2019

Does Medicaid Increase Emergency Room Use: Evidence from Oregon Health Program?

Md Fourkan
South Dakota State University

DOES MEDICAID INCREASE EMERGENCY ROOM USE: EVIDENCE FROM OREGON HEALTH PROGRAM?

BY

MD FOURKAN

A thesis submitted in partial fulfillment of the requirements for the

Master of Science

Major in Economics

South Dakota State University

2019

THESIS ACCEPTANCE PAGE

MD FOURKAN

This thesis is approved as a creditable and independent investigation by a candidate for the master's degree and is acceptable for meeting the thesis requirements for this degree. Acceptance of this does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Myoung-Jin Keay
Advisor Date

Eluned Jones
Department Head Date

Dean, Graduate School Date

This thesis is dedicated to my parents, to whom I am indebted for my whole life and for everything, and my wife, who is a beacon of light for me after frustration over life.

ACKNOWLEDGEMENTS

First, I am expressing my gratitude to Almighty Allah for giving me the courage to endeavor such an intricate work of writing a master's thesis with a very short time, and with scarcity of knowledge. I would like to mention the incredible help of my professor and thesis advisor Dr Myoung-Jin Keay, an excellent mentor and guider, who escorted me in the whole process of generating this master's thesis with lots of encouragement and enthusiasm. The more I bothered him by knocking at his office door without any appointment, the more he inspired me to work with multiple enthusiasm to become excellent. Moreover, I was a fledgling in research field with zero knowledge in my econometrics, and it is Dr Myoung-Jin Keay, who has taught me the ABCs of research and showed me how to produce a thesis paper without prior research knowledge. Later, for understanding what I am doing, i had to resort to Briggs library for hours and hours. Besides, my parents have continuously gave me motivation to work during my depressed time when I was struggling to continue this esoteric project of thesis writing. Besides, my wife constantly fueled me to prosper not only by helping in continuing writing this project but also in my life since her arrival in my life last year. I candidly express my gratitude to my thesis committee members -Dr David Davis and Dr Anna Sadovnikova who supported me by giving corrections, and guidance in this work and throughout my two-year master's degree. I have annoyed Dr Joseph Santos innumerable times, though, not directly for my thesis work but for several academic issues which I couldn't solve myself. So, I express my heartfelt appreciation and respect for him for his enduring help in my two-year master's degree.

TABLE OF CONTENTS

ABBREVIATIONS

ED	emergency department
EEV	endogenous explanatory variable
IV	instrumental variable

LIST OF TABLES

ABSTRACT

DOES MEDICAID INCREASE EMERGENCY ROOM USE: EVIDENCE FROM OREGON HEALTH PROGRAM?

MD FOURKAN

2019

This thesis paper strives to identify the relationship between Medicaid expansion and Emergency Department use. I use a Monte Carlo simulation for demonstrating the endogeneity problem and a copula model using the Oregon Health Program (OHP) data to show the previous literature has exaggerated the causal relation between Medicaid expansion and Emergency Department use. This paper can be divided into two parts. First, it tries to focus on the under-identification of multiple endogenous variables problem in typical econometrics papers, where researchers correct for a single endogenous variable but intentionally or unintentionally ignore the endogeneity of one or more other independent regressors. So, the motivation for first part of this thesis comes from the fact that the previous literature does correct the multiple endogeneity issue. Second, I endeavored to solve this under-identification problem of multiple endogeneity by incorporating a copula regression, along with OLS and 2SLS. The new approach to solve the under-identification problem is a copula method where we have flexibility of using different distribution methods to choose the best one. Using a copula method, we have found that Medicaid does indeed increase the emergency department use, however, not at the rate as the previous literature showed. This is the major contribution of this thesis.

CHAPTER ONE: INTRODUCTION

1.1 Introduction:

To correct the endogeneity problem, researchers include a suitable instrumental variable in a model to get a result, which is at least less biased than not including an instrument. This issue is resolved here if the regression model is simple regression model with one endogenous variable in the regression function. However, in real life, econometrics problems do not allow us to be restricted in model with one endogenous variable. For analysis purposes, we include as many variables in the model as it is required for persuading us and our readers that we didn't deliberately exclude any variable just for the case it is endogenous or other problems such as variable is not observable, or measurement problem prevails among others. We can't just do that for making our work facile and circumventing the critic and scholars in the field to earn accomplishment and contentment. That is quite impossible in this highly competitive world where there is someone in another corner who is doing some addition to the existing work that I am going to evade. So, how to solve endogeneity problem in multiple regression model with more than one endogenous variable. This is exasperated when the multicollinearity exists among the regressor variables in the model of interest. So, our motivation to investigate multiple endogeneity problem by doing a simulation of a dataset of 10000 observations created randomly using Stata software resembling the Oregon Heath Program data to satisfy our quest that having more than one endogenous variable with less than required instrumental variable can cause biasedness, even if we have instruments for some variables. That is what we have shown with our analysis of 1000 simulations by intentionally omitting instruments for one endogenous variable and found biasedness in the instrumented variable. We have found significantly biased results in instrumented variable which is similar to the omitted variable bias as shown in several econometrics books including Wooldridge (2013). In most of the cases, with some exceptions, when the correlation between instrumented variable and dependent variable is positive and correlation between two endogenous variables also positive then biasedness is also positive and vice versa. This is the significant result that i have been able to prove with the simulation study. I am happy to claim it as a significant identification of my master's thesis study. Furthermore, we have also demonstrated in the simulation results, the same biasedness is expunged when we include an instrument for the other endogenous variable. This means when the other endogenous variable is instrumented then biasedness is removed from the results. The significant result is shown in the results and discussions part for the readers with empirical proof. In the second part of this thesis, I made another contribution showing how to improve under identification problem of second endogenous explanatory variable (EEV). In order to solve the under-identification problem that is overlooked by Taubman, Allen et al. (2014) with Oregon Health Program data, we have incorporated a copula bivariate regression model. This copula bivariate regression, which is our main contribution in this paper allow me to use various copula distributions such as Gaussian, Clayton, and Frank to generate the best results. We find that Medicaid expansion causes more use of Emergency department as it was found by Taubman, Allen et al. (2014). Although i get positive relation with Medicaid expansion to ED use, the coefficient is lower than that found by Taubman et al (2014). So, it enhances the acceptance of our copula results since it suggests lower ED use. These are the main contribution of my analysis. Furthermore, in the next sections we will sequentially write a literature review, conceptual framework, data section, method and procedure, results of analysis including Monte Carlo simulation and copula model, and then findings and conclusion of our study.

Keywords: endogeneity, Emergency Department use, biasedness, copula.

1.2 Background on Medicaid

Medicaid is a federal and state government program in the USA that provides medical and health-related services for U.S. Americans with a low income and limited resources. It was initiated in 1965 by signing into law in order to expand the health care facility to indigent Americans and children with poor financial conditions. Statistics show just under 1 billion dollars were spend on Medicaid in the following year of 1966. However, it has expanded a lot and in 2018 a total of 629.3 billion U.S. dollars were expended on the Medicaid public health insurance program. Among this Federal contribution was 393 billion dollar and state contributions was 236.3 billion dollars. Medicaid holds the third largest position after private insurance and Medicare by providing around 17% of total health care bill in the year 2017. In 2008 the number of Medicaid enrollees was around 48 million and in the year 2018, just after 10 years, the total number if Medicaid enrollment has risen to around 75 million (statistica.com).

1.3 Background on Emergency Room use:

Annually there is on an average 136.3 million ED visits in the United States. Among them the number of ED visits from injury related issue was 40.2 million annually. The number of ED visits that result in hospital admission is 16.2 million annually. Only 2.1 million ED visits each year are admitted into critical care facility. In the United States, there are around 42 ED visits per 100 persons. Top five more populous states in the ED visit in per 100 persons are: California, Florida, Illinois, Texas, And New York. In the year 2016 the percentage of ED visits resulting in hospital admission is just 8.7%. In the same year, 39% of ED patients are seen in less than 15 minutes. (beckershospitalreview , Statistics).

1.4 Objective of this Paper:

In this thesis paper, I show that overlooking of endogeneity causes the coefficients of the variable of interest to provide an inconsistent estimate. This is caused by under-identification- a problem related to not being able by researchers to find the possible endogeneity of one or more variable in an econometrics model.

1. By Monte Carlo simulation of Oregon Health Program data, I show that endogeneity of a variable, if remain unidentified, can cause the overall result of a model to be inconsistent.
2. I have tried to find an alternative to 2SLS, to address this under-identification problem and generate more accurate estimates. By using trivariate copula regression I show that under-identification problem of Oregon Health program data can be at least reduced.

1.5 Research Question for objective 1:

Are estimates biased in instrumented variable methods when we have one instrumented endogenous variable and one or more un-instrumented endogenous variable in a model? If yes, then how serious is the bias?

1.6 Null Hypothesis for objective 1:

$H_{0:}$ There is no presence of biasedness arising from endogeneity in a Two Stage Least Square Method when we lack a required instrumental variable.

1.7 Alternate Hypothesis for objective 1:

H_1: There is moderate to high biasedness arising from endogeneity problem in a Two Stage Least Square Method if we do not have the required number of Instrumental variables.

CHAPTER TWO: LITERATURE REVIEW

According to Krochmal and Riley (1994) Emergency Department (ED) overcrowding causes increased cost per patient. This over cost per patient is because of emergency department overcrowding with inpatient admission which causes an increased average stay of ED patients in hospital. They made an analysis on five different medical diagnosis related groups for three years-1988, 1989, 1990 and found that patients admitted via ED spent more than 1 day on average in the hospital. In those three years, 26020 people were admitted via ED, and for 19% of those the length of stay in ED was 11% longer than for the group admitted in inpatient bed on the first day.

According to Salway, Valenzuela et al. (2017) a list of reasons leading to ED overcrowding includes the poor and uninsured who lack primary care, needless visits, the lower social safety net, seasonal illness among others. They indicated overcrowding in ED is a big problem for several countries in the world, including United States of America. They mentioned that ED overcrowding causes for many problems for patients and staff, including extended waiting times, more medical errors and increased patient mortality, and enhanced overall financial losses.

Pines, Zocchi et al. (2016) referred that about seventeen million previously uninsured people got insurance facility under Medicaid in 24 states and District of Columbia in 2014. In addition, federal and state-based marketplace, a service that helps people buy and enroll in affordable health insurance, provided subsidized private health insurance to qualified individuals. After examining 478 hospitals in 36 states in 2014, they found that Medicaid expansion causes 27.1% increase in ED visit, and 31.4% decrease in uninsured ED visit and 6.7% decrease in privately insured ED visit during that year. Overall, total ED visit grew by only 3% in 2014, compared to previous year.

Nikpay, Freedman et al. (2017) tried to identify whether Medicaid expansion through Affordable Care Act (ACA) results in any differences in Emergency Department (ED) use or ED payer mix. They used a difference-in-difference method to compare changes in ED visits per person and the share of ED visits by payer mix (Medicaid, uninsured and private insurance) in treatment groups across 14 states, which expanded Medicaid and control group across 11 states, which did not expand Medicaid. They found that expansion states experience average increase of 2.5 visit per 1000 population than in non-expansion states after 2014. Among the visit types, the increase mentioned to be the largest in injury related visits and states with largest changes in Medicaid enrollment. They also mentioned, in comparison to non-expansion states, there was an increase in share of ED visits covered by Medicaid and there was a decrease in uninsured share in expansion states. Their major finding was that Medicaid expansion under Affordable Care Act (ACA) has made changes in payer mix. In addition, they also found the same result that Medicaid expansion increase ED visits.

Oregon health insurance experiment has so far been used in a few studies, to reveal the effect of health insurance on emergency room use. In his article, Keay (2018) described the justification of using Average Treatment Effect (ATE) estimates and included the ATE estimates along with Local Average Treatment Effect (LATE) estimates which the other authors, doing research emergency room, have used. Taubman, Allen et al. (2014) used discrete random variables for the variable of ED use and pre-randomization (Jan 2007 – March 2008) emergency room (PED) variables. However, by dropping observations with number over 17 for variables PED, which is emergency room use in pre-randomization period, and ED, which is emergency room use in experiment period, the author creates one sub sample. Then followed it by dropping observations over 10 and 7 for ED and PED for creating additional two subsamples. Then the author conducted OLS, IV (LATE), and ATE estimates. In OLS, the author found a significant positive coefficient of .431 which means having insurance still increases the number of emergency room use. Nevertheless, insurance variable is infected with endogeneity since the author suspects those who are willing to go to emergency room are more willing to accept Medicaid. So, in order to escape this drawback, the author used an IV method that offers LATE estimates. Although the LATE estimates are a little smaller than OLS, but they give us a positive estimate of Medicaid on ED use, which means insurance does not necessarily reduce emergency room use.

According to Lowe, Localio et al. (2005) an impaired access to primary care can cause more emergency care use. In summary, they have explored the causal relation between characteristics of primary care practice and the emergency department use. The authors' motivation to work with this causal relation has derived from the argument in the literature that many patients use ED as a substitute for primary care. These characteristics of primary care practices include accessibility for urgent care, administrative characteristics, and availability of specialized equipment. Under these three practice site characteristics there are a few predictor variables that they included in their model. The outcome variable of "ED visits" were counted using claims data, and they included visits even if Health Maintenance Organization (HMO), a group of health care providers which limit care provided through doctors and hospitals who are under contract of HMO, denied payment. They also identified two categories of ED visits- “potentially avoidable” and “probably unavoidable”. The first one is the case in which ED visit could be averted if prompt primary care access had occurred, and the second one is the category in which the patients really need emergency admit. The number of patients included was 57850 who were assigned to 353 primary care practices affiliated with a Medicaid HMO. They found that on average patients made 0.80 emergency department visits per person per year. They also found that practices with more than 12 evening hours per week have used ED 20% less than the patients from practices who do not provide evening hours services. The practice sites which have greater ratio of active patients per clinician-hour of practice time also saw more emergency department visit than those who have the lower ratio. Besides, more Medicaid patients in a practice were associated with more emergency department use. Lack of availability of specialty equipment in practice site also associated with more emergency department use. They concluded that increased access to primary care can facilitate decreased ED use.

Hoot and Aronsky (2008) have tried to identify the causes, effects and solutions of emergency department crowding by making a systematic review of existing literature. According to them, quality of health care and access to it are affected negatively by the crowding in the emergency department. For this robust and substantial systematic review, they have first identified 4271 abstracts of articles and then retrieved 188 full articles from this whole. From these 188 articles only 93 relevant articles included for review and 95 were excluded. After summarizing all 93 articles they have come up to some findings of the causes, effects and solutions of the ED crowding. After analyzing 33 articles which studied causes of ED crowding the authors found that three themes play vital role that cause the use of ED: input factors, throughput factors, and output factors. Nonurgent visits, so-called frequent-flyer patients, and the influenza season are the commonly studied factors that cause crowding. Inadequate staffing is said to be throughput factor that may cause crowding. Also, inpatient boarding and hospital bed shortages are said to be common output factors that may cause crowding. According to the authors the common effects of emergency department crowding that we may observe are adverse outcomes, reduced quality, impaired access, and provider losses. Additional personnel, observation units, and hospital bed access are commonly found solutions of crowding which are related to increased resources. Nonurgent referrals and ambulance diversion are thought to be solution to crowding related to demand management.

In their paper, Finkelstein, Taubman et al. (2012) used the Oregon 2008 experiment project which undertook randomized lottery to provide Medicaid to low income people in Oregon to see the impacts of expanding insurance to uninsured. The whole experiment was done by a randomized controlled project. In many papers which tried to measure the difference of health and heath care utilization between insured and uninsured people tend to overlook that there are differences in terms of demographics characteristics such as income, age, education, previous health condition and employment etc. They have measured the effect of insurance one year after the initiation of the Oregon Medicaid lottery. In their paper they used a randomized trial by selecting a group of people who are both uninsured and who are similar in some demographic characteristics. They have strictly maintained the check for balance among the treatment and control group. By using this randomized trial, they have been able to avoid the discrepancy in the outcomes of the treatment and control group by making an experiment which is rare in social science. This lottery has increased the total insurance covered people in Oregon area during that study period by 25 percent with evidence that not significant changes come from private insurance in that period. It is found in this paper that the subsidized insurance policy for uninsured people should increase the motivation of more heath care facilities use which can enhance health care cost. However, access to more health care services can reduce the burden of emergency department use by uninsured people which can offset the cost of more heath care use. They have gauged and compared their outcome of first year of Medicaid impact on health and cost of expanding this service by using two data set: one from administrative data from different hospitals in Oregon area and the other from survey data from those who are included in lottery in order to compare the effects of insurance on treatment and control group. And one year after expanding this Medicaid lottery their analysis found that the treatment group who were offered Medicaid by lottery has shown greater heath care utilization, lower out of pocket cost for treatment, and medical debt, and better self-reported health than control group (Finkelstein, Taubman et al. 2012). Their result shows that the coverage of insurance (by using lottery as an instrument) has caused this treatment group to a 30 percentage point increase in probability of having hospital admission, a 15 percent increase in the probability of taking prescription drugs, and a 35 percent increase in the probability of having an outpatient visit. However, it does not show any reduction in emergency room use because of expanding this insurance. The result of Finkelstein, Taubman et al. (2012) also show a decrease of 25 percent in the probability of an unpaid bill taken care by collection agency. And a 35 percent decline in the probability of out of pocket medical expense by this treatment group.

CHAPTER THREE: CONCEPTUAL FRAMEWORK

One of the five important econometrics methods, called by some scholars as Furious Five[1], is randomized trials, which means channeling the path for true identification of the causal effect in an econometrics analysis (Angrish and Pischke 2014). But when we try to estimate the causal effect of one variable on certain outcomes, we need to ensure that the *ceteris paribus* condition is met. *Ceteris paribus* means measuring the causal effect of the variable of interest while keeping all other relevant factors controlled or fixed. However, can we really ensure that truly *all other things are fixed*? If not, then this is a real problem for the research will provide misleading results. This is caused by selection bias or the self-selection problem which means deciding an action based on the likely benefits, or costs of taking that action (Wooldridge 2013). Now, the question is how to get rid of this very sticky problem which is prevalent in many econometric analyses. The answer is incorporating random assignment to reduce the selection bias (Angrish and Pischke 2014). The term "Randomization" means creating a subsample from an underlying population by fluke, or by lottery, or by a coin toss or by any other means to create two groups- treatment group and control group. Then, we can run a quasi-experimental design.

When we attempt to implement randomized trials, it is crucial to observe checking for balance which means checking whether treatment and control group are indeed similar particularly in terms of their demographic characteristics (Angrish and Pischke 2014). One excellent example of this randomization approach is the Oregon Health Experiment. The purpose of the Oregon Health trial was to expand Medicaid to a limited number of currently uninsured low income household and to examine whether the insurance coverage has any benefit for the health sector by reducing costly and extravagant emergency room uses

(Finkelstein, Taubman et al. 2012). Certainly, the Medicaid expansion will reach to those who are uninsured. But who among the uninsured will get the coverage for free or for a very small amount of fee? That was a conundrum for the state governments. However, the state of Oregon made this Medicaid expansion experiment possible by publicly offering a health insurance lottery, and randomly choosing the lottery winner from the underlying population. Although the lottery winners and losers were random, the coverage was not automatic for lottery winners. The winners must fulfill requirements, including not having insurance for the last six months, being in age group of 19-64, not being qualified under current Medicaid plan, living below the federal poverty line, and having assets below $2000. They also must be a US citizen or legal immigrant.

3.1 Instrumental Variable (IV) & Multiple Endogeneity

Now we can derive the necessary identification conditions of instrumental variable (IV) estimations for any econometrics research. “In an equation with an endogenous explanatory variable, an IV is a variable that does not appear in the equation , is uncorrelated with error term, and is (partially) correlated with the endogenous explanatory variable”(Wooldridge 2013). IV regression enters into the analysis when we face the difficulty of endogeneity problems which means “an explanatory variable in a multiple regression model is correlated with the error term, either because of an omitted variable, measurement error or simultaneity”(Wooldridge 2013). There are two identification conditions for an instrumental variable to be considered in the IV regression. First, the IV should be uncorrelated with the error term in the model; that means IV is uncorrelated with omitted variables in the model. Second, the IV must be satisfactorily correlated with the endogenous variable. In other words, the IV must be related with endogenous variable either positively or negatively(Wooldridge 2013). In this study, the IV used is a randomized lottery assignment and it fulfils the assumptions of an IV. Taubman, Allen et al. (2014) also showed that emergency room use increases even after providing insurance to the currently uninsured people, which counters to the assumption that insurance might reduce emergency room use. However, in their regression model, there is more than one regressor variable which might be correlated with error term. But they have defined and found an instrumental variable for only one regressor. This term is referred to, in statistical language, as endogeneity which means correlation exists between a regressor variable and the outcome variable of the model. This causes a larger standard error for Two Stage Least Square (2SLS) as compared with Ordinary Least Square (OLS). Taubman, Allen et al. (2014) have satisfied these IV assumptions, which is apparent that, Winning the Lottery (IV) does not have any correlation with any health background of the participants. The IV is uncorrelated with error terms of the model. Besides, to satisfy the other IV assumption condition it is correlated with insurance coverage since only the winner who satisfy some criteria can buy the OHP insurance. According to Wooldridge (2013) it is possible to have more than one endogenous variable in a regression model. If so, for two endogenous variables two Instrumental variables are needed. Nonetheless, Taubman et al (2014) having more than one endogenous variable did not include more than one instrumental variable in their regression model. To be specific about the issue, they did not incorporate an exogenous variable in the model with variable X which is the emergency room use in the pre-randomization periods. The variable X might be endogenous; the people who regularly visited the emergency room earlier will likely have more tendency to visit again in the post-randomization period even if their disease is not of emergency standard.

Human psychology is shaped and affected by the repetition of the same behavior. The people who have experience of frequently visiting the emergency room in pre-randomization period are more likely prone to some physical health problems, those with health problems have more chance to use emergency room during the study period. So, PED might also be endogenous variable in terms of definition of endogeneity which is mentioned already in the paper. Because of not including another exogenous variable in the model for the pre-randomization emergency room use variable, we argue that the coefficient of β_2 in equation 1 might lead to an endogeneity issue. Furthermore, the variance of β_1, the variable of interest, var (β_1) = $\frac{\sigma 2}{SST(1-\mathrm{R1sq})}$ should be larger too. Here, in the equation R_1^2 is the R squared from the simple regression of independent variable of interest on other independent variables in the model. R1sq is a proportion of total variation in an independent variable based on the other independent variables in the model (Wooldridge 2013). When R_1^2 gets closer to one, the var (β_1) gets larger and larger. The issue of larger variance is same for the β_2. A value of 0 in R_1^2 means the smallest value of var (β_1). This is the best case to sustain but it rarely happens. So, although we get a coefficient of β_1 that is not substantially biased, the variance of β_1 should be large if pre-randomization emergency room use and the variable of interest, whether a respondent has insurance or not, are correlated. Second, as we claim that the pre-randomization period emergency room use is substantially correlated with the pre-existing health conditions of the participants; when included in the model, it is endogenous and also requires a valid instrument to ensure unbiased results. The claim of having at least equal number of exogenous variable for the a given number of endogenous variable is evident from the study of Wooldridge (2013) in the discussion of multiple endogenous explanatory variable. So, we are suspicious about the conclusion offered from this analysis of Taubman, Allen et al. (2014), and we are furthermore doing some analysis for this paper to check whether leaving a second endogenous variable un-instrumented in the model can create biasedness in the overall model results. That is one of urgings which led us to invest in this work. In order to prove our suspicion with experiment we create a simulation data set resembling the model of Taubman, Allen et al. (2014). In the next section, we explicate the result of which is congruent to our assumption of the biasedness in the model.

CHAPTER FOUR: DATA SECTION

4.1 Data and Oregon Health Program Details:

Moving on, now, we can focus on the details of data and a brief synopsis of Oregon Health Program. The reference of details can be attributed to Taubman, Allen et al. (2014). The authors of the paper Taubman, Allen et al. (2014) and Finkelstein, Taubman et al. (2012) collected the field-level data for all emergency department visits from twelve hospitals from 2007 to 2009. These twelve hospitals are the ones most residents in Portland area and neighboring suburbs use. They collected the data on emergency department visits with details of name, date of birth, and gender and then matched this information from the Oregon health insurance experiment study with same information mentioned above. Then they were provided data on Medicaid Lottery by State of Oregon. They also collected data on pre-randomization demographic information from the State of Oregon which people provided when they applied for Medicaid to the state. Finally, they conducted survey on respondents across Oregon for almost one year after the lottery draw. And collected data through in-person interview in only Portland area for two years after the lottery. This experiment can be thought of as divided in two distinct stages- pre-randomization stage, and a certain time range during and after expanding the Medicaid to the eligible people through a comprehensive lottery. As we already mentioned there were some criteria that were strictly maintained about whether an individual can be included for the lottery draw. There are variety of discrepancies between insured and uninsured people which invalidate the outcome of any analysis. Thus, the experiment was done by random assignment of lottery which facilitated to isolate the impact of insurance on emergency room use. Jan 2007 to March 2008 was considered as pre-randomization period. In 2008 the experiment drew about 30000 names from a total pool of applications of 90000 for the lottery to measure the effect of Medicaid coverage on emergency room use in post-March 2008 for about 18 months. They not only tried to measure the overall effect of Medicaid on emergency room use but also did the analysis for several types of visits, conditions, and groups. The model that has been used in the Oregon Health Plan (OHP) study by Taubman, Allen et al. (2014) is as follows:

$$y_i = \alpha + \beta_1 X_{1i} + \beta_2 Z_{2i} + \beta_3 Z_{3i} + \mu_i \qquad (1)$$

Here, we have used the variable names in the population model that resembles the original model used in analysis by Taubman, Allen et al. (2014) in their study. y is the emergency room use in the study period from 10 March 2008 to 30 September 2009, and X_1 is insurance coverage by Oregon Health program which they included lottery assignment as an instrument for insurance coverage and Z_2 is the emergency room use history for the pre-randomization. Z_3 includes any other covariates relevant to the model. So, the authors in the paper Taubman, Allen et al. (2014) get credit for devising a randomized trial for a social insurance program, which was unprecedented. They have come up with an idea of matching Oregon health trial data with the Emergency Department use data to figure out causal effects of Medicaid on emergency room uses by using Oregon lottery as an instrument for Medicaid. They have reflected the impact of Medicaid expansion on two aspects. First one is the impact of Medicaid on health of the respondents and the other is the influence on the intensity of emergency department use. The result shows us that Medicaid expansion increases emergency department use. They have found that Medicaid increases access to health care other than emergency room use such as outpatient physician visits, prescriptions and recommended preventive care. It is also reported that self-reported access to and quality of care also improved because of Medicaid expansion. The result also demonstrates that although Medicaid improved self-reported health and decreased depression, different measures of physical health did not produce statistically significant results (Finkelstein, Taubman et al. 2012).

CHAPTER FIVE: SIMULATION AND ENDOGENEITY PROBLEM

5.1 Monte Carlo Simulation and Multiple Endogeneity Problem:

We have started our study very naively and from very simple perspective. So, we have started from very inchoate stage of analyzing an existing paper and to find out any new and valid findings better than Taubman, Allen et al. (2014) paper. Moreover, that is a very recognized paper published in Science magazine where the authors have delineated the use of randomized controlled trial in examining the impact of Medicaid expansion on Emergency department use. This is the crucial paper which motivated us to further explore this study and create a sense of research interest in our mind. As we have demonstrated the model used in the data section part, we have mimicked the same structural data so that we can make distinction of our claim with the contrasting paper. Our model is resembling to the Taubman, Allen et al. (2014) model as provided below:

$$yi = \alpha + \beta_1 W_{1i} + \beta_2 W_{2i} + \mu_i \qquad (2)$$

$$W_{1i} = 1*Z_{1i} + .75*Z_{2i} + \mu_{1i} \qquad (3)$$

$$W_{2i} = .5*Z_{1i} + .75*Z_{2i} + u_{2i} \qquad (4)$$

Using Stata software, we have tried to create data set of 10000 observations in order to make it representative of population and created equations for y, W_1 *and* W_2 where both the latter two are reduced form equations. And equation y is our equation of interest. In equation 2 above y can be thought of emergency room use in the current period. W_1 is whether a person has enrolled in any insurance or not after notification of winning the lottery and W_2 is whether the respondents have used any emergency room pre-randomization period. We have produced two instrumental variables Z_1, Z_2 with random distribution. Then we can express two reduced from equation for explaining W_1 and W_2 as a function of instrumental variable and some error terms. With a view to maintaining the focus of study, we have manufactured the reduced form equation by establishing relation of instrumental variable with endogenous variables. In our simulation work we have tried to maintain the correlation matrix of error of population regression and reduced form equations at different magnitude. We have run each simulation for 1000 times with each error correlation and found out the following results which satisfy our suspicion about the study of (Taubman, Allen et al. 2014). The results of our simulation are shown below to make evidence of our claim that the lack of instrumental variable in a model can cause biasedness even if we have instrument of some other endogenous variable. Finally, these are the summary of the coefficients of both W_1 W_2 with 1000 times simulation where $\beta 1$ is substitute for coefficients of W_1 and $\beta 2$ is substitute for coefficients of W_2.

Table 1: Description of variables used in the simulation analysis

Variable Name	Variable Description
y	Dependent variable
W_1	first endogenous explanatory variable
W_2	second endogenous explanatory variable
Z_1	Instrumental variable used for W_1
Z_2	Instrumental variable used for W_2

note: Our analysis is based on randomly generated data using STATA software. The data set follows normal distribution.

Table 2: Results of Monte Carlo simulation for various error correlations

Coefficients	Model[1]		Model[2]		Correlation among errors[3]
	Mean	Std. Dev.	Mean	Std. Dev.	$(u,u_1), (u,u_2), (u_1,u_2)$[4]
β_1	1.000***	.0148	1.000***	.023	0, 0, 0
β_2	.999***	.012	1.000***	.032	
β_1	1.112***	.014	1.000***	.023	-.2,.2, 0
β_2	.775***	.012	1.000***	.032	
β_1	.888***	.014	1.000***	.032	.2,-.2, 0
β_2	1.224***	.011	.999***	.032	
β_1	.865***	.014	1.000***	.023	.2,-.2,.2
β_2	1.269***	.012	.999***	.032	
β_1	.904***	.014	1.000***	.023	.2,-.2,-.2
β_2	1.191***	.011	1.000***	.032	
β_1	.775***	.013	1.000***	.023	.3,-.3,.3
β_2	1.450***	.012	.999***	.032	
β_1	.866 ***	.013	.999***	.023	.3,-.3,-.3
β_2	1.268***	.010	1.000***	.032	
β_1	1.338***	.011	.999***	.023	-.4,.5,.3
β_2	.323***	.010	1.000***	.032	

*, ** and *** indicate the significance at 10%, 5% and 1% levels, respectively.

1. We have used instrumental variable for one endogenous variable.
2. We have used instrumental variable for both endogenous variables.
3. In column six above a correlation .2,-.2,-.2 means correlation between (u,u1) is .2, (u,u2) is -.2, (u1,u2) is -.2
4. u , u1 and u2 are error term used for respectively population model, reduced form equation for W1 and reduced form equation W2.

5.2 Interpretations of Monte Carlo Simulation:

In the above table of IV estimation, the column 2 and 3 are representing respectively the coefficients and standard errors for β1 and β2 when we have lack of required instrumental variable. Again, column 4 and 5 is showing respectively the coefficients and standard errors for β1 and β2 for the case when we have all required instrumental variable for all endogenous variables. Column 6 is used for showing the different correlation for the error terms based on which I have measured the biasedness of the instrumented variable because of not having instrument for another endogenous variable. So, it is evident from the above simulation project that when we have two instrumental variables for two endogenous variables in the IV regression then the coefficients of both instrumented variables are not showing any biasedness. In all eight variety of error correlations we have seen no deviation of coefficients from *one* which is our population model coefficients. However, when we have lack of one instrumental variable for one endogenous variable in the model then the coefficients of our estimation show deviant behaver which depicts itself that is shown in column 2 where in almost all cases the coefficients of β1 which is instrumented is showing either positive or negative biasedness because of not having instruments of another endogenous variable in the model. Interestingly, there has been observed from this analysis that there is trend of biasedness in the instrumented variable which is inversely correlated between the correlation of u and $u_{2.}$ When correlation of u and u_2 is positive the biasedness is also positive for the instrumented variable. That is the case for example when u and u_2 correlation is .2 then the coefficient of β1 is 1.112834 that is positively biased. Similarly, as has been shown in five correlation the coefficients of β1 is negatively biased when the correlation between u and u2 is negative.

CHAPTER SIX: METHODS AND PROCEDURE

6.1 Problem in Taubman et al (2014) Methodology:

In the Taubman, Allen et al. (2014) paper they have censored two variables -the outcome variable of emergency room use in the lottery period and the regressor variable of emergency room use in previous period of the lottery. The paper has censored the outcome variable for the number of emergency room use to the study period to 22 and the number of emergency department use to pre-randomization period is 17. According to the economic theory, censoring the extreme value of outcome variable will affect variable-of-interest coefficients both in terms of biasedness and consistency. However, when we censor the extreme value of regressor variable then the effect on biasedness is small and no effect on consistency. What makes the Taubman, Allen et al. (2014) paper doubtable is that they censored both the outcome variable y which is" the number of emergency room use in lottery period" and regressor variable X1 which represents " the number of emergency room use in previous period". That is why we are proposing copula regression by using a binary indicator of dependent and independent variables excluding only for FAMSIZE for which we are using discrete random distribution. By using copula regression, we eliminate this censoring issue since we are using the binary response data with non-normal distribution for this analysis and copula regression is well suited for binary response data. According to Keay (2016), partial copula method provides a flexible parametric approach to deal with various non-normal distributions. He also showed partial copula method can be broadly applied to models with discrete endogenous explanatory variables and sample selection.

6.2 Methods and Procedure

In this paper we have tried to identify the effect of insurance on emergency room use. Below are the three equations which we have constructed to identify the relationship of Medicaid on emergency room use.

$$ED_{ih} = 1[\beta 0 + \beta 1\ MEDICAID_{ih} + \beta 2\ PED_{ih} + \beta 3 X_{ih} - \mu_{ih} > 0\ ] \quad (5)$$

$$MEDICAID_{ih} = 1\ [\ \delta 10 + \delta 11\ LOTTERY_{h} + \delta 12\ X_{ih}\ - \varepsilon_{ih}\ > 0\ ] \quad (6)$$

$$PED_{ih}\ = 1\ [\ \delta 20 + \delta 21\ X_{ih} - e_{ih} > 0] \quad (7)$$

Here, the subscript i denotes a respondent and h denotes the household where a respondent belongs to. Also, ED, PED and MEDICIAD denote ED visit in the study period, ED visit in pre-randomization period and Medicaid or Insurance, respectively. Our main interest rests in the first equation which is the outcome equation. This is a binary choice model equation where an individual visits ED if the right-side function in the bracket is greater than zero. Lottery is an indicator variable; if a person wins the lottery takes the value of 1 and zero otherwise. Taubman, Allen et al. (2014) keep the ED and PED variables that show the number of times an individual visit the ED along with the indicators that only indicate whether an individual visits ED or not. We use the binary indicators only because the numerical variable is censored and can cause inconsistency.

Taubman, Allen et al. (2014) use 2SLS in order to find the causal effect of medicaid on ED visit, and use Lottery as an instrumental variable (IV) for Medicaid. They have used a model considering only one endogenous variable and included two equations for their analysis. What they might have overlooked is PED can potentially be an endogenous varaible. Therefore, we are considering three separate equations, equation two and three for two endogenous variable and the one we are mostly interested in, the outcome equation or equation one.

We have also included that family-size in the model following Taubman, Allen et al. (2014) in order to get rid of the sample selection issue. In oregon health program the whole family members are awarded the lottery if one person in that family wins. Thus a member from a large family has higher chance to be awarded than a small family's member. So, family-size is included in order to control for this.

Family size is assumed to be exogenous so it enters in all three equations of our model. Because lottery has been used as an instrumental variable (IV) for Medicaid, it must be highly correlated with Medicaid and independent of all the errors. Then it can be used as an IV for Medicaid. Since it can not be used as an IV for PED, it doesn't enter equation three. The identification condition for such a model is offered by Keay (2019). We will use them as assumptions for our model.

Assumption 1: $\text{Medicaid} \perp \text{PED} \mid \text{ED}, X$, where X stands for vector of all covariates and $\perp$ means statistical independence.

Assumption 2: According to Han and Vytlacil (2017) a bivariate probit model with a dummy endogenous regressor can be idenfified if the errors have a copula that is stochastically increasing in joint distribution and, additinally, an IV is available for the reduced form equation.

Assumption 1 explains that dependence among the endogenous explanatory variables (EEV) should be eliminated conditional on dependent variable and other covariates , equivalently on the error. We know each endogenous explanatory variable is correlated with error by defination and they might be correlated with each other. We need some mechanism to eliminate this dependence among EEVs. This assumption can be easily verified. According to Keay (2019) an OLS regression of one of EEVs on the other EEV and the dependent variable along with the covariates can be the treatment in this situation. The regression of one EEVs on other EEVs along with other covariates is shown in results and discussion section.

Assumption 2 applies to a bivariate probit model. However, we are working with a trivariate probti model since the population distribution follows a multivariate distribution. Fortunately, Keay (2019) shows that by using copula decomposition we can split the trivariate probit into two bivariate probit models. We can identify a trivariate probit model if each bivariate probit model follows the same assumption as above and there is one valid IV. A few well known copula support this property: Gaussian, Plackett, Clayton And Frank. We have used R software for the copula analysis and since we did not have a package for Plackett in R Software, we have used other three than Plackett copula. Although the number of IV is less than the number of EEVs, the additional information provided by joint distribution by the copulae can be used to supplement the lack of information.

CHAPTER SEVEN: RESULTS AND DISCUSSION

7.1 Variable Identification

Table 3 Description of Variables

Dependent Variable	**Description**	**Data Type**
ED	Ever visited ED in the study period	Binary (Yes=1, No=0)
Independent Variables	**Description**	**Data Type**
Lottery	Selected in the lottery or not	Binary (Selected=1, Not Selected=0)
MEDICAID	Ever enrolled in Medicaid from matched notification date to 30sep2009	Binary (Enrolled=1, Not enrolled=0)
PED	Ever visited ED in the pre-randomization period	Binary (Yes=1, No=0)
FAMSIZE	Number of people in household on lottery list	Number (signed self up=1, signed self up + 1 additional person=2, signed self up + 2 additional people=3)
SNAP_2	Ever personally on SNAP between 01jan2007 and 10mar2008 if in 12m sample	Binary (Enrolled=1, Not enrolled=0)
SNAP_4	Ever personally on SNAP between 10mar2008 and 30sep2009	Binary (Enrolled=1, Not enrolled=0)
Joint Snap	SNAP_2 & SNAP_4 added together	

In table three above most of the variables, either dependent or independent, are explained and defined, except the Lottery variable and the two versions of SNAP. Lottery is included as an independent variable which has been used in this study as an instrumental variable for Medicaid; it is found out to be an endogenous variable. The Lottery variable has met the two vital assumptions of instrumental variables. SNAP is a federal nutrition program that provides nutrition benefits to supplement the food budget of families who need assistance to purchase healthy food and can be self-sufficient (Food and Nutrition services). SNAP_2 explains whether they have used SNAP in pre-randomization period or not. And SNAP_4 indicates whether they have used SNAP benefit during study period. As mentioned in assumption one in the methods and procedure chapter, the additional covariates can be used to make one EEVs independent from other EEVs.

7.2 Regression Results of an endogenous explanatory variable (EEV) on other endogenous explanatory variables (EEVs)

Table 4 Regression of EEV on Other EEV Including Dependent Variable

Dependent Variable: MEDICAID								
	Model 1		Model 2		Model 3		Model 4	
	Coeff.	t stat	Coeff.	t stat	Coeff.	t stat	Coeff.	t stat
PED	.034*** (.006)	5.51	.007 (.007)	1.07	.006 (.007)	0.83	.004 (.007)	0.64
ED	.130*** (.006)	21.32	.111*** (.007)	15.51	.095*** (.007)	13.34	.095*** (.007)	13.28
FAMSIZE	.039*** (.006)	5.91	.023*** (.007)	3.25	.025*** (.007)	3.61	.026*** (.007)	3.74
SNAP_2	N/A	N/A	.148*** (.006)	22.72	N/A	N/A	.018** (.008)	2.13
SNAP_4	N/A	N/A	N/A	N/A	.204*** (.006)	31.39	.191*** (.008)	21.48
cons	.137 (.009)	15.23	.122*** (.010)	11.31	.084*** (.010)	7.81	.081*** (.010)	7.51

Note: Coeff. stands for coefficients. The numbers report the estimates and standard errors in parenthesis. *, ** and *** indicate the significance at 10%, 5% and 1% levels, respectively.

Discussion:

In table 4, we have presented the results of the regression of one endogenous explanatory variable to the other endogenous explanatory variable, along with the dependent variable and other explanatory variables. The reason of including this regression is explained using the two assumptions included in the method and procedure section of this paper. First, we run a regression of Medicaid with PED as an independent variable and the family size as the only other covariate. The coefficient of PED on Medicaid is .034 and with a t statistics value of 5.51. It is easily observable that they are not independent. So, this model does not meet independence assumption of one EEV on other EEVs, so our first assumption fails. Then we have conducted the regression by adding the SNAP_2 in with all other existing variables and found that it changes the coefficient of PED on Medicaid to .007 with a t statistics value of 1.07. So, this evidently shows their dependence has been fed by the addition of another covariate SNAP_2. Similarly, we have modified the regression model 1 by adding another covariates SNAP_4 and found the coefficient value of .006 with a t statistics value of 0.83. So, this covariate also satisfies the assumption 1 of our model. Finally, we have added the mentioned two covariates, SNAP_2 and SNAP_4, and conducted a regression including this new covariate and found that joint SNAP also satisfies as a successful covariate to facilitate the assumption one of our model.

7.3 Regression Results of Model 1

Table 5 Regression Results, Model 1: with SNAP_2

Dependent variable: Emergency Room use (ED)											
	OLS	**2SLS**	**CF**	**CG**	**CC**	**FF**	**FC**	**FG**	**GF**	**GG**	**GC**
MEDICAID	.113*** (.006)	.056** (.023)	0.152** (0.061)	0.151** (0.061)	0.166*** (0.060)	0.148** (0.069)	0.149** (0.069)	0.147** 0.069	0.177*** 0.069	0.178*** 0.069	0.177*** 0.068
PED	.327*** (.006)	.330*** (.006)	0.628*** (0.181)	0.699*** (0.178)	0.808*** (0.129)	0.733*** (0.259)	0.852*** (0.141)	0.850*** 0.230	0.781** 0.315	0.925*** 0.254	0.870*** 0.149
FAMSIZE	-.072*** (.006)	-.069*** (.006)	-0.265*** (0.029)	-0.257*** (0.029)	-0.244*** (0.025)	-0.254*** (0.036)	-0.240*** (0.026)	-0.241*** 0.034	-0.251*** 0.042	-0.234*** 0.036	-0.240*** 0.026
SNAP_2	.119*** (.005)	.128*** (.006)	0.451*** (0.037)	0.436*** (0.038)	0.411*** (0.028)	0.435*** (0.051)	0.409*** (0.029)	0.412*** 0.048	0.420*** 0.062	0.392*** 0.054	0.402*** 0.030
CON	.235*** (.009)	.240*** (.009)	-0.611*** (0.071)	-0.634*** (0.071)	-0.680*** (0.050)	-0.646*** (0.094)	-0.689*** (0.051)	-0.686*** 0.086	-0.665*** 0.111	-0.714*** 0.093	-0.697*** 0.053
R1			0.226*** (0.079)	0.225*** (0.080)	0.207*** (0.078)	0.674*** (0.245)	0.675*** (0.247)	0.680*** 0.246	0.099** 0.041	0.099** 0.041	0.099** 0.041
R2			0.859 (0.585)	0.115 (0.104)	0.085 (0.133)	0.523 (0.829)	0.042 (0.139)	0.026 0.137	0.370 1.012	-0.017 0.153	0.024 0.145
APE of MEDICAID			0.050**	0.049***	0.053***	0.048**	0.047**	0.047**	0.057***	0.056***	0.056***

			(0.020)	(0.019)	(0.019)	(0.022)	(0.022)	0.022	0.022	0.021	0.021
Likelihood			-40529.49	-40529.93	-40530.25	-40530.68	-40530.82	-40530.85	-40531.42	-40531.46	-40531.46

Notes: Here, in the table header CF=Clayton-Frank, CG=Clayton-Gaussian, CC=Clayton-Clayton, FF=Frank-Frank, FC=Frank-Clayton, FG=Frank-Gaussian, GF=Gaussian-Frank, GG=Gaussian-Gaussian, GC=Gaussian-Clayton. The first and second letter are the copulae used for joint distribution of error (μ,ε) and (μ,e), respectively. The numbers report the estimates and standard errors in parenthesis. *, ** and *** indicate the significance at 10%, 5% and 1% levels, respectively.

Discussion:

Table five represents the regression results of OLS, 2SLS, and copula regression results of Gaussian, Clayton and Frank copula. We included PED in the right hand side of our model as Taubman, Allen et al. (2014) did. They claim that inclusion of PED in the right-hand side of model improves the precision of estimates without changing the estimation results. However, there might be correlation between Medicaid and PED as the people who used ED in previous period might have more incentive to buy Medicaid. Controlling for PED might help us to reach to the true effect of variable of interest, but PED might be endogenous too. They have overlooked the endogeneity issue of PED which we assume to be endogenous. Therefore, we are short of one IV as we have only one IV for another endogenous variable Medicaid. We are using copula decomposition to solve the problem. With nine different copula distribution we have shown our results along with the OLS and 2SLS results. Our OLS and 2SLS results support the same result Taubman, Allen et al. (2014) found. We have found an OLS coefficients of 11.3% increase in ED visit with Medicaid. By using 2SLS we have seen the coefficients of Medicaid decrease to 5.6% whereas Taubman, Allen et al. (2014), using 2SLS without using SNAP as covariate, found Medicaid to increase ED visit by 7%. Since OLS is giving us dubious result due to not properly checking endogeneity issue of Medicaid, the result is not satisfactory. Although the current 2SLS provides us the coefficient of Medicaid being lower than previous literature, but the model can face under-identification because of PED is assumed to be endogenous. So, we tried to resolve the issue by doing a model where one IV is enough for two EEVs.

As the main contribution of this paper, we have presented in this table the copula results from nine different copula family. In the header of the table, the first and second letter are the copulae used for joint distribution of error *(μ,ε) and (μ,e),* respectively. Mentionable, *μ, ε and e* are the errors for the equation five, equation six, and equation seven respectively. For example, CF means Clayton and Frank copula are used. The serial of copula results is ordered in terms of the likelihood value. Since we are not comparing the nine copula models over the covariates, likelihood value is used for ranking them. That means Clayton copula provides the best estimates and Gaussian does the worst estimates for *(μ,ε)*. The results depict that coefficients estimates of Medicaid on ED visit are around .147-.178. These are not the partial effect of Medicaid on ED visit. This indicates the sign of correlation whether they are positively correlated or negatively. However, the actual partial effect of Medicaid on ED visit still follows the same sign, but one can see that the APE of the Medicaid is around 4.7-5.7%. What it implies is with the copula regression we cannot reject the fact that Medicaid increases the ED use.

There are two dependence parameters R_1 and R_2 which shows the dependency of error of *(μ,ε) and (μ,e)*. Through this dependency estimates, we have tried to find out the endogeneity of Medicaid and PED, respectively. Our first dependency estimate shown in R_1 which gives us estimates results that are significant. But R_2, the dependency of *(μ,e)*, which is the correlation of PED, and the error, shows insignificant results. Therefore, although we could confirm the endogeneity of Medicaid, but we could not confirm it for PED.

7.4 Regression Results of Model 2

Table 6 Regression Results, Model 2: With SNAP_4

Dependent variable: Emergency Room use (ED)								
	OLS	**2SLS**	**GG**	**GF**	**FG**	**FF**	**CF**	**CC**
MEDICAID	.110*** (.006)	-.060** (.028)	-.169** (0.076)	-.169** (.076)	-.107 (.076)	-.107 (.076)	.074 (.068)	.086 .067
PED	.339*** (.006)	.347*** (.006)	.923*** (.105)	.912*** (.105)	.896*** (.113)	.893*** (.113)	.762*** .137	.867*** .117
FAMSIZE	-.091*** (.006)	-.087*** (.006)	-.289*** (.026)	-.290*** (.027)	-.297*** (.027)	-.298*** (.027)	-.324*** .029	-.307*** .025
SNAP_4	.097*** (.005)	.133*** (.008)	.403*** (.028)	.404*** (.029)	.393*** (.029)	.393*** (.029)	.371*** .030	.352*** .026
CON	.277*** (.009)	.295*** (.009)	-.521*** (.055)	-.516*** (.054)	-.513*** (.059)	-.512*** (.058)	-.481*** (.067)	-.531*** .052
R1			.303*** (.044)	.303*** (.044)	1.59*** (.292)	1.598*** (.291)	.317*** (.095)	.300*** .093
R2			-.012 (.063)	-.032 (.343)	.005 (.067)	.041 (.367)	.515 (.441)	.053 .117
APE of MEDICAID			-.054** (.024)	-.054** (.024)	-.034 (.024)	-.034 (.024)	.024 (.022)	.028 .021
Likelihood			-40848.22	-40848.24	-40853.19	-40853.19	-40861.44	-40861.98

Notes: Here, in the header GG=Gaussian-Gaussian, GF=Gaussian-Frank, FG=Frank-Gaussian, FF=Frank-Frank, CF=Clayton-Frank, CC=Clayton-Clayton. The first and second letter are the copulae used for joint distribution of error (μ,ε) and (μ,e), respectively. The numbers report the estimates and standard errors in parenthesis. *, ** and *** indicate the significance at 10%, 5% and 1% levels, respectively.

Discussion:

In the table 6 above, we have demonstrated the results of OSL, 2SLS and nine different copula regression results**.** Here the model is same as the previous one except that we have used the new covariates of SNAP_4 instead of SNAP_2 in this model. The justification of using this covariate in the model is already given in the regression of one EEVs on other EEVs along with dependent variable and other covariates. The coefficient of OLS regression to measure the effect of Medicaid on ED visit is 0.11 which is slightly lower than the previous model. However, the 2SLS regression gives us the opposite correlation of Medicaid and ED visits. Our 2SLS results displays that Medicaid and ED visit are negatively correlated which contradicts our previous model estimation. However, the conceptual model maintains that the expansion of Medicaid by state or federal authority in a city or county should reduce the ED visit since it treats patients with nonemergency disease in regular visit. So, the need for ED visit get reduced. That is what general sense or economic theory says us. Furthermore, we got the negative coefficients with four of six copula distribution in this model. The last two copula distribution give us positive coefficients of Medicaid and ED visits as out previous model provided. So, we got this result when we included SNAP as covariate which is SNAP benefit status during study period. One explanation for this can be that the use of SNAP during study period makes people healthier or more mentally satisfied with food consumption which makes them physically healthy and they use ED less than those without SNAP. One thing to mention here is that we could not find the nine different copula results as we got in our first model. We have followed the same way to represent the six different copula results in terms of likelihood value. The larger likelihood value is of GG- Gaussian Gaussian and the worst results is shown by CC-Clayton Clayton. The best estimate of *(μ,ε)* is given by Gaussian and the same Gaussian is best for *(μ,e) too.* We cannot report the copula regression result of FC, GC, and CG since we could not successfully convert them.

CHAPTER EIGHT: DISCUSSION AND FINDINGS

8.1 Discussion:

According to Taubman, Allen et al. (2014) there is hitherto no theory prevailing to conclude that providing insurance to uninsured either increases or decreases the emergency department use. One reason why uninsured patients seek emergency department service is irrespective of whether a patient has insurance or not is entitled to get emergency room service (U.S Code). One reason Medicaid increases emergency room use might be the out of pocket cost reduction by insurance. For instance, say a person does not have insurance then he needs to pay, $100 out of pocket since having treatment from emergency care is not completely free. On the other hand, having insurance or Medicaid can reduce this out of pocket cost a little bit to, say $80. This might be one reason to explain why insured people's motivation for using more emergency room when they have less a severe case which can be treated by physician visit or regular hospital visit. Getting Medicaid coverage might also enhance the access to emergency department service which can be a reason of increased emergency room use (Taubman, Allen et al. 2014). The former reason of more emergency room use can be explained by moral hazard theory of economics. Moral hazard is defined in economics as one party involving in a risky activity by knowing that if any risk occur the other party is going to pay for the cost. This can be explained by an analogy of home insurance such as when a person did not insure his/her home, they might take more cautionary action not to pose it for burglary or fire, however, when he has insurance for the house he might be careless about the safety of the house because of the assurance that someone else will take care of financial loss if any damage or burglary occur. Similarly, when a person has insurance in terms of Medicaid coverage, he/she might be more willing to reach emergency room for treatment which is curable by taking other nonemergency room medical treatment. Another reason of using more emergency room use can be people think that in emergency department they might get quick and prompt service with better medical equipment or prognosis machines than the regular physician office visit. According to (Statistics) in 2016 a total of 145.6 million visit occurred in emergency room. Out of which 12.6 million patients has resulted in hospital admission which is just 8.7% of total visits. Number of patients resulting in admission to critical care unit is 2.2 million. The percent of total visit who have been treated within 15 minutes is 39.0%. So, this also justify out assertion of moral hazard issue of Medicaid covered people who use more emergency room even if those case could be treated in person physician office visit.

8.2 Findings:

In this paper, we tried to find the causal effect of Medicaid expansion on emergency department use. We first found out some loopholes in existing literature review. First, we found that if we have a multiple endogeneity issue in our model, but we cannot still find required number of valid IV, we might get spurious coefficients of our variable of interests. We also might face criticism due to the under identification of the endogeneity of probable endogenous variables. As we have projected that the independent variable PED, which previous literature included as an exogenous variable might be endogenous, although with our copula regression of error terms we couldn't finally confirm it since we didn't get significant result of R_2 as we have shown in regression results of table 5 and 6. However, through Monte Carlo Simulation we have demonstrated that if more than one endogenous variables are in the model with a single IV variable the model estimates can give us biased estimates. As the simulation results show that the model with less than the required number of IV doesn't give a result close to the population parameters. Finally, we have run two separate copula regression each nine different copula distribution of errors. We have not been able to reject the hypothesis that Medicaid increases emergency room use. We did some background study about this and our study convinces us that people are facing the issue of moral hazard. Because of this moral issue people do not like to go through the regular channel of treatment for treatable disease rather they seek the path of ED visit which is free in most cases.

REFERENCES:

Angrish, J. D. and J.-S. Pischke (2014). Mastering 'Metrics. Scottsdale, Princeton University Press.

beckershospitalreview. "25 facts and statistics on emergency departments in the US." Retrieved 10/28/2019, 2019, from https://www.beckershospitalreview.com/hospital-management-administration/25-facts-and-statistics-on-emergency-departments-in-the-us.html.

Finkelstein, A., et al. (2012). "The Oregon health insurance experiment: evidence from the first year." The Quarterly journal of economics **127**(3): 1057-1106.

Food and Nutrition services, U. D. o. A. "Supplemental Nutrition Assistance Program." Retrieved 10/26/2019, 2019, from https://www.fns.usda.gov/snap/supplemental-nutrition-assistance-program.

Han, S. and E. J. Vytlacil (2017). "Identification in a generalization of bivariate probit models with dummy endogenous regressors." Journal of Econometrics **199**(1): 63-73.

Hoot, N. R. and D. Aronsky (2008). "Systematic review of emergency department crowding: causes, effects, and solutions." Annals of Emergency Medicine **52**(2): 126-136. e121.

Keay, M.-J. (2016). "Partial copula methods for models with multiple discrete endogenous explanatory variables and sample selection." Economics Letters **144**: 85-87.

Keay, M.-J. (2018). "A Two-Regime Correlated Random Coefficients Model with Nonnegative Dependent Variable." working paper: 21-26.

Keay, M. J. (2019). Identification and Estimation by a Flexible Parametric Method for Models with Discrete Endogenous Explanatory Variables. N/A, South Dakota State University. **N/A**.

Krochmal, P. and T. A. Riley (1994). "Increased health care costs associated with ED overcrowding." The American journal of emergency medicine **12**(3): 265-266.

Lowe, R. A., et al. (2005). "Association between primary care practice characteristics and emergency department use in a Medicaid managed care organization." Medical Care: 792-800.

Nikpay, S., et al. (2017). "Effect of the Affordable Care Act Medicaid expansion on emergency department visits: evidence from state-level emergency department databases." Annals of Emergency Medicine **70**(2): 215-225. e216.

Pines, J. M., et al. (2016). "Medicaid expansion in 2014 did not increase emergency department use but did change insurance payer mix." Health Affairs **35**(8): 1480-1486.

Salway, R., et al. (2017). "Emergency department (ED) overcrowding: evidence-based answers to frequently asked questions." Revista Médica Clínica Las Condes **28**(2): 213-219.

statistica.com. "Total Medicaid enrollment 1966-2018." Retrieved 10/28/2019, 2019, from https://www.statista.com/statistics/245350/total-medicaid-federal-and-state-expenditure-in-the-us-since-1966/.

Statistics, N. C. f. H. "Emergency Department Visits." Retrieved 10/28/2019, 2019, from https://www.cdc.gov/nchs/fastats/emergency-department.htm.

Statistics, N. C. n. d. (2016). "National Hospital Ambulatory Medical Care Survey: 2016 Emergency Department Summary Tables, tables 1, 4, 11, 14, 24, 25."

Taubman, S. L., et al. (2014). "Medicaid increases emergency-department use: evidence from Oregon's Health Insurance Experiment." Science **343**(6168): 263-268.

U.S Code, t., THE PUBLIC HEALTH AND WELFARE Examination and treatment for emergency medical conditions and women in labor**:** Part E. Miscellaneous Provisions Section 1395dd.

Wooldridge, J. M. (2013). Introductory Econometrics: A modern Approach. Mason,OH, Cengage Learning.