td>
<td>Prob &gt; chi2</td>
</tr>
</tbody>
</table>

| doc | Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|-----|--------|-----------|-------|------|---------------------|
| meaninc | 6.54e-06 | 1.21e-06 | 5.39  | 0.000 | 4.16e-06 to 8.91e-06 |
| medinc | -8.14e-06 | 2.36e-06 | -3.44 | 0.001 | -0.0000128 to -3.51e-06 |
| pov  | -1.259799 | .2781198 | -4.53 | 0.000 | -1.804904 to -0.716943 |
| pop  | .0000207 | .122e-06 | 16.94 | 0.000 | .0000183 to .0000231 |
| black | .0003619 | .0918514 | 0.00 | 0.997 | -0.1796636 to .1803874 |
| _cons | 2.571638 | .1075555 | 23.91 | 0.000 | 2.360833 to 2.782443 |

The independent variables are listed in the following order: mean income, median income, percentage of the population below the poverty level, population and the percentage of the population that is black. All independent variables except black (referring to the percentage of black people in the zip-code) were significant in the model based on a level of significance of 5%. Mean income had a positive coefficient indicating that doctors indeed have a higher tendency to
locate to zip-codes that have a higher income. This is also demonstrated through the median income variable which has a negative coefficient. It might seem counterintuitive but a negative median income coefficient with a positive mean income coefficient could signify that doctors locate in high wealth areas where the distribution of wealth is negatively skewed. However, in reality, doctors would most likely choose a location that has either a high mean or median income. The percentage of people in poverty had the highest negative coefficient indicating that doctors take poverty very seriously when they make their location decisions. Doctors do not want to locate to an area that has extreme poverty because it indicates bad living conditions and less revenue. The population variable also had a positive coefficient. Doctors favor higher population rates because more people ultimately means more potential income within a specific zip-code area. Prior assumptions on the percentage of black people influencing doctors’ decisions was wrong as the variable was insignificant in doctors’ decisions. A few other variables such as Occupancy and Housing were not included in the final regression but can be seen in other regressions in Appendix B. The results from these two variables were either insignificant or do not have a strong impact on the study and were therefore left out.

A way to test the validity of the model is to calculate the differences in the predicted number of doctors for the entire sample set of 80 zip-codes. This test investigates if the non-linearity of the Poisson model proves to be a problem in predicting the data. For the 2010 data, this yielded a difference of 12.42 doctors. In total, the model was only off by 0.01 percent in its prediction of total doctors in the 80 zip-codes. For the 1982 data, the difference in actual versus predicted doctors was higher at roughly 254 doctors or about 15 percent off the actual number of doctors. The non-linearity of the Poisson model is therefore a non-issue with the data used and indicates that the Poisson model is useful to the study.
To understand the sensitivity of the data, the zip-codes at 25th and 75th quartiles of the mean income, percentage of people in poverty and population were calculated and then compared to see the change in predicted number of doctors. To do this, the variable under focus, such as mean income would be assigned the values at the 25th and 75th quartile while the other three variables would be set to equal the average of the whole dataset for both the 25th and 75th quartiles. This allows for isolation of just one variable to understand how sensitive doctors are to the changes of that variable. The following two tables represent the quartile doctor predictions for both 2010 and 1982.

<table>
<thead>
<tr>
<th>Table 2: Sensitivity Analysis at the 25th and 75th Quartiles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Quartiles</td>
</tr>
<tr>
<td>Mean Income</td>
</tr>
<tr>
<td>Poverty Level</td>
</tr>
<tr>
<td>Population</td>
</tr>
</tbody>
</table>

The results from looking at the quartiles are to be expected. The 1982 data yields typical results with the 75th quartiles having a greater number of predicted doctors than the 25th quartiles. The largest difference in predicted doctors between the 25th and 75th quartiles was with the Population quartiles, yielding a difference of 1.03 doctors. This is a very large change compared to the other variables which only yield a difference of 0.28 doctors for mean income and 0.12 doctors for poverty level. The same type of results are seen in the 2009 data. Population changes the predicted doctors the most, moving from 17.45 doctors to 18.23 doctors. However, mean income has a slightly larger effect on the prediction of doctors with a movement of 0.53 between the 25th and 75th quartiles. Referring back to Figure 1, the coefficient for population is bigger than the mean income coefficient indicating that it could have a larger effect on the number of predicted doctors depending on how large the population. Overall, this sensitivity analysis demonstrates
that even with a high income and low poverty level, doctors might not locate to a particular area if the population is not high enough to warrant the location.

**Policy Implications**

Returning to Bindman et al (1996), this study indicates very different results from their conclusion. Communities with high proportions of blacks were more likely to have a shortage of physicians regardless of income. The results of the study indicated that race had no effect on doctors’ location decisions at all. However, unlike Bindman et al (1996), community income was proven to be a major factor in deciding how many doctors located to a particular zip-code area. In fact, income measured as the mean, median or percentage of people below the poverty line all showed significance in the ability to sway doctor location decisions.

The results indicate that policies could be changed that would dramatically affect areas with low access to doctors. First, policymakers should not try to micromanage policies based on demographics. Doctors do not locate based on the racial makeup of the community and therefore any policies toward such a tactic will yield very little to no results. Second, policymakers should not consider subsidizing doctors to go to poor communities in an effort to better distribute the level of care based on where doctors practice. Ultimately this leads to a transfer of the subsidy from the government to the wealthy. Let’s consider an example where the government decides to subsidize doctors in order to get them to offer care to low-income communities. Doctors, being profit-maximizing people, are always trying to compete for the wealthy patients who have better insurance and more profit to offer. Therefore, a likely scenario is doctors will continue to compete against each other for wealthy patients and because of the new subsidy, will lower their prices to entice the wealthy patients. This will ultimately crowd out the poor and provide slow and inadequate service if even treated. Basically, the government subsidized healthcare to the wealthy.
One proposal to get around this problem is to directly subsidize patients to get healthcare such as the Affordable Care Act. This would follow the results of the study and effectively increase the “income” of low-income communities and drive more doctors to the area. The demand for care and doctors is already in place in these communities but gaining the income to make the demand possible is the key to changing the distribution of doctors. The concept of basic supply and demand applies here, if demand increases, the supply of doctors should increase as well.

The key to any proposed policy to help redistribute doctors based on socioeconomic factors is to focus on the demand side, not the supply side. The Affordable Care Act achieves this goal by giving the uninsured in the country a chance to become insured, effectively increasing demand in many communities where it previously did not exist. According to the results of the model, this should be followed by practice locations popping up in communities that would never have been noticed before.

**Unresolved Problems and Issues**

This study could have some factors that make for a less compelling story and should be investigated further. The number of doctors calculated within the study could potentially be inaccurate because classification of practice in the American Medical Association survey was done by the doctors themselves and could reflect personal opinion and bias. The study does not necessarily take into account the ability of doctors to serve areas outside their practice zip code for example, a doctor lives on the border of one zip-code however half his patients actually come from the zip-code next to the one the practice is technically in. This fails to account for a doctor locating in a specific zip-code because of the surrounding zip-codes.

Further research should be done on this study with varying parameters. A period of 30 years might not be enough to accurately predict a large change in how doctors decide to set up
their practice location. Doctors that started their practice in 1980 might still be working in the area, not truly reflecting the arrival of new doctors to the area when older more established practices decide to retire. Also, the area set as Non-Philadelphia was not scientifically chosen and might not be a comparable area to Philadelphia. The area used represents a good understanding on the Philadelphia area and surrounding suburbs but multiple areas should be used in order to achieve consistency in the results. Finally, one city might not tell the entire story of physicians’ location decisions for the United States and the average American city. A multiple city study should be used in order to determine the applicability of these results to areas other than Philadelphia.
Appendix A

Zip Code Maps

1. Map of Northern Zip Codes

![Map of Northern Zip Codes](image1)

2. Map of Southern Zip Codes

![Map of Southern Zip Codes](image2)
3. Map of Areas Not Included
Appendix B

Additional Regression Tables

1. STATA Output: xtpoisson doc medinc pov pop black, pa

|                          | Coef. | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|--------------------------|-------|-----------|-------|------|----------------------|
| **doc**                  |       |           |       |      |                      |
| medinc                   | 3.41e-06 | 9.99e-07 | 3.41  | 0.001| 1.45e-06 to 5.37e-06 |
| pov                      | -.7517031 | .2605137 | -2.89 | 0.004| -1.262301 to -.2411657 |
| pop                      | .0000208 | 1.22e-06 | 17.04 | 0.000| .0000184 to .0000232  |
| black                    | .0076798 | .0917316 | 0.08  | 0.933| -.1721108 to .1874703 |
| _cons                    | 2.307825 | .0966766 | 23.87 | 0.000| 2.118343 to 2.497308  |

2. STATA Output: xtpoisson doc meaninc pov pop black

|                          | Coef. | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|--------------------------|-------|-----------|-------|------|----------------------|
| **doc**                  |       |           |       |      |                      |
| meaninc                 | 2.66e-06 | 5.20e-07 | 5.12  | 0.000| 1.64e-06 to 3.68e-06 |
| pov                     | -.8115897 | .2454698 | -3.31 | 0.001| -1.292702 to -.3304777 |
| pop                     | .0000213 | 1.21e-06 | 17.61 | 0.000| .0000189 to .0000236  |
| black                   | .0069779 | .0919923 | 0.08  | 0.940| -.1733237 to .1872795 |
| _cons                   | 2.30171 | .0728801 | 31.61 | 0.000| 2.159000 to 2.444411  |

3. STATA Output: xtpoisson doc meaninc medinc pov pop occup, pa
4. \textbf{STATA Output: xtpoisson doc meaninc medinc pov pop black occup, pa}

\begin{verbatim}
GEE population-averaged model  Number of obs  =  160
Group variable:               zipcodes  Number of groups  =  80
Link:                         log       Obs per group:  min =  2
Family:                       Poisson    avg =  2.0
Correlation:                  exchangeable max =  2
                                 Wald chi2(5)  =  348.39
Scale parameter:             1         Prob > chi2   =  0.0000

|    Coef. |       Std. Err. |       z |     P>|z|     [95% Conf. Interval] |
|----------|----------------|---------|---------|-----------------------------|
| meaninc  |  4.95e-06      |  1.26e-06 |  3.94   |     0.000       |  2.48e-06     |  7.41e-06 |
| medinc   | -6.21e-06      |  2.39e-06 | -2.59   |     0.010       | -0.000109     | -1.52e-06 |
| pov      | -2.180134      |  0.311694 | -6.99   |     0.000       | -2.791043     | -1.563225 |
| pop      |  0.000214      |  1.22e-06 |  17.53  |     0.000       |  0.000019     |  0.000237 |
| occup    | -2.954438      |  0.644103 | -4.59   |     0.000       | -4.216850     | -1.692019 |
| _cons    |  5.431204      |  0.631839 |  8.60   |     0.000       |  4.192822     |  6.669586 |
\end{verbatim}

5. \textbf{STATA Output: xtpoisson doc meaninc medinc pov pop black housing, pa}

\begin{verbatim}
GEE population-averaged model  Number of obs  =  160
Group variable:               zipcodes  Number of groups  =  80
Link:                         log       Obs per group:  min =  2
Family:                       Poisson    avg =  2.0
Correlation:                  exchangeable max =  2
                                 Wald chi2(6)  =  348.86
Scale parameter:             1         Prob > chi2   =  0.0000

|    Coef. |       Std. Err. |       z |     P>|z|     [95% Conf. Interval] |
|----------|----------------|---------|---------|-----------------------------|
| meaninc  |  4.95e-06      |  1.26e-06 |  3.94   |     0.000       |  2.49e-06     |  7.41e-06 |
| medinc   | -6.23e-06      |  2.39e-06 | -2.60   |     0.009       | -0.000109     | -1.54e-06 |
| pov      | -2.089665      |  0.340605 | -6.14   |     0.000       | -2.757241     | -1.422089 |
| pop      |  0.000214      |  1.23e-06 |  17.49  |     0.000       |  0.000019     |  0.000238 |
| black    | -0.0596402     |  0.926432 | -0.65   |     0.510       | -4.214176     |  1.217371 |
| occup    | -2.905106      |  0.643570 | -4.65   |     0.000       | -4.250909     | -1.761272 |
| _cons    |  5.464695      |  0.631043 |  8.66   |     0.000       |  4.227932     |  6.701461 |
\end{verbatim}
GEE population-averaged model

Group variable: zipcode

Link: log

Family: Poisson

Correlation: exchangeable

Scale parameter: 1

Number of obs = 160
Number of groups = 80
Obs per group: min = 2
avg = 2.0
max = 2

Wald chi2(6) = 324.54
Prob > chi2 = 0.0000

| doc    | Coef.  | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|--------|--------|-----------|-------|-----|----------------------|
| meaninc | 6.23e-06 | 1.24e-06 | 5.02  | 0.000 | 3.80e-06 to 8.66e-06 |
| medinc  | -7.46e-06 | 2.44e-06 | -3.06 | 0.002 | -.0000122 to -.2.68e-06 |
| pov     | -1.245967 | .2794337 | -4.46 | 0.000 | -1.793586 to -.6382479 |
| pop     | .0000157  | 4.48e-06 | 3.51  | 0.000 | 6.94e-06 to .0000245  |
| black   | -.0104961 | .0924147 | -0.11 | 0.910 | -.1916255 to .1706334 |
| housing | .0000133  | .0000110 | 1.13  | 0.250 | -9.73e-06 to .0000364 |
| _cons   | 2.542994  | .111584  | 22.75 | 0.000 | 2.324294 to 2.761695  |
Works Cited


http://www.ij-healthgeographics.com/content/3/1/3


LUKE T. STUBY

Temporary Address: 228 East Foster Ave State College, PA 16801
Permanent Address: 6476 Warden Road New Tripoli, PA 18066

luke.stuby@gmail.com 610-704-3503

EDUCATION

The Pennsylvania State University
University Park, PA
The Schreyer Honors College, The Smeal College of Business Anticipated May 2014
- Bachelor of Science in Finance and Bachelor of Arts in Economics

The Maastricht University Summer Study Abroad Maastricht, Netherlands June 2011-July 2011
- Expanded knowledge of the European Union through coursework on the Economics of European Integration

EXPERIENCE

Mercer Philadelphia, PA
Talent Division Intern June 2013-August 2013
- Developed a holistic five-year people strategy in the healthcare provider industry, creating changes based on market trends and the Affordable Care Act legislation with the goal of lowering labor costs to 45% of operating revenue
- Formed a national employee pay strategy for a consumer goods company by analyzing geographic pay differentials for over 1,000 employees throughout the United States and segmenting the employees into 5 pay-adjusted tiers
- Analyzed executive compensation for many companies, providing insight into current pay practices based on market pricing and peer group analysis, and recommending pay in coordination with the companies’ strategy

Johnson and Johnson Inc. (McNeil Nutritionals LLC) Fort Washington, PA
Sales Finance Co-Op December 2011-June 2012
- Led a 5 person ecommerce team to address critical issues within the business, proposing SKU rationalization and implementing marketing promotions
- Improved the forecasting accuracy of the $500,000 ecommerce business by 10% by analyzing historical buying trends, current market consumption and past to current year sales and sales trends year to date
- Managed a $15M+ selling budget, actively determining potential favorability and risks, and worked closely with the Director and VP of Sales to ensure proper fiscal control

Penergy Solutions (Renewable Energy Startup) University Park, PA
Finance Team Member August 2011-December 2011
- Created startup business plan including five-year financial projections and presented findings to the client, shortening time to market with client’s patent-pending technology

Nittany Lion Fund, LLC University Park, PA
Associate Fund Manager, Materials Sector January 2011-August 2011
- Managed a $150k equity portfolio in a $4M private fund, leveraging comparable and ratio analyses
- Led training in valuation techniques to grow Nittany Lion Fund capabilities within a team of 5 student analysts

ACTIVITIES

- **Schreyer Consulting Group:** Established and managed relationships with consultants to provide students with an in-depth view of the consulting field through numerous events as the Networking and Programs Chair
- **Students Consulting for Non-Profit Organizations:** Part of a 5 person team that created an on-campus marketing plan for a local nonprofit’s storefront to better capture the student population
- **Penn State Dept. of Labor and Employment Relations Tutor:** Recruited by Econ. Dept. to instruct a 5 student group on the fundamentals of labor economics, improving test scores by an average of 7% above mean
- **Sapphire Leadership Program:** Recruited to apply and accepted to 80-member cohort, representing top 8% academically of Smeal College of Business incoming class

INTERESTS AND ACHIEVEMENTS

- Placed 1st out of over 20 teams in the 2012 Deloitte Penn State Case Competition
- Placed 2nd out of 12 university teams in the 2012 Deloitte National Case Competition in Dallas, Texas