Senior Project
Department of Economics

"SES and Health: Income the Weapon Against CHD."

Julian Avale
May, 2014

Advisor: *Dr. Francesco Renna*
Dr. Musriel Fang

Abstract:

I test if higher income has a negative effect on the probability of being diagnosed with Coronary Heart Disease (CHD), in adults aged 18 to 64 in the United States. I use a probit model to describe the association between income and CHD. The results indicate a weak negative association of income with CHD. Higher income leads to better health but the effect is small. Because the model suffers from reverse causality, I run a 2 Stage Least Square (2SLS) which corrects for endogeneity. Reverse causality happens because income has an impact on CHD, and CHD has an effect on income. To run this model I use an Instrumental Variable (IV) which affects income and CHD through income but not CHD but not through other channels. The result (-0.036) indicates that income has a negative impact on CHD, but the impact is weak.

I would like to thank Dr. Francesco Renna and Dr. Zheng Fang for their assistance over the course of this project.

I. Table of Contents:

II. Introduction:

Coronary Heart Disease is the most common kind of heart disease killing 380,000 men and women in the United States each year (Murphy et al., 2013). The condition is developed partially because of genetic predisposition, and partially because of behavior. This behavior is linked to our SES and studies such as Cutler et al. (2007.) have found that higher SES leads to better behavior, which in turn could lead to better health. A number of behavior choices such as smoking, drinking, diet and exercise are well documented to impact our general health and more specifically our cardiovascular health.

Coronary Heart Disease has become a substantial problem for the health care system, in that the high cost of curing and controlling this condition has become a burden on the industry, which has led to higher costs for the consumer (Heidenreich et al., 2011). The treatment and care of coronary heart disease is a $108.9 billion a year cost in the United States (Heidenreich et al., 2011). Identifying a causal link between income and CHD, would be helpful to create the appropriate policies. This study could positively influence policies dealing with wage increase and affordable care. Higher wages would help individual have better health if the study concludes that higher income leads a lower probability of developing CHD. Higher wages would also help individuals gain access to better health care, which in turn would lead to better health outcome.

The findings by Smith (2007) and others like Cutler, Lleras-Muney and Vogl (2008) have been overwhelmingly significant in the development of the field because they have been able to establish causality. Education has a more significant causal relationship with diabetes, hypertension, and stroke. The consensus of these studies is that lower SES is strongly correlated to worse health outcomes.

Adams et al. (2004) attempt to prove that causality does not exist between SES and health. This study is ambitious while at the same time limiting. It is only looking at retired individuals, and how SES affects their health after retirement, therefore the findings cannot be applicable to the whole population.

In this paper I attempt to confirm a causal relationship between SES and CHD, focusing on which SES variables are most closely related to CHD. The goal is to confirm the findings in Smith (2007), where causality exists between SES and health. I will test whether higher Socioeconomic Status (SES) has a negative effect on the prevalence of Coronary Heart Disease (CHD) in adults aged 18 to 64 in the United States. I will account for potential reverse causality: the effects of CHD on SES. I will estimate a probit model and compute the marginal effects. The signs on the estimated coefficients are as expected. Income, education, exercise and alcohol consumption are negative. This indicates that these variables lead to lower probability of being diagnosed with CHD. The study suffers from endogeneity, where CHD impacts SES. To correct for endogeneity a "two stage least square" (2SLS) model is run. Employment is used as the instrumental variable (IV), with the individual being employed=1 or not employed=0.

III. Literature Review:

Smith (2007), and Johnson et al. (2009) amongst others have been very interested in the relationship between health and Socioeconomic Status (SES). In his paper, Smith (2007) looks at possible causation between SES indicators (income, wealth, and education) and health, and finds that there is a causal relationship between SES and health. Education has the largest impact on an individual's health over their lifetime. Income also impacts health. The study found that reduction in income level is correlated with health deterioration. Individuals in lower income brackets are 30% more likely to suffer from poor health than individuals in higher income brackets (Smith, 2007).

Kiuila and Mieszkowsk (2007) look at the effect of socioeconomic studys (SES) and health. Data from the National Health Interview Survey and the Multiple Cause of Death files were used. The Cox hazard model and the probit model were used to test the hypothesis. The authors found a stronger correlation between income and mortality in people with good health and a weaker correlation in people with poor health.

Cutler, Lleras-Muney and Vogl (2008) examined the effects of education, financial resources, rank, and race/ethnicity on the health of individuals. Looking at the mechanisms of SES and Health, they concluded that, "some dimensions such as education cause health, others like income are caused by health, and some such as wealth are mutually determined with health" (Cutler et. al., 2008). Their research confirmed the findings by Smith (2007), that education has the most impact on health. Education is also the most consistent of all SES factors because it doesn't fluctuate through the years. Higher income has a negative effect on health in the short run but a positive impact over the long run. This happens because the change in income would lead to a change in human behavior, such as date, consumption and leisure. They conclude that

the mechanism used for the study is not perfect. The authors believe it would be more appropriate to focus on the impact that SES has on children. Higher parental SES (income, education) has a positive impact on children health during childhood and into adulthood. It would be safe to assume that children coming from families with higher income and education have better health in childhood and therefore higher SES in adulthood.

Johnson et al. (2009) examine the possible causal relationship between income and health (hypertension). They also evaluate if there is a difference between self-reported and objective measures for hypertension. The authors found no connection between income and hypertension in the self-reported data, but found a significant connection in objective measures. They suspect that self-reporting data could lead to results that under estimate the effect of income on health.

Lei, Yin and Zhao (2011) looks at China because it has recently undergone a "rapid Epidemiological transition from infectious to chronic disease" (LEI et. al, 2011). In the paper the authors examine whether there is a causal relationship between SES and hypertension. They focus on prevalence, awareness, treatment and control of hypertension. The study found no relationship between income and education on hypertension but a relationship was found between education and diagnosis and treatment of hypertension in urban areas but not rural areas.

Adams et al. (2004) unlike the other studies focus their efforts in identifying the absence of causality between SES and health. The authors target group are individuals aged seventy and older. The data used come from the Asset and Health Dynamics among the Oldest Old (AHEAD). It consists of 8,222 individuals who are interviewed every 24 months, between 1923 and 1993. A probit model is initially used, but the main focus of this study is the first-order

Markov process. The study confirmed the absence of causality for mortality and acute, sudden onset diseases but found a causal link between mental problems and SES.

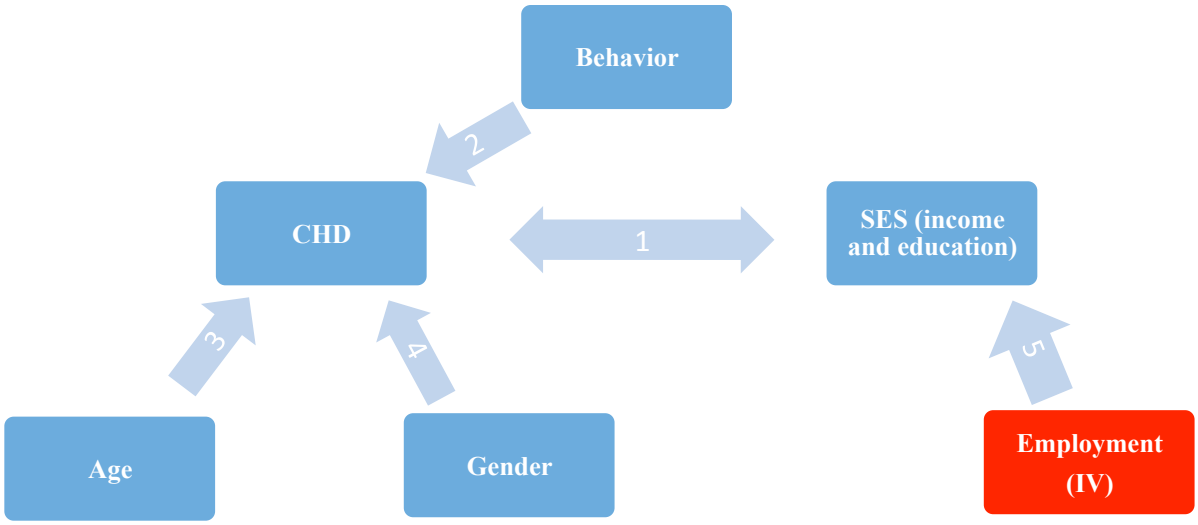
Deaton (2002) studied the policy implications of income and health. He conclude that higher income would lead to better health, because healthcare would become more affordable. These could be reached through an increase in wages or through tax credits by the government.

IV. Theoretical Model:

This study follows the theoretical model in Adams et al. (2004). Behavior (2)[1], which includes smoking, alcohol consumption, exercise and obesity is related to CHD. The poor behavior choices we make such as smoking and drinking regularly, and not exercising lead to poor health. The model also controls for age (3)[2] , which is a demographic variable. As we get older there is a higher risk of developing health problems, but the correlation between age and SES decreases as we age. This happens because over time genetic factors become more predominant than behavioral factors (Cutler et al., 2008). Gender (4)[3] can also affect the chance of developing CHD. Adams et al. (2004) found that males have a higher chance than woman of developing CHD. Employment (5)[4]is chosen as the IV in this study. It affects SES and most importantly income directly, because employed individuals will have higher income than their unemployed counter parts. In this case employment does not affect CHD directly but rather does so through income. Employment does affect health but only through the type of employment and not by whether you are employed or not.

[1] Link between behavior and CHD in Diagram 1.
[2] Link between age and CHD in Diagram 1.
[3] Link between sex and CHD in Diagram 1.
[4] Link between employment and SES in Diagram 1.

Diagram 1: SES and CHD diagram.



V. Model Specifications:

The model being used for this study is the standard simultaneous-equation model (SEM):

SES => Health

1. $CHD = \beta 0 + \beta 1(income) + \beta 2(educa) + \beta 3(age) + \beta 4(gender) + \beta 5(obese) + \beta 6(smoke) + \beta 7(dalc) + \beta 8(exercise) + e$

Health => SES

2. $Income = b_0 + b_1(CHD) + b_2(educa) + b_3(age) + b_4(gender) + b_5(obese) + b_6(smoke) + b_7(dalc) + b_8(exercise) + b_9(emp) + e$

This model examines the causal relationship between Socioeconomic Status and Coronary Heart Disease. I use the Behavioral Risk Factor Surveillance System (BRFSS) 2012 Survey Data, which collects information such as age, gender, behavior, to financial and health statuses. In 2012 over 470,000 individuals took the survey.

I first estimate a probit model. CHD is binary variable which equals 1 if the respondent has responded YES to having Coronary Heart Disease, 0 otherwise. I use the 2 Stage Least Square (2SLS) which takes account for the presence of reverse causality in which CHD affects income. The 2 Stage Least Square (2SLS) is used when models violate Ordinary Least Square regression assumption that there is no correlation between one or more predictor variables and the disturbance term of the dependent variable.

I hypothesize that CHD is affected by income and other variables listed in the model (1). In model (2) it is expected to inversely affect income. The *income* variable is categorical, ranging between one and five, with five being the highest income level. As an individual gets older, health deteriorates and there is a larger risk of developing CHD. Therefore, model (1) controls for age. Age could also be correlated with increased income through accumulated work experience. Higher education is typically correlated with increased income, and lower chance of being diagnosed with CHD. The education variable is also categorical ranging between one and four, with four representing highest education level. Adams et al. (2004) showed that males are more likely to develop CHD than females. So model (1) controls for gender differences in CHD

risk. Model (1) also controls for risk factors such as *smoking, obese and, alcohol consumption.* They are all binary variables with YES=1 or else 0.

VI. Interpretation of results:

The initial results of the probit model in Table 3 show that individuals with higher income (-0.145) and higher education (-0.023) are less like to have coronary heart disease CHD. The results also show that people that exercise regularly (-0.0171) are less likely to be diagnosed with CHD. As expected age (0.037), obesity (0.218) and smoking (0.146) lead to a higher chance of developing this condition Alcohol consumption in moderation is beneficial to health. This is confirmed by the sign in the results.

The marginal effect estimators from the probit model confirm what other studies have found. The results are shown in Table 4. The signs of the variables in the marginal effect are the same as in the initial results of the probit model. The low values of income (-0.0097) and education (-0.0016) in the marginal effect, indicate that the effect is weak.

The results from the 2SLS are shown in Table 5. The sign on the parameter estimate are as expected. The SES factors, income (-0.036) and education (-0.004) negatively impact CHD, while behavior factor such as obesity (0.01) and smoking (0.011) positively impact CHD. The results indicate that older individuals are more likely to suffer from CHD, then younger individuals. They also show that males (0.019) suffer from CHD more than woman. The parameter estimates are low, therefore we can conclude that the effect of income on CHD is weak.

VII. Conclusion

The goal in this paper is to show that higher Socioeconomic Status (SES) leads to lower Coronary Heart Disease (CHD). The first model probit measured the effect of each variable on CHD. The results indicate a negative impact by income and education. Behavior factor such as obesity and smoking have a negative impact. The second model 2SLS is used to correct for endogeneity, which happens when explanatory variable is related to the dependent variable. This results show a weak effect of income on CHD.

The findings of this study are helpful to develop policies that can make a positive impact in the health of individuals. Because the effect of income on health is positive, where higher income leads to better health, increasing wages should lead to better health. Deaton (2002) points out that the wealthy, top 5 percent of the population have a 25 percent lower rate of mortality as compared to the poor, bottom 5 percent of the population. In the study Deaton (2002) also found that an increase income leads to a proportional increase in health.

Wealthy individuals have better access to healthcare than people in the lower income bracket (Deaton, 2002). An assumption can be made, where making healthcare more affordable/accessible should help individuals in the lower income bracket live a healthier life.

The findings of this study are not representative of the entire population. It is only focused on the 18-64 age group. One difference between this two groups is the access to health services and the type of healthcare this group uses (Adams et al., 2004). Medicare is the primary health coverage for individuals aged 65 and older, while younger individuals do not have access to these plan (Adams et al., 2004). Adams et al. (2004) concluded that individuals aged 65 and older are affected differently by the SES and CHD dynamic.

Potential problems with this study is the IV and the data set. A strong IV is needed to produce strong and significant results. Finding the appropriate IV is challenging. It needs to be backed by theoretical or logical evidence. The data set can also be limiting to the study. It is a cross-sectional data set looking at only one particular point in time. A better data set would be a panel data set, because models other than the 2SLS can be used, which do not require an IV.

Works Cited:

1. Adams, Peter et al. “Healthy, Wealthy, and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status.” *National Bureau of Economic Research* (2004) 415-526.

2. Cutler, M.D. “Socioeconomic Status and Health: Dimensions and Mechanisms.” *National Bureau of Economic Research* (2008) 1-53.

3. Deaton, Angus. “Income and Health: Policy Implications of the Gradient of Health and Wealth.” *Health Aff* 21 (2002):213-30.

4. Heidenreich PA, Trogdon JG, Khavjou OA, et al. “Forecasting the future of cardiovascular disease in the United States: a policy statement from the American Heart Association.” *Circulation*. 123 (2011):933-44.

5. Johnson, W.D “Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient.” *Journal of Health Economics* 28 (2009): 540-552.

6. Kiuila, Olga and Mieszkowski, Peter. “The effects of income, education and age on health.” *Health Economics* 16 (2017): 781-798.

7. Lei, Xiaoyan et al. “Socioeconomic staus and chronic diseases: The cause of hypertension in China.” *China Economic Review* 23 (2012): 105-121.

8. Murphy SL, Xu JQ, Kochanek KD. “Deaths: Final data for 2010.” *Natl Vital Stat Rep*. 2013;61(4).

9. Smith P.J. “The Impact of Socioeconomic Status on Health over the Life-Course.” *The Board of Regents of the University of Wisconsin System* 42 (2007): 739-764.

Table 1.

Variable Definitions, Summary Statistics and Data Sources.		
Variable	**Definition**	**Source**
Educa	Education Level 1 – 5 with 5 representing high education.	**Behavioral Risk Factor Surveillance System (BRFSS) 2012 Survey** (http://www.cdc.gov/brfss/annual_data/annual_2012.htm)
Age	Age	
CHD	Dummy variable for Coronary Heart Disease Yes=1 else 0	
Income	Income 1 – 5 with 5 representing high income.	
Obese	Binary variable for obesity. Yes =1 else 0.	
Dalc	Binary variable for alcohol consumption. Yes = 1 else 0.	
Smoke	Binary variable for smoking. Yes=1 else 0.	
Exercise	Exercise variable. Yes=1 else 0.	
Gender	Male=1 Female=0	
Emp	Employment Ye=1 else 0	

Table 2.

Data Means

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
chd	Coronary Heart Disease	262267	0.0344954	0.1824982	0	1.0000000
income	Income	262267	3.4522872	1.3758026	1.0000000	5.0000000
EDUCA	Education	262267	4.9755593	0.9995144	1.0000000	6.0000000
AGE	Age	262267	46.1437314	12.5444771	18.0000000	64.0000000
gender	Gender	262267	0.4413022	0.4965436	0	1.0000000
obese	Obesity	262267	0.6574445	0.4745652	0	1.0000000
smoke	Smoking	262267	0.2016037	0.4011985	0	1.0000000
dalc	Alcohol consumer	262267	0.5751581	0.4943199	0	1.0000000
exercise	Exercise	262267	0.7931421	0.4050535	0	1.0000000
emp	Employment	262267	0.6860261	0.4641068	0	1.0000000

Table 3.

Probit model

Analysis of Maximum Likelihood Estimates					
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	1	-3.2416	0.0411	6234.4124	<.0001
income	1	-0.1450	0.00428	1148.7766	<.0001
EDUCA	1	-0.0239	0.00550	18.9065	<.0001
AGE	1	0.0371	0.000565	4312.3652	<.0001
gender	1	0.2636	0.0105	631.5910	<.0001
obese	1	0.2185	0.0124	308.5576	<.0001
smoke	1	0.1468	0.0123	143.3277	<.0001
dalc	1	-0.1769	0.0108	266.0633	<.0001
exercise	1	-0.1716	0.0115	223.8414	<.0001

Table 4.

Marginal Effect

Variable	Label	Mean
income	Marginal effect of income on the probability of chd=1	-0.0097558
educa	Marginal effect of educa on the probability of chd=1	-0.0016090
age	Marginal effect of age on the probability of chd=1	0.0024987
gender	Marginal effect of gender on the probability of chd=1	0.0177379
obese	Marginal effect of obese on the probability of chd=1	0.0147005
smoke	Marginal effect of smoke on the probability of chd=1	0.0098761
dalc	Marginal effect of dalc on the probability of chd=1	-0.0119052
exercise	Marginal Effect of exercise on the probability of chd=1	-0.0115499

Table 5.

Parameter Estimates 2SLS						
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**3**
Intercept	1	0.100334	0.003329	30.14	<.0001	Intercept
income	1	-0.03619	0.000735	-49.22	<.0001	Income
EDUCA	1	-0.00400	0.000386	-10.37	<.0001	Education
AGE	1	0.001796	0.000029	61.85	<.0001	Age
gender	1	0.019999	0.000738	27.11	<.0001	Gender
obese	1	0.010818	0.000775	13.96	<.0001	Obesity
smoke	1	0.011113	0.000924	12.02	<.0001	Smoking
dalc	1	-0.01305	0.000755	-17.29	<.0001	Alcohol consumer
exercise	1	-0.01831	0.000921	-19.87	<.0001	Exercise

SAS Code:

```
data chd;
set work.chd;
if cvdcrhd4=1 then chd=1;
else chd=0;
if income2 in (1,2) then income=1;
if income2 in (3,4) then income=2;
if income2 in (5,6) then income=3;
if income2 in (7) then income=4;
else if income2 in (8) then income=5;
if sex=1 then gender=1;
else gender=0;
if _rfbmi5=2 then obese=1;
else obese=0;
if _rfsmok3=2 then smoke=1;
else smoke=0;
if drnkany5=1 then dalc=1;
else dalc=0;
if _totinda=1 then exercise=1;
else exercise=0;
if employ in (1,2) then emp=1;
else emp=0;
run;
```

Data Means:

```
proc means;
var chd income educa age gender obese smoke dalc exercise emp;
run;
```

Probit:

```
proc qlim data=chd;
model chd = income educa age gender obese smoke dalc exercise / discrete
(dist=normal);
output out=chd22 marginal;
run;
```

Probit marginal effects;

```
proc means data=chd22 mean std;
var Meff_P2_:;
run;
```

2SLS

```
proc syslin data=chd 2sls;
endogenous income;
instruments emp;
model chd = income educa age gender obese smoke dalc exercise;
run;
```