# Impact of economic recession on cardiovascular health across 30 major American cities

**By**

**Abinaya Arulselvan**

**June, 2014**

A Community Based Master's Project presented to the faculty of Drexel University School of Public Health in partial fulfillment of the Requirement for the Degree of Master of Public Health.

## ACKNOWLEDGEMENTS

I would like to thank to my advisor, Professor Longjian Liu, for supporting me for the past one year. He is someone you would instantly love and never forget once you meet him. He is the funniest advisor and one of the smartest people I know. I hope that I could be as lively, enthusiastic, and energetic as he is someday. He has been always supportive and given me the freedom to pursue the project without objection. He has also provided insightful discussions about the research and his smile is contagious. I am also thankful to my friend Pooja Rangan for her help and guidance throughout the project. She has always been there for me to happily discuss about the research project.

I also thank my mom, dad and sister who have sacrificed their lives for my sister and myself and provided unconditional love and care. I love them so much, and I would not have made it this far without them.

I especially thank my best friends Navish Raj, Arvindh Francis and Nagamani for being my good critiques. They have made my life colorful and meaningful. Without them nothing seems possible.

Special thanks to my roommates and my other friends for tolerating me during stressful periods. I am grateful to god for this beautiful life.

**TABLE OF CONTENTS**

Abstract……v

Introduction ……1

Statement of the Problem……2

Overall Significance Of The Study……3

Specific Aims……3

Background and Significance……4

Research Design and Methods……8

Data Analysis……15

Results……16

Discussion……19

Strengths and Limitations……20

Conclusions and Recommendations……21

References……23

## APPENDICES

Appendix A …… Figures

Appendix B …… Tables

Appendix C …… Institutional Review Board Approval

Appendix D …….. SAS Codes

## LIST OF FIGURES

Figure 1 & 2: Age adjusted CVD prevalence across all 30 American cities

Figure 3: Age adjusted CVD prevalence with respect to gender

Figure 4: Age adjusted CVD prevalence with respect to race/ethnicity

Figure 5: Age adjusted CVD prevalence and per capita real GDP across all 7 years (2006 – 2012)

Figure 6 & 7: Detroit – CVD prevalence and economic revenue

Figure 8 & 9: San Francisco - CVD prevalence and economic revenue

## LIST OF TABLES

Table 1: Division of cities into three risk zones

Table 2: Association (Odds ratios and 95% confidence intervals) of CVD and economic recession using BRFSS 2006-2012.

## ABSTRACT

**The impact of economic recession on cardiovascular health among urban population across thirty cities in the U.S.**

Abinaya Arulselvan BDS, MPH, and Longjian Liu, MD, PHD, MSC, FAHA

**Background:** The economic recession from 2007 to 2009 created great fiscal challenges for American cities that have made them slow to recover. Economic instability causes increased unemployment rates forcing many to opt for compromised unhealthy lifestyles and increased mental stress. Therefore, the impact of economic recession on population's health has become a major concern.

**Objective:** To study the impact of economic recession on cardiovascular disease (CVD) among urban population aged 18 years and older in 30 major cities in United States.

**Methods:** The study with an ecological design was conducted using data from 2006-2012 Behavioral Risk Factor Surveillance System (BRFSS). Thirty largest cities by population were included in the analysis. Participants' health conditions including prevalence of CVD were defined on the basis of self-reported physician-diagnosis of disease. Differences in CVD prevalence across the 30 cities and associations of CVD with socioeconomic and behavior risk factors were examined using univariate and multivariate analysis approaches. All analysis was conducted using SAS 9.3 software.

**Results:** Among 997161 participants of the total study sample, the prevalence of coronary heart disease (CHD) and stroke was 8.6% and 4.2%.Increasing age was a significant predictor for both the cardiovascular disease. Males had increased odds of developing CVD than females. Native Americans/Alaskan Indians followed by Blacks and had greater odds for CVD. Detroit, St. Louis and San Antonio had highest prevalence of CVD, while San Francisco, Seattle and Minneapolis had the lowest. Age adjusted results showed increased prevalence during economic recession period in all cities. Per capita real GDP, cigarette smoking, diabetes, depression, obesity, education, employment and income were significantly associated with CVD.

**Conclusion:** The study confirms that in addition to the study of well-established CVD risk factors, such as smoking and obesity, economic recession has a significant impact on cardiovascular health among urban population. Further studies are needed to test the potential mechanisms by which economic recession causes the risk of CVD.

**INTRODUCTION**

About 30% of the all deaths across the world are due to Cardio Vascular Disease (CVD) which makes it the leading cause of death worldwide. It accounts for around 16.7 million deaths each year which is about 1 in 3 deaths, primarily from heart attacks and strokes, and is 3 times more in comparison to the combined deaths caused by infectious diseases including HIV/AIDS, tuberculosis and malaria (Dahlöf, 2010). It is the leading cause of death in both the sexes as well as in most of the ethnicities in United States except American Indians and Alaskan Natives where it is second to cancer. It is predicted that deaths due to CVD especially cerebrovascular disease (stroke) and ischemic heart disease (IHD) will increase to 23.4 million per year by 2030, still retaining its position of being the number one cause of death if the trend continues (Dahlöf, 2010).

Annually about 600,000 lives are claimed due to some form of CVD in United States and one in four person die every 33 seconds. The CVD risk factors such as hypertension, high cholesterol, tobacco use, metabolic syndrome poor nutrition, physical inactivity, alcohol intake, socioeconomic status and poor lifestyle choices are increasing at an alarming rate that can put people at a higher risk of CVD (CDC, 2005).

Although, research in this field is carried out since 1947 (Minnesota Business Men Study), the recent trend analysis shows that the mortality from CVD has decreased only by 50-60%. This could be due to unidentified risk factors or potential risk factors on which we have less information that may be playing a major role in causing the disease (CDC, 2013). One of these is the effect of psychophysiological factors on CVD. Among all the major social stresses

that may be related to the incidence of CVD are those that originate during adverse time periods, especially when an individual's economic status changes. Adverse changes in economic status will restrict the degree of dependence of an individual on basics of life such as food, clothing and shelter. This will ultimately affect his/her health with little economic resources to invest on health. Periods of economic depression will force a number of people into unemployment and reduced income in the society (Clark et al., 2009). The effect of this economic instability will substantially reduce the ability of segments of population to afford fundamental as well as highly valued products or services of the society, including health care. This will result in individuals being depressed and force them to increase their smoking or alcohol consumption in order to relieve stress. This will only add to the burden of the disease especially when little is spent by the government on health care during such economic downturns (Brenner, 1971). Also, research has proven that even a healthy individual can increase his CVD risk due to increased mental stress when his economic condition worsens (O'Connor, 2008). Therefore, a need for further research on other less known risk factors is necessary to reduce the economic burden of the disease.

**STATEMENT OF THE PROBLEM**

The objective of the study is to assess the impact of economic recession of 2007 to 2009 on CVD in age groups above 18 years of age, among the urban population living in 30 major cities of United States; by analyzing the health data provided by the Centers for Disease Control and Prevention through Behavioral Risk Factor Surveillance System (BRFSS) 2006 - 2012.

## OVERALL SIGNIFICANCE OF THE STUDY

The significance of this study lies in the unique opportunity it affords to develop a better understanding of the regional differences in cardiovascular disease risk across different age groups of the urban population living in 30 major cities of United States. Predominantly, we are interested in evaluating the association of the country's economic trends against the cardiovascular disease risk, particularly during the economically volatile years of 2007 through 2010.

The importance of this study is to understand the process of economic disaster that a country experiences and its effects on the health of its citizens, where additional economic stress and mental depression can act as potential confounders or effect modifiers. This study is significant as very little research is done comparing the cardiovascular disease risk against potential risk factors such as economic status, especially among the urban population; and to establish its trend during the economic roller-coaster period, the period before, during and immediately after the recession period.

## SPECIFIC AIMS

The purpose of the study is to study the effect of economic recession on cardiovascular health of urban population across 30 major American cities. Specific aims of the study include:

- To determine if a nation's economy impacts its citizen's cardiovascular health.
- To investigate the impact of economic recession on American population in 30 major populous cities from 2006 to 2012.

- To evaluate which city, race/ethnic group and sex is more resistant and prone to cardiovascular diseases with respect to economic downturn.

**BACKGROUND AND SIGNIFICANCE**

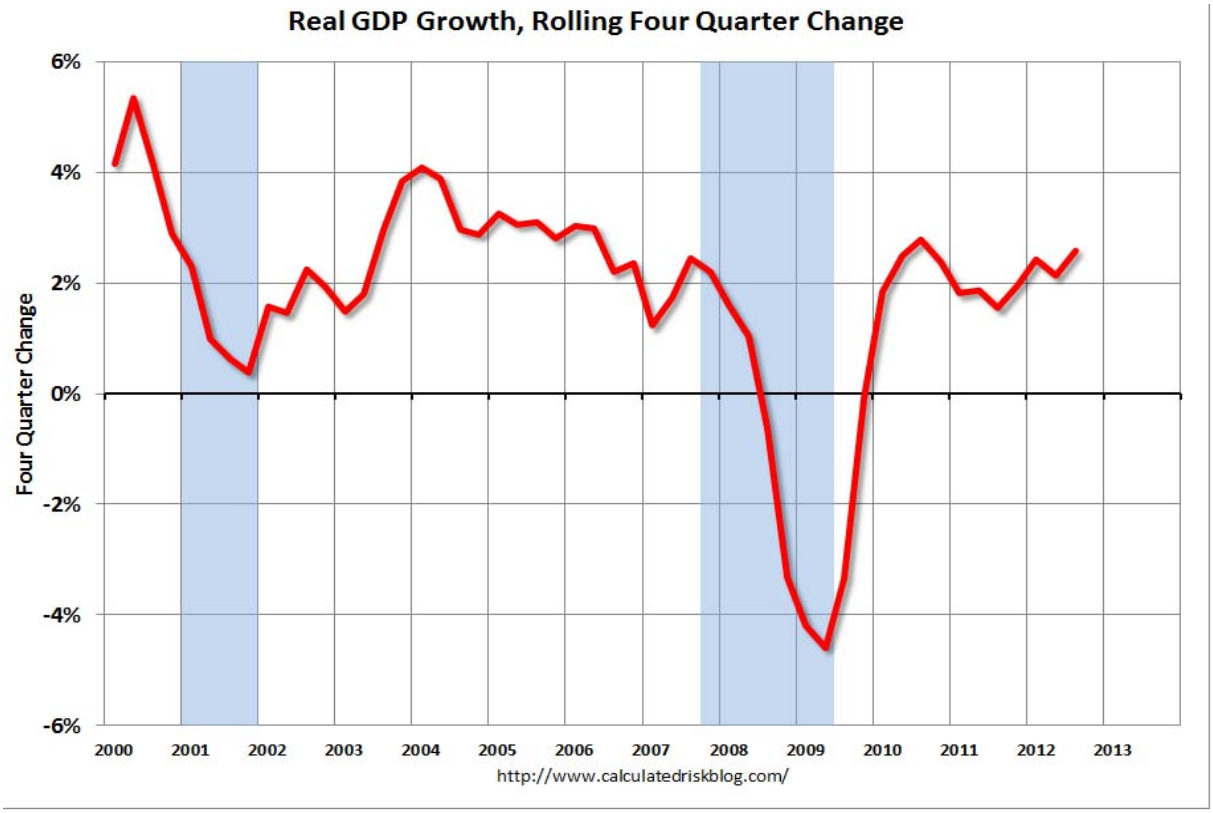
According to the 2014 updated Heart Disease and Stroke Statistics, about 83.6 million adults in U.S suffer from atleast 1 (or more than 1) form of CVD. Of these, 42.2 million are 60 years (or above) of age. The first cardiovascular event escalates from 3 per 1000 men at 35 - 44 years of age to 74 per 1000 men at 85-94 years of age. Similar rates occur for women, 10 years later in life. Men most commonly suffer from CVD events and women have a higher proportion of stroke (CDC).

Coronary heart disease is responsible for 48.2 % of deaths due to cardiovascular diseases followed by stroke (16.4%), high blood pressure (8%) and heart failure (7.3%) (Go et al, 2014). Race also plays a major role in the unequal distribution of the disease in America. The Blacks or African Americans are more 30% likely to die from a heart disease than non-Hispanic whites. They also tend to suffer from high blood pressure at an earlier stage of life than others (CDC). Some of the classical risk factors of CVD include hypertension, high ratios of LDL with HDL, physical inactivity, obesity, uncontrolled diabetes, smoking, alcohol and depression. All these risk factors are also increasing at an alarming rate (CDC).

Cities are the vital organs of a nation. They are responsible for significant prosperity of a country. Any economic crisis will drastically affect the cities than any other regions of a country (PEW charitable trust, 2013). According to WHO, 40% of the global population lived in cities in

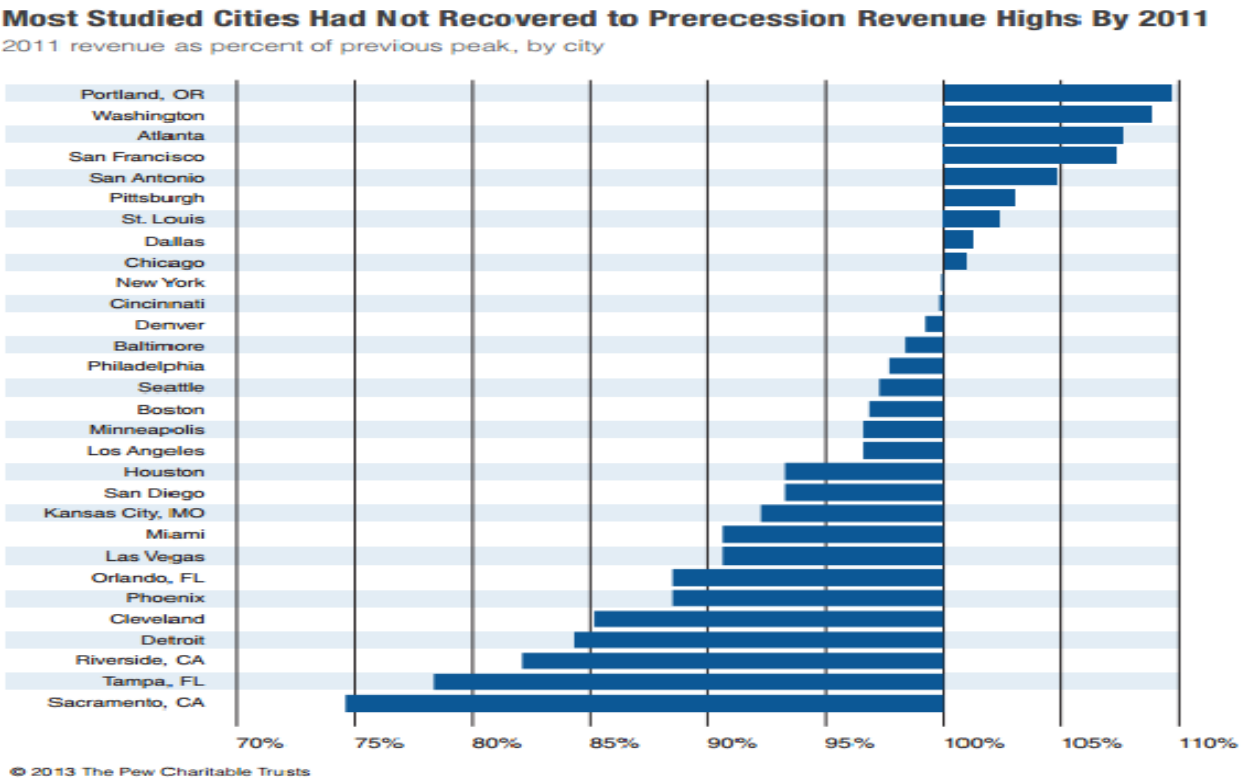
the year 2010. By 2030, 6 in 10 and by 2050, 7 in 10 are expected to be living in cities (CDC, 2011). Therefore, the study was primarily interested in analyzing the cardiovascular health of the city dwellers.

Economic recession is a period of falling economic activity spread across the economy. The worldwide recession took place from late 2008 to early 2009. But in United States, the recession started in December 2007 and ended officially only in June 2009, spanning over 18 months. The Gross Domestic Product (GDP) of U.S did not improve until the later 2010 and only 9 cities had recovered to their previous revenue peak by 2011(PEW charitable trust, 2013).




American cities have a significant impact on the economies and long-term prosperity of states and the nation. The economic recession from 2007 to 2009 created great fiscal

challenges for American cities that have made them slow to recover(PEW charitable trust, 2013). The details on how each city navigated the worst economy during and after the great depression is provided in the online article "America's Big Cities in Volatile Times" published by PEW Charitable Trusts. This study was primarily influenced by this article. The 30 cities in this study and their metropolitan areas account for 49 percent of the nation's gross domestic product. Collectively, they have nearly 34 million residents i.e. more than 1 in 10 Americans with an additional 108 million who temporarily anchor here. These populations depend on the core functions of municipalities (PEW charitable trust, 2013). When the inflow of government revenue is low and demands of public increases then local leaders have to make tough choices. This will have serious implications on the nation's economy and basic infrastructure. This can in turn affect the health of its citizens (Leiseca, 2013).



Previous studies have shown that during recession there was increase in obesity and overweight as many opted for cheaper fast food meals than preparing healthy meals at home which many could not afford. Although economic pressures is expected to increase smoking and alcohol consumption, but increasing their taxes during such periods has increased the call to quit lines by 369% compared to the year before the recession. Also, it is reported that 1 in Americans skipped medications or doses and 19% postponed their preventive health care visits. Many resorted to home remedies or over the counter drugs than visiting a physician. Additionally, women were more likely than men to report indulging in unhealthy behaviors such as eating poorly due to stress (Hughes et al., 2009)

Very few studies have focused on the effect of economic recession on cardiovascular health. Many CVD studies focus on general population or identify the prevalence across different states. But this research provides a detailed relationship of common CVD risk factors along with the per capita real GDP of 30 major American cities where many working classes reside. Any impact on economic growth will first hurt the people in these cities more than others. It is important to evaluate this relationship among this specific population as it will help the healthcare provider to remain vigilant during such tough economic periods, to ensure that patients receive evidence based prevention and treatment strategies. Also, policies can be made to reduce the screening costs or medication costs of cardiac diseases during economic recessions to save the lives of citizens already under stress. Additionally, it will help establish a multidisciplinary approach to combat CVD in the future.

## RESEARCH DESIGN AND METHODS

### Overview

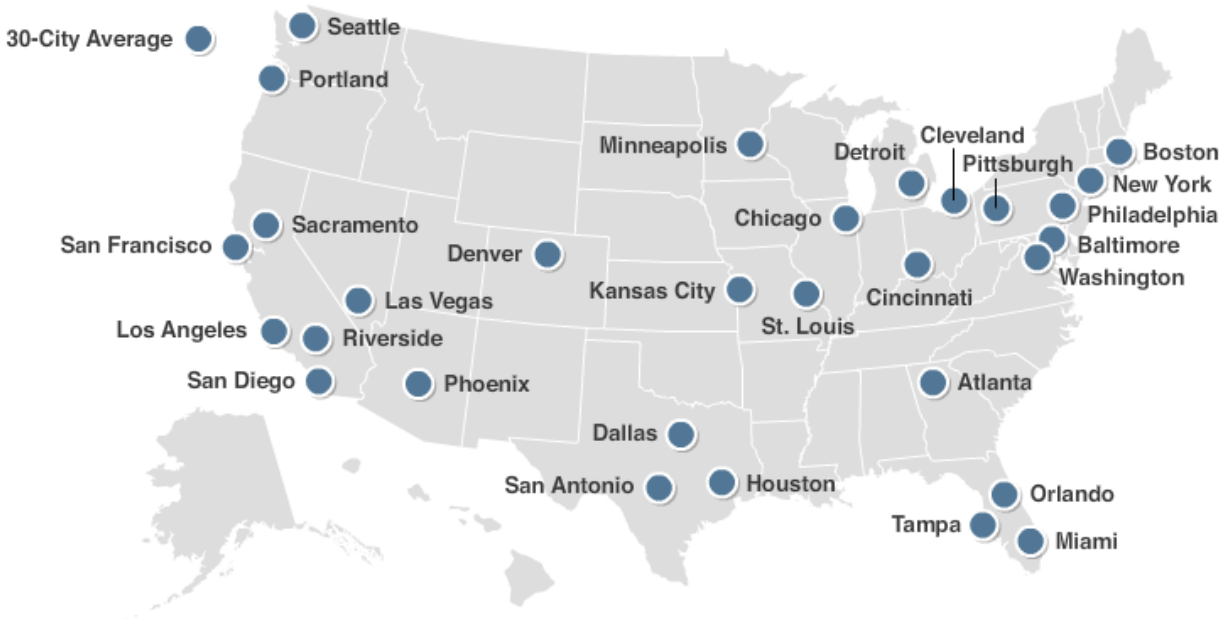
A cross sectional study with an ecological component comparing CVD risk in 30 major cities in United States was conducted using publically available cross sectional datasets from Behavioral Risk Factor Surveillance System (BRFSS) and the U.S Bureau of Economic Analysis (BEA) between the years 2006 and 2012.

### Data Sources

The 30 major cities were selected based on the article "America's Big Cities in Volatile Times" published by PEW Charitable Trusts. The PEW's American cities project conducts research and analyses to help policymakers to understand the economic challenges of the cities as they consider policy choices that could affect these localities significantly. The 30 cities are as follows (Leiseca, 2013).

- Atlanta, GA
- Baltimore, MD
- Boston, MA
- Chicago, IL
- Cincinnati, OH
- Cleveland, OH
- Dallas, TX
- Denver, CO
- Detroit, MI
- Houston, TX
- Kansas City, MO
- Las Vegas, NV
- Los Angeles, CA
- Miami, FL
- Minneapolis, MN
- New York, NY

- Orlando, FL
- Philadelphia, PA
- Phoenix, AZ
- Pittsburgh, PA
- Portland, OR
- Riverside, CA
- Sacramento, CA
- San Antonio, TX
- San Diego, CA
- San Francisco, CA
- Seattle, WA
- St. Louis, MO
- Tampa, FL
- Washington, D.C.




**Behavioral Risk Factor Surveillance System (BRFSS)**

The BRFSS surveys are conducted annually since 1984 by the state departments of health, with assistance from the Centers for Disease Control and Prevention (CDC). It is the world's

largest ongoing telephone health surveillance system. It gathers information on participant's demographics, socioeconomic status, dietary, personal health behaviors and health status of adults of all ages above 18 years, races, sex and socio-economic status from all over the country (CDC, 2013). Data is available for all the 50 states and District of Columbia and at the county level. The survey relies on random-digit dialing to interview adults who are part of the civilian, noninstitutionalized population. The survey has 3 components – core component, optional questionnaire and state added questions. The core component is a fixed set of questions asked by all states which includes questions on demographics and current health behaviors such as smoking. The rotating core is made up of two different sets of questions, each asked in alternating years by all states. About a set of five questions are added to the existing ones every year which are called the emerging core. Questions on specific topics are called the optional modules that states elect to use on their questionnaire (CDC, 2013).

Only questions from the core component were included in the study for all the 7 years. The variables included in the study were based on the self – reported physician's diagnosis of health conditions i.e. Stroke, Angina pectoris, Coronary Heart Disease and Diabetes Mellitus. The variables Stroke, Angina pectoris and Coronary Heart Disease were combined together to form the independent variable - cardiovascular disease (CVD). The dependent variables included from BRFSS were age, gender, race/ethnicity, city, education, employment, income, marital status, insurance status, alcohol intake, smoking status, depression, diabetes and body mass index (BMI). Since the information was obtained through interviewer administered telephonic survey, the information provided by the respondents had to be relied upon without an opportunity to cross verify with their medical or legal records.

**U.S Bureau of Economic Analysis (BEA)**

The U.S Bureau of Economic Analysis is a world's leading statistical agency under the Department of Commerce and is a part of the Department's Economics and Statistics Administration, along with Census Bureau (Commerce, 2014). It produces economic statistics for the government, researchers and public to understand the economic performance of the country. It produces a detailed quantitative economic data at the national, regional, state, metropolitan area and county level (Commerce, 2014). The gross domestic product (GDP) which is the preferred measure of economic activity is updated every quarter. The economic data on each city's per capita real GDP was obtained from the BEA for all 7 years and merged with the BRFSS datasets.

## Subjects

**Sample definition:**

The sample consisted of all men and women above 18 years of age and of all racial groups from 30 major American cities who were included in BRFSS survey, conducted by the Centers for Disease Control and Prevention (CDC) in all the states of the United States through the telephone survey. Thus, the sample is representative of the urban population of United States (2006 – 2012).

**Inclusion criteria**

The participants in the study should meet the following criteria to be included in the study:

- All individuals in the study sample should be above 18 years of age.
- The employment status of these individuals should be available.
- All the individuals with any previous history of cardiovascular diseases at the time of data collection.
- The individuals should belong to any one of the 30 major cities of United States as mentioned above.

After defining all the inclusion and exclusion criteria, the total sample size of the study inclusive of all the 30 cities and 7 consecutive years consisted of 997161 participants.

**Institutional Review Board consideration**

This is a secondary data analysis of a publically available de-identified dataset – BRFSS and BEA, 2006-2012 to assess the association between CVD and economic depression. The investigators conducting this research do not have access to any codes or links that may be used to re-identify the subjects in the data set. No personal information or identifiers were collected and responses cannot be traced back to the source subject. Therefore, this study was approved by the Drexel University's Office of Research Internal Review Board (IRB) as a non-human research for not involving human subjects as defined by United States Department of Health and Human Services (DHHS) and U.S Food and Drug Administration (FDA) regulations. IRB submission was approved in February 2014.

**Variable definition and measurement**

The primary variables of interest that will be considered during in this analysis are cardiovascular diseases and per capita real GDP along with other covariates that are potential risk factors of CVD.

Cardiovascular Diseases

It is a group of disorders of the heart and blood vessels which has the following major forms (Liu, 2014).

Coronary heart disease: It is the disease of the blood vessels supplying the heart muscle. It occurs when blood supply to the heart muscle is partially or completely blocked. It is of five types - Acute myocardial infarction/ heart attack (410), other acute and sub-acute forms of ischemic heart disease (411), old myocardial infarction (412), Angina pectoris (413) and other forms of chronic ischemic heart disease (414) (American Heart Association, 2011).

- Cerebrovascular disease: It is the disease of the blood vessels supplying the brain. Blockage or bursting of the blood vessel in the brain causes stroke. It is a medical emergency and third cause of death in U.S. It is of two major types - Ischemic stroke (ICD: 433-434) and Hemorrhagic stroke (ICD: 430-431) (American Heart Association, 2011).

Gross Domestic Product (GDP): The GDP is the most commonly used indicator of the economic status of a geographical region. It gauges the region's standard of living. It is defined as the monetary value of all finished goods and services produced within a region 's borders in a

specific time period, though GDP is usually calculated on an annual basis. It includes all of private and public consumption, government outlays, investments and exports less imports that occur within a defined territory (Investopedia).

**GDP = C + G + I + NX**

where:

"**C**" is equal to all private consumption, or consumer spending

"**G**" is the sum of government spending

"**I**" is the sum of all the region 's businesses spending on capital

"**NX**" is the nation's total net exports, calculated as total exports minus total imports. (NX = Exports - Imports)

But GDP excludes imports, foreign income from American companies and people and does not account for inflation. Therefore, it is not an accurate measure for comparison (Investopedia).

Per capita Real GDP: It is the measure of the total output of a region and takes the gross domestic product (**GDP**) and divides it by the number of people in the region. It adjusts for the price changes or inflation and is the accurate economic indicator for comparing economic indicators of different regions with very different population sizes (US Economy, 2014).

**Per capita Real GDP = Real GDP of the region/ Total population of the region**

Other variables such as age, race/ethnicity, gender, education, employment, income, marital status, exercise, insurance, smoking status, alcohol consumption, depression, diabetes, and BMI will be considered as potential confounders and will be adjusted for in the analysis.

**Data analysis**

All the extraction and merging of datasets from BRFSS was done using SAS 9.3. Three cardiovascular outcomes were obtained from the BRFSS dataset from the year 2006 – 2012, namely Stroke, Angina pectoris and Coronary Heart Disease. Since Angina is a type of CHD, both the variables were combined together to form the coronary heart disease. The CVD outcome binomial variable was obtained by combining the stroke and coronary heart disease. Based on the 2000 US standard projected population, age was divided into 6 categories (18 to 24, 24 to 34, 35 to 44, 45 to 54, 55 to 64 and above 65 years). All the 7 years of BRFSS dataset were merged together along with the economic data obtained from the BEA. For the ease of analysis, the cities were categorized into 3 zones based on their age adjusted prevalence rates – high, medium and low prevalence zones consisting of 10 cities in each zone. The list of city zones is given in Table 1. The 7 years were also divided into three phases – Phase1, Phase 2 and Phase 3 representing the period before, during and after recession. Phase 1 consists of 2006 and 2007Similarly, the per capita real GDP was also categorized into low, medium and high GDP. 20,000 – 40,000 dollars/person/year was considered as low GDP, 40,000 – 60,000 dollars/person/year as medium and 60,000 – 80,000 dollars/person/year as high GDP.

Chi – square tests were carried out for categorical variables. Figure 1 and 2 shows the age standardized estimates of CVD prevalence rates across all 30 American cities. Figure 3 and 4

presents the age standardized CVD prevalence rates with respect to gender and race/ethnicity. Figure 5 depicts the age standardized CVD prevalence rates and per capita real GDP across all the 7 years. The association between CVD risk and per capita real GDP was evaluated over all the years and cities using a multivariate survey logistic regression along with other covariates. A full model with all the covariates and a reduced model only with the democratic variables against the CVD were performed. In the model, we used low prevalence zone as the baseline. All analyses were done that accounted for the complex survey design of BRFSS dataset. The results of the multivariate analysis are provided in Table 1.

**Results**

The final analytic sample with all observations including missing values was a total of 997161 out of which 68.68% were females and 31.32% were males. Off these only 7.8% of females and 4.88% of males were diagnosed with CVD. The calculated BRFSS overall prevalence of stroke and CHD in all 30 cities were 8.6% and 4.2%. The age adjusted prevalence of CVD was highest in Detroit with 164.82 cases/1000/year followed by 152.19 cases/1000/year in St. Louis and 145.12 cases/1000/year San Antonio. The lowest prevalence was about 59.34 cases/1000/year in San Francisco, 73.33 cases/1000/year in Seattle and 74.703 cases/1000/year in Minneapolis. Based on the prevalence rates, the cities were divided into three zones – high, medium and low risk zones, with each zone containing 10 cities. The high risk zone contains about 123 to 164 cases/1000/year and medium risk zone contains 123 to 102 cases/1000/year. The low risk zone contains 60 to 102 cases/1000/year. The prevalence of CVD was highest among males (155.8 cases/1000/year) compared to that of females (113.72 cases/1000/year).

With respect to race, American Indians/ Alaskan Natives (70.14 cases/1000/year) followed by Whites (58.51 cases/1000/year), Blacks (49.10 cases/1000/year) and Hispanics/Latinos (48.56 cases/1000/year) had higher prevalence of CVD.

When CVD prevalence and per capita real GDP was plotted against years, there was a linear increase in CVD rates over the years. But the greatest peak was seen during the year 2010 (138.46 cases/1000/year). This was right after the economic recession ended officially in 2009 although the cities did not recover completely after the recession. Similar trend were seen for individual cities as well. For example, when the CVD rates of a high risk city such as Detroit was compared against the economic data; there was an increase in the CVD rates during the recession period (181.22 cases/1000/year) and decreased almost to the baseline rates (146.1 cases/1000/year) after the recession phase(152.69 cases/1000/year). For a low risk city such as San Francisco, the economic status was almost similar before and during the recession periods. It increased dramatically after the recession phase. Its effect was reflected on the CVD rates where there was an increase in the CVD prevalence before the recession (88.23 cases/1000/year) and decreased once economy improved (52.63 cases/1000/year). These results are best explained through illustrations in Figure 6, 7, 8 and 9.

Statistically significant differences in CVD risk was observed by age, gender, race/ethnicity, education levels, employment, income, marital status, insurance status, exercise, smoking status, alcohol consumption, city zone, diabetes, BMI, mental depression and per capita real GDP. The results of the multivariate analysis are provided in Table 2.

As age increases the odds of CVD risk also increases after adjusting for all other covariates. Almost all categories of age were found to be statistically significant except the 25 – 35 years in a full model during multivariate analysis. The older age group of 65 and above had the highest prevalence of CVD of 38 cases/1000/year. Males were found to be at greater risk for CVD than females. Disparities were seen with respect to races as well. American Indians/ Alaskan Natives were significantly more prone to CVD followed by Blacks and Hispanics/ Latinos with Whites as the baseline. But Blacks and Hispanics/ Latinos were not significantly associated with CVD risk holding all other variables constant. Significant association was observed with respect to current smokers and former smokers with never smokers as the baseline.

Alcohol consumption, exercise, income, education and per capita real had a protect effect with respect to CDV risk. Drinking more than 1 glass (for females) and more than 2 glasses (for males) of alcohol everyday improved the cardiac health which is mimicking the results of the existing literature. Increasing educational levels and income decreased the CVD risk. More importantly, compared to low GDP, medium and high GDP had a protective effect against CVD risk. This proves that the economic status of a country and individual significantly impacts the cardiac health of a population.

Insurance increased the odds of developing CVD. But this could be due to individuals with insurance visiting the doctor more often and hence detected with CVD than those without any insurance. Retired, unemployed, homemaker, self-employed and those unable to work had greater odds of developing CVD compared to wage earners. Majority of retired persons are above 65 years of age who have more prone to CVD and those unable to work might be

physically or mentally handicapped with many disadvantages such as limited exercise. This might be the reason for their increased risk for CVD.

Mental depression, diabetes mellitus and increasing BMI all were found to be significantly associated with CVD risk holding all other variables constant.

**Discussions**

This study exclusively deals with the city dwellers which is the new emerging population gaining more attention. This report provides the prevalence of CVD among different American cities and the impact of economic recession on CVD risk. Previous studies have conducted similar type of prevalence analysis for different states in United States. But this study deals with cities and the urban population only. Additionally, the results obtained studying the impact of economic recession on CVD is similar to the previous studies in literature. A study by Hughes and Dennison has proven the relationship between CVD risk and economic downturn. They formulate that depression, stress and behavioral habits are factors that play a role in the pathogenesis and expression of CVD (Hughes & Dennison, 2009). We have comparable results where there was a significant relationship between CVD and smoking as well as depression which could have been one of the reasons for increased risk during recession periods.

According to the study conducted by M.H Brenner o economic changes and heart disease mortality, there was a lag in the heart disease mortality behind changes in employment status by two years (Brenner, 1971). Similar trend was seen in our study, where a lag in CVD was seen behind economic recession period with increased unemployment rates. All these results helps us to conclude that in addition to well established risk factors of CVD, economic

recession has a significant impact on the cardiovascular health of urban population as it has the potential to trigger all the risk factors of CVD.

**Conclusions**

The study confirms that in addition to well established risk factors of CVD, economic recession has a significant impact on the cardiovascular health of the urban population. Economic recession has the potential to trigger all the risk factors of CVD such as unemployment, stress, depression, poor diet, change in socio economic status and increased smoking.

**Strengths**

As a population based, cross sectional survey performed across the United Stated every year, the BRFSS offers the strengths of a large sample size as well as power. The stratified random sample provides the ability to examine data at various geographical levels. This study is exclusively based on the cardiovascular health of the urban population on which very few studies have been carried out. Also, it illustrates the prevalence of risk factors and disease for different cities. Moreover, it compares the effects of different phases of economic crisis (before, during and after) on CVD, thereby filling the gap in the current literature.

**Limitations**

The greatest limitation of BRFSS in terms of accuracy is the data being dependent on self-reports. Self-reporting may cause desirability bias and certain measures that had to be

measured objectively were calibrated by the participants' responses, thereby underestimating the results. Additionally, BRFSS includes only data from those who own a landline and excludes individuals admitted in institutions and long term care facilities. Moreover, the BRFSS is a prevalence data that includes only those individuals with a favorable survivorship.

The study was cross sectional and did not have any evidence of temporal association between the exposure and outcome of interest. Since there was an ecological component to this study design, the problem of ecological fallacy exists as well. Additionally, few variables or possible confounders such as diet, hypertension and lipid profile could not be included as they were available only on alternative years. An exclusive mental depression variable was included only since 2011. Therefore, an alternative variable for depression that was common in all 7 years' datasets were included. Furthermore, the income could not be adjusted for the rise in dollar value over the years.

**Recommendations**

The data can be used by the national and state level CVD prevention programs as a baseline to assess the CV health at the city level. The study highlights its flexibility to document disparities by the economic, socio demographic status, geographic region. The study supports the need to continue advocating for tobacco, diabetes and obesity control to further reduce the burden of the disease. It recommends for vigilant health care providers during economic distressed periods and to establish a multidisciplinary approach such as clinical psychologists to care for cardiac patients in the future. There is also a need to incorporate mental depression,

employment status and socio-economic status into CVD risk calculations and screening tools due to its profound effect on the disease.

Our findings of association of CVD with economic recession and prevalence of high CVD rates in certain cities are consistent with the findings from the literature. Therefore, findings from this report can be used by stakeholders to initiate and support programs to improve CV health among the urban population. Additionally, further research is required to study CVD in detail and to overcome the limitations of the present study.

**References**

American Heart Association. (2011, 12 28). *What is Cardiovascular Disease (Heart Disease)?* Retrieved from http://www.heart.org/HEARTORG/Caregiver/Resources/WhatisCardiovascularDisease What-is-Cardiovascular-Disease_UCM_301852_Article.jsp#

Brenner, M. H. (1971). Economic changes and heart disease mortality.*American Journal of Public Health*, *61*(3), 606-611.

Carney, R. M., Freedland, K. E., Miller, G. E., & Jaffe, A. S. (2002). Depression as a risk factor for cardiac mortality and morbidity: a review of potential mechanisms. *Journal of psychosomatic research*, *53*(4), 897-902.

Centers for Disease Control and Prevention. (2011, March 31). An Estimated 1 in 10 U.S. Adults Report Depression. Retrieved from http://www.cdc.gov/features/dsdepression/

Centers for Disease Control and Prevention CDC. 2011, Prevalence of coronary heart disease United States, 2006-2010. *MMWR. Morbidity and mortality weekly report*, *60*(40), 1377.

Centers for Disease Control and Prevention. (2013, march 19). *Behavioral Risk Factor Surveillance System*. Retrieved from http://www.cdc.gov/brfss/about/about_brfss.htm

Centers for Disease Control and prevention. (2013, August 28). *Heart Disease Facts*. Retrieved from www.cdc.gov: http://www.cdc.gov/heartdisease/facts.htm

Clark, A. M., DesMeules, M., Luo, W., Duncan, A. S., & Wielgosz, A. (2009). Socioeconomic status and cardiovascular disease: risks and implications for care. *Nature Reviews Cardiology*, *6*(11), 712-722.

Centers for Disease Control and Prevention CDC. (2005). Racial/ethnic and socioeconomic disparities in multiple risk factors for heart disease and stroke--United States, 2003. *MMWR. Morbidity and mortality weekly report*, *54*(5), 113.

Commerce, U. D. (2014). Retrieved from Bureau of Economic Analysis: http://www.bea.gov/iTable/iTable.cfm?reqid=70&step=1&isuri=1&acrdn=2#reqid=70&step=1&isuri=1&7003=1000&7035=-1&7004=naics&7005=1&7006=00998&7036=-

1&7001=21000&7002=2&7090=70&7007=2012,2011,2010,2009,2008,2007,2006&7093=percentchange

Dahlöf, B. (2010). Cardiovascular disease risk factors: epidemiology and risk assessment. *The American journal of cardiology*, *105*(1), 3A-9A.

Goldston, K., & Baillie, A. J. (2008). Depression and coronary heart disease: a review of the epidemiological evidence, explanatory mechanisms and management approaches. *Clinical psychology review*, *28*(2), 288-306.

Gonzalez, O., Berry, J. T., McKnight-Eily, L. R., Strine, T., Edwards, V. J., Lu, H., & Croft, J. B. (2010). Current depression among adults: United States, 2006 and 2008. *Morbidity and Mortality Weekly Report*, *59*(38), 1229-1235

Hughes, S., & Dennison, C. R. (2009). Progress in Prevention: Economic Downturn = Cardiovascular Disease Risk Uptick? *The Journal of Cardiovascular Nursing*.

*Investopedia*. (n.d.). Retrieved from 2014: http://www.investopedia.com/terms/g/gdp.asp

Klein, R. J., & Schoenborn, C. A. (2001). Age adjustment using the 2000 projected US population. Centers for Disease Control and Prevention CDC.

Leiseca, S. (2013, November 11). *America's Big Cities in Volatile Times*. Retrieved from PEW Charitable Trust: http://www.pewstates.org/research/data-visualizations/americas-big cities-in-volatile-times-85899509284

NIMH. (2013). *What Is Depression?* . Retrieved from National Institute of Mental Health: http://www.nimh.nih.gov/health/topics/depression/index.shtml

O'Connor, A. (2008, september 25). *How the Financial Crisis Could Break Your Heart, Literally*. Retrieved from Health.com:http://www.health.com/health/article/0,,20411265,00.html

The PEW charitable trust. (2013). *America's Big Cities in Volatile Times.* Washington: The PEW charitable trust.

Suzanne Hughes, C. R. (2009). Progress in Prevention: Economic Downturn = Cardiovascular Disease Risk Uptick? *The Journal of Cardiovascular Nursing*. doi:10.1097/JCN.0b013e3181a9bf87

*US Economy*. (2014). Retrieved from About.com: http://useconomy.about.com/od/grossdomesticproduct/g/Real-GDP-Per-Capita.htm

## APPENDIX A: FIGURES

**Figure 1: Age adjusted CVD prevalence across all 30 American cities**

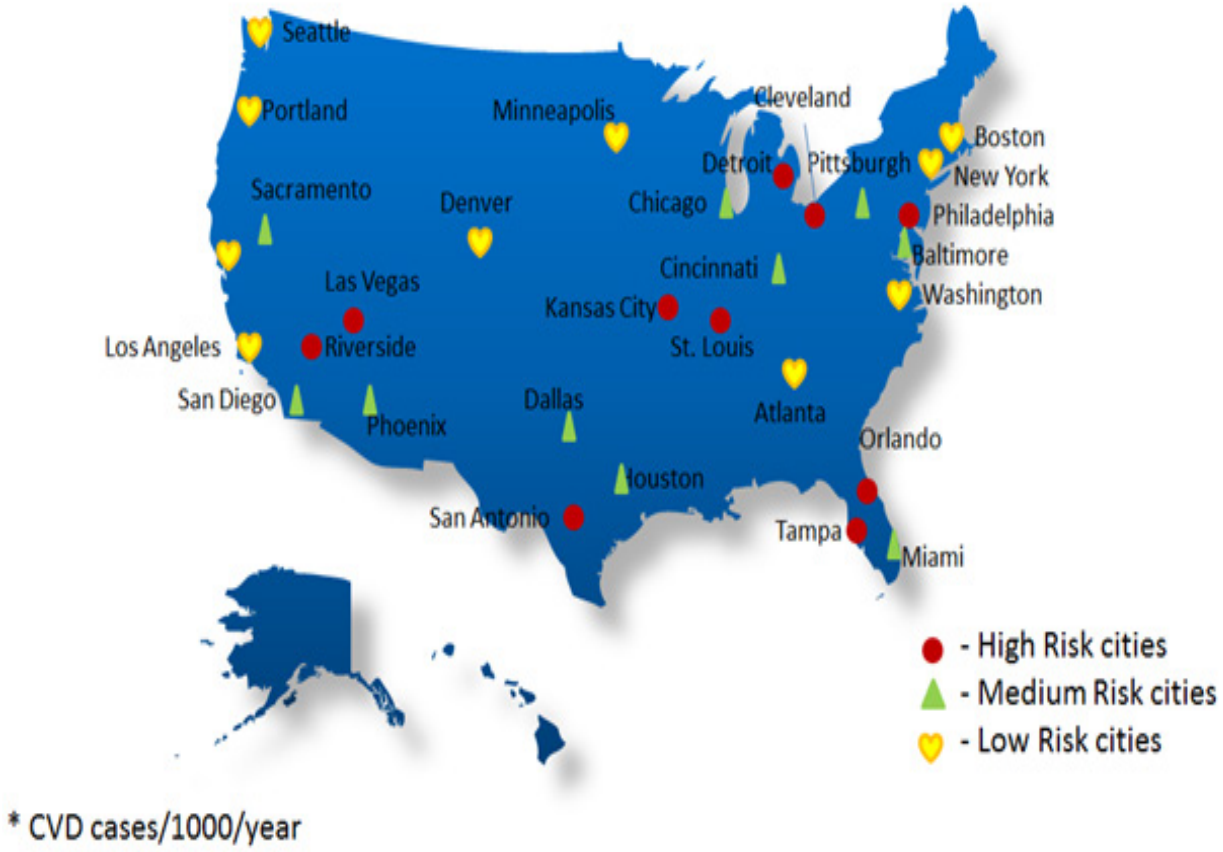


**Figure 2: Age adjusted CVD prevalence across all 30 American cities**

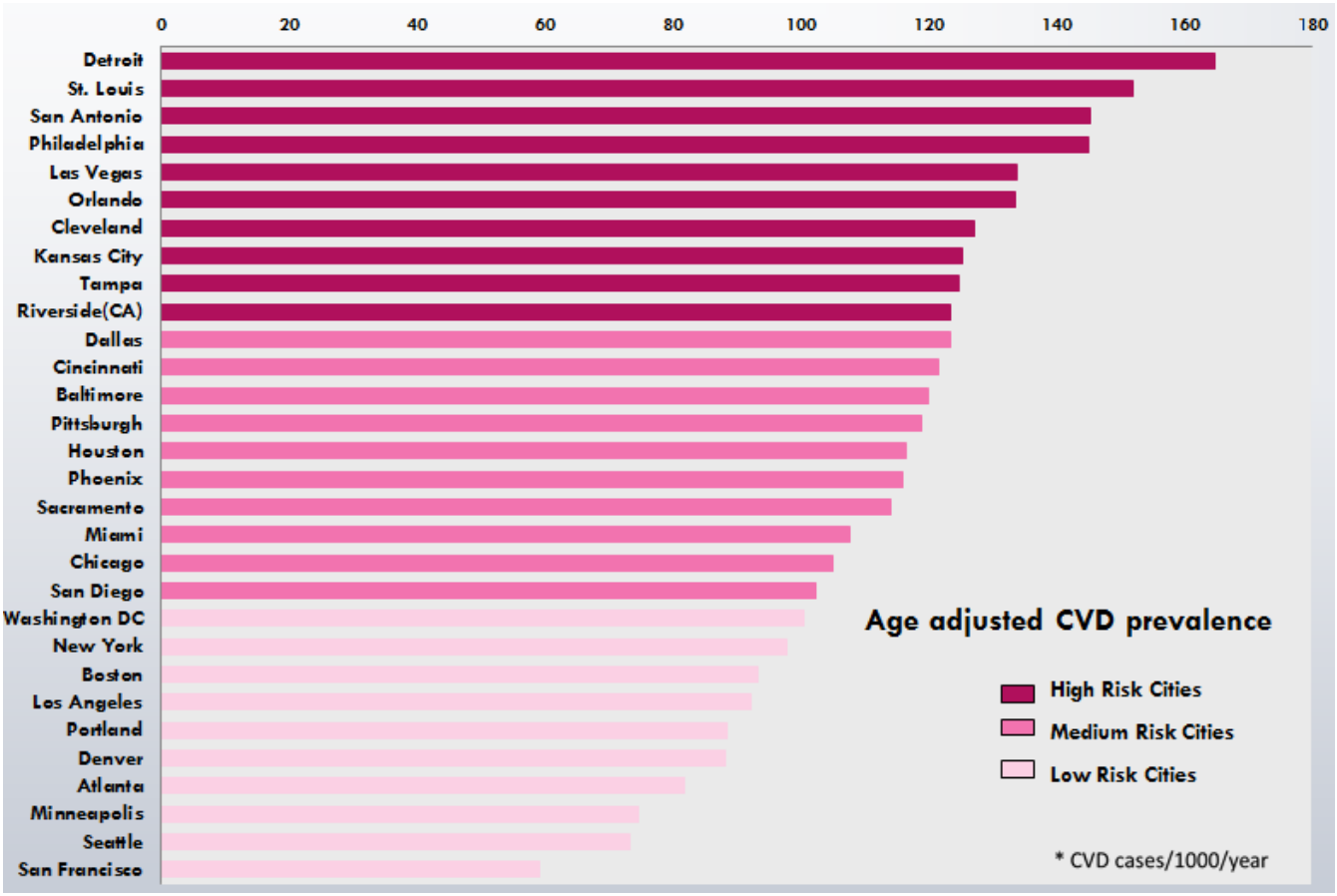


**Figure 3: Age adjusted CVD prevalence with respect to gender**

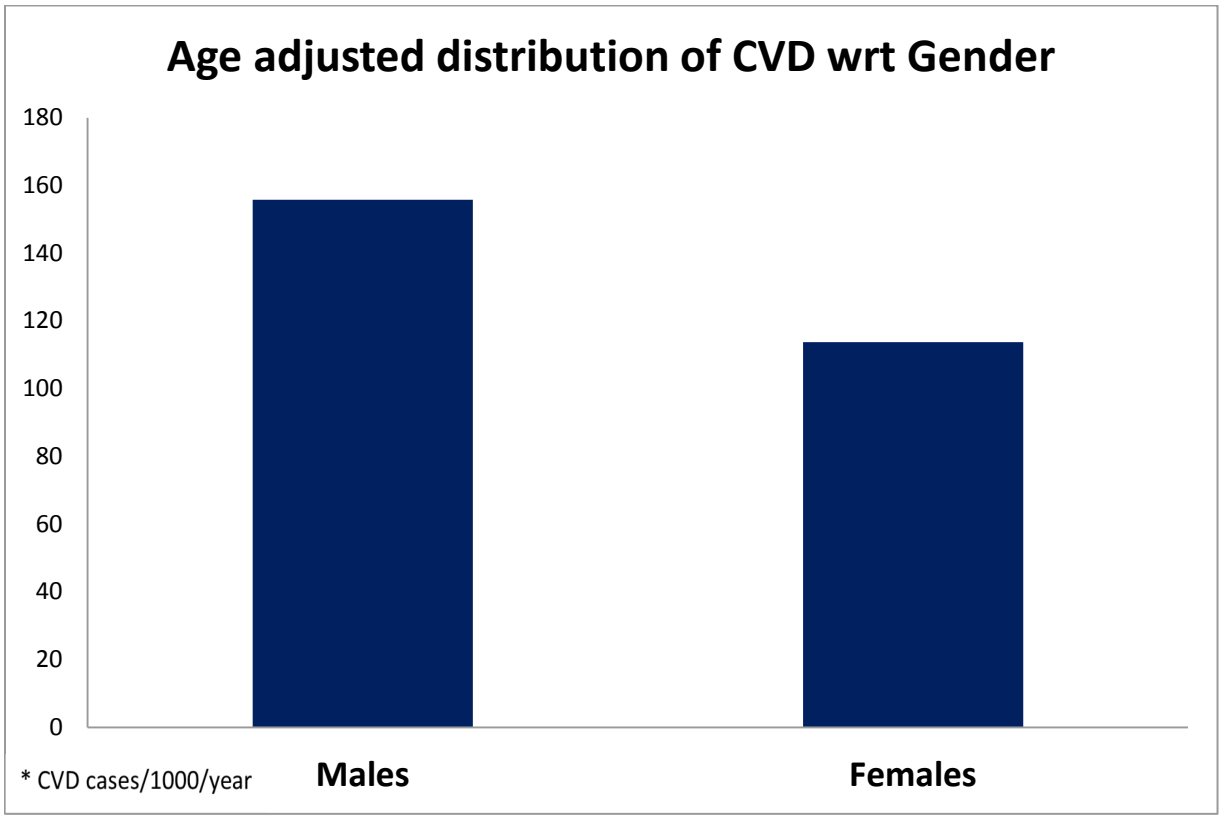


**Figure 4: Age adjusted CVD prevalence with respect to race/ethnicity**

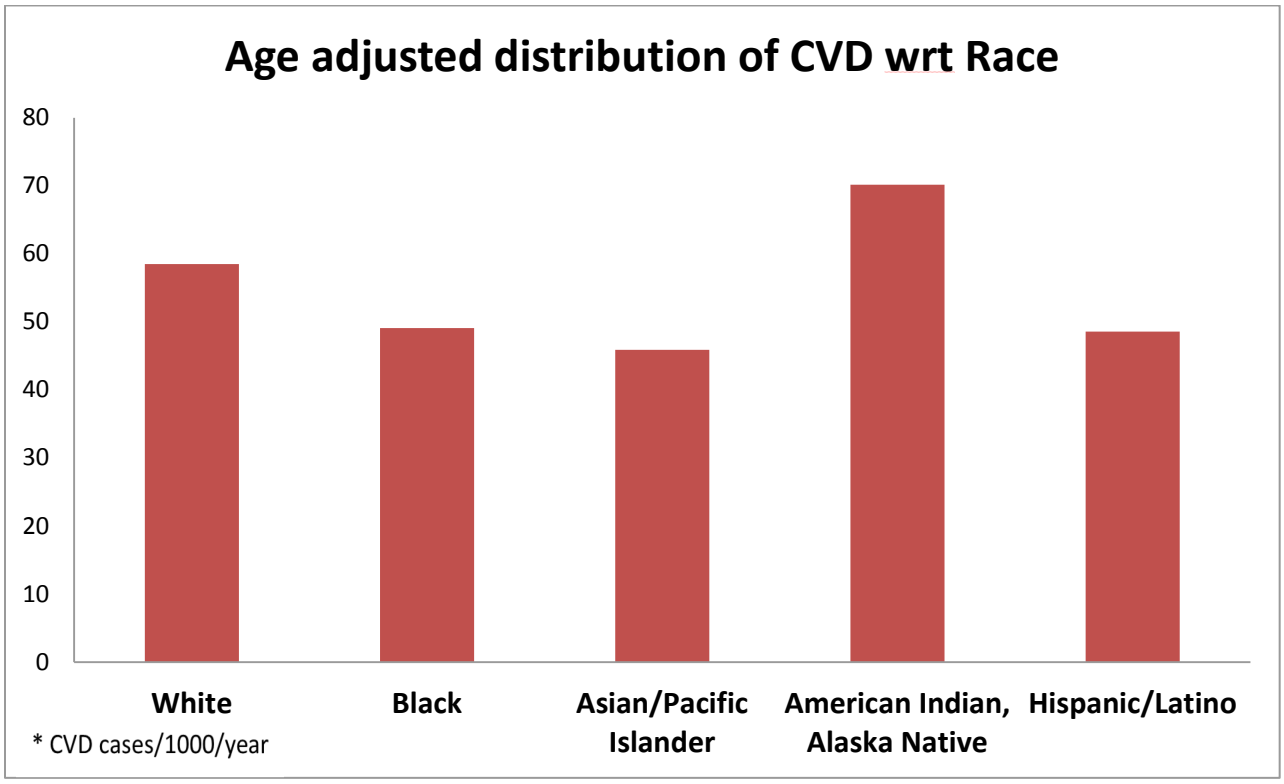
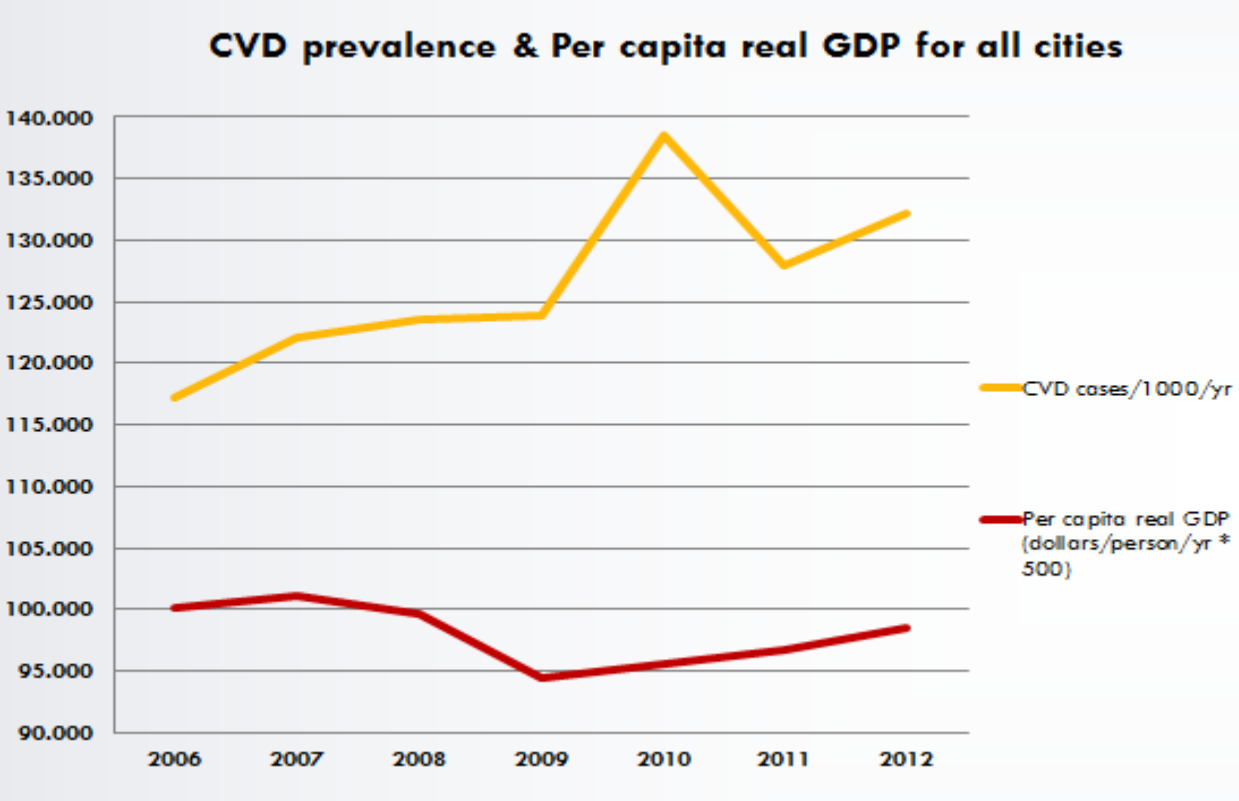


**Figure 5: Age adjusted CVD prevalence and per capita real GDP across all 7 years (2006 – 2012)**




*

**Figure 6: Detroit – Economic revenue data**

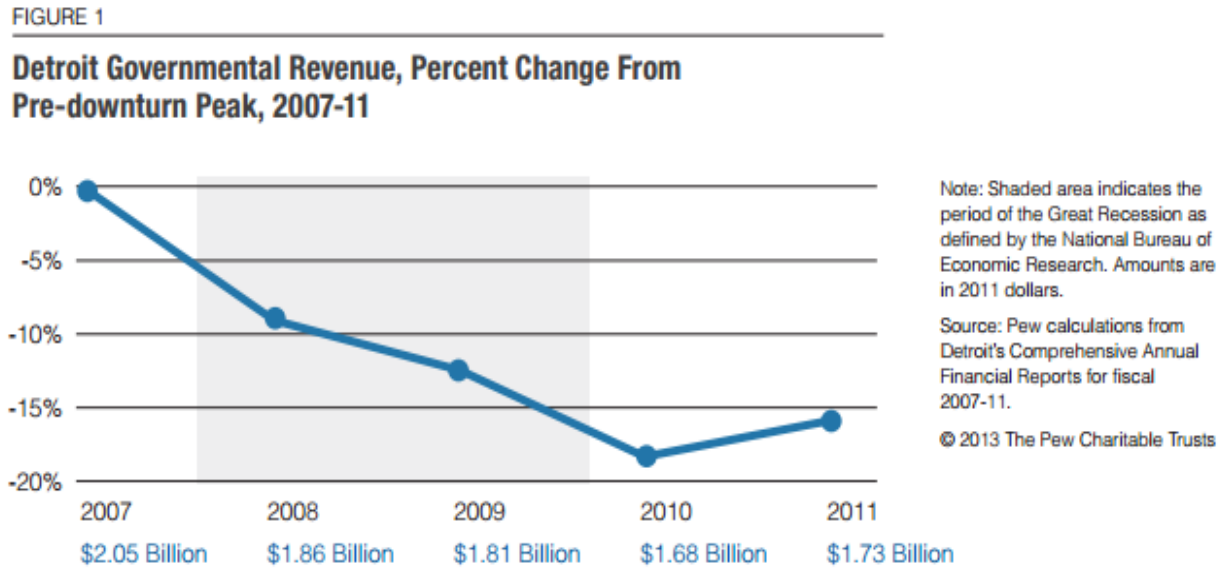



**Figure 7: Detroit – CVD prevalence**

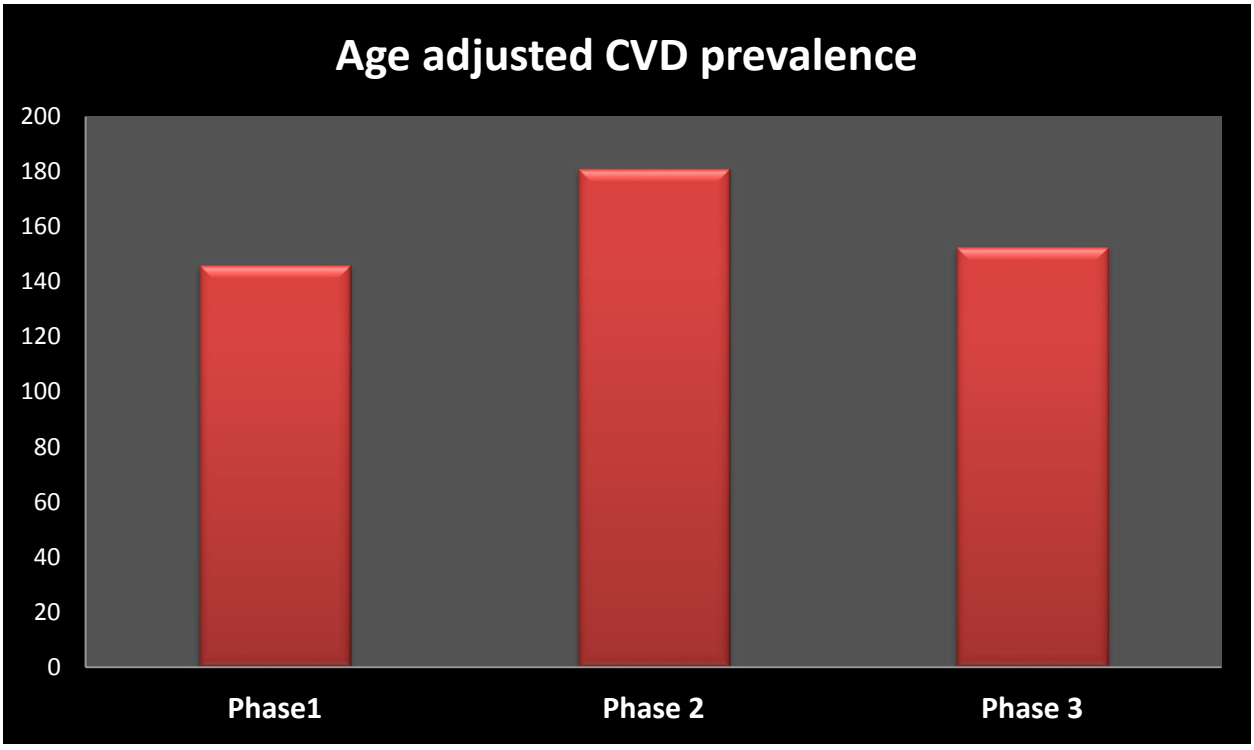


**Figure 8: San Francisco - Economic revenue data**

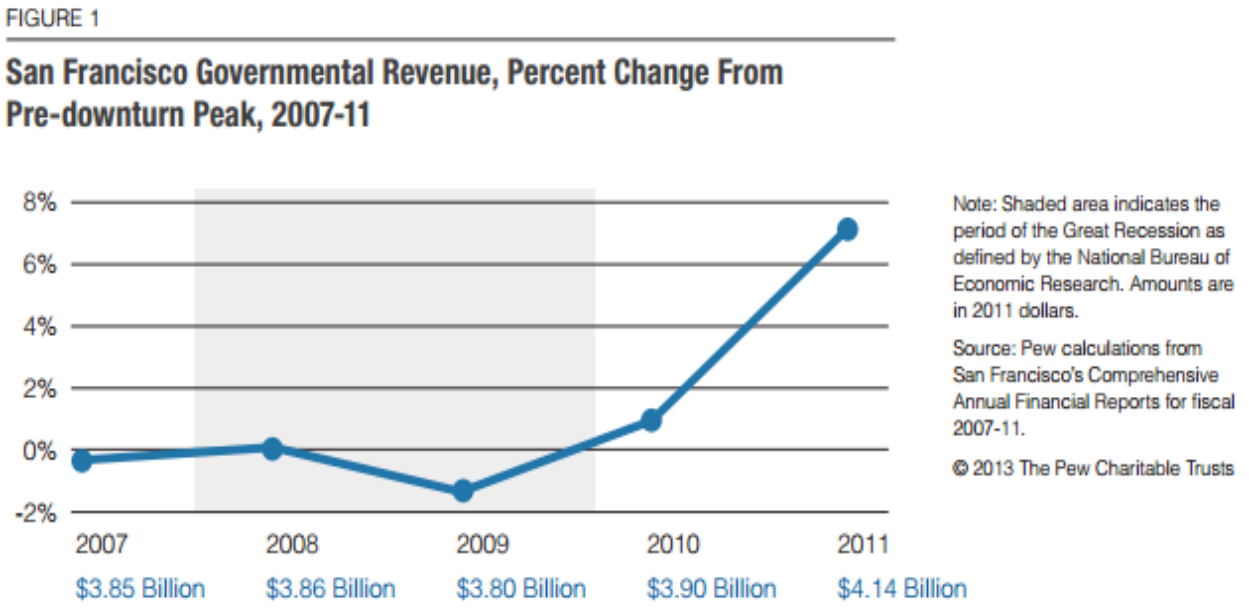
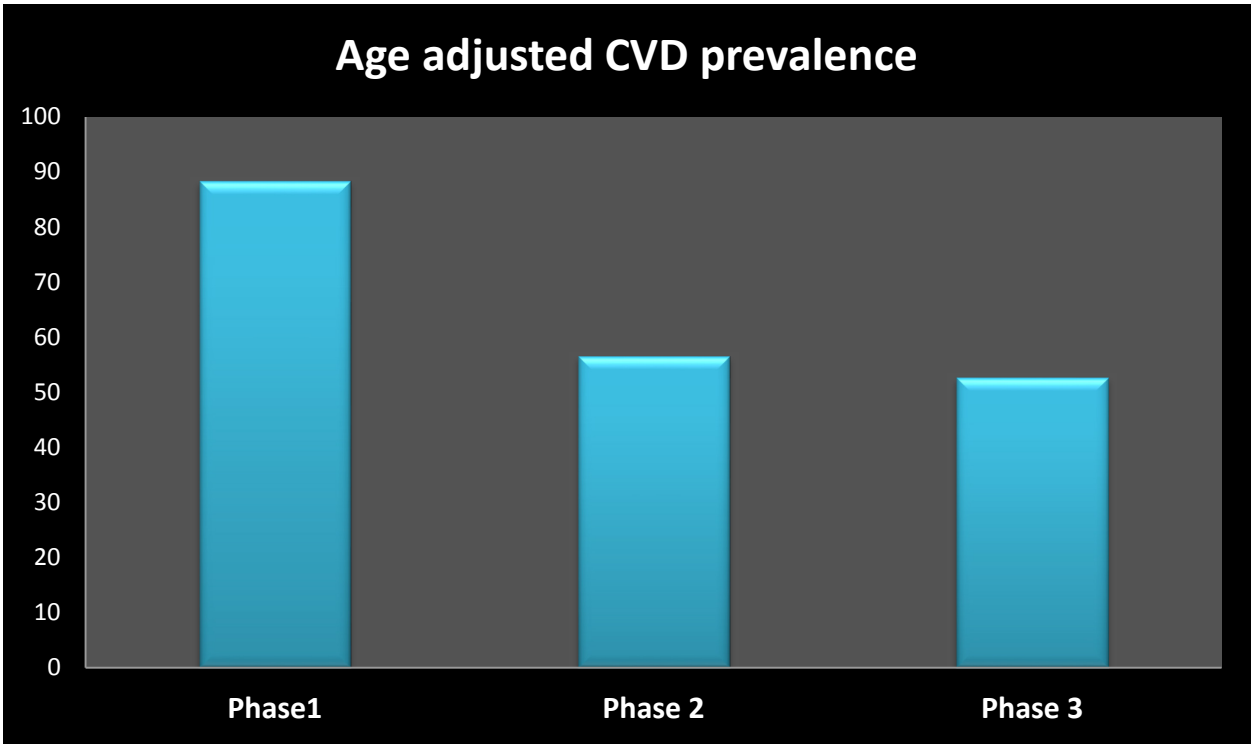



**Figure9: San Francisco - CVD prevalence**



## APPENDIX B: TABLES

**Table 1: Division of cities into three risk zones**

| High Risk | Medium risk | Low risk |
|---|---|---|
| • Detroit | • Dallas | • Washington DC |
| • St. Louis | • Cincinnati | • New York |
| • San Antonio | • Baltimore | • Boston |
| • Philadelphia | • Pittsburgh | • Los Angeles |
| • Las Vegas | • Houston | • Portland |
| • Orlando | • Phoenix | • Denver |
| • Cleveland | • Sacramento | • Atlanta |
| • Kansas City | • Miami | • Minneapolis |
| • Tampa | • Chicago | • Seattle |
| • Riverside (CA) | • San Diego | • San Francisco |

**Table 2: Association (Odds ratios and 95% confidence intervals) of CVD and economic recession using BRFSS 2006-2012.**

| COVARIATES | Model 1 (OR, 95% CI) | P value | Model 2 (OR, 95% CI) | P value |
|---|---|---|---|---|
| **Age** | | | | |
| 18 – 25 years | 1 | | 1 | |
| 25 – 35 years | 1.177 (1.009 -1.375) | 0.0386 | 1.151 (0.982 - 1.350) | 0.0833 |
| 35 – 45 years | 2.310(2.007 – 2.659) | <0.0001 | 2.172 (1.879- 2.511) | <.0001 |
| 45 – 55 years | 4.245 (3.710 -4.858) | <0.0001 | 3.813 (3.319 - 4.382) | <.0001 |
| 55 – 65 years | 6.812 (5.951 – 7.798) | <0.0001 | 6.009 (5.226- 6.910) | <.0001 |
| Above 65 years | 11.079 (9.662 – 12.704) | <0.0001 | 10.190 (8.846- 11.738) | <.0001 |
| **Gender** | | | | |
| Female | 1 | | 1 | |
| Male | 1.659(1.611 – 1.709) | <0.0001 | 1.637 (1.587 - 1.689) | <.0001 |
| **Race** | | | | |
| White | 1 | | 1 | |
| Black | 1.150 (0.888 – 1.489) | 0.2811 | 1.089 (0.833- 1.425) | 0.5326 |
| Asian/ Pacific Islander | 0.910 (0.698 – 1.185) | 0.4840 | 0.856 (0.645- 1.136) | 0.2812 |
| American Indian/Alaskan Native | 1.508 (1.249 – 1.821) | <.0001 | 1.407 (1.153- 1.718) | 0.0008 |
| Hispanic/Latino | 1.010 (0.955 -1.068) | 0.7249 | 0.958 (0.903- 1.016) | 0.1523 |
| **Education** | | | | |
| Graduate | 1 | | 1 | |
| Less than high school | 1.382 (1.319 – 1.448) | <.0001 | 1.289 (1.226-1.354) | <.0001 |
| High school graduate | 1.199 (1.154 – 1.245) | <.0001 | 1.151 (1.106-1.198) | <.0001 |
| Technical College | 1.233 (1.187 -1.281) | <.0001 | 1.183 (1.137- 1.231) | <.0001 |
| **Employment** | | | | |
| Wage earner | 1 | | 1 | |
| Self-employed | 1.120 (1.045 – 1.202) | 0.0014 | 1.126 (1.050- 1.209) | 0.0009 |
| Unemployed | 1.602 (1.502 – 1.709) | <.0001 | 1.511 (1.413- 1.615) | <.0001 |
| Homemaker | 1.581 (1.494 – 1.672) | <.0001 | 1.540 (1.453- 1.632) | <.0001 |
| Student | 1.000 (0.823 – 1.216) | 0.9984 | 1.022 (0.841 - 1.242) | 0.8261 |
| Retired | 2.076 (1.985 – 2.170) | <.0001 | 1.979 (1.889- 2.072) | <.0001 |
| Unable to work | 3.914 (3.742 – 4.095) | <.0001 | 3.324 (3.169- 3.486) | <.0001 |
| **Income** | | | | |
| Less than $15,000 | 1 | | 1 | |
| $15,000 - $25,000 | 1.002 (0.968 – 1.038) | 0.8899 | 1.001 (0.964- 1.039) | 0.9717 |
| $25,000 - $35,000 | 0.822 (0.786 – 0.860) | <.0001 | 0.827 (0.789- 0.867 ) | <.0001 |
| $35,000 - $50,000 | 0.710 (0.678 – 0.743) | <.0001 | 0.733 (0.699 - 0.768 ) | <.0001 |

| | | | | |
|---|---|---|---|---|
| More than $50,000 | 0.575 (0.551 – 0.601) | <.0001 | 0.612 (0.585 - 0.641 ) | <.0001 |
| **Marriage** | | | | |
| Married | 1 | | 1 | |
| Divorced | 1.011 (0.974 – 1.049) | 0.5576 | 1.019 (0.980- 1.059) | 0.3516 |
| Widowed | 1.125 (1.086 -1.165) | <.0001 | 1.130 (1.088- 1.173) | <.0001 |
| Separated | 1.123 (1.045 – 1.206) | 0.0015 | 1.100 (1.021- 1.184) | 0.0118 |
| Single | 0.738 (0.701 – 0.778) | <.0001 | 0.756 (0.716- 0.798) | <.0001 |
| Unmarried in a relationship | 0.965 (0.869 – 1.072) | 0.5071 | 0.983 (0.884- 1.093) | 0.7529 |
| **Exercise** | 0.745 (0.7225– 0.766) | <.0001 | 0.804 (0.782 - 0.828) | <.0001 |
| **Insurance** | 1.061 (1.011 – 1.115) | 0.0170 | 1.035 (0.984- 1.089) | 0.1870 |
| **Alcohol** | | | | |
| < 2 drinks/day | 1 | | 1 | |
| > 2drinks/day | 0.712 (0.663 – 0.765) | <.0001 | 0.763 (0.708- 0.822) | <.0001 |
| **Smoking** | | | | |
| Never smoked | 1 | | 1 | |
| Current smoker | 1.402 (1.351 – 1.455) | <.0001 | 1.442 (1.387- 1.500) | <.0001 |
| Former smoker | 1.350 (1.310 – 1.391) | <.0001 | 1.333 (1.292- 1.376) | <.0001 |
| **City Zone** | | | | |
| High risk | 1.256 (1.148 – 1.373) | <.0001 | 1.227 (1.119- 1.347) | <.0001 |
| Medium risk | 1.132 (1.028 – 1.246) | 0.0113 | 1.115 (1.010-1.231) | 0.0305 |
| Low risk | 1 | | 1 | |
| **Per capita real GDP** | | | | |
| Low GDP | 1 | | 1 | |
| Medium GDP | 0.841 (0.784 – 0.903) | <.0001 | 0.848 (0.789 – 0.912) | <.0001 |
| High GDP | 0.781 (0.706 – 0.864) | <.0001 | 0.764 (0.696 -0.838) | <.0001 |
| **Depression** | | | 1.360 (1.320 – 1.400) | <.0001 |
| **Diabetes** | | | 2.011(1.946 - 2.077) | <.0001 |
| **BMI** | | | | |
| Normal weight | | | 1 | |
| Under Weight | | | 1.130 (1.013 -1.260 ) | 0.0286 |
| Over Weight | | | 1.128 (1.087-1.170) | <.0001 |
| Obese | | | 1.245 (1.199- 1.293) | <.0001 |

# APPENDIX C: INSTITUTIONAL REVIEW BOARD APPROVAL

Drexel UNIVERSITY
*Human Research Protection*

NOT HUMAN RESEARCH DETERMINATION

February 24, 2014

Longjian Liu, MD, PhD, MSC, FAHA
School of Public Health
Epidemiology and Biostatistics
MS 1033
Email: hrp22@drexel.edu

Dear Dr. Liu,

On February 24, 2014, the IRB reviewed the following protocol:

| | |
|---|---|
| Type of Review: | *Initial / Letter of Determination* |
| Title: | Analysis of the Regional Differences in Cardiovascular Disease Risk Across the Lifespan, in the Population of Thirty Big American Cities |
| Investigator: | *Liu, Longjian* |
| IRB ID: | *1311002626* |
| Funding: | *Internal* |
| Grant Title: | *None* |
| Grant ID: | *None* |
| IND, IDE or HDE: | *None* |
| Documents Reviewed: | *Letter of Determination Application; Email from PI* |

The IRB determined that the proposed activity is not research involving human subjects as defined by DHHS and FDA regulations.

IRB review and approval by this organization is not required. This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these activities are research involving human in which the organization is engaged, please submit a new request to the IRB for a determination.

Sincerely,

John Medendorp, MS, BSN, RN, CIP
IRB Manager

## APPENDIX D: SAS CODES

```
********* Editing the data for all 7 years 2006 - 2012 **************;


libname cbmp12 "C:\Users\ABINAYA\Desktop\CBMP\BRFSS\2012\LLCP2012XPT" ;

data cbmp12.fresh12(keep = _STATE CTYCODE1 CVDINFR4 CVDCRHD4 CVDSTRK3 ADDEPEV2
DIABETE3 _BMI5CAT AGE HISPANC2 ORACE2 _RACE_G MARITAL _EDUCAG EMPLOY
_INCOMG SEX _SMOKER3 _RFDRHV4 _RFDRMN4 _RFDRWM4 MENTHLTH FRUIT1 FVGREEN
FVORANG VEGETAB1 _TOTINDA HLTHPLN1 _strwt _ststr _psu _finalwt) ;
set cbmp12.rawfile ;
run ;

data cbmp12.cities12;
set cbmp12.fresh12;
* city codes ;
if _STATE = 36 and CTYCODE1 in (005, 047, 061, 081, 085) then city = 0 ;*NY-BASELINE;
else if _STATE = 13 and CTYCODE1 in (121, 089) then city = 1 ;
else if _STATE = 24 and CTYCODE1 = 510 then city = 2 ;
else if _STATE = 25 and CTYCODE1 = 025 then city = 3 ;
else if _STATE = 17 and CTYCODE1 in (031, 043) then city = 4 ;
else if _STATE = 39 and CTYCODE1 = 061 then city = 5 ;
else if _STATE = 39 and CTYCODE1 = 035 then city = 6 ;
else if _STATE = 48 and CTYCODE1 in (085, 113, 121, 257, 397) then city = 7 ;
else if _STATE = 08 and CTYCODE1 = 031 then city = 8 ;
else if _STATE = 26 and CTYCODE1 = 163 then city = 9 ;
else if _STATE = 48 and CTYCODE1 in (157, 201, 339) then city = 10 ;
else if _STATE = 29 and CTYCODE1 in (037, 047, 095, 165) then city = 11 ;
else if _STATE = 32 and CTYCODE1 = 3 then city = 12 ;
else if _STATE = 06 and CTYCODE1 = 037 then city = 13 ;
else if _STATE = 12 and CTYCODE1 = 086 then city = 14 ;
else if _STATE = 27 and CTYCODE1 = 053 then city = 15 ;
else if _STATE = 12 and CTYCODE1 = 095 then city = 16 ;
else if _STATE = 42 and CTYCODE1 = 101 then city = 17 ;
else if _STATE = 04 and CTYCODE1 = 013 then city = 18 ;
else if _STATE = 42 and CTYCODE1 = 003 then city = 19 ;
else if _STATE = 41 and CTYCODE1 in (005, 051, 067) then city = 20 ;
else if _STATE = 06 and CTYCODE1 = 065 then city = 21 ;
else if _STATE = 06 and CTYCODE1 = 067 then city = 22 ;
else if _STATE = 48 and CTYCODE1 in (029, 091, 325) then city = 23 ;
else if _STATE = 06 and CTYCODE1 = 073 then city = 24 ;
else if _STATE = 06 and CTYCODE1 = 075 then city = 25 ;
```

```
else if _STATE = 53 and CTYCODE1 = 033 then city = 26 ;
else if _STATE = 29 and CTYCODE1 = 510 then city = 27 ;
else if _STATE = 12 and CTYCODE1 = 057 then city = 28 ;
else if _STATE = 11 and CTYCODE1 = 001 then city = 29 ;
else city = . ;

* heart attack, HRT = 1 = YES, 0 = N0;
if CVDINFR4 = 1 then HRT = 1;
else if CVDINFR4 = 2 then HRT = 0;
else if CVDINFR4 = 7 or CVDINFR4 = 9 then HRT = . ;
else HRT = CVDINFR4 ;

* Angina or CHD, CHD = 1 = YES, 0 = N0;
if CVDCRHD4 = 1 then CHD = 1;
else if CVDCRHD4 = 2 then CHD = 0;
else if CVDCRHD4 = 7 or CVDCRHD4 = 9 then CHD = . ;
else CHD = CVDCRHD4 ;

* Stroke, STRK = 1 = YES, 0 = N0;
if CVDSTRK3 = 1 then STRK = 1;
else if CVDSTRK3 = 2 then STRK = 0;
else if CVDSTRK3 = 7 or CVDSTRK3 = 9 then STRK = . ;
else STRK = CVDSTRK3 ;

* Depression, DEP = 1 = YES, 0 = N0;
if ADDEPEV2 = 1 then DEP= 1;
else if ADDEPEV2 = 2 then DEP = 0;
else if ADDEPEV2 = 7 or ADDEPEV2 = 9 then DEP = . ;
else DEP = ADDEPEV2 ;

* Diabetes, DM = 1 = YES, 0 = N0 ;
if DIABETE3 in (1, 2) then DM = 1;
else if DIABETE3 in (3, 4) then DM = 0;
else if DIABETE3 = 7 or DIABETE3 = 9 then DM = . ;
else DM = DIABETE3 ;

* age categories;
if AGE >= 18 and AGE < 25 then agecat = 0 ;
else if AGE >= 25 and AGE < 35 then agecat = 1 ;
else if AGE >= 35 and AGE < 45 then agecat = 2 ;
else if AGE >= 45 and AGE < 55 then agecat = 3 ;
else if AGE >= 55 and AGE < 65 then agecat = 4 ;
else if AGE >= 65 then agecat = 5 ;
else if AGE = 7 or AGE = 9 then agecat = . ;
```

```
else agecat = AGE ;

* sex, GENDER  = 1 = MALE, 0 = FEMALE ;
if SEX = 1 then GENDER = 1 ;
else if SEX = 2 then GENDER = 0 ;
else GENDER = SEX ;

*alcohol ;
if _RFDRHV4 = 1 then alco = 0 ;
else if _RFDRHV4 = 2 then alco = 1 ;
else if _RFDRHV4 in (9, .) then alco = . ;
else alco = _RFDRHV4 ;

*race;
if ORACE2 = 1 then race = 0 ;
else if ORACE2 = 2 then race = 1 ;
else if ORACE2 in (3,4) then race = 2 ;
else if ORACE2 in (5, 6) then race = 3 ;
else if ORACE2 in (7, 8, 9, .) then race = . ;

*BMI ;
if _BMI5CAT = 1 then  BMI = 1;
else if _BMI5CAT = 2 then BMI = 0 ;
else if _BMI5CAT = 3 then BMI = 2 ;
else if _BMI5CAT = 4 then BMI = 3 ;
else BMI = _BMI5CAT ;

*Education ;
if _EDUCAG = 1 then EDU = 1 ;
else if _EDUCAG = 2 then EDU = 2 ;
else if _EDUCAG = 3 then EDU = 3 ;
else if _EDUCAG = 4 then EDU = 0 ;
else if _EDUCAG = 9 then EDU = . ;
else EDU = _EDUCAG ;

* Employment ;
if EMPLOY =  1 then EMPLO = 0 ;
else if EMPLOY = 2 then EMPLO = 1 ;
else if EMPLOY = 3 then EMPLO = 2;
else if EMPLOY = 4 then EMPLO = 3 ;
else if EMPLOY = 5 then EMPLO = 4 ;
else if EMPLOY = 6 then EMPLO = 5 ;
else if EMPLOY = 7 then EMPLO = 6 ;
else if EMPLOY = 8 then EMPLO = 7 ;
```

```
else if EMPLOY = 9 then EMPLO = . ;
else EMPLO = EMPLOY ;

* Income ;
if _INCOMG = 1 then INCOME = 0 ;
else if _INCOMG = 2 then INCOME = 1 ;
else if _INCOMG = 3 then INCOME = 2 ;
else if _INCOMG = 4 then INCOME = 3 ;
else if _INCOMG = 5 then INCOME = 4 ;
else if _INCOMG = 9 then INCOME = . ;

*Marital Status ;
if MARITAL = 1 then MARRY = 0 ;
else if MARITAL = 2 then MARRY  = 1 ;
else if MARITAL = 3 then MARRY  = 2 ;
else if MARITAL = 4 then MARRY  = 3 ;
else if MARITAL = 5 then MARRY  = 4 ;
else if MARITAL = 6 then MARRY  = 5 ;
else if MARITAL = 9 then MARRY  = . ;
else MARRY  =  MARITAL ;

* smoking;
if _SMOKER3 = 1 then SMOKE = 1 ;
else if _SMOKER3 = 2 then SMOKE = 2 ;
else if _SMOKER3 = 3 then SMOKE = 3 ;
else if _SMOKER3 = 4 then SMOKE = 0 ;
else if _SMOKER3 = 9 then SMOKE = . ;

* Depression days ;
if MENTHLTH gt 0 and MENTHLTH le 14 then DEPN = 0 ;
else if MENTHLTH ge 15 and MENTHLTH le 30 then DEPN = 1 ;
else if MENTHLTH = 88 then if DEPN = 0 ;
else if MENTHLTH in (77, 99) then DEPN = . ;
else DEPN = MENTHLTH ;

*exercise;
if _TOTINDA = 1 then exer = 1;
else if _TOTINDA = 2 then exer = 0 ;
else if _TOTINDA = 9 then exer = . ;
else exer = _TOTINDA ;

*health insurance;
if HLTHPLN1 = 1 then INSUR= 1;
else if HLTHPLN1 = 2 then INSUR = 0;
```

```
else if HLTHPLN1 = 7 or HLTHPLN1 = 9 then INSUR = . ;
else INSUR = HLTHPLN1 ;

run ;

*************** Merging all 7 years datasets ********************;

libname combi "C:\Users\ABINAYA\Desktop\CBMP\BRFSS" ;

proc format ;
                value city  0 = 'New York ' 1 = 'Atlanta'
                                        2 = 'Baltimore' 3 = 'Boston'
                                        4 = 'Chicago'   5 = 'Cincinnati'
                                        6 = 'Cleveland' 7 = 'Dallas'
                                        8 = 'Denver'    9 = 'Detroit'
                                        10 = 'Houston' 11 = 'Kansas City'
                                        12 = 'Las Vegas' 13 = 'Los Angeles'
                                        14 = 'Miami'    15 = 'Minneapolis'
                                        16 = 'Orlando'  17 = 'Philadelphia'
                                        18 = 'Phoenix'  19 = 'Pittsburgh'
                                        20 = 'Portland' 21  = 'Riverside(CA)'
                                        22 = 'Sacramento' 23 = 'San Antonio'
                                        24 = 'San Diego' 25 = 'San Francisco'
                                        26 = 'Seattle'   27 = 'St. Louis'
                                        28 = 'Tampa'   29 = 'Washington DC' ;

                value HRT  1 = 'YES'  0 = 'N0'  . = 'MISSING';
                value CHD  1 = 'YES'  0 = 'N0'  . = 'MISSING';
                value STRK 1 = 'YES'  0 = 'N0'  . = 'MISSING';
                value DEPN  1 = 'YES'  0 = 'N0'  . = 'MISSING';
                value DM   1 = 'YES'  0 = 'N0'  . = 'MISSING';

                value BMI               0 = 'Normal Weight'
                                                1 = 'Underweight'
                                          2 = 'Overweight'
                                                3 = 'Obese'
                                                . = 'Missing' ;

                value agecat    0 = '18 - 24'
                                                1 = '25-34'
                                                2 = '35-44'
                                                3 = '45-54'
                                                4 = '55-64'
                                                5 = '65 and above'
```

```
                                        . = 'Missing' ;

        value GENDER  1 = 'MALE' 0 = 'FEMALE' . = 'MISSING';
        value HISLAT  1 = 'YES'  0 = 'N0'  . = 'MISSING';

        value race      0 = 'White'
                                        1 = 'Black'
                                        2 = 'Asian/Pacific Islander'
                                        3 ='American Indian, Alaska Native'
                                        4 = 'Hispanic/Latino'
                                        . = 'Missing' ;

        value  EDU     1  =   "Did not graduate High School"
        2  =   "Graduated High School"
                                3  =   "Attended College or Technical School"
                                0  =   "Graduated from College or Technical School"
. =   "Missing" ;

        value EMPLO    0  =   "Employed for wages"
                       1  =   "Self-employed"
                       2  =   "Out of work for more than 1
                                year"
                                3  =   "Out of work for less that 1 year"
                       4  =   "A homemaker"
                       5  =   "A student"
                       6  =   "Retired"
                       7  =   "Unable to work"
                       .  =   "Missing" ;

        value INCOME   0  =   "Less than $15,000"
                       1  =   "$15,000 to less than $25,000"
                       2  =   "$25,000 to less than $35,000"
                       3  =   "$35,000 to less than $50,000"
                       4  =   "$50,000 or more"
                       .  =   "Missing" ;

        value MARRY           0   =   "Married"
                       1   =   "Divorced"
                       2   =   "Widowed"
                       3   =   "Separated"
                       4   =   "Never married"
                  5   =   "A member of an unmarried couple"
                       .   =   "Missing" ;
```

```
                    value alco 1 = 'YES'  0 = 'N0'  . = 'MISSING';

                    value SMOKE          0 = "Never smoked"
                                         1 = "Current smoker - everyday"
                                         2 = "Current smoker - some days"
                                         3 = "Former smoker"
                                         . = "Missing" ;

                    value exer 1 = 'YES'  0 = 'N0'  . = 'MISSING';
                    value INSUR 1 = 'YES'  0 = 'N0'  . = 'MISSING';

                    value year      0 = 'Yr 06' 1= 'Yr 07' 2= 'Yr 08' 3='Yr 09' 4='Yr 10' 5='Yr 11'
                          6='Yr 12' ;
                    value CAD       1 = 'YES'  0 = 'N0'  . = 'MISSING';
                    value CVD       1 = 'YES'  0 = 'N0'  . = 'MISSING';
                    value zone  1 = 'High prevalence' 2 = 'Mid prevalence' 3='Low prevalence'
;
                    value phase 1 = 'Before recession' 2='During recession' 3='After
recession';


run ;


data combi.data06 ;
set cbmp06.cities06 ;
run ;

data combi.data07 ;
set cbmp07.cities07 ;
run ;

data combi.data08 ;
set cbmp08.cities08 ;
run ;

data combi.data09 ;
set cbmp09.cities09 ;
run ;

data combi.data10 ;
set cbmp10.cities10 ;
run ;
```

```
data combi.data11 ;
set cbmp11.cities11 ;
run ;

data combi.data12 ;
set cbmp12.cities12 ;
run ;

proc sort data = combi.merge ;
by city ;
run ;

data combi.merge (keep = year city AGE HRT CHD STRK DEP DEPN DM BMI agecat GENDER
HISLAT
race EDU EMPLO INCOME MARRY alco SMOKE exer INSUR CAD CVD phase zone _strwt _ststr );
length _TOTINDA 8 _SMOKER3 8 _RACE_G 8 _BMI4 8 _INCOMG 8 _RFDRHV3 8 ;

set combi.data06 (in=yr6)
        combi.data07 (in=yr7)
        combi.data08 (in=yr8)
        combi.data09 (in=yr9)
        combi.data10 (in=yr10)
        combi.data11 (in=yr11)
        combi.data12 (in=yr12) ;
if yr6 then year = 0 ;
else if yr7 then year = 1 ;
else if yr8 then year = 2 ;
else if yr9 then year = 3 ;
else if yr10 then year = 4 ;
else if yr11 then year = 5 ;
else if yr12 then year = 6 ;

*CAD ;
if HRT = 1 or CHD = 1 then CAD = 1 ;
else if HRT = 0 or CHD = 0 then CAD = 0 ;
else CAD = . ;

* Phase ;
if year = 0 then phase = 1 ;
else if year in (1, 2,3)then phase = 2;
else if year in (4, 5,6)then phase = 3 ;

* CVD ;
if CAD = 1 or STRK = 1 then CVD = 1 ;
```

```
else if CAD = 0 or STRK = 0 then CVD = 0 ;
else CVD = . ;

*emplo ;
if emplo = 3 then emplo = 2 ;

*zone ;
if city in (9, 27, 23, 17, 12, 16, 6, 11, 28, 21)then zone = 1;
else if city in (7,5,2,19,10,18,22,14,4,24)then zone = 2;
else if city in (29,0,3,13,20,8,1,15,26,25)then zone = 3 ;

run ;

proc contents data = combi.merge ;
run ;

proc univariate data = combi.merge ;
var age ;
where CVD = 1 ;
run ;

title "Prevalence of CVD in major cities" ;
proc gchart data = combi.merge ;
where CVD = 1;
hbar city/discrete descending ;
label CVD = 'CVD Prevalence'
                city = 'CITIES' ;
                format city city. ;
run ;

* CVD - stratified by age for all city ;
proc freq data = combi.merge ;
tables agecat *city  /CHISQ missing NOPERCENT nocum norow nocol ;
title " CVD age stratify " ;
run;

* CVD - statified by year   ;
proc freq data = combi.merge ;
tables  CVD * year * agecat   /CHISQ missing NOPERCENT nocum norow nocol ;
title " cvd year stratify " ;
run;

Proc freq data = combi.merge ;
*where CVD = 1 ;
```

```
table race * agecat /CHISQ missing NOPERCENT nocum norow nocol ;
title "gender vs race" ;
run ;

Proc freq data = combi.merge ;
table  (agecat GENDER race DM BMI DEPN SMOKE alco EDU EMPLO INCOME MARRY exer insur
zone) * CVD /CHISQ missing NOPERCENT nocum norow nocol ;
title "gender vs race" ;
run ;

* unemployment and cvd risk ;
proc freq data = combi.merge ;
*where emplo = 2  ;
tables  EMPLO * CVD * year /chisq nocum norow nocol ;
run;

proc freq data = combi.merge ;
*where emplo = 2  ;
table CVD * year/chisq nocum norow nocol;
run ;

data combi.table ;
set combi.merge ;
*DUMMY VARIABLES;
if city = 9 then Detroit = 1 ;else Detroit = 0 ;
if city = 17 then Philly = 1 ;else Philly = 0 ;
if city = 25 then Sanfan = 1 ;else Sanfan = 0 ;

if zone = 1 then highcity = 1 ; else highcity = 0 ;
if zone = 2 then midcity = 1 ;else midcity = 0 ;
*if zone = 3 then lowcity = 1 ;*else lowcity = 0 ;

if race = 1 then race1 = 1; else race1 = 0;
if race = 2 then race2 = 1; else race2 = 0;
if race = 3 then race3 = 1; else race3 = 0;
if race = 4 then race4 = 1; else race4 = 0;

if agecat = 1 then agecat1 = 1; else agecat1 = 0;
if agecat = 2 then agecat2 = 1; else agecat2 = 0;
if agecat = 3 then agecat3 = 1; else agecat3 = 0;
if agecat = 4 then agecat4 = 1; else agecat4 = 0;
if agecat = 5 then agecat5 = 1; else agecat5 = 0;

if edu = 1 then edu1 = 1; else edu1 = 0;
```

```
if edu = 2 then edu2 = 1; else edu2 = 0;
if edu = 3 then edu3 = 1; else edu3 = 0;

if emplo in (2,3) then emplo = 2 ;

if emplo = 1 then emplo1 = 1; else emplo1 = 0 ;
if emplo = 2 then emplo2 = 1; else emplo2 = 0 ;
if emplo = 4 then emplo3 = 1; else emplo3 = 0 ;
if emplo = 5 then emplo4 = 1; else emplo4 = 0 ;
if emplo = 6 then emplo5 = 1; else emplo5 = 0 ;
if emplo = 7 then emplo6 = 1; else emplo6 = 0 ;

if income = 1  then income1 = 1; else income1 = 0 ;
if income = 2  then income2 = 1; else income2 = 0 ;
if income = 3  then income3 = 1; else income3 = 0 ;
if income = 4  then income4 = 1; else income4 = 0 ;

if marry = 1 then marry1 = 1; else marry1 = 0 ;
if marry = 2 then marry2 = 1; else marry2 = 0 ;
if marry = 3 then marry3 = 1; else marry3 = 0 ;
if marry = 4 then marry4 = 1; else marry4 = 0 ;
if marry = 5 then marry5 = 1; else marry5 = 0 ;

if phase = 1 then phase1 = 1 ; else phase1 = 0;
*if phase = 2 then phase2 = 1 ; *else phase2 = 0;
if phase = 3 then phase3 = 1 ; else phase3 = 0;

If bmi = 1 then bmi1 = 1; else bmi1 = 0;
If bmi = 2 then bmi2 = 1; else bmi2 = 0;
If bmi = 3 then bmi3 = 1; else bmi3 = 0;

run ;



proc freq data = combi.table ;
*where cvd = 1 ;
tables philly * phase * agecat/chisq missing nocum norow nocol nopercent ;
run;

proc freq data = combi.table ;
*where cvd = 1 ;
tables sanfan * phase * agecat/chisq missing nocum norow nocol nopercent ;
run;
```

```
proc freq data = combi.table ;
where cvd = 1 ;
tables detroit * phase * agecat/chisq missing nocum norow nocol nopercent ;
run;

/************************ Economic data merging************************/

data combi.excel ;
infile "C:\Users\ABINAYA\Desktop\Per capita income of cities.csv" dlm=',' dsd missover
firstobs=2 ;
length city 8 eco_06   8 eco_07       8 eco_08 8 eco_09     8 eco_10 8 eco_11 8   eco_12 8 ;
input city eco_06       eco_07eco_08eco_09eco_10eco_11eco_12 ;
run ;

proc sort data = combi.excel ;
by city;
run;

proc sort data = combi.table ;
by city;
run;

data combi.econo ;
merge combi.table combi.excel ;
by city ;

if eco_06 ge 20000 and eco_06 lt 40000 then mic = 1 ;
else if eco_06 gt 40000 and eco_06 lt 60000 then mic = 2 ;
else if eco_06 gt 60000 and eco_06 lt 80000 then mic = 3 ;

if mic = 1 then mic1 = 1 ; else mic1 = 0 ;
if mic = 2 then mic2 = 1 ; else mic2 = 0 ;
if mic = 3 then mic3 = 1 ; else mic3 = 0 ;

if year = 0 then year0 = 1 ; else year0 = 0 ;
if year = 1 then year1 = 1 ; else year1 = 0 ;
if year = 2 then year2 = 1 ; else year2 = 0 ;
if year = 3 then year3 = 1 ; else year3 = 0 ;
*if year = 4 then year4 = 1 ; *else year4 = 0 ;
if year = 5 then year5 = 1 ; else year5 = 0 ;
if year = 6 then year6 = 1 ; else year6 = 0 ;

if smoke in (1,2) then smoke1 = 1; else smoke1 = 0;
```

```
if smoke = 3 then smoke3 = 1; else smoke3 = 0;

run ;

/******************** Modelling ***************************/


title " survey log - reduced model";
proc surveylogistic data = combi.econo ;
weight _strwt ;
strata _ststr ;
model cvd (ref = "0") = agecat1 agecat2 agecat3 agecat4 agecat5 gender race1 race2 race3
race4 edu1 edu2 edu3 alco smoke1 smoke3 highcity midcity marry1 marry2 marry3 marry4
marry5 income1 income2 income3 income4 emplo1 emplo2 emplo3 emplo4 emplo5 emplo6
exer insur mic2 mic3 phase1 phase3 ;
run;

title " survey log - full model";
proc surveylogistic data = combi.econo ;
weight _strwt ;
strata _ststr ;
model cvd (ref = "0") = agecat1 agecat2 agecat3 agecat4 agecat5 gender race1 race2 race3
race4 edu1 edu2 edu3 alco smoke1
smoke3 highcity midcity marry1 marry2 marry3 marry4 marry5 income1 income2 income3
income4 emplo1 emplo2 emplo3 emplo4 emplo5 emplo6
exer insur phase1 phase3 mic2 mic3 depn dm bmi1 bmi2 bmi3 ;
run;
```