UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Income Inequality and Mortality: The Costa RIcan Case

Permalink
https://escholarship.org/uc/item/44r5s4x9

Author
Modrek, Sepideh

Publication Date
2009

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Income Inequality and Mortality: The Costa Rican Case

by

Sepideh Modrek

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Health Services and Policy Analysis

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor William H. Dow, Chair

Professor Ralph Catalano

Professor Jennifer Ahern

Fall 2009

Table of Contents

List of Tables ..... **iii**

List of Figures ..... **iv**

Glossary ..... **v**

Abstract **1**

Executive Summary **2**

Chapter 1: Background and Significance **6**
1.1 Literature Review **7**
1.1.1 Income and Health Relationships 7
1.1.2 Mechanisms and Spurious Correlation for the Income Inequality Hypothesis 8
1.1.2 Empirical Review: Review of Reviews 11
1.2 Institutional Setting **18**
1.3 Study Aims **21**
1.4 Contributions: **21**

Chapter 2: Descriptive Analysis of the Measurement of Income Inequality in Costa Rica 27
2.1 Introduction **27**
2.2 Economic Landscape in Modern Costa Rica **27**
2.3 Data **29**
2.4 Description and Analyses of Constructed Measures **31**
2.4.1 National Trends 31
2.4.2 Regional Trends 33
2.5 Other Analyses **36**
2.6 Discussion **38**

Chapter 3: Ecological Relationship Between Income Inequality and Mortality **55**
3.1 Introduction **55**
3.2 Data **55**
3.3 Methods **56**
3.4 Results **59**
3.4.1 All-Cause Mortality 59
3.4.2 Infectious Disease 60
3.4.3 Cancer Mortality 61
3.4.4 Lifestyle or Behavior Related 61
3.4.4 External Causes 62
3.5 Sensitivity Analysis **62**
3.5.1 Pooled OLS Models and Longer-Term Changes Models 62
3.5.2 Model Specification 63
3.5.3 Measurement Error 64
3.6 Discussion **65**
3.7 Limitations **66**

Chapter 4: Individual-level Association of Income Inequality and Mortality **83**
4.1 Data Collection **83**
4.2 Measures **85**
4.2.1 Individual-Level Covariates 85

4.2.2 Time Varying Variables 86
4.2.3 Outcomes 86
4.3 Analysis Methods 87
4.4 Results 89
4.4.1 Income Inequality Hypotheses 89
4.4.2 Relative Deprivation Hypothesis 91
4.5 Discussion 91
4.6 Limitations 93
Chapter 5: Conclusions 108
5.1 Summary of Results 108
5.2 Pathways and Mechanism 109
5.2.1 Psychosocial Pathway 109
5.2.2 Resource Distribution Pathway 110
5.2.3 Neomaterial Pathway 111
5.3 Interpretations 111
5.4 Contributions 112
5.5 Limitations 113
5.6 Further Research 114

List of Tables

Table 1.1: Stratification of Multi-Level Studies with Small Geographical Units by Health Outcome 23
Table 1.2: Selected Health and Economics Indicators Comparing Costa Rica to Other Countries 24
Table 1.3: Comparison of Ethnic, Linguistic and Religious Fractionalization 25
Table 1.4: Comparison of Selected Political Indicators 25
Table 2.1: Year Groupings Used to Generate Area-Level Measures from July Household Survey 39
Table 2.2: Correlation Coefficient for Canton Gini Coefficients in 1987/1988 to All Other Years 39
Table 2.3: Correlation Coefficient for Canton Gini Coefficients in 2005/2006 to All Other Years 40
Table 2.4: Association Between Canton-Level Growth Rate in Gini from 1989-2005 and Labor Force Characteristics 41
Table 2.5: Mean Number of Observations that Canton -Level Gini Coefficient were Calculated over by Year 42
Table 2.6: Mean Estimated Gini coefficient Based on Simulations with Varying Number of Observations Used to Calculate Gini Coefficient 42
Table 2.7: Association of Social Trust and Income Inequality at the Canton Level, 2002, 2004, 2006 43
Table 3.1: Weighted Mean Mortality Rate by Year and Cause 69
Table 3.2: Summary Statistics of Independent Variables 71
Table 3.3: Correlation Coefficients of Main Independent Variables 72
Table 3.4: Association of Inequality and All-cause Mortality Rates 74
Table 3.5: Association of Inequality and Infectious Disease Mortality Rates 75
Table 3.6: Association of Inequality and Cancer Related Mortality Rates 76
Table 3.7: Association of Inequality and Lifestyle Related Mortality Rate 77
Table 3.8: Association of Inequality and External Related Mortality Rates 78
Table 3.9: Comparison of Pooled OLS, 15-year Change and 5-year Change Models 79
Table 3.10: Summary of Sensitivity Analysis 80
Table 4.1: Census Mortality Linkage Sample Creation 95
Table 4.2: Means of Individual-level Variables 96
Table 4.3: Mapping Time Varying Variables 97
Table 4.4: Categorization by ICD-9 & ICD-10 Cause of Death 97
Table 4.5: Number of Deaths by Cause in Sample 98
Table 4.6: Summary Statistics of Weighted Canton-Level Measures 99
Table 4.7: Gompertz Models of Risk of Mortality and Income Inequality 100
Table 4.8: Gompertz Models of the Risk of Heart Disease Mortality and Income Inequality 101
Table 4.9: Comparison of Gompertz and Cox Proportional Hazard Models 102
Table 4.10: Comparison of Different Parametric Models 103
Table 4.11: Gompertz Models of the Risk of Mortality and Individual-Level Relative Deprivation 104
Table 5.1: Summary of Hypotheses 116

List of Figures

Figure 1.1: Conceptual Diagram of Mechanisms Relating Inequality Health Outcomes 26
Figure 2.1: GDP Growth Rate in Costa Rica from 1987-2006 44
Figure 2.2: Labor Force Trends in Costa Rica from 1980-2006 45
Figure 2.3: Income Inequality Trends in Costa Rica from 1980-2006 46
Figure 2.4: Household Poverty Trends in Costa Rica from 1980-2006 47
Figure 2.5: Percentage change in Household Equivalent Income from 1987 to 2005 by Income Percentile 48
Figure 2.6: Scatter Plot of Income Inequality by Canton and Year – Adjacent Years 49
Figure 2.7: Scatter Plot of Income Inequality by Canton and Year 50
Figure 2.8: Map of Estimated Gini Coefficient in 1989 and 2005 51
Figure 2.9: Map of estimated 5-year growth rates of Income Inequality from 1989-2005 52
Figure 2.10: Map of estimated 5-year growth rates of Unemployment, In-Migration, Household Equivalent Monthly Income (HHEMI), and Extreme Poverty from 1989-2005 53
Figure 3.1: Scatter Plot of Canton Income Inequality and Mortality Rate in 2005 81
Figure 3.2: Scatter Plot of Change in Mortality and Income Inequality between 1989-2005 81
Figure 4.1: Wealth Index Box Plots by Province 105
Figure 4.2: Kaplan-Meier Survival Curves for All-Cause Mortality 106
Figure 4.3: Kaplan-Meier Survival Curves for Heart Disease Mortality 107

Glossary

CCP-The Central American Population Center is a research center at the University of Costa Rica.

CCSS-Costa Rican Social Security Fund, which provides universal health care insurance to employed Costa Ricans

CRHS- Costa Rican Household Surveys for Multiple Purposes

HHEMI- Household equivalent monthly income

ICD- International Classification of Disease

ICE- Instituto Costarricense de Electricidad

IMF- International Monetary Fund

INEC -Instituto Nacional de Estadisticos y Censos or National Institute of Statistics and Censuses in Costa Rica

SES- Socio-economic status

Cedula- Similar to the United States Social Security Number

Civil Registry- The Costa Rican Civil Registry database contains the unique c*edula* identification number and individual names. The *cedula* identification number is matched based on the name and demographic information from the 1984 census.

Segmento- Similar to a United States census tract

Abstract

Income Inequality and Mortality: The Costa Rican Case
by
Sepideh Modrek
Doctor of Philosophy in Health Services and Policy Analysis
University of California, Berkeley
Professor William H. Dow, Chair

In recent decades there has been much debate on the extent to which income inequality affects health outcomes. The majority of studies in this literature have examined the relationship between income distribution and health outcomes in the contexts of high-income nations. This dissertation evaluates both the ecological and individual-level relationship between income inequality and mortality in the context of a middle-income country, Costa Rica. The Costa Rican case is particularly interesting because its strong social safety net, political organization, and history limit the effects of income inequality on health through indirect channels. Using 20 years of household data, this study begins by documenting changes in income inequality at the canton level. These data reveal that on average inequality has been growing at an annual rate of 0.5%. Using these inequality data, Vital Statistics Mortality records, and longitudinal methods, this study next examines whether cantons with greater growth in inequality also experienced increasing mortality. The analyses indicate that there is no relationship between changes in income inequality and changes in all-cause mortality; however, a relationship does exist for specific causes of death. Area-level inequality was found to be associated with an *increase* in suicide mortality rates, and a *decrease* in cardiovascular, respiratory cancer, and breast cancer mortality rates. At the individual-level survival analysis was applied to a new census-mortality linked dataset that allows for a 20-year prospective follow-up. These individual level analyses also test whether income inequality disproportionately affects the poor. In the case of all-cause mortality, the results suggest that there is no association between income inequality and time to death for either the entire sample or for the poor. Contrary to expectations the magnitudes of the estimated effects indicate that residents living in areas with increasing income inequality had a survival advantage for cardiovascular mortality, which further confirms the results found at the ecological level. Overall, this study finds that while inequality has been increasing in Costa Rica, it had limited effect on mortality, indicating that in an environment where safety net services are available, the negative effects of income inequality may be limited to suicide mortality.

Keywords: Costa Rica, Socio-economic Inequality, Mortality, Health Inequities, Relative Deprivation, Social Trust

Executive Summary

Hypotheses:

The hypothesis that income inequality adversely affects health outcomes has been extensively debated in the economics, epidemiology, and public health literature in recent decades. The majority of studies have examined the relationship between income distribution and health outcomes in the contexts of the United States and other high-income nations. In order to gain insight into the mechanisms underlying this hypothesis, the present study evaluated the relationship between income inequality and mortality in the context of a middle-income country, Costa Rica, whose social structure and history confer the unique benefit of being less susceptible to common sources of confounding. Four main propositions, motivated by the larger literature on the income inequality hypothesis, were tested.

1. Was there a cross-sectional association between income inequality and health as observed in other countries?
2. Given that previous studies argue that income inequality has been a determinant of population health, did areas with increasing income inequality exhibit increases in all-cause mortality rates?
3. Assuming that income inequality affects health through stress pathways, did areas with increasing income inequality have increases in mortality rates from stress related diseases such as cardiovascular disease?
4. Assuming that these relationships exist, did individuals living in areas with increasing inequality, and particularly the poor, have lower survival?

Past Research:

An enormous amount of literature documents the relationship between measures of socio-economic status (SES) (correlates of income/wealth) and mortality, as well as other health outcomes (Adler, Boyce, Chesney, Cohen, & Folkman, 1994). This relationship reflects the absolute income hypothesis, which posits that income and SES correlates dictate the amount of resources available to buy better nutrition and healthcare treatments, and may affect the time available for health producing activities.[1] A more contentious hypothesis argues that an individual's relative rank or the dispersion of ranks within a community also directly affects health. This latter hypothesis, known as the relative income hypothesis, is the focus of this study.

While there are several formulations of this relative income hypothesis, supporters of the hypothesis generally contend that holding individual income fixed, greater income inequality, as a measure of class differences within a community, leads to poorer health (Richard G. Wilkinson, 1996). This hypothesis posits that people who feel more economically disadvantaged than their peers may experience chronic stress, which then leads to illness (Baum, Garofalo, & Yali, 1999). Others have argued that support for this

[1] Economists also argue that individuals with better health earn higher incomes, and thus, the relationship between income and health is likely to work in both directions.

hypothesis has been based on flawed analyses. Indeed numerous methodological shortcomings, such as the lack of a natural reference group over which to define communities, the cross-country ecological design of most studies, insufficient controls for correlates of inequality (e.g. race and poverty), and the subjective nature of the dimensions of health considered have plagued studies that support this hypothesis (A. Deaton, 2003). Moreover, most of the studies in this literature are based in wealthy nations, especially the United States, where local taxation laws or other historical processes may confound the results.

Research Design and Significance:

Given the above complexity and the limited number of studies in middle-income countries, this study examined the relationship between area-level income distributions and mortality in Costa Rica, both in the cross-section and by examining *changes* overtime. Costa Rica has several distinct features that make it an ideal setting to analyze this hypothesis. It has had universal health insurance since the 1970s, a highly centralized political system, a vast public health infrastructure, and a highly homogenous population. These characteristics help limit potential confounders, such as access to healthcare, local distribution of public health goods, and the historical racial inequality that may exist in other settings. Additionally, in the cross-national literature, others have pointed to Costa Rica as an example that supports the income inequality hypothesis. As such, Costa Rica was an important case to examine.

The current project had three distinct components. First, a longitudinal dataset of inequality measures from over 20 years of household data were constructed. These data were used to document and compare regional inequality over this time. Area-level measures of inequality and other labor force variables were also created, validated, and mapped. Second, these measures were used to examine the ecological association of mortality and income inequality in levels and changes for 18 distinct causes of death. Third, the ecological analyses were supplemented with individual-level analyses of the income inequality hypothesis employing survival analysis for all-cause and heart disease mortality. The study also considered poverty as a moderator of the income inequality and health relationship.

This study employed several methodological strategies in response to previous criticisms in the literature:

- Longitudinal methods, which account for time-invariant factors that are correlated with income inequality and mortality are used,
- As opposed to self-reported health, mortality was the objective health outcome,
- Income inequality in relatively small administrative areas was employed as the reference group over which income inequality was measured,
- An area without taxation power constituted the reference group, and
- Ecological and individual-level results were compared.

Main Results:

- Areas with higher income inequality tended to be more affluent and have higher household equivalent mean incomes in the cross-section. Likewise, increases in income inequality were due to increasing shares of income for those at the top of the income distribution at both the national and regional levels. Similarly, areas with increasing income inequality were also areas that were undergoing economic development more rapidly.

- In the cross-section, areas with higher income inequality did not have higher all-cause mortality after controlling for mean income. The standard errors on these cross-sectional estimates were large, and the other area-level variables were highly related to all-cause mortality suggesting that there may be other unobserved area-level determinants of mortality that may not have been accounted for in the cross-sectional models.

- At the regional level, areas with increasing income inequality did not generally experience increasing mortality rates. Of the 18 causes of death examined, only increases in suicide mortality rates were associated with increases in income inequality; however, the magnitude of the effect size was small and only marginally significant. Moreover, there was some evidence that areas with increasing income inequality experienced decreasing mortality rates for cardiovascular disease, respiratory cancers, and breast cancer. These results were robust to several sensitivity analyses for model specification. Furthermore, simulations suggested that attenuation bias due to measurement error in the inequality measures was likely to be small and could not explain the lack of results generally.

- At the individual level, results from survival analyses were generally consistent with those from the ecological analyses. In areas with increasing inequality, there was some evidence for higher all-cause mortality survival, but no correlation between income inequality and survival once controls for area-level fixed-effects were included. The conflicting results between models without and with area fixed-effects may have been due to other area-level unobserved characteristics such as economic development correlated with both higher survival and higher income inequality. Furthermore, residents of areas with increasing income inequality had higher survival from cardiovascular mortality. Contrary to the initial predictions, there were no additional adverse effects of income inequality for poor residents for either mortality outcome.

Implications:

This study generally failed to find support for the initial hypotheses, and in fact, found evidence contrary to the premise of the income inequality hypothesis. In a setting where strong social safety nets are available and income inequality is increasing due to increases in income at the top of the income distribution, there may be no deleterious effect caused by rising income inequality on population health. These results question the notion that inequality *itself* is the root cause of health disparities in other countries where this

relationship has been found. While future studies should continue to assess how hierarchy or feelings of deprivation affect our health, equal attention should be given to social systems that may compensate for the negative impacts of increased social hierarchy.

Chapter 1: Background and Significance

The extent to which income inequality affects population health is a controversial issue despite decades of research. Since the 1970's scholars have hypothesized that community-level socioeconomic inequalities may negatively influence health over and above the effects determined by an individual's absolute socioeconomic level (Preston, 1975; Rodgers, 1979). Early studies of this income inequality hypothesis focused on cross-country analyses which have since been widely criticized, yet interest in the topic continues (R. G. Wilkinson, 1992; Richard G. Wilkinson, 1996). More recent studies address the shortcomings of these earlier efforts by using multi-level statistical methods, but their results have failed to reach a consensus. The remaining controversy in the field is evident since even reviews within this literature come to different conclusions while considering similar evidence (A. Deaton, 2003; J. Lynch, Smith, Harper, & Hillemeier, 2004a, b; S. Subramanian, 2004; S. V. Subramanian & Kawachi, 2004). Given this ambiguity, most recent multi-level studies have focused on country-level analyses to help enrich this literature. One country that has been a notable international outlier and deserves more attention is Costa Rica.

Although for decades observers have noted that Costa Rica has remarkably high life expectancy (higher than the United States) despite its limited resources[2], until recently there has only been limited in-depth study of the determinants of Costa Rica's exceptional longevity. While many hypotheses have been suggested to help explain this outlier, one recurring untested notion is that Costa Rica is an unusually equitable society in certain dimensions, and that this has contributed to its good health. In fact, supporters of the income inequality hypothesis often point to Costa Rica as evidence for their views. Unlike the United States or Brazil where inequality is highly correlated with ethnic composition and historical discrimination, the Costa Rican setting allows us to assess the relative income hypothesis with less concern for these types of confounders. Moreover, Costa Rica's universal health insurance, highly centralized political system, and public health infrastructure make it an unlikely place to find an association if it is indeed due to political or structural area-level pathways. Thus, if we do find an association here, where we might not expect to, the results are even more compelling.

This study aims to advance the income inequality hypotheses by examining it in the context of Costa Rica. First, I create a new dataset with detailed geographic identifiers in order to describe inequality patterns in Costa Rica at a sub-national level. Second, I examine how changes in the 20-year panel of area-level income inequality relate to changes in age-sex and cause specific mortality rates at the ecological level. This allows the use of longitudinal modeling techniques that can control for the confounding of fixed unmeasured characteristics at the area-level. Third, I will look at a prospective cohort to examine the association of income inequality and mortality at the individual-level. Again, here I am able to control for potential confounders of other area-level characteristics. By synthesizing the results of the related studies, this thesis will examine the Costa Rican setting in detail to help shed light on the broader debate of the income inequality hypothesis.

---

[2] Costa Rica's per capita income is roughly one-fifth that of the United States'.

1.1 Literature Review

1.1.1 Income and Health Relationships

The robust cross-country association between income inequality and population health has long been documented, yet the interpretation of these associations have been quite contentious. There are several competing/complementary hypotheses about the income and health relationship that could produce an association between income inequality and population health at the ecological-level. In order to distinguish between these hypotheses and interpret these associations, Wagstaff and Doorslaer categorize several hypotheses regarding the income and health relationship and present detailed models to help discriminate between them (Wagstaff & van Doorslaer, 2000). I begin with a review of these hypotheses below, focusing on the relative income hypothesis, as it will be the main topic of this dissertation.

1.1.1.1 Absolute Income hypotheses

The absolute income hypothesis is the simplest of these hypotheses. This hypothesis states that one's absolute income dictates the amount of resources available to buy better healthcare treatments and nutritious foods and may affect the time available for health producing activities and thus leads to better health outcomes. It further argues that the relationship between health and income is concave, and thus each additional dollar spent on health raises an individual's health less than the previous dollar. The shape of the income and health relationship implies that if you take X dollars away from a poorer person and give it to a richer person, the combined health of the two people will be worse because those X dollars can buy more health for the person of lower income. Mechanically, this exchange also increases inequality at the same time that it lowers overall health. Consequently, we would see a negative association with income inequality and health in data, but this reflects the underlying relationship of absolute income and health, not income inequality and health *per se*.

Another version of the absolute income hypothesis is the so-called Deprivation or Poverty Hypothesis. Here as before the underlying relationship between health and income is concave, but after a certain income level the relationship becomes flat. Assuming that after a certain level of income, say the poverty line, there is no relationship between income and health, then it does not matter for the health of a population to have ever-increasing income for a select few. Rather the goal should be to ensure that everyone gets past this minimum income threshold. Many economists argue that these simple versions of the absolute income hypothesis fail to discuss how the causal direction can run in the opposite direction in which case people of poor health are less able to work and thus have lower income/wealth.

1.1.1.2 Relative Income Hypotheses

A large, but more contentious literature argues that one's relative rank or the dispersion of incomes within a community directly affects health. There are also several versions of these hypotheses. The relative income hypothesis assumes that one's health depends on the deviation of one's income from a population mean, rather than one's absolute income level. If people identify more closely with those in their community

over those in the broader population, then the hypothesis would compare the individual's income to the community mean. The comparison population is important to highlight because depending on the reference population the deviation could change dramatically. For example, $60,000 may be higher than the mean income in California, but lower than the mean for the San Francisco Bay Area. Another version of this hypothesis is the relative position hypothesis. Here it is not necessarily the deviation from some mean income, but rather one's place in the income distribution. Here again, the population used in the reference distribution becomes very important.

A notable variant of these relative income hypotheses is the so-called income inequality hypothesis. This hypothesis claims that holding individual income fixed, income inequality, as a measure of the dispersion of ranks, class differences, or social hierarchy within a group, leads to poor health (Richard G. Wilkinson, 1996). This hypothesis assumes that people who feel more economically disadvantaged than their peers may experience chronic stress or depression, which then leads to illness. Furthermore, two subversions have been proposed. The strong version states that all members of a group are equivalently hurt by the unequal distribution, whereas the weak form suggests that only those who are least well off or most vulnerable are affected by income inequality, or are hurt disproportionately so.[3] The weak form of the income inequality hypothesis parallels yet another variant, the relative deprivation hypothesis, which posits that individuals' health is adversely affected when they perceive themselves as more economically disadvantaged than their peers.

While all these variations have been proposed for the income and health relationship, empirically at the population level they lead to the same aggregate relationship, and as such it has been difficult to differentiate between the absolute income hypothesis and the various relative income hypotheses. The income inequality hypothesis has been particularly contested, as early studies in this literature did not account for the relationship between absolute income and health. Even those that did only examined ecological relationships, which Wagstaff argues are not particularly useful in distinguishing between the absolute income and relative income hypotheses. In order to assess the relative income hypothesis, which is the focus of the remainder of this study, he argues that individual-level study designs are necessary.

1.1.2 Mechanisms and Spurious Correlation for the Income Inequality Hypothesis

In addition to the absolute income and health relationship that the association between income inequality and health may capture inadvertently, the underlying mechanisms of how income inequality could possibly affect health are also highly debated. Three general groups of pathways have been proposed to explain the associations between health and income distribution: psychosocial, resource allocation, and neomaterial.

The psychosocial pathway is hypothesized to be the only direct channel through which inequality affects health. There are two main components of the psychosocial mechanism. The first component posits the possibility that the mere evidence of social hierarchy causes stress (Richard G. Wilkinson, 1996), which activates neuroendocrine

[3] The definition of "vulnerable" is not well stated in the literature and is often in terms of wealth or age.

production (increased cortisol, epinephrine, or norepinephrine) and acute-stress response that ultimately increases susceptibility to disease (Bruce S. McEwen, 2004). This view of the relationship between inequality and health is based on the direct extensions of the primate studies to humans. A recent review of primate studies links status, stress, and health outcomes (Sapolsky, 2004). For animals, the adaptiveness of the acute stress response is thought to be pathogenic when repeatedly and unnecessarily activated. Studies show that when incorrectly activated, the acute-stress response leads to higher circulating levels of glucocorticoids, which can lead to insulin (in)sensitivity, elevated blood pressures, atherosclerotic plaque, repressed reproduction, and inhibited long-term immune response.[4] As a corollary to this view, it is hypothesized that some people can buffer against inequality-induced stress (those with high incomes, good genetics, good health behavior, and strong social support), while others cannot, leaving them vulnerable to disease. Again, numerous animal studies demonstrate that social support, such as grooming practices or having a source of distraction or displacement, can blunt this stress response (Sapolsky, 2004). The strongest evidence in the animal literature comes from Shively and Clark who experimentally regroup female monkeys and show that lower status monkeys that gain status have health gains, while high status monkeys who fall in status exhibit worse health. These perturbation studies in the animal literature have been the most compelling as to the potential effects of status on health (Shively & Clarkson, 1994).

While compelling, extrapolating results from these studies to humans is exceptionally difficult. First and foremost, humans can simultaneously belong to multiple social groups and thus the negative impact from being of low status in one group may be offset by being of high status in another group. Moreover, we can choose our level of involvement in a group based on our perceived status within the group. Despite these limitations, there is some support for these hypotheses in humans. A recent meta analysis of 208 cortisol related studies in healthy adults found that tasks characterized by the threat of social-evaluation and uncontrollability lead to the highest cortisol responses in lab settings, which would be consistent with the psychosocial mechanism outlined above (Dickerson & Kemeny, 2004).

The second component of the psychosocial mechanism posits that income inequality weakens social ties within a community, which buffer against stress and disease. The weakening of social ties is often conceptualized to affect health by limiting social capital, defined as the level of interpersonal trust between citizens, the norms of reciprocity, and the vibrancy of civic associations within a community (Putnam, 2000). Social capital is thought to be important because of its potential to help individuals within a community buffer against stress and consequent diseases (Kawachi & Kennedy, 1999a). However, the measurement of the construct of social capital is also contentious. Social capital is usually conceptualized according to the level of measurement, individual or community-level, and then further categorized into two groups, cognitive and structural. The cognitive component includes perceptions of trust, reciprocity and

[4] Sapolsky (2004) provides theoretical, experimental and anthropological evidence that animal health is affected by status through the proposed stress pathway. However, he also discusses how these relationships differ across species, by the hierarchy stability, and by mating patterns. His discussion highlights the difficulty of finding consistent results in the relationship between status and health even for animals.

sharing, whereas the structural component includes the density of associational links within society (Kawachi & Kennedy, 1999a; R. M. Scheffler, Brown, Syme, Kawachi, Tolstykh, & Iribarren, 2008b).

Moving beyond these psychosocial pathways, others argue that income inequality affects health through resource allocation. Some argue that income inequality is in fact a proxy for greater poverty in an area and that this poverty or lack of resources is what impacts health (A. Deaton, 2003). Others posit that the greater the income inequality in a community, the more infrastructural and political interests might diverge. For example, a family at the 95th income percentile may pay substantially more taxes than a family at the 50th percentile, but does not receive a correspondingly higher benefit from public services, such as public health services. If these two families live in the same community, this income gap creates pressure from economic elites to reduce local public services, including those related to health. In contrast, in more equal areas, economic interests may be more aligned and investments in human capital infrastructure represent a more common interest (Krugman, 1996). This competing interests model for public good investments has been used to look at how inequality affects economic growth (Alberro Alesina & Perotti, 1994) and institutions (Chong & Gradstein, 2006). This mechanism could similarly explain the lack of public resources in relatively unequal areas, causing poor health outcomes. In this case, it is not the inequality *per se* that is causing the poor health, but rather the lack of public resources.

Finally, the neomaterial view of the inequality and health relationship contends that the underlying contextual process that generates inequality is also what causes poor health. Under this interpretation, the effects of income inequality on health are not inevitable and are contingent on the distribution of other social resources in that society. Here income inequality proxies a combination and accumulation of other negative exposures, lack of individual resources, and systematic underinvestment in health and social infrastructure. Proponents of this model see income distribution as a result of historical, cultural, and political-economic processes (such as slavery and discrimination in the United States). Some of these processes are thought to influence both the private resources available to individuals and to shape the nature of the public infrastructure for the provision of education, health services, transportation, environmental controls, or quality housing (J. Lynch et al., 2004a). The processes that generate the income inequality and the resulting maldistribution of resources or stress are thought to be the cause of poor health outcomes. This interpretation views the observed association between income inequality and health as a spurious correlation. For example, if increases in inequality are due to unemployment, and unemployment causes psychosocial stress or depression (Dooley, Catalano, & Wilson, 1994) as opposed to the inequality itself, then one might find a spurious correlation between income inequality and depression. A related argument maintains that changes in income distribution are often associated with societal structural changes (such as trade liberalization) and that poor health outcomes (especially in perceived health) have more to do with unhappiness about these changes or an individuals' inability to adapt to these changes in their environment than to inequality itself.

Figure 1.1 gives a conceptual outline of how each of the mechanisms relates inequality to health. The level of variation and how it interacts with the other pathways

are diagramed. The direction of the action of the neomaterial pathway is outlined in black dotted lines, the maldistribution of resources pathway is outlined in the black dashed lines, and the psychosocial pathway (the only direct pathway) is outlined in grey dashed lines. In this setup, these three mechanisms are not mutually exclusive, and hence researchers need to think about how to distinguish between different mechanisms based on theory and empirical analysis. For example, the neomaterial pathway is upstream with regard to the other two pathways making it difficult to distinguish it from the other two pathways. Theoretically, if we could find a way or a place that would break the link between income inequality and community-level resources then the problem would be slightly more tractable. Later I argue that Costa Rica may be such a place.

1.1.2 Empirical Review: Review of Reviews

The number of empirical studies in this literature and related literature is enormous, and among this extensive literature the methodological quality, claims, and settings vary substantially. Instead of critiquing each study, I summarize areas where there is most consensus and areas of contention in the literature in terms of methods, results, and setting, while noting results from certain key studies.

1.1.2.1 Methods

1.1.2.1.1 Controls

As mentioned above, the observed association between income inequality and health can reflect multiple hypotheses regarding the income and health relationship that are indistinguishable at the population level (Wagstaff & van Doorslaer, 2000). Furthermore, findings from these aggregate level studies have been criticized both for inadequately accounting for the possibility that a disproportionate number of poor people live in high-inequality areas (A. Deaton, 2003) as well as for not properly addressing the concavity of the relationship between individual income and health which significantly changes the interpretation of the results.[5] Thus, testing the effects of income inequality requires the simultaneous testing of the effects of individual income/wealth, or area-level income in addition to the effect of area income inequality. Accordingly, studies must account for these issues.

The other controls necessary for testing this hypothesis have also been debated. Area-level studies may suffer from an ecological fallacy, especially if they do not account for unobserved community-level factors that related to both income distribution and health. At the area-level, variables such as ethnic segregation (Angus Deaton & Lubotsky, 2003; Galea, Ahern, Vlahov, Coffin, Fuller, Leon et al., 2003; Shi, 2005; S. V. Subramanian & Kawachi, 2003), social capital (Drukker, Feron, & van Os, 2004; Franzini & Spears, 2003; Pattussi, Marcenes, Croucher, & Sheiham, 2001), access to health (Carlos Larrea & Kawachi, 2005b; Li & Zhu, 2006; S. V. Subramanian & Kawachi, 2003), and unemployment (Shi, 2005; S. V. Subramanian & Kawachi, 2003) have been used as controls.

[5]Any study that fails to control for income may mechanically find a spurious correlation between income inequality and health. However, the relationship may be due to the concave shape of the income and health relationship rather than an actual association between income inequality and health.

In the United States, race has been a particularly controversial confounder. Deaton and Lubotsky (Angus Deaton & Lubotsky, 2003) find that once the proportion of African Americans are added to their model, the ecological relationship between income inequality and health disappears or even reverses for both states and Metropolitan Statistical Areas. They argue that since African Americans tend to have a higher mortality rate and lower incomes than whites, in areas with a higher concentration of African Americans inequality and mortality are both higher. Therefore, the observed relationship between income inequality and health is a consequence of this mechanical relationship. Yet, Wilkinson argues that variables, such as race, are proxies for social hierarchy or mediators of the income inequality relationship as opposed to confounders, and subsequently should not be included as controls (R. G. Wilkinson & Pickett, 2006).

While there is no consensus, area-level income and/or poverty controls are considered a minimum and other controls should be considered depending on the setting. For example, given the legacy of slavery in the United States, percent African American is thought to be an important control for ecological studies in the United States. At the individual-level, multi-level studies may also suffer from omitted variable bias if crucial individual-level variables are missing. Most researchers agree that at a minimum individual-level income as well as education should be included.[6]

1.1.2.1.2 Inequality Measures

Even the measures of inequality vary across studies. There are four characteristics, known collectively as the Dalton Criteria, that inequality measures generally should exhibit: (1) scale invariance, (2) population size independence, (3) insensitivity to anonymous transfers, and (4) sensitivity to transfer from the poor to the rich. While many inequality measures exhibit these characteristics, they differ systematically in the different aspects of the underlying distributions they capture.[7] Some measures used have been the Robin-Hood index[8] (Kennedy, Kawachi, & Prothrow-Stith, 1996), Income deciles (Skalli, Theodossiou, & Etienne), and the Coefficient of Variation (CV) (Mellor & Milyo, 2002), however, the most widely used inequality measure is the Gini coefficient. From a theoretical perspective, while the Gini coefficient does meet the Dalton criteria, it is not necessarily the best. First, it is not decomposable, at least not in the sense that it will satisfy consistency requirements for arbitrary partitions of the population (it is decomposable for non-overlapping partitions, but that is rarely the case).[9] Second, its statistical properties are far less tractable than other inequality measures, as it is not differentiable. Finally, the Gini coefficient does not emerge naturally from welfare-economics although its graphical interpretation makes it fit well into this literature.

Despite these drawbacks, the Gini coefficient's ease of interpretation and its comparability make it a desirable measure. First, it has an intuitive graphical

[6] Others also argue that race must be included if the study takes place in a highly heterogeneous society where race is a marker of social stratification.
[7] For instance, the coefficient of variation tends to be more sensitive to disparities at extremes of a distribution, while the Gini coefficient is more sensitive in the middle of the distribution.
[8] This measure does not meet the Dalton Criteria.
[9] The total Gini coefficient of a society is not equal to the sum of the Gini coefficients for its sub-groups.

interpretation based on the Lorenz curve. Second, it captures the idea of the "average distance" between incomes in the population according to a particular definition of distance, namely ranking. For instance, a Gini of 0.2 indicates that the average difference in two random observations is 40 percent of the mean. Other readily accessible measures of inequality such as the CV are sensitive to changes at the extremes and are thus more sensitive to outliers. Finally, Gini measures are widely available for most countries and are the standard measure of inequality in most datasets. For these reasons it is now standard to use the Gini despite its theoretical drawbacks.

1.1.2.1.3 Reference Groups

To model the effects of inequality on health, we have to first address the proper way to construct a relevant reference group for any individual (A. Deaton, 2003; J. Lynch et al., 2004a). As discussed above, human groupings are based on selection, unlike animals whose groupings can be created based on experimental manipulation. Moreover, humans may belong to multiple groups simultaneously (especially in mobile societies), and within each group the same individual may have a different status. This could further affect the weight that the individual places on membership in the group or the duration of exposure that individual chooses within the group. With the exception of a few studies where natural reference groups exist (Anderson & Marmot, 2007; D. A. Redelmeier & Singh, 2001a; Donald A. Redelmeier & Singh, 2001b), relevant reference groups are generally assumed instead of inferred by researchers.

Without sufficient research to establish the “correct” reference group and given the difficulty of collecting data on an individual’s social network, many researchers have focused on geographical area as a context in which people make comparisons and in which status is evident. The social psychology literature suggests that members of one's reference group are typically selected based on either demographic similarity (age, sex, education, and race) or geographic proximity, thus this practice is not entirely atheoretical (Singer, 1981). Several studies have also created reference groups based on demographics. Deaton and Paxson (A. Deaton & Paxson, 1998), define reference groups in terms of birth-cohort. In doing so, they find no relationship between income inequality and mortality at the cohort level. Eibner and Evans (Eibner & Evans, 2005), address the related relative deprivation hypothesis by using state of residence, race, education, and age to create one’s reference group. Here they find that even after controlling for individual income, men who are more deprived relative to their demographic-state reference group have higher mortality and poorer health behaviors. Overall, one’s proximate environment is one context in which people experience daily or repeated exposure, and this makes it especially relevant in the context of the proposed psychosocial mechanism (Ahern & Galea, 2006).

If one accepts spatial area as a basis for forming a reference group, another complication is the size of the area over which income inequality is measured. Based on theories that posit biological stress as the mechanism causing poor health (B. S. McEwen & Seeman, 1999), one would surmise that the spatial area should be of a scale such that one has repeated and regular exposure to the reference group. This would point most

naturally to a small area where one has repeated interactions with the environment.[10] Furthermore, Lynch (J. Lynch et al., 2004a) invokes Messner and Tardiff's (1986) discussion of "frames of reference" to argue that local comparisons may be most salient, and therefore local deprivation might be most important. However, Wilkinson and Pickett argue that the degree of residential segregation is not the right unit of analysis because individuals can still perceive inequality within a larger social context and thus feel deprived even if little income variation is observed within their own neighborhoods (R. G. Wilkinson & Pickett, 2006).[11] Yet, if one rarely interacts with that larger context, then the assumed chronic exposure may be missing. Consequently, a natural geographic unit of analysis is larger than one's immediate neighborhood, and yet one that remains salient. Thus, many use county or Metropolitan Statistical Areas as units over which to calculate income inequality.

1.1.2.1.4 Life Course and Time Dynamics

The timing of the ways in which inequality may affect health has also been considered important. In view of the psychosocial mechanism mentioned above, it seems that inequality would require some time to exert negative effects. It is thought that the chronic nature of stress exposure to an unequal area eventually wears down the body's stress adaption mechanisms. Inequality may also have latent effects on adult health that may not materialize into observable effects until they are older. Thus looking at a single cross-section of contemporaneous inequality does not seem to fit the theoretical premises of the psychosocial pathway. Accordingly, some researchers have explored potential lag effects of income inequality on health (Blakely et al., 2000; Subramanian and Kawachi, 2004; (Milyo & Mellor, 2003) by considering 15 to 25 year lagged income inequality to current health outcomes, (S. V. Subramanian & Kawachi, 2006). To date, there is no consensus about whether or not lagged or contemporaneous inequality measures are better, or how migration affects these lagged measures over larger geographical area. In general, examining the income inequality hypothesis at (any) one point in time may be misleading. Ideally, we would assess the relationship of income inequality to health over time and attempt to link changes in income inequality to changes in health, or to changes in risk factors. Within a causal framework, this method would be far more convincing and have the advantage of limiting area-level confounding.

1.1.2.1.5 Outcomes

Debate also remains as to which health outcomes income inequality affects. The theory relating status and health in primates outlined by Sapolsky (Sapolsky, 2004) posits that status affects many types of health outcomes. If income inequality does indeed affect health through the stress mechanisms, then a variety of health outcomes could be affected adversely, which would be reflected in all-cause mortality. At the same time,

[10] Depending on the age and mobility of the individual, neighborhoods might not be the group that one interacts with the most (Lynch et al., 2004). For example for the young and the elderly, a neighborhood might be more appropriate than for a middle-aged man.

[11] If there is little income inequality within neighborhoods but much income inequality in a nation as a whole then this implies that there are high levels of disparities between neighborhoods in income inequality because individuals have segregated into small homogenous pockets. A high inequality measure for a neighborhood thus may correspond to more overall integration.

humans can respond with an array of health behaviors and buffering resources, so some health outcomes may be affected more than others.[12] Therefore choosing among health outcomes in humans can be more difficult. Many initial studies focused on life expectancy and paid little attention to the psychological dimension of health, even though researchers often assumed that income inequality exerts its deleterious impact through psychological mechanisms.

In a recent review, Wilkinson (R. G. Wilkinson & Pickett, 2006) claims that the weakest support for the income inequality hypothesis comes from studies in which inequality is measured over a small area, such as a county or neighborhood. However, Wilkinson did not stratify these studies by outcome. Doing so, I show that certain outcomes, such as perceived health or health behavior, are more sensitive to an inequality effect. Table 1.1 displays studies that were either supportive or unsupportive of the income inequality hypothesis in Wilkinson's review (2006) stratified by design, geographical area, population and outcome. Table 1.1 presents only studies that use a multilevel model and a small geographic area (county or smaller). From this synthesis, it is clear that studies that find an income inequality effect have either perceived health or health behaviors as their health outcome, which is consistent with the psychosocial mechanism proposed. Studies that do not find an inequality effect tend to have area level all-cause mortality as the health outcome measure. Area level all-cause mortality includes many heterogeneous causes of death and so may be less sensitive than other measures of health, such as cause-specific mortality or subjective health. While far from perfect, a cause specific analysis may be useful in understanding mechanisms, since the contribution of income inequality to different causes of death could vary across conditions, especially those thought to be particularly sensitive to economic condition such as cardiovascular disease (Adler et al., 1994).

1.1.2.1.5 Populations

Finally, many studies that have mixed support for the income inequality hypothesis find effects for specific subpopulations (J. Lynch et al., 2004a). While there is already a fair amount of literature that posits how income inequality can affect the poor or vulnerable more than other groups, there is insufficient theory on why the effects differ by group and how exactly to define vulnerability (by age, income, status, mental health, social ties, etc). Where studies find mixed results, there is also the possibility of making ad hoc or arbitrary judgments on who constitutes a vulnerable population. Moreover, most studies use populations of convenience rather than explicitly modeling different effects for specific populations, making comparisons across studies difficult.

The differential effects of income inequality across populations have critical implications for navigating the complex web of pathways. The main argument of the weak form of the income inequality hypothesis suggests that certain vulnerable subpopulations are more acutely affected by income inequality. One such analysis would

[12] For instance, disadvantaged individuals (compared to their peers) may feel depressed and disgruntled leading to poor perceived health (depressed affect, or low self-reported health), and these conditions could further affect health directly (via heart disease, high blood pressure) or indirectly (via poor behaviors, such as bad eating habits, tobacco & alcohol abuse). The timing of this process makes certain outcomes more sensitive than others.

be to treat the elderly as vulnerable. This assumes that as people age and develop health problems and become less mobile, their residential areas can become their primary social contexts and sources of social support. While social interactions and informal social support may occur in the home and within neighborhoods, sources of institutional support, such as government transfers and health services span larger areas. Thus, there could be a different impact of inequality at different levels of aggregation for the elderly (Muramatsu, 2003). Insofar as inequality affects social support and perceived social standing, one could expect to see effects at a proximate level.[13] In contrast, working adults tend to be more mobile, less vulnerable and less dependent than the elderly on local resources. For this group, it may be that inequality has little impact on their health as they may have more options in responding to and buffering against it. Another natural comparison group is the poor and the non-poor as the poor may not have the necessary resources the buffer against any negative of impact income inequality. Beyond the elderly and the poor, it is difficult and often arbitrary how vulnerable subpopulations have been defined.

While there is still debate on the above issues, there is some consensus on methodological issues such as the need for longitudinal multi-level models with appropriate controls and Gini as the measure of income inequality. Beyond the methodological debates, I argue that from a theoretical point of view an administrative area in which one has had repeated social contacts is the appropriate unit of analysis. While ideally we would like to assess the income inequality and health relationship across several dimensions of health for a nationally representative population over time, such measures are not yet available. Consequently, a compromise starting point could be to examine mortality by cause. This too should provide insight on pathways. Finally, in accounting for different populations a natural comparison would be working aged adults to the elderly and the comparing the poor to the non-poor.

1.1.2.2 Results

Beyond the methodological debate, some consistent results have been found. First, there is a consistent relationship between income inequality and violent causes of death, namely homicide and suicide (J. Lynch et al., 2004a; R. G. Wilkinson & Pickett, 2006), especially in the United States. Income inequality is thought to produce distrust and hostility, which in turn undermine social cohesion and lead to violence. In Wilkinson's review, of the 15 country-level studies where homicide was the outcome, 14 were supportive of the notion that income inequality was positively associated with homicide rates (R. G. Wilkinson & Pickett, 2006). The most persuasive recent study, analyzed changes in homicide rates across the 10 Canadian provinces over 10 years while controlling for changes in mean income. This study found a robust positive relationship between increases in homicide mortality rates and increases in income inequality (Daly, Wilson, & Vasdev, 2001). These results are particularly interesting because Canada has much lower inequality and homicide rates than the United States, but even at these lower levels of inequality, the relationship was present.

[13] In an ongoing study using 80 Chicago neighborhoods, a positive and robust association between a measure of collective efficacy (a measure that is similar to cognitive social capital) and health outcomes was found for the elderly (Cagney & Brown, 2008 unpublished manuscript presented at PAA 2008)

In addition, there is stronger evidence of a relationship between income inequality and self-reported health than all-cause mortality in many studies (Angus Deaton & Lubotsky, 2003; Mellor & Milyo, 2002; S. V. Subramanian & Kawachi, 2004). For example of the 12 multi-level studies included in Kawachi's review, 7 found support for the income inequality hypothesis when self-reported health was the outcome. In contrast, only 1 of the 7 multi-level studies supported the income inequality hypothesis when mortality was the outcome. This suggests that income inequality could affects perceptions of health, which may or may not affect the *realizations* of health in the long-run.

1.1.2.3 Setting

Since cross-country methods have been largely invalidated, multi-level country specific studies are the new standard in this literature. These studies reveal a pattern that deserves attention. Regardless of the methods utilized, countries in Western Europe, Canada, and New Zealand tend not to display relationships between income inequality and mortality or other health outcomes (Blakely, Atkinson, & O'Dea, 2003; Drukker et al., 2004; Hou & Myles, 2005). In contrast, some of the most robust and recent relationships have been found in developing countries such as, Brazil, Ecuador, Chile, and China (C. Larrea & Kawachi, 2005a; Li & Zhu, 2006; Pattussi et al., 2001; S. V. Subramanian, Delgado, Jadue, & Vega, 2003; Szwarcwald, 1999).

Two studies looking at the association between adult health and income inequality are particularly relevant to this study.[14] Using a nationally representative sample and looking across 285 communities in Chile, Subramanian et al. (S. V. Subramanian, Delgado, Jadue, & Vega, 2003) find that income inequality across communities is related to a higher probability of reporting poor health. This study carefully controls for individual income and ethnicity, but it does not include area-level fixed effects, arguing that random-effects are more appropriate. Likewise, Li & Zhu (Li & Zhu, 2006) also find that multiple measures of inequality, including lagged and contemporaneous measures of income inequality are associated with poorer self-reported health. In addition, Li & Zhu find that income inequality is related to worse health behaviors. Both of these studies examine mostly cross-sectional associations.[15] Accordingly, more studies in less developed countries or middle-income countries that account for area-level confounders are needed.

In addition, the primary outcome measure of both studies is self-reported health. Yet, as discussed above when small-area studies were stratified by outcome, there was a stronger relationship between income inequality and perceptions of health, than objective measures of health (see Table 1.1). Additionally, Fiscella & Franks find a relationship between income inequality and poor self-reported health but do not find effects on mortality in the Current Population Survey in the United States (Fiscella & Franks,

[14] The other two studies mentioned, Larrea & Kawachi (2005) & Swarcwald et al. (2002), evaluate childhood stunting and infant mortality respectively, and thus are not directly comparable to the proposed study.

[15] Li & Zhu do have some models with individual fixed effects where they find that lagged Gini coefficient has an inverted-U shape relationship with self-reported health, but the significance on health behaviors disappears. They argue that there is too little variation in the self-reported health and health behavior outcomes and as such they express more confidence in the models without the individual fixed effects.

2000). Thus, the association between inequality and objective measures of morbidity and mortality should also be examined in middle-income countries.

In light of the mixed findings across developed and developing countries, Kawachi further suggests that countries with low inequality lack the necessary variation in inequality to have an income inequality effect after controlling for individual or area-level income (S. V. Subramanian & Kawachi, 2006). He argues that it is best to look at countries that are at least as unequal as the United States to ensure that true variation exists. However, countries that tend to be more unequal are generally less developed with poorer healthcare access and infrastructure, and they also fail to offer other social safety nets. In order to test his hypothesis, we must find a setting where a social safety net exists, access to health care is not related to income inequality, and there is sufficient variation in income inequality. Costa Rica is one such place.

1.2 Institutional Setting

As an outlier in terms of health achievements and given its GDP, Costa Rica provides an interesting setting to explore income inequality hypothesis. Situated in the most unequal region of the world, Costa Rica has lower inequality than its neighbors (Table 1.2 Panel A). Meanwhile it also has a lower rate of infant mortality and a higher life expectancy than even the United States, but its GDP per capita is one fifth and its health spending is one tenth that of the Unites States (Table 1.2 Panel B). Consequently, Costa Rica has been touted as the quintessential example of the income inequality hypothesis at the cross-country level and hailed as the exception to Latin America (Daniels et al. 2000).

Given the complex pathways through which income inequality can affect health described above, Costa Rica's characteristics make it an ideal settings to examine the income inequality hypothesis. First, it is a setting where the link between income inequality and area-level resource allocation is thought to be weak. Second, it's highly centralized political system, near universal health insurance and very homogeneous population, enables us to look at the income inequality hypothesis with less concern for confounding based on local political capture, basic access, or ethnic heterogeneity compared to other well-studied cases such as the United States.

With a highly centralized political structure, Costa Rica enables the examination of the income inequality hypothesis with limited confounding due to local politics and skewed resource allocation. Sewastynowicz outlines the level of political control as follows:

> "There are three levels of government in Costa Rica: the central government composed of executive, legislative, and judicial branches; the Municipal Council which functions at the cantonal level; and the District Council which serves as a subsidiary arm of the Municipal Council. Of these, only the first constitutes an independent locus of authority, since it alone has meaningful powers of taxation. The Municipal Council can levy only miniscule taxes, and the District Council none at all. Both are subsidized by the country's Legislative Assembly [SIC]. All major officials are chosen by election every four years. Suffrage is universal with citizens 18 years of age and older legally required to vote." (Sewastynowicz, 1983)

This centralized political structure is even reflected in the education system in Costa Rica. A recent report by Latin American Consortium for the Evaluation of Education Systems described Costa Rica's education system as highly regulated in a highly centralized structure where organization seems to be more important than autonomy or innovation (Iaies, 2005). Again the limited political discretion over local public resources makes the Costa Rican context less susceptible to *systematic* differences across areas. This is in contrast to what happens in countries like the United States where education and other local public goods are funded by local property taxes and thus are *systematically* related to the income distribution and preferences in the local area.

Costa Rica has also had near universal healthcare coverage since the mid-1970s, reducing possible confounders from basic healthcare access.[16] The Costa Rican Social Security Fund (CCSS) provides universal health insurance to employed Costa Ricans. The CCSS covers 80% of the population through either the compulsory or the voluntary system, or as pensioners or their dependents. Of the remaining 20%, 10% are insured through state subsidies and the other 10% can request public services when necessary and pay for them directly.[17] Physical access to medical care is also exceptional as about 50% of Costa Rican's live less than 1 km away from a primary health care facility (Luis Rosero-Bixby, 2004). Moreover, this public health system is highly centralized, implying that the local political economy may play less of a role in mediating or causing inequality than in other contexts (Clark, 2002). [18] Taken together Costa Rica seems to be a place where we can be less worried about the local resource allocation mechanism mentioned above.

Costa Rica's particular colonial history and relatively racially homogeneous population also helps to eliminate potentially confounding factors related to historically institutionalized inequities, which plague studies of inequality in other settings such as those based in United States. While historical data on Costa Rican population is limited, it is thought that prior to Spanish conquest in the 1500s, there were at least 20 distinct indigenous cultures, with an overall population of about 100,000. After the conquest in 1600 there remained only 10,000 indigenous people (Palmer & Molina, 2004b). The natives who did survive the initial colonization did so by staying in remote and inaccessible areas (Quiros, 2004). The racial homogeneity of today slowly set in as the remaining natives mixed with the Spaniard colonists over the next two centuries. According to figures from 1777-78, the Central Valley was then comprised of 60 percent *mestizos* (mixed Spaniard and natives ancestors), 18% blacks and mulattoes, 12% natives, and 10% full Spaniard (born in both the new world and Spain). This intermixing continues today where Costa Rica's population is 94% mixed white. Table 1.3 shows the Ethnic, Linguistic and Religious fractionalization in Costa Rica. This measure is a concentration index where a score of 0 indicates perfect homogeneity and a score of 1

[16] This is an important advantage in studying Costa Rica because in other settings, such as the United States, individuals go on and off insurance often so that concurrent insurance status does not fully account for basic access to health care.
[17] "Profile of the Health Services System of Costa Rica." PAHO. May 27, 2002. Pg. 6. http://www.lachealthsys.org/index.php?option=com_docman&task=doc_download&gid=162
[18] According to the 2000 census, 96% of all births took place in a CCSS hospital (Rosero-Bixby 2004).

indicates perfect heterogeneity.[19] Costa Rica scores quite low on all three indices, demonstrating its homogeneity.

Beyond the racial mixing, another consequence of the indigenous population decline is thought to be the formation of more equal labor relations. With very limited labor in the Costa Rican Central Valley, the Spaniards could not establish large plantation style colonies, as was done elsewhere in Latin America. The limited labor in the Central Valley not only impaired the formation of large plantations, but also enabled the expansion of rights for different groups as small population was forced to work together to protect the colony. For instance in 1676 after the mulattoes, free blacks and lower class *mestizos* helped ward off pirate attacks in the capital of Cartiago, they were granted property rights and some political recognition (Palmer & Molina, 2004a).

Yet, the description above is an over simplification of Colonial Costa Rica as there was slave trade and slavery. Beyond the Central Valley, black slaves did work on cacao farms in the Atlantic region and the cattle ranches of Pacific Coast. The Hacienda style of development along the Pacific Coast was particularly different as this area was initially a part of the Nicaraguan colony, and the black slaves that worked and lived in on these farms were isolated from the center of the colony.[20] Yet, it was the culture and liberalism of the Central Valley that provided the foundation for Costa Rica's national institutions after gaining independence from Spain in 1821 instead of the culture of the Caribbean or the Pacific north (Palmer & Molina, 2004b). The modern Costa Rican state was founded with suffrage for women, full citizenship for blacks, established a presidential term limit, an independent Electoral Tribunal to oversee future elections, and no standing army.

International indices confirm the notion that Costa Rica is more democratic society than it's neighbors. Table 1.4 compares Costa Rica's scores on international indices of civil and political freedom, ranking in woman parliamentarians, and Human Development to its neighbors. Column 1 shows the overall Freedom House score, which is based on a checklist of 10 political rights and 15 civil liberties designated by academics and analysts. Higher scores indicate greater democratic freedom and Costa Rica scores higher than all its neighbors and most Latin American countries. Its score is on par of that of the United States. Column 2 shows that Costa Rica ranks 29th in terms of the number of women in parliament. This is far better than its neighbors and even higher than the United States. Finally column 3 shows that Costa Rica ranks 44th in the Human Development Index. This is a composite score based on life expectance, education, and GDP. Though it's GDP is two-thirds that of Chile's, it score is very similar indicating its high life expectancy and high educational attainment. Together these indices paint a picture of a democratic, gender equal and educated society.

---

[19] Fractionalization is calculated as $Frac_j = 1 - \sum_{i=1}^{N} s_{ij}^2$ where s is the share of group i in country j. If there is country of a 100 people and they all speak the same language, then the score would be Frac=1-$\sum_1(1)^2$=1-1=0. If each person speaks a different language, then Frac=1-$\sum_1^{100}(.01)^2$=1-100(.0001)=.99

[20] The pacific north of Costa Rica was under the influence of the then Nicaraguan colony and hence administered quite differently. It is still the most unequal region of Costa Rica today.

1.3 Study Aims

Given Costa Rica's distinct advantages, this study seeks to assess the income inequality hypothesis in this setting. The first aim of the study will be to construct, validate and map measures of income inequality at the sub-national level as no such income based measure exist. While mostly descriptive, this exercise is necessary to create the explanatory variables of interest. It will also help in understanding what forces are driving the increases in income inequality in Costa Rica at a sub-national level over the past 20 years. These constructed inequality measures will also be used to evaluate the ecological association between social capital and income inequality, which has been posited as one of the channels through which inequality affects heath (refer to Figure 1.1).

H1: Income inequality is negatively associated with measures of social capital.

Next, the association between increases in income inequality and increases in mortality rates will be examined. Taking into account the consensus around longitudinal multi-level design, this study will explicitly test the main argument of the income inequality hypothesis in the Costa Rican context with longitudinal data for all-cause mortality and 18 other causes of death at the ecological level.

H2: Higher income inequality is associated with higher all-cause mortality.

H3: Higher income inequality is associated with higher mortality for stress related causes of death (such as cardiovascular disease).

The literature suggests that income inequality has a stronger adverse impact on specific subpopulations. Some argue that the elderly may be particularly affected (Muramatsu, 2003), therefore separate the elderly, 60 and over, and adults in this analyses.[21] Moreover, according to the weak form of the income inequality hypothesis individuals with low SES are thought to encounter more daily stressors and have fewer resources to optimally buffer against them in more unequal areas. Thus, health consequences may be more pronounced for these more vulnerable populations. The second study will investigate the relationship between mortality and income inequality at the individual level and have the advantage of differentiating effects for specific sub-populations, especially the poor.

H4: The individual-level association between area-level income inequality and health may be stronger for more vulnerable populations such as the poor.

1.4 Contributions:

This study will first document regional variations in income inequality over time in Costa Rica. Using detailed geographic identifiers, a panel of ecologically linked inequality, poverty, income, migration and employment measures will be constructed, thus allowing the use of longitudinal modeling techniques in assessing the income inequality hypothesis. These data will then be exploited to assess the effects of income inequality on health outcomes through two sets of additional analyses. First, data from

[21] By separating those aged 60 and over, the results from this study can later be compared to a nationally representative sample of Costa Rican's for whom detailed data on health and health behaviors are available.

the Costa Rican Vital Registry will be used to assess the ecological relationship between area-level income inequality and age specific and cause specific mortality for adults (aged 30-59) and the elderly (aged 60 and over). Second, this study will use a unique dataset which links 16,000 adults from the 1984 Costa Rican census with the national death registry to examine the relationship between mortality and inequality, controlling for individual-level characteristics including wealth.

The main contribution of this study is its longitudinal design, long-term follow-up, objective health measures, and the unique institutional setting from which the data were gathered. The longitudinal design is key to understanding if there is indeed a causal relationship and helps reduce the number of area-level covariates required. The long-term follow- up also allows us to assess the impact of income inequality over decades instead of years, and this is more consistent with the time frame over which income inequality is thought to affect health. In addition, this study is the first one to examine the relationship between income inequality and adult mortality in a middle-income country. By looking at both ecological and individual-level relationships within the same country, this project also aims to give a more holistic view of the relationship between income inequality and mortality in Costa Rica, a country with a highly centralized political system, near universal healthcare coverage, and rising inequality. This combination of characteristics, coupled with its ethnically homogenous population and small geographical size makes Costa Rica an ideal setting to eliminate several sources of confounding mentioned in the literature.

This dissertation proceeds in 5 chapters. Chapter will 2 present a descriptive picture of inequality and labor trends in Costa Rica over the past 25 years. The section will also explore the correlates of changing inequality in Costa Rica. Chapter 3 will present findings based on the ecological relationship between income equality and health in Costa Rica. Chapter 4 will present survival analysis based on a census –mortality linkage. Chapter 5 concludes.

Table 1.1: Stratification of Multi-Level Studies with Small Geographical Units by Health Outcome

Authors, Year	Outcome Type	Outcome	Area	Population
Supportive				
M Soobader, F LeClere, 1999	Perception	Self-reported health	County	Adult White Males Americans
N Muramatsu, 2003	Perception	Elderly depression (CES-D) 8 item	County	Elderly 70+ Americans
S Subramanian, I Delgado, L Jadue, J Vega, 2003	Perception	Self-reported health	285 communities nested within 13 regions	Adult Chileans
M Pattussi, W Marcenes, R Croucher, A Sheiham, 2001	Behavioral	Dental Caries	19 Administrative Areas around Brazilia	6-12 year old Brazilian Children
S Galea, J Ahern, D Vlahov, P Coffin, C Fuller, 2003	Behavioral	Drug overdose	NYC is divided into 59 residential community	New York City residents
Unsupportive				
T Blakely, J Atkinson, DO'Dea, 2003	Objective	3 year mortality	Regional areas	Adults New Zealanders aged 25-64
K Fiscella, P Franks, 1997	Objective	Mortality	Nationa Health Interview Survey PSU	Adult Americans
L Franzini, W Spears, 2003	Objective	Excess mortaliy	County	Heart disease deaths in Texas
C Larrea, I Kawachi, 2006	Objective	Child malnutrition	Parish and census zone	5 years and younger, Ecuadorian children
M Drukker, F Feron, J van Os, 2004	Both	Quality of Life	Postal code and neighborhood	Adults in Maastricht, the Netherlands.
F Hou, J Myles, 2005	Perception	SRH-ordinal logit	Canadian neighborhoods in metropolitan areas	12 years and older

Source: Author's categorization based on Wilkinson's review (R. G. Wilkinson & Pickett, 2006).

Table 1.2: Selected Health and Economics Indicators Comparing Costa Rica to Other Countries

Panel A: Selected Economic Indicators

	Tax Revenue (% of GDP) 2001	Gini coefficient 2002	2005 GDP/cap (PPP adjusted international $)
Brazil		60.7	8,473
Chile		56.7	12,248
Costa Rica	13.02	45.9	8,712
Guatemala	9.69	55.8	4,919
Nicaragua	12.99	60.3	2,639
United States	12.74	40.8	41,813

Panel A: HDI=Human Development Index, GDP=Gross Domestic Product. Sources: http://hdr.undp.org/en/statistics/, World Bank http://econ.worldbank.org/

Panel B: Selected Health Sector Indicators

	Spending on Health % of GDP (Public)	Spending on Health % of GDP (Private)	Infant Mortality	Life Expectancy at Birth
Brazil	3.6	3.8	21.2	72.1
Chile	3.9	2.1	7.6	78.4
Costa Rica	4.8	3.2	9.8	78.7
Guatemala	2.2	4.8	39	70
Nicaragua	3.5	4.1	33	72.5
United States	8.1	4.6	6.9	78.1

Panel B Sources: http://hdr.undp.org/en/statistics/, World Bank http://econ.worldbank.org/

Table 1.3: Comparison of Ethnic, Linguistic and Religious Fractionalization

Ethnic, Linguistic and Religious Fractionalization			
	Ethnic Fractionalization	Linguistic Fractionalization	Religious Fractionalization
Brazil	0.5408	0.0468	0.6054
Chile	0.1861	0.1871	0.3841
Costa Rica	0.2368	0.0489	0.241
Guatemala	0.5122	0.4586	0.3753
Nicaragua	0.4844	0.0473	0.429
United States	0.4901	0.2514	0.8241

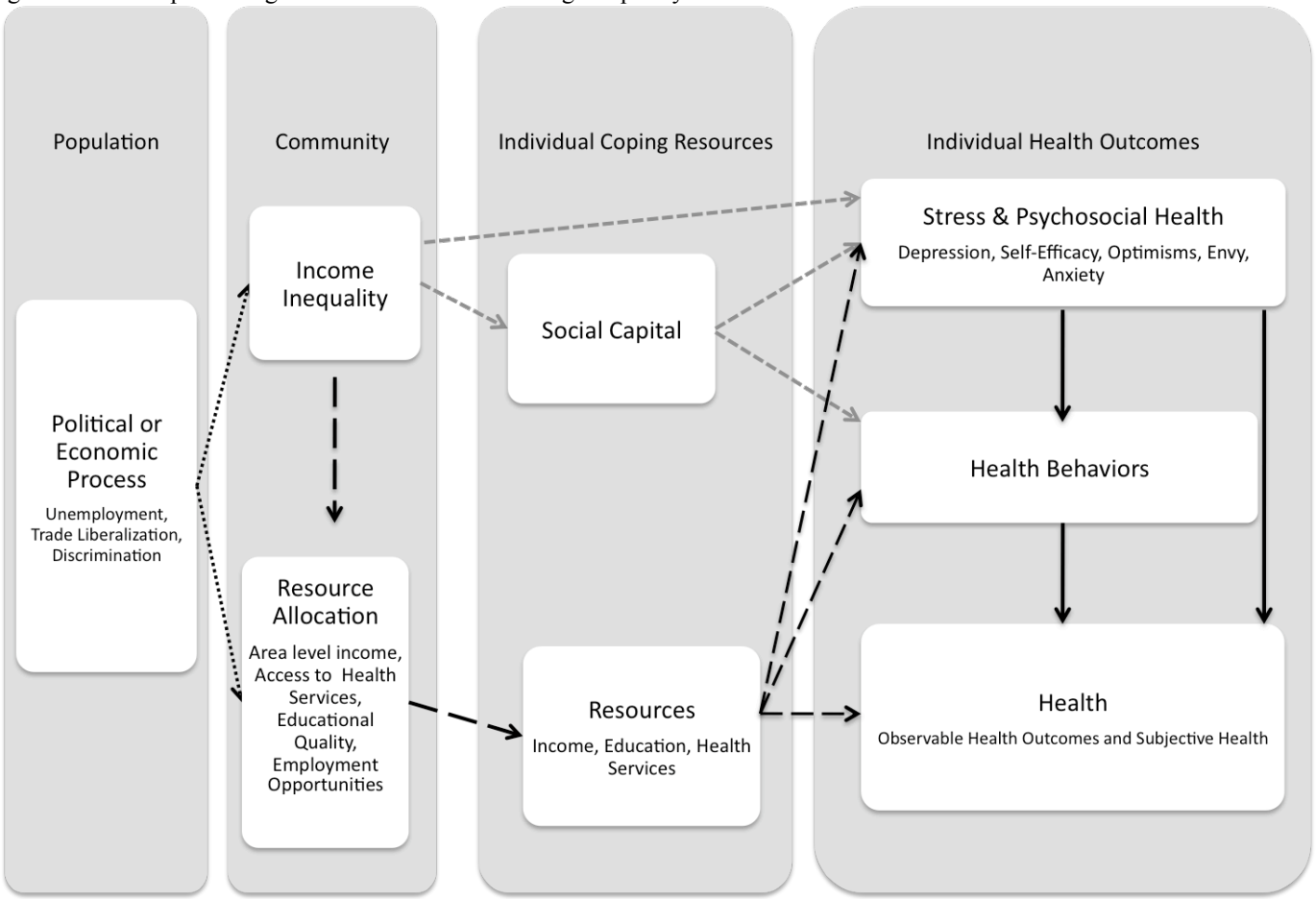
Fractionalization is $Frac_j = 1 - \sum_{i=1}^{N} s_{ij}^2$ calculated as where s is the share of group i in country j. Higher values indicate greater heterogeneity.
Source: (A. Alesina, Devleeschauwer, Easterly, Kurlat, & Wacziarg, 2003)

Table 1.4: Comparison of Selected Political Indicators

Selected Political Indicators			
	Standardized Freedom House Score 2001	Ranking for # women in parliament	HDI Rank 2002
Brazil	71.4	89	65
Chile	85.68	60	43
Costa Rica	92.82	29	42
Guatemala	64.26	72	119
Nicaragua	71.4	66	121
United States	99.96	50	7

Source: Pippa Norris. 2006. *Democracy Time-series Dataset*.
http://www.pippanorris.com/. Higher values of Freedom House scores indicate higher functioning democracy.

Figure 1.1: Conceptual Diagram of Mechanisms Relating Inequality Health Outcomes



Psychosocial pathway
Resource allocation pathway
Neomaterial pathway
All pathways

Chapter 2: Descriptive Analysis of the Measurement of Income Inequality in Costa Rica

2.1 Introduction

Costa Rica is a notable exception in the Latin American development literature. Compared to other Latin American countries, Costa Rica has had consistently more political stability, better social services, higher life expectancy, and less income inequality than its regional counterparts (see Table 1.2 and 1.4 as well as the discussion in Chapter 1). Many have speculated that this lower income inequality has contributed to the nation's exceptional health outcomes.

The goal of this chapter is describe the changes in income inequality and other key labor force measures from the late 1980s to 2005. In doing so, it will then be possible to assess hypotheses relating income inequality and health outcomes. The descriptive analysis that ensues will build on previous work by documenting (and validating) the national and regional trends of various labor force and income indicators across Costa Rica over time. The first goal is to document labor force, income inequality, and poverty trends at a national level. The second goal is to map the income inequality indicators at a regional level, to assess their stability, and to map changes and patterns over time. These inequality measures will then be used to examine possible associations with the other area-level indicators, including social capital. This section will examine a broad time period but will ultimately focus on 1987 to the present (for data reasons explained below).

Before turning to the inequality data, a brief outline Costa Rica's recent macro-economic policies and conditions is necessary. This account is not meant to be comprehensive, but rather it is meant to help validate the descriptive analysis from the data that follows.

2.2 Economic Landscape in Modern Costa Rica

Between 1950-1979, the Costa Rican economy was financed principally by foreign capital and national agro-business elites, leading some scholars to describe it as a dependent capitalist state, a form of capitalism in which the trade of raw materials to core countries is the main sector of the economy (Mesa-Lago, 2000).[22] In this period, agro-business as a whole, and particularly coffee and banana crops, served as the main sector of the economy and the chief source for all foreign currency, which is consistent with the depiction of Costa Rica's dependency. From the 1950s to the 1970s, the Costa Rican economy began to shift away from agricultural exports towards protected industrialization where an import substitution model of trade was encouraged. The incorporation of Costa Rica into the Central American Common Market (CACM) in 1963 helped stimulate this shift. In the 1970's, the country followed a strategy of high investment in state-owned firms and other policies that expanded government expenditure, which led to high fiscal deficits and a strong increase in public debt (A

[22] Mesa-Lago (2000) refers to Costa Rica as a "dependent capitalist state with social considerations" to account for the social safety nets established under the 1949 constitution and the universal access to social security and health insurance mandated in 1961.

Rodríguez-Clare, M Sáenz, & Trejos, 2002). Meanwhile the nascent consumer products industry (versus heavy industry) was developing, and foreign exchange was necessary to buy inputs for the production process. Varying agricultural prices often could not support the need for foreign exchange. Together, these factors led to a growing external debt imbalance. At the same time, the expanding social programs led to increasing internal debt. By 1980, the national debt had grown so much that the government could not make interest payments, and a debt crisis ensued. In September of 1980, Costa Rica decided to devalue its currency, which led to even greater debt and the economy contracted substantially (Mesa-Lago, 2004).

From 1980-1983, Costa Rica found itself in the midst of a severe recession. Unlike Chile and Argentina, which chose to take on rapid structural adjustments to get loans from the IMF (known as the "shock therapy approach"), Costa Rica chose a much more gradual acceptance of IMF loan terms for both trade liberalization and privatization. For example, Costa Rica only began to undertake trade liberalization in the mid to late 1980's (1986-1988). Privatization of state industries was also enacted slowly, partly due to strikes of telephone and electricity workers against the privatization of ICE (Instituto Costarricense de Electricidad) in 1988 (Eduardo, 2009). There were also protests against the expansion of multinational corporations, such as Alcoa, (Palmer & Molina, 2004c). Nevertheless, this gradual approach is thought to have mitigated some of the negative long-term impacts of these reforms such as high transitional unemployment, lower government subsidies, and higher inflation (Gingrich & Garber., 2009).

As a consequence of the structural adjustments required by the IMF, Costa Rica's GDP has recovered, but has also varied greatly. From 1986 onward, Costa Rica has expanded free trade, established a high-tech free-trade zone, and experienced a substantial increase in tourism. Figure 2.1 shows Costa Rica's yearly GDP growth since the late 1980's. The graph shows two distinct periods of GDP expansion, one in the early 1990s and the second at the end of the 1990s. The early 1990s increase in the GDP may have been the consequence of the first round of trade liberalization and privatization; in 1993, the fastest growing sector was the newly privatized financial intermediaries sector (Eduardo, 2009). The late 1990s increase in GDP may have been due to the establishment of the free-trade zones that attracted direct foreign investment such as the Intel™ campus, which accounted for 40% of Costa Rica's exports in 1999.[23]

Since the late 1980s, Costa Rica has also experienced an increased earnings inequality. In a recent article, Gindling argues that the main contributors to the growth in earnings inequality in Costa Rica from 1990-2005 have been the increase in dispersion of educational attainment and higher returns to education. Less educated Costa Ricans have mitigated this by increasing the number of hours worked, including more females working part-time, which has lead to a decrease in income inequality from what it would have been otherwise. Some have speculated that the increase in inequality is also due to the influx of Nicaraguans since the late 1980s. However, Gindling argues against this notion and finds that there is little evidence of any wage gap between Nicaraguans and Costa Ricans after accounting for observable characteristics (Gindling, 2009; Gindling & Trejos, 2005). Nonetheless, it might be argued that the influx of Nicaraguans could have lowered wages for native Costa Ricans, which in turn could have broadly contributed to

[23] Detailed accounts of Costa Rica's macroeconomic situation can be found at http://dosfan.lib.uic.edu/ERC/economics/trade_reports/1994/CostaRica.html

increases in inequality. This issue remains to be resolved. Given these broad national developments, the next section will detail the data required for understanding income, labor force, and income inequality trends in Costa Rica over the last 20 years.

2.3 Data

Three different data sources were used to create the dataset employed by this descriptive analysis:

(1) The Costa Rican Household Surveys for Multiple Purposes (CRHS) are used to generate area-level measures of income, income inequality, migration, unemployment, and extreme poverty.

(2) 1973 and 1984 census data are used to match area-level identifiers to CRHS.

(3) Population projection data are used to reweigh the CRHS data.

The main source of the data used in this section will be the Costa Rican Household Surveys for Multiple Purposes (CRHS), conducted in July of each year from 1976 until the present (except for 1984) by the Costa Rican Institute of Statistics, Instituto Nacional de Estadisticos y Censos (INEC). The CRHS surveys are nationally representative household surveys of approximately 1% of the population. While the sample changes from year to year, 75% of the selected primary sampling units are maintained in any two consecutive years. These surveys are the only source of comparable yearly data on the earnings and personal characteristics of workers that is available in Costa Rica.

While these surveys are one of the most consistent data sources available, they have three major drawbacks. First, the CRHS has undergone two main revisions that, for practical purposes, make the annual measures incomparable across some years. Between 1976 and 1979 the survey asked about the earnings of only salaried employees, whereas from 1980 to the present, the earnings question was asked of salaried, hourly and self-employed workers. There were also substantial changes in the survey sample, design, and questionnaire between 1986 and 1987 rendering the pre- and post- era measurements not fully comparable (Gindling & Trejos, 2005). Second, the canton and district level identifiers are included in the survey only for 1987-1991 and from 2005 forward. For the 1976-1986 data these identifiers must be linked in from the 1973 census based on a "*segmento*" identifier (census tract). For a small portion of the data, area identifiers cannot be matched because certain census tracts did not exist in the prior census. More importantly, the area-level identifiers between 1992-2004 are scrambled in the dataset. Instead of listing the *segmento* identifier, a consecutive number is listed to mask the census tract. Thus the canton and district identifiers must be inferred from other variables in the dataset, whereas if the *segmento* identifier had been listed, it could then have been partitioned to create canton and district identifiers. Finally, the data are only weighted at the regional level; thus in order to make area-level measures, these data need to be reweighed based on area-level population projections (canton and district level).

Several data processing steps have been necessary to address these problems. First, for the 1976-1986 household surveys, a crosswalk file based on the 1973 and 1984 censuses are used to match *segmento* identifier area-level identifiers. Second, to address

the problem of missing identifiers for the years 1992-2004, two algorithms were used to approximate the canton identifiers that were recoded in the data. The first algorithm exploits changes in urban/rural and regional identifiers to approximate the consecutive *segmento* number in cases of new cantons. Since the data are ordered by consecutive *segmento* numbers, and shifts in the urban/rural or regional identifiers usually indicate a canton change, it follows that this information can be compared with flanking years' data to identify the consecutive *segmento* values where canton shifts have occurred. This approach was applied in the construction of canton identifiers for the 1992, 1993, 2003 & 2004 CRHS data. A second and more robust method was also used. The variables "Migracio" and "Migrante" in the data are used to determine how *segmentos* map to cantons. When a person did not move, his canton of residence is enumerated in the Migracio variable. By limiting the data to those who did not move, the Migracio and consecutive *segmento* number can be manipulated to construct a crosswalk file to map canton codes to consecutive *segmento* codes. This second algorithm was used to construct canton identifiers for the 1992, 1993, 1995, 1996, 2001 and 2002 CRHS. When comparing these two algorithms for 1992 and 1993, the results agree 98% of the time, indicating that either method is probably adequate to use to generate the necessary canton identifiers.[24] Third, the dataset was reweighed using population projection data for people 10 years or older. [25] Fourth, inequality measures may be prone to measurement error if they are calculated over too few observations due to outliers.[26] Thus, to increase the number of observations in any given time-year cell, the inequality measures were constructed based on two or three years of surveys (Table 2.1 shows these year groupings). Finally, the problem of revisions in the sample and design of the Household surveys between 1986 and 1987 is harder to address. While inequality measures and labor force trends at the national level are presented for 1980 forward, the inequality measures used in the regional analysis will focus on the post-1987 surveys due to the changes in sample and design.[27]

Before using CRHS data, several data cleaning steps were also performed. First, for the 1976-1979 data, all observations were not used in the inequality calculation where: a) the hours worked in primary job is 0 or 99; b) the person was not employed; or c) the salary was equal to 0. Finally, all observations younger than 12 years old were deleted in the 1976-1986 datasets as these data should not have surveyed those younger than 12.

---

[24] Since the canton level identifiers are missing for all the years between 1992-2004, the algorithms were only used to create the canton level identifiers for some, but not all the years. Specifically, measures for 1987/1988, 1989/1900/1991, 1992/1993, 1995/1996, 2000/2001, 2003/2004 and 2005/2006 are available. Data for 1994, 1997, 1998, 1999 and 2002 are not constructed and will not be necessary in the subsequent analyses. It was not possible to get down to the district level identifiers using either of these algorithms. Thus there are seven year groupings for which we have canton level identifiers between 1987-2005, but only three year groups for which we have this data at the district level. Ultimately, for reason described above canton data will be employed.

[25] The number of individuals aged 10 years and older is used as the weights because the survey only used those 12 years and over for the employment questions across all years.

[26] This is a particularly important limitation to acknowledge as fixed effects models are generally more biased when there is measurement error. Some simulations will be conducted in subsequent chapters to find bounds on this measurement error problem.

[27] The graphs to follow show clearly that for a number of measures there is an unexpected discontinuity between the 1986 and 1987 measures.

Even though from 1987 onward, datasets did include this age group, no observations from respondents under 12 were used in any of the inequality calculations.[28]

Generally, when making an inequality scale based on a household measure, some adjustments are necessary to account for household size. To ensure reliable comparison, it becomes necessary to account for poorer families coming together to form larger households to share resources and ensure basic consumption. Since inequality measures are based on monthly household income, a household equivalent monthly income variable is first created. The square root method, which divides total monthly household income by the square root of the number of household members, was employed. This method has the lowest "equivalence elasticity" of the methods suggested by the OECD, and the smaller values for this elasticity suggest higher the economies of scale in consumption for the household (Foster, 2009). Given that households tend to share more resources in developing countries, this method seems reasonable.

The measure of inequality used in these analyses, the Gini coefficient of household equivalent monthly income, is applied at both the canton and district level.[29]

2.4 Description and Analyses of Constructed Measures

2.4.1 National Trends

First, the overall labor force trends in Costa Rica over the past 25 years are presented. Figure 2.2 shows the trends in unemployment rate, labor force participation rate, female labor force participation rate, and the percentage of workers who are salaried from 1980-2006. There are a few notable features of the graph. First, there is a marked jump in the labor force participation rate, the female labor force participation rate, and the percentage of workers who were salaried between 1986 and 1987. This reflects the changes in the survey sample, design, and questionnaire from one year to the next. Second, unemployment rates have been steady over this period except for a brief increase in the early 1980s when unemployment rates were estimated at about 10% during the Latin American debt crisis. Third, throughout this entire period, female labor force participation increased rapidly from 25-30% in the mid-1980 to 41% in 2005. This rise in female labor force participation is also reflected in the increase in overall labor force participation. Finally, the trend in the percentage of salaried workers is hard to interpret, because there was a change in the sampling frame to include more *segmentos* with migrants following the 2000 census.

Figure 2.3 shows the Gini coefficients over this same period. These results are consistent with Gindling and Trejos, showing that Costa Rica's inequality fell throughout the mid-1980s. After 1987 this pattern of declining inequality changed, first stabilizing between 1987 and 1992 and then increasing from 1992 to 2006 (Gindling & Trejos, 2005).[30] There are several other marked trends. First, there was a spike in household equivalent income inequality in the early 1980s that coincided with the Costa Rican debt

[28] Those under 12 were never used in inequality calculations.
[29] Refer Chapter 1 for a discussion of the nuances of creating an inequality measure and which measure may be best.
[30] The Gini measures use here are weighted and replicate those found by Gindling & Tejos's who use the same dataset. However, Gindling & Tejos do not include a household equivalent measure of inequality.

crisis described above. Using non-equivalized data this spike is not as pronounced. Second, from 1987 forward household equivalent measures of inequality are quite similar to the unequivalized measure and the correlation coefficient between these measures is 0.98 (p-val< 0.001). Finally, the Gini coefficients for salaried workers and all workers follow a similar trend throughout this period. The inequality is consistently higher in the Gini measure that includes all workers. The correlation coefficient comparing these two measures of income inequality is 0.93 (p-val< 0.001). Based on this analysis, the household equivalent inequality measures from 1987 forward will be used throughout the remainder of the study. Moreover, since the measures of the Gini coefficients using all workers or salaried workers are highly correlated, the measure based on salaried workers will be used to ensure consistency, especially due to the changes in sampling in 2001.[31]

The trends for poverty are also interesting in that they have been steadily decreasing in since the 1980s. Figure 2.4 presents poverty rates in Costa Rica from 1980-2006. The household survey has two measures of poverty: extreme poverty and not being able to pay for basic necessities. Total poverty and extreme poverty are graphed separately by year. In Figure 2.4 the 1980 debt crisis is clearly observable as a spike in both measures of poverty in 1981 as well as in 1982. The figure shows the downward trend in both poverty measures since 1986 with a single minor increase in the early 1990s.

To understand which segments of society are driving the increases in income inequality, the percentage increases in income at specific percentiles of the income distributions are presented from 1987-2005. Figure 2.5 shows the percentage increase in Consumer Price Index adjusted monthly income thresholds at the 10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}, and 99^{th} percentiles of the income distribution with 1987 as the base year. In the top corner of Figure 2.5, yearly growth rates are calculated for the overall secular trend. Both the growth rates and standard errors are presented. The graph shows that those in the 10^{th}, 25^{th}, and 50^{th} percentiles saw very little growth in household equivalent income over this period where as those at the top of the income distribution, the 75^{th}, 90^{th}, and 99^{th} percentiles, saw substantial and statistically significant growth (even if this growth tapered off in the last few years of the data). The divergence is most clear in 1991 and 1997: the same years that Costa Rica experienced significant GDP growth (see Figure 2.1), which suggests that the rich are disproportionally favored in these growth years. Even more importantly, those at the 99^{th} percentile had the largest yearly growth rate of 2.5%.[32] Figures 2.4 and 2.5 together show that at a national level, those at the top of the income distribution are driving increases in income inequality. This trend is similar to those that have been observed in the United States where increases in income inequality have been due to those at the top of the income distribution.

---

[31] When the generated Gini coefficients of salaried workers versus all workers are compared by canton across all years a correlation coefficient of 0.76 (p-val< 0.001) is found. Also, in Figure 2.3 there is a larger jump in income inequality for the all-workers measures between 2000 and 2001 than in the salaried workers measures.

[32] A least-squares growth rate, r, is estimated by fitting a linear regression trend line to the logarithmic annual values of the household equivalent income in each percentile group for each year across 21 time points. The regression equation takes the form, $\ln X_t = a + b_t$, which is equivalent to the logarithmic transformation of the compound growth equation, $X_t = X_o (1 + r)t$. If b* is the least-squares estimate of b, then the average annual growth rate, r, is obtained as [exp(b*) – 1] and is multiplied by 100 to get a percentage.

These national trends are not only consistent with previously published trends but are also consistent with the macro economic situation that is widely accepted. It was crucial that these results match previously published accounts of the Costa Rican economy because these same underlying data will be used to construct regional level measures that have yet to be explored.

2.4.2 Regional Trends

Costa Rica is comprised of 7 provinces that are subdivided into 81 cantons and these cantons are further subdivided into approximately 460 districts. The second goal of this analysis is to create meaningful measures of inequality at sub-national levels in order to explore changes in inequality in an area over time. Based on the same data described above, a household equivalent Gini coefficient for each canton (or district) is created based on either 2 or 3 years of the CRHS data in order to ensure there are sufficient observations for each canton (or district).[33] Table 2.1 presents the years of the CRHS that were grouped to create the measures. The tables and figures below use measures of the Gini coefficient based on salaried workers, but using the Gini for all workers gives similar results.

2.4.2.1 Canton Analysis

A canton is the second-level administrative division of Costa Rica. Cantons range in size from approximately 8 square kilometers to 3,350 square kilometers with an average of about 700 square kilometers. They have an average population of about 50,000 with varying population densities.[34] Finally, cantons are the only administrative divisions in Costa Rica with municipalities. These local government bodies are an extension of the central government.

Income inequality measures were constructed for each canton as described above over several years. In order to assess the internal validity of these inequality measures, the relationship of these measures across time is examined. Figure 2.6 shows a scatter plot for income inequalities by canton for adjacent years. The lines represent the weighted mean of the Gini coefficients across all cantons for those years, and each marker size reflects the mean income of that canton. All the graphs show that our measures are relatively consistent from year to year. This is reinforced by the correlation coefficients of about 0.5 (p-val<0.001) for each set of adjacent years. In contrast, looking over a longer period, the Gini coefficients have changed considerably. Figure 2.7 shows a scatter plot of income inequalities by canton comparing the 1989/1990/1991 to 2005/2006 and 1992/1993 to 2003/2004. Here the previously observed correlation disappears. There are several cantons with low Gini coefficients (below the mean) in the early 90's whose subsequent Gini coefficients in 2003/2004 are high (above the mean).

[33] On average the Gini coefficients are calculated over 500+ observations at the canton level. Only 3% of the canton Gini measures were created based on less than 100 obs. 9.3% of the canton Gini measures were created based on less than 150 observations and 17% were created based on less than 200. For district Gini coefficients, only those with greater than 50 observations were used to calculate the Gini.

[34] The size of cantons (or district) is particularly important in that the income inequality hypothesis requires that the reference group over which inequality is measured be a group that one interacts with regularly. If most social and economic interaction take place within one's canton (district) then this local environment becomes more salient and hence a likely reference group. Here this reference group is assumed to be reasonable, but it cannot be empirically validated.

Table 2.2 presents the correlation coefficient for canton Gini coefficients in 1887/1988 to all other years. As expected, a gradual decline over the correlations is found (and significance of the correlation) when observations get further apart in time. Table 2.3 presents the correlation coefficient for canton Gini coefficients in 2005/2006 to all other years. Here again the same graded decline in correlation (and significance) is apparent as observations get further apart. This suggests that while there is some reasonable stability in these measures, the study period is a time of changing inequality internally.[35]

To further validate inequality measures at the canton level, the relationship of these changing Gini coefficients with the changes in other canton level measures is assessed. The results show that areas with increasing mean incomes also have increasing Gini coefficients during the same period. This was verified using a regression across 4 years with the unit of observation taken to be the canton-year. After including canton year fixed effects, a positive relationship between the Gini and mean income is apparent. Moreover, the Gini coefficients are not highly associated with the proportion in extreme poverty, mean level of education, unemployment, or migration at the canton level once control for canton and time fixed effects were included (regressions not presented).

Given that the analysis will ultimately use the variation in *changes* within the Gini coefficients, the characteristics of cantons with increases in inequality are explored. First, a least-squares growth rate is estimated for each canton by fitting a linear regression trend line to the logarithmic Gini values of the household equivalent income over four time points.[36] This method estimates a growth rate per canton, which is used as the dependent variable to assess the association with lagged and leading labor force characteristics in the areas.[37] Table 2.4 presents the weighted regressions for the 78 cantons for which growth rates could be estimated. The growth rate in the Gini coefficient is associated with areas that had a higher proportion of workers in manufacturing jobs according to the CRHS in 1987. These growth rates are also associated with areas that had a higher proportion of workers in all sectors of the economy in 2005. This implies that the Gini coefficients are increasing in areas that are developing economically.

Finally, Gini coefficients in levels and changes for each canton are mapped in 1989 and 2005 allowing for an examination of the spatial patterning. Figure 2.8 shows a map of these measures. Panel A presents the Gini coefficients in 1989/1990/1991 and Panel B shows these same data for 2005/2006.[38] Areas in the northwest of Costa Rica in the province of Guanacaste show consistently high inequality, which has been documented elsewhere from census wealth measures (per L. Rosero-Bixby, Professor and Director of the Central American Population Center, personal communication, September 14, 2009).[39] Likewise, some areas on the central Caribbean coast, such as Limón, also exhibit consistently high inequality. A more relevant map exhibits areas that have experienced

[35] While the sampling design of the CRHS, which refreshes 25% of the *segmentos* each year, could explain some of this internal consistency it cannot explain all of it.
[36] Only 78 cantons have an estimated growth rate because only cantons with 4 Gini measures were used.
[37] The population in the area was used to weight observations in these OLS regressions.
[38] Only cantons with greater than 250 observations in 1987 were included in these maps to ensure greater comparability. See discussion below.
[39] This is an area with historically high inequality rooted in the Hacienda style of development referred to in Chapter 1.

changes in their Gini coefficients. Figure 2.9 shows the cantons' estimated 5-year growth rates of inequality.[40] The first figure includes all the cantons for which data were available, and the second figure includes only areas that have more than 250 observations over which the Gini was calculated (see discussion below). While there are some small differences, the overall trend seems to remain.

Three general areas have exhibited changing inequality. Cantons to the east of San José and those on Nicaraguan border have had increasing inequality, while several costal cantons to the west of San José have had decreasing inequality. These maps show a clear pattern of changes in inequality along the central Costa Rican axis from Limón to Puntarenas. This is also the area that has been developing most quickly in the past several decades (per L. Rosero-Bixby, Professor and Director of the Central American Population Center, personal communication, September 14, 2009). International migration may be driving some of the increase in income inequality because cantons that have had the highest proportion of recent Nicaraguan immigrants are also those with changing inequality.[41] This issue may not be directly apparent in these data analyses because the measure of migration includes both internal and foreign migrants.

Figure 2.10 also presents the maps of changes in the other area-level measures used in these analyses.[42] Panel A presents the estimated 5-year growth rates for unemployment rates in each canton. This unemployment rate was calculated using the CRHS and where unemployment was defined as those not working and looking for work in the last week. There seems to be an increasing rate of unemployment in the southwestern area of the country. This may be due to the closing of many of the banana plantations in the area. Panel B presents the estimated 5-year growth rates of in-migration for each canton. The measure captures the proportion of those surveyed who moved in the last two years and includes both internal and foreign migrants. Panel C presents the estimated 5-year growth rates for HHEMI. Together Panel B and C show that areas with increasing income tend to have increasing migrants and vice versa. Panel D presents the estimated 5-year growth rates for the proportion of those surveyed in each canton who were considered extremely poor. INEC designates the poverty line in each year of data by total household income, number of people in the household, and whether one lives in an urban or rural area. Those below the poverty line are considered to be in extreme poverty. Panel C and D together show that these data have *prima facie* validity, since the areas that had increasing HHEMI also had decreasing poverty.

2.4.2.2 District Analysis

[40] Here again a least-squares growth rate is estimated by fitting a linear regression trend line to the logarithmic Gini coefficient values of the household equivalent income over 4 time points.
[41] This is consistent with the 2000 census where cantons La Cruz (510), Upala (213), Los Chiles (214), and Sarapiqui (410) had some of the highest proportion of recent Nicaraguan immigrates.
[42] Here again a least-squares growth rate is estimated by fitting a linear regression trend line to the logarithmic Gini coefficient values of the household equivalent income over 4 time points. This method allows to both the growth rate and the confidence intervals of the growth rate. The confidence intervals are used to decide on the cut-off for the magnitude grouped.

As argued in Chapter 1, the appropriate geographical unit of analysis over which to assess the income inequality hypothesis is ambiguous. Accordingly, the possibility of examining a smaller unit of analysis, districts, is considered. Districts are the smallest administrative units and somewhat comparable to zip codes in the United States in size and scale. There are approximately 460 districts in Costa Rica with an average population of 8,000. Districts, like cantons, vary considerably in size, and therefore the number of observations per district in CRHS was quite small for several districts. Consequently, only districts with at least 50 observations in CRHS and those that are represented in all years of the CRHS were considered.[43] There are only 3 time periods in the CRHS that included the geocodes necessary to create measures of district level inequality. Figure 2.11 shows income inequality by district for adjacent years and for two years with a 15-year gap. Here the graph shows that the Gini measures are not highly correlated even in adjacent years. This suggests that these measures are highly variable from year to year. Since this study relies on survey data instead of census data, the inequality measure depends largely on who is surveyed from year to year, and in these smaller geographic areas this may lead to inconsistent (and invalid) measures of inequality. Thus, district-level inequality measures are not used in the remainder of this thesis.

2.4.2.3 Measurement Error

As mentioned the Gini measures created above are based on survey samples instead of census data. These samples may have too few observations in any given area and in any given year to generate an accurate measure of an area's income distribution (see Table 2.5). In order to understand the number of observations required to create reliable Gini measures, simulations were done. Using the data for two cantons, 50 observations were sampled and used to create the Gini coefficient. This was repeated one hundred times. From these hundred observations, an estimate of the mean and standard deviation of these Gini coefficients was calculated. This procedure was repeated 7 times, each time increasing the size of the subsample by increments of 50 until 350. Table 2.6 Panel A & B presents the results from these simulations for Heredia (canton 401) and Goicoechea (canton 108).

Comparing the Gini measures over the entire survey sample versus the subsamples shows that indeed the estimated magnitude of the Gini coefficients increase with subsample size. When only 50 observations are sampled, the variability in the estimated Gini coefficient is quite large. Yet, as the number of observation increases past 250 the variability becomes much smaller. This confirms that measurement error is likely to be more common for district measures. Again this suggests that for data limitation reasons, cantons are a more reliable unit of analysis as they all have at least 50 observations. The number of cantons that have fewer than 250 observations ranges from 9 in 2005 to 27 to in 1987, and thus measurement error and attenuation biases will be assessed in models presented in the subsequent chapters.

2.5 Other Analyses

2.5.1 Social Capital

[43] Not all districts are represented in the CRHS data in all years whereas most cantons (79 or 80 of 81) were in the CRHS dataset.

One pathway in which income inequality may affect health is through the erosion of social capital (described in Chapter 1). Recently, researchers have argued that certain components of social capital may be related to population health outcomes (Kawachi & Kennedy, 1999b; R. Scheffler, Brown, Syme, Kawachi, Tolstykh, & Iribarren, 2008a). While measures of social capital vary substantially in unit of observation and scope, measures that tap into levels of interpersonal trust, often referred to as cognitive social capital, are particularly interesting for health outcomes. In one study in the United States, low cognitive social capital as measured by the percentage of survey respondents in each state answering that "most people can't be trusted" was correlated with higher mortality rates ($r=0.79$, $P< 0.05$). This study also suggested that those living in states characterized by high income inequality tend to be more mistrustful of each other ($r=0.71$, $P< 0.05$) (Kawachi, Kennedy, Lochner, & Prothrow-Stith, 1997). In addition, this study also documented a significant relationship between the measures of interpersonal trust and mortality even after controls for state median income and poverty rates were included.

Accordingly, the association between income inequality and social trust was examined in Costa Rica using the Latin American Public Opinion Project.[44] The survey is a nationally representative questionnaire on a range of issues. In survey years 2002, 2004, 2006, the survey asks "Now talking about the people from around here, would you say that the people of your community are very trustworthy, somewhat trustworthy, little trustworthy or not at all trustworthy?"[45,46] This question is meant to tap into people's interpersonal trust and measure what many have termed cognitive social capital. If the psychosocial mechanism proposed in Chapter 1 is working through social capital, then areas with increasing income inequality may also exhibit decreasing social trust.

The LAPOP questionnaire was administered in 56 of the 81 cantons in Costa Rica. The raw correlations indicate that there is no relationship between the area-level social trust measure (percent who answer little or no trustworthy) and income inequality in the 2002, 2004 or 2006 data. These associations were then reexamined in a multivariate context using weighted least squares in Table 2.7. Again no correlation between changes in income inequality and social capital was found.[47] However, since the standard errors are quite large, an association cannot be ruled out. In addition, the regressions consider changes in social trust over 2-years, but it is unlikely that social trust changes in response to changes in income inequality in such a short time period. In contrast, there is an association between interpersonal trust with unemployment and migration. Areas with growing unemployment exhibit an increasing number of people reporting low-trust (marginally significant). Likewise, areas with greater in-migration exhibit higher trust (see Table 2.4).

---

[44] Microdata were not available and must be obtained from Vanderbilt University, web page: http://www.lapopsurveys.org. Canton aggregates were constructed based on data available from the Central American Population Center (CCP).

[45] In 2002 and 2004 the question left out the "Now talking about the people around here" part.

[46] Using this same data and comparing eight Latin American countries, Rosero-Bixby finds that Costa Rica has the highest value of social trust. This is thought to be consistent with its democratic tradition and confidence in institutions (Luis Rosero-Bixby, 2006).

[47] Since the outcome variable is a proportion, a logit transformation is also examined to see if the results and standard errors improved in using this different functional form, but they did not.

2.6 Discussion

These descriptive analyses show that canton-level income inequality measures from household equivalent monthly income for salaried workers are satisfactory measures of income inequality at the sub-national level. The created measures are consistent with those previously published at the national level and are stable from year to year at the canton-level. Additionally, at both the national and canton level, the results demonstrate that these measures are not merely proxies for area-level poverty. In fact, there is some evidence that income inequality is higher in more affluent areas and that areas with rapidly changing inequality are also the areas that have been developing most rapidly during this time period. Moreover, these measures clustered on the maps, which would not have been the case if the Gini measures were completely random. In creating these measures several years of data were pooled to ensure sufficient observations at the canton level, yet even with these precautions the possibility remains that these measures are still too noisy for a valid temporal analysis, as fixed effects models tend to exacerbate measurement error and cause attenuation bias.[48] This is a key limitation that will be empirically assessed in subsequent analyses.

These analyses also attempted to assess the arguments that social capital is highly related to income inequality. In the Costa Rican setting, this relationship was tested with the limited nationally representative data that are available. Yet, the results do not support the notion that there is a relationship between inequality and social capital, at least in terms of cognitive social capital. That is not to undermine the potential for a relationship in other measures of social capital. Also considering that increases in inequality are due to the widening of the income distribution at the top at the same time as poverty is decreasing, the consequences for social capital are also less clear. This question deserves much more scrutiny than is presented here.

[48] Initially, district level inequality measures were examined in preliminary analysis as the motivation for this project before longitudinal measures were constructed. For a single measure it is difficult to determine whether the measure is flawed. Having multiple measures for the same unit of observation allows for the analyses above, which indicates that there is reason to be skeptical of the district level measures. The analyses also show that while canton measures are not perfect, they are superior.

Tables

Table 2.1: Year Groupings Used to Generate Area-Level Measures from July Household Survey

Canton Group	Years	District Group	Years
1	1987, 1988	1	1987, 1988
2	1989, 1990, 1991	2	1989, 1990, 1991
3[&]	1992, 1993		
4+	1995, 1996		
5+	2001, 2002		
6[&]	2003, 2004		
7	2005, 2006	3	2005, 2006

[&]Canton codes were not in the original dataset. They were generated using a mapping algorithm. + Canton codes were not in original dataset. They were generated using migration algorithm.

Table 2.2: Correlation Coefficient for Canton Gini Coefficients in 1987/1988 to All Other Years

Correlation Coefficient- Canton Gini coefficients		
Group	Year	1987, 1988
2	1989, 1990, 1991	0.533**
3[&]	1992, 1993	0.363**
4+	1995, 1996	0.263*
5+	2001, 2002	0.288*
6[&]	2003, 2004	0.1614
7	2005, 2006	0.0714

&,+ See Table 2.1 notes. * significant at 5%; ** significant

Table 2.3: Correlation Coefficient for Canton Gini Coefficients in 2005/2006 to All Other Years

Correlation Coefficient- Canton Gini coefficients		
Group	Year	2005, 2006
1	1987. 1988	0.0714
2	1989, 1990, 1991	0.0980
3[&]	1992, 1993	0.0529
4+	1995, 1996	0.2845[*]
5+	2001, 2002	0.4357[**]
6[&]	2003, 2004	0.5286[**]

&,+ See Table 2.1 notes. *significant at 5%; ** significant

Table 2.4: Association Between Canton-Level Growth Rate in Gini from 1989-2005 and Labor Force Characteristics

Population Weighted Least Squares Regressions of Gini Growth Rate		
Gini Coefficient in 1989	**-0.05**	**-0.05**
	[0.022]*	**[0.019]****
	1989 Values	*2005 Values*
Mean HHEMI (10,000 colones)	0.002	0.006
	[0.006]	[0.005]
Unemployment Rate	**-0.102**	0.001
	[0.049]*	[0.063]
Migration	-0.015	0.019
	[0.034]	[0.061]
Proportion with Secondary Education or Above	0.003	-0.019
	[0.025]	[0.023]
Proportion Professional	-0.075	**0.136**
	[0.107]	**[0.064]***
Proportion Technical	-0.003	**0.109**
	[0.033]	**[0.053]***
Proportion Agricultural	-0.051	**0.133**
	[0.049]	**[0.058]***
Proportion Manufacturing	**0.298**	**0.122**
	[0.127]*	**[0.051]***
Proportion Service	0.070	**0.111**
	[0.104]	**[0.056]+**
Constant	0.023	-0.095
	[0.017]	[0.051]+
Observations	78	78
R-squared	0.25	0.22

Robust Standard errors in brackets
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 2.5: Mean Number of Observations that Canton -Level Gini Coefficient were Calculated over by Year

Number of Observations that Gini is Calculated Over				
Year	Mean	Std. Dev.	Freq.	Cantons with <250 obs
1987	570	634	80	27
1989	880	985	80	17
1992	641	710	80	25
1995	695	764	80	22
2001	653	604	79	16
2003	818	756	79	10
2005	847	777	79	9

Table 2.6: Mean Estimated Gini coefficient Based on Simulations with Varying Number of Observations Used to Calculate Gini Coefficient

Panel A: Estimated Gini for Canton 401 Heredia

Number of Obs used in Gini calculation	Estimated Gini	Standard Deviation
50	0.341	0.050
100	0.347	0.035
150	0.349	0.028
200	0.349	0.024
250	0.350	0.018
300	0.350	0.018
350	0.352	0.015
	Actual Gini	Actual Obs
	0.353	690

Panel B: Estimated Gini for Canton 108 Goicoechea

Number of Obs used in Gini calculation	Estimated Gini	Standard Deviation
50	0.355	0.051
100	0.361	0.039
150	0.363	0.030
200	0.364	0.024
250	0.363	0.021
300	0.363	0.018
350	0.364	0.018
	Actual Gini	Actual Obs
	0.366	1060

Table 2.7: Association of Social Trust and Income Inequality at the Canton Level, 2002, 2004, 2006

	Proportion who report low trust in canton [0-1]		
Gini coefficient	-0.192	-0.202	0.389
	[0.456]	[0.413]	[0.689]
(Mean HHEMI)/100,000	0.008	**0.050**	-0.004
	[0.02]	**[0.019]****	[0.065]
Mean Years of Education		-0.010	-0.005
		[0.008]	[0.008]
Unemployment Rate [0-1]		**2.100**	1.638
		[0.599]***	[1.065]
In- migration [0-1]		-0.792	**-1.266**
		[0.704]	**[0.63]****
Constant	0.327	0.198	0.114
	[0.149]**	[0.155]	[0.331]
Observations	110	110	110
Canton Fixed effects	NO	NO	YES
R-squared	0.01	0.13	0.72

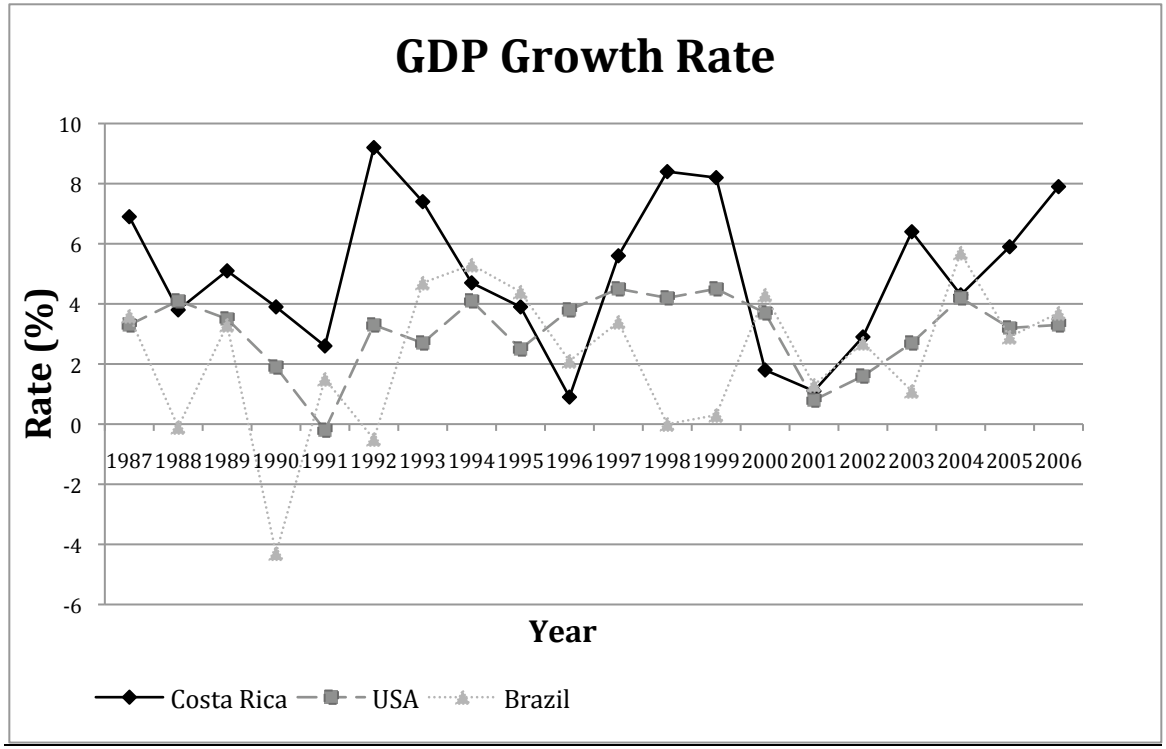
Robust standard errors in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%

HHEMI=Household Equivalent Monthly Income

Population weighted least-squares regressions.

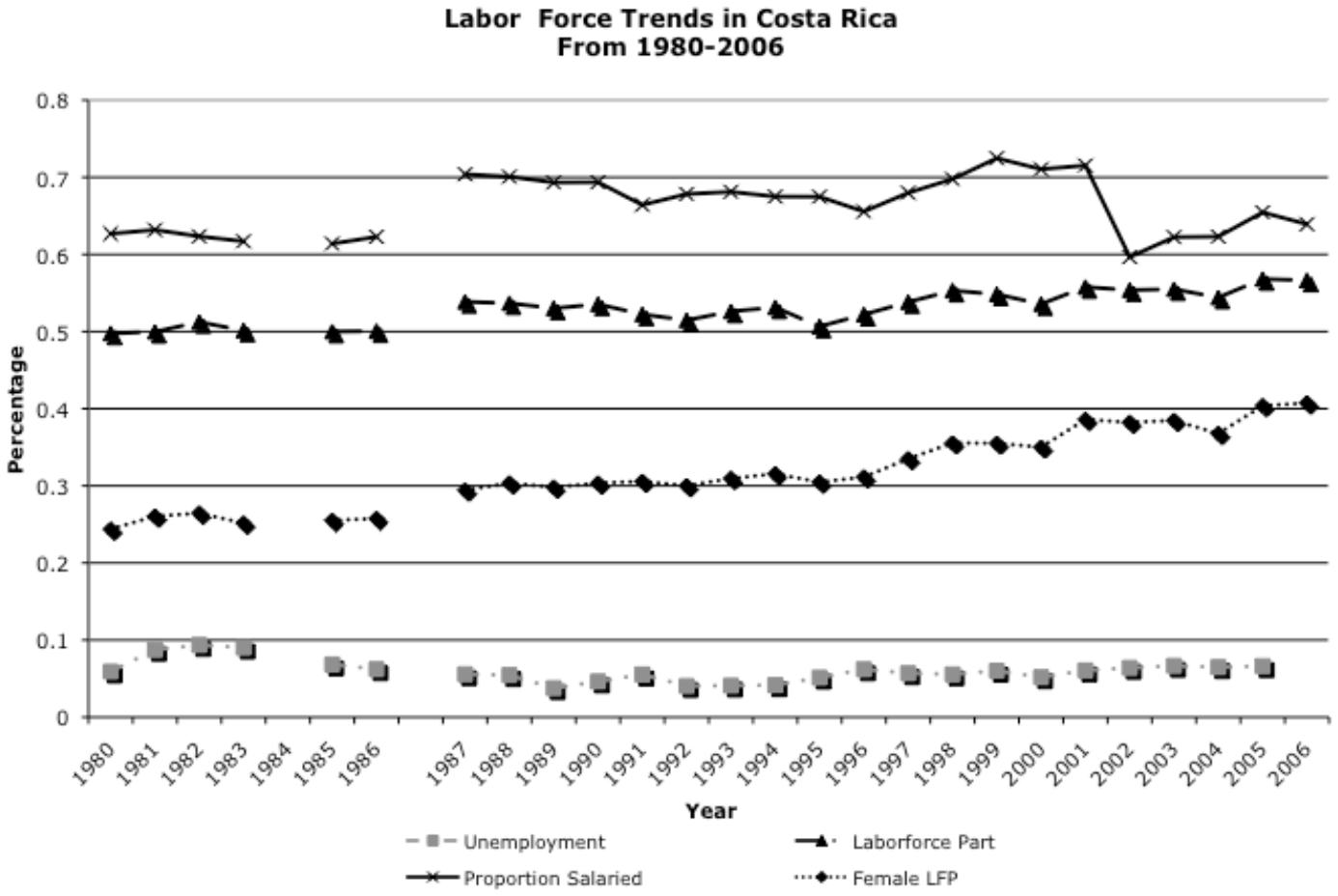
Figures:

Figure 2.1: GDP Growth Rate in Costa Rica from 1987-2006



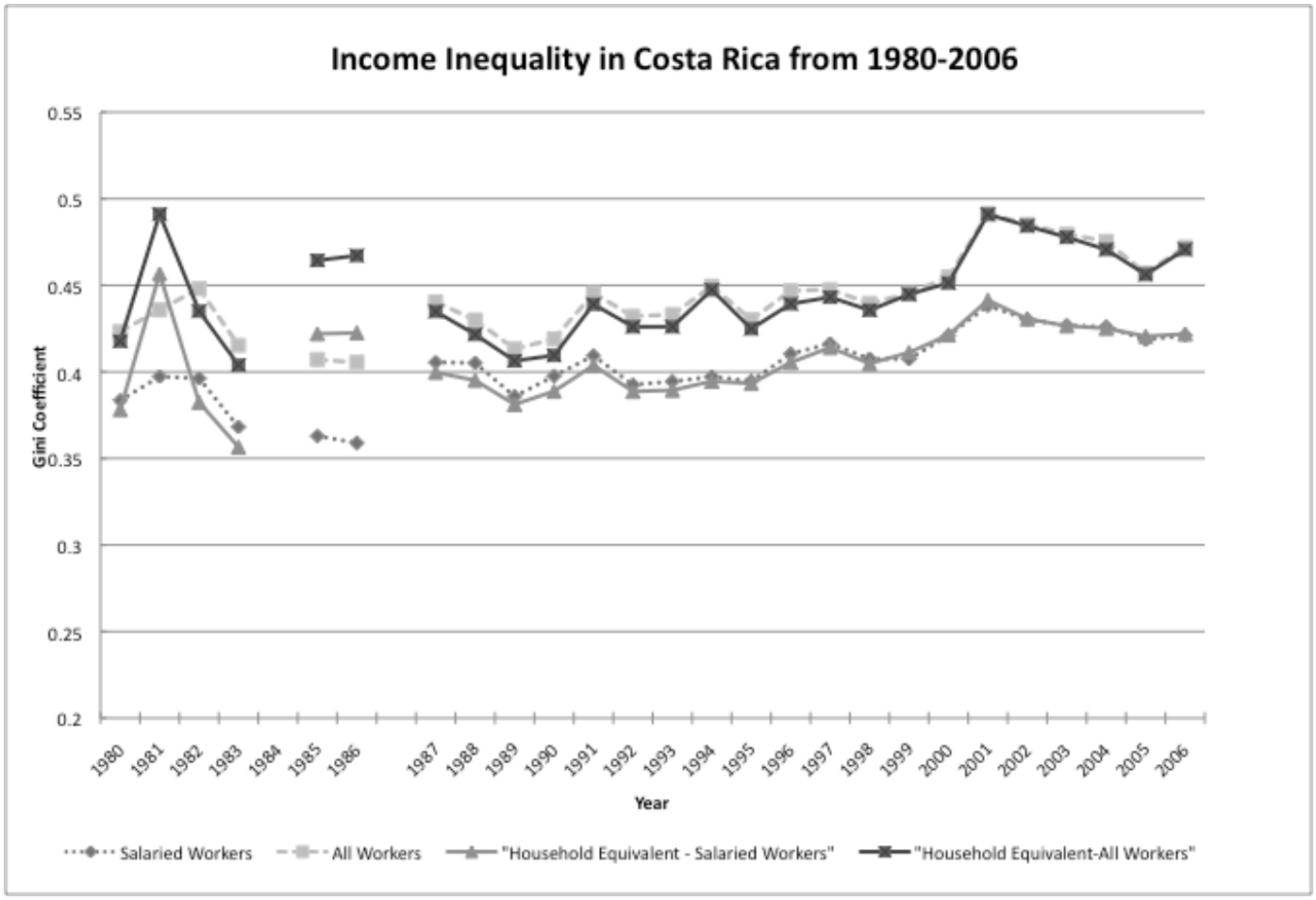
Source: World Bank Development Indicators. http://go.worldbank.org/1SF48T40L0

Figure 2.2: Labor Force Trends in Costa Rica from 1980-2006



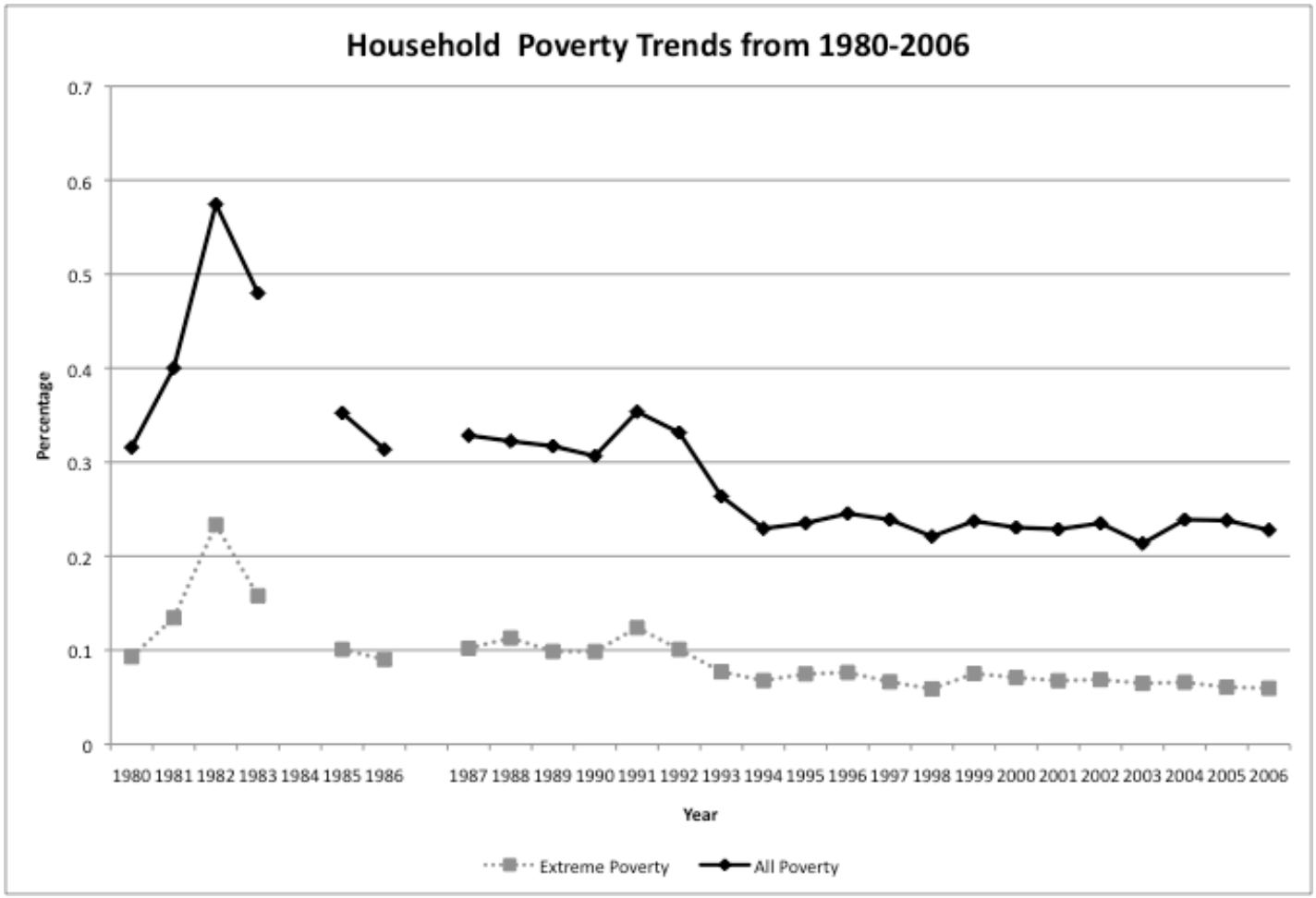
Source: Author's calculation based on Costa Rican Household Survey for Multiple Purposes. Employment characteristics are for individuals aged 12+.

Figure 2.3: Income Inequality Trends in Costa Rica from 1980-2006



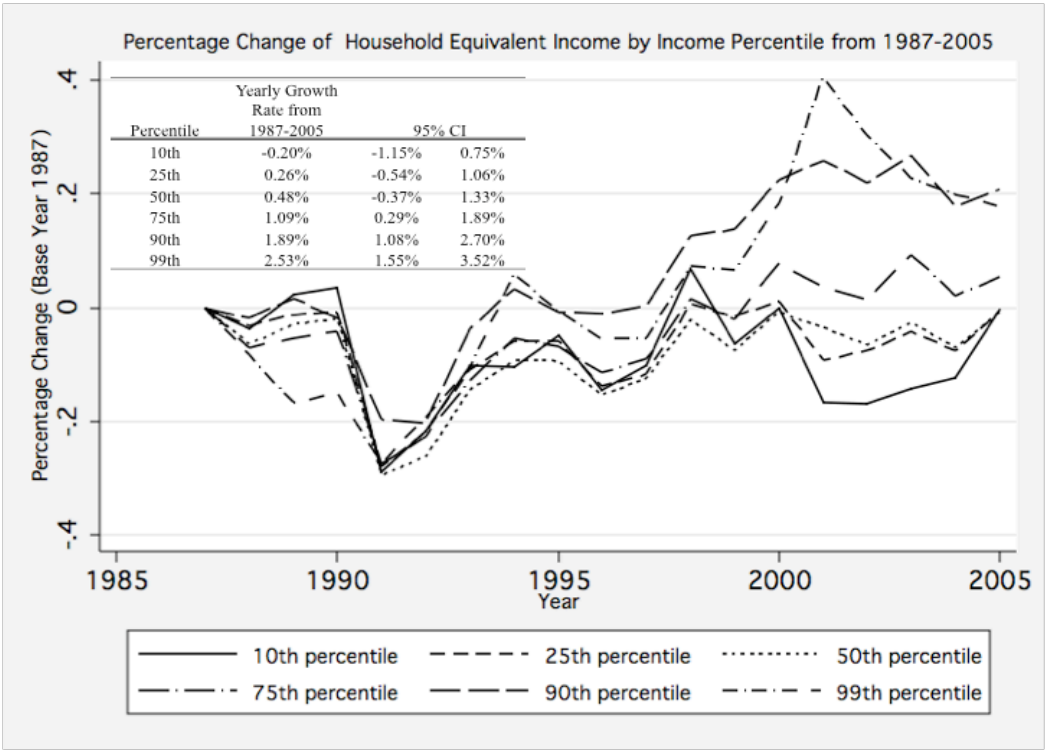
Source: Author's calculation based on Costa Rican Household Survey for Multiple Purposes.

Figure 2.4: Household Poverty Trends in Costa Rica from 1980-2006



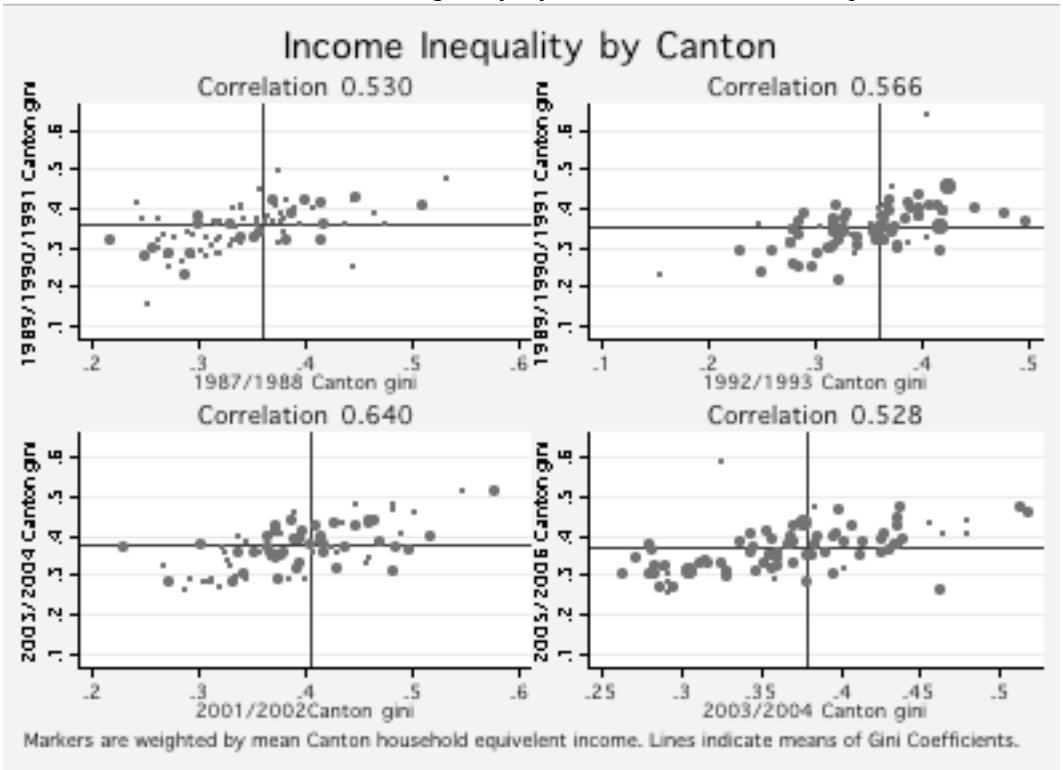
Source: Author's calculation based on Costa Rican Household Survey for Multiple Purposes.

Figure 2.5: Percentage change in Household Equivalent Income from 1987 to 2005 by Income Percentile



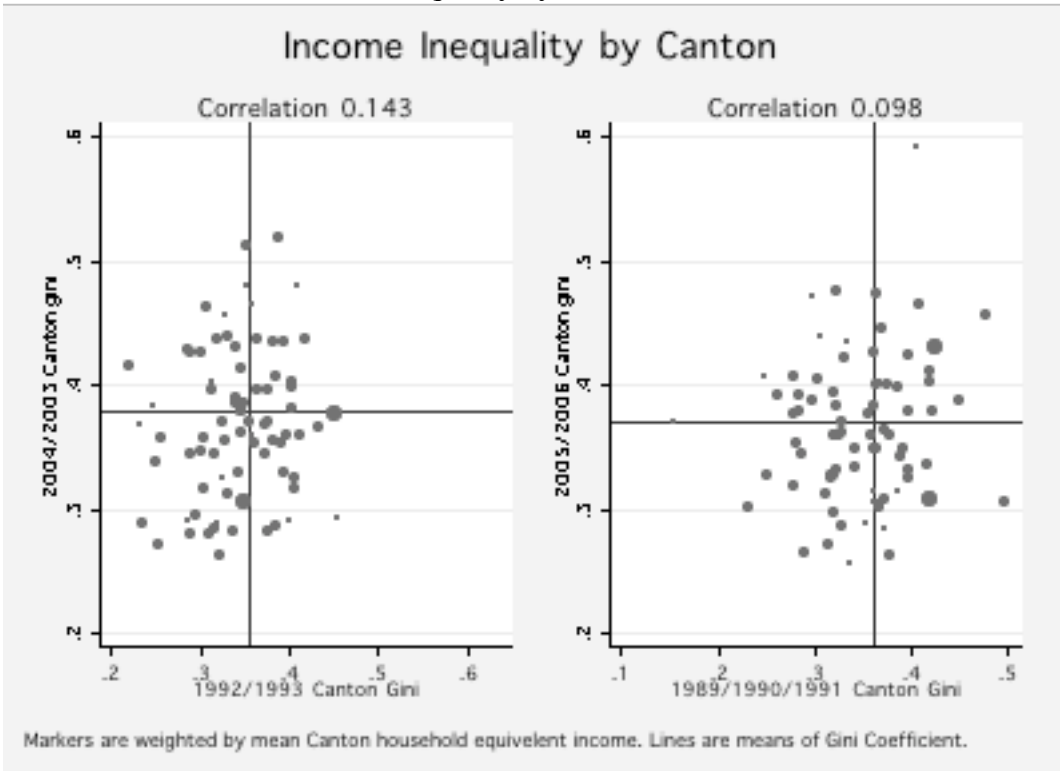
Source: Author's calculation based on Costa Rican Household Survey for Multiple Purposes and Costa Rican CPI deflators found at Banco Central De Costa Rica (http://indicadoreseconomicos.bccr.fi.cr). All colones adjusted to 2006 colones.

Figure 2.6: Scatter Plot of Income Inequality by Canton and Year – Adjacent Years



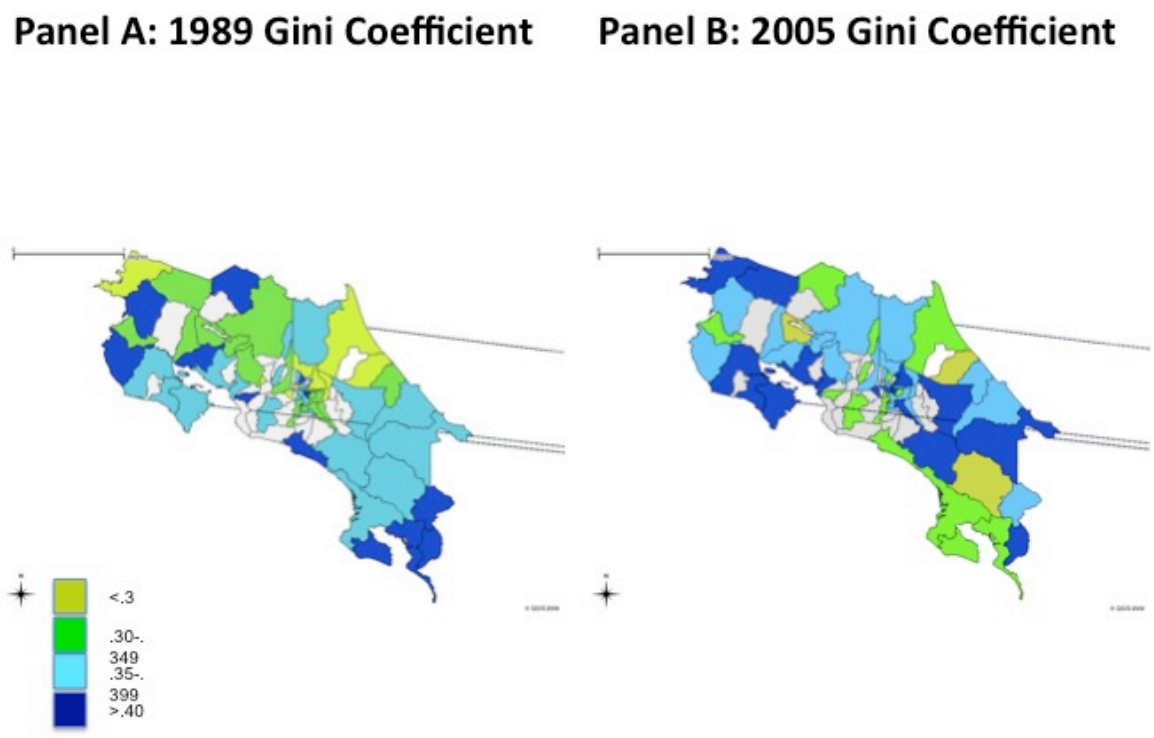
Source: Author's calculation based on Costa Rican Household Survey for Multiple Purposes.

Figure 2.7: Scatter Plot of Income Inequality by Canton and Year



Source: Author's calculation based on Costa Rican Household Survey for Multiple Purposes.

Figure 2.8: Map of Estimated Gini Coefficient in 1989 and 2005

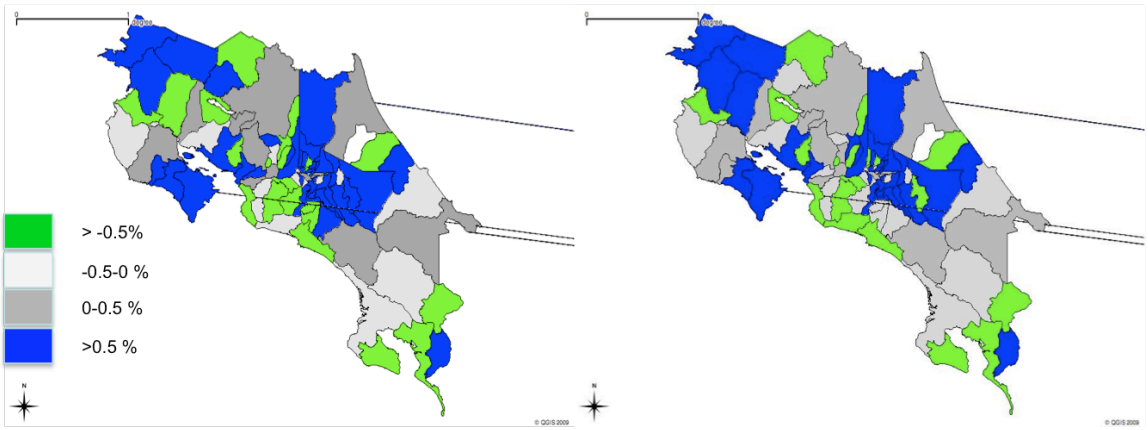


Source: Estimated Gini Coefficients are based on author's calculations from Costa Rican Household Survey for Multiple Purposes. Cantons in grey had less than 250 observations in 1989.

Figure 2.9: Map of estimated 5-year growth rates of Income Inequality from 1989-2005

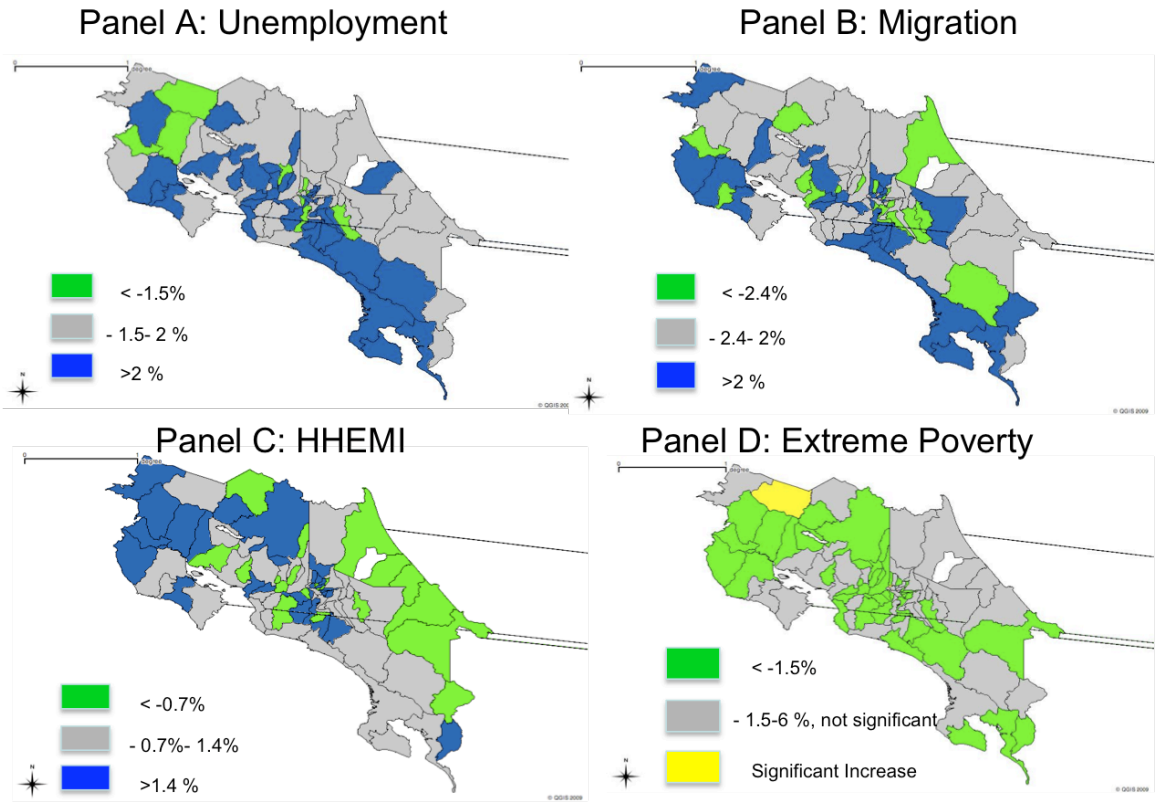
Panel A: Estimates using all data

Panel B: Estimates using Gini Coefficients made with >250 observations



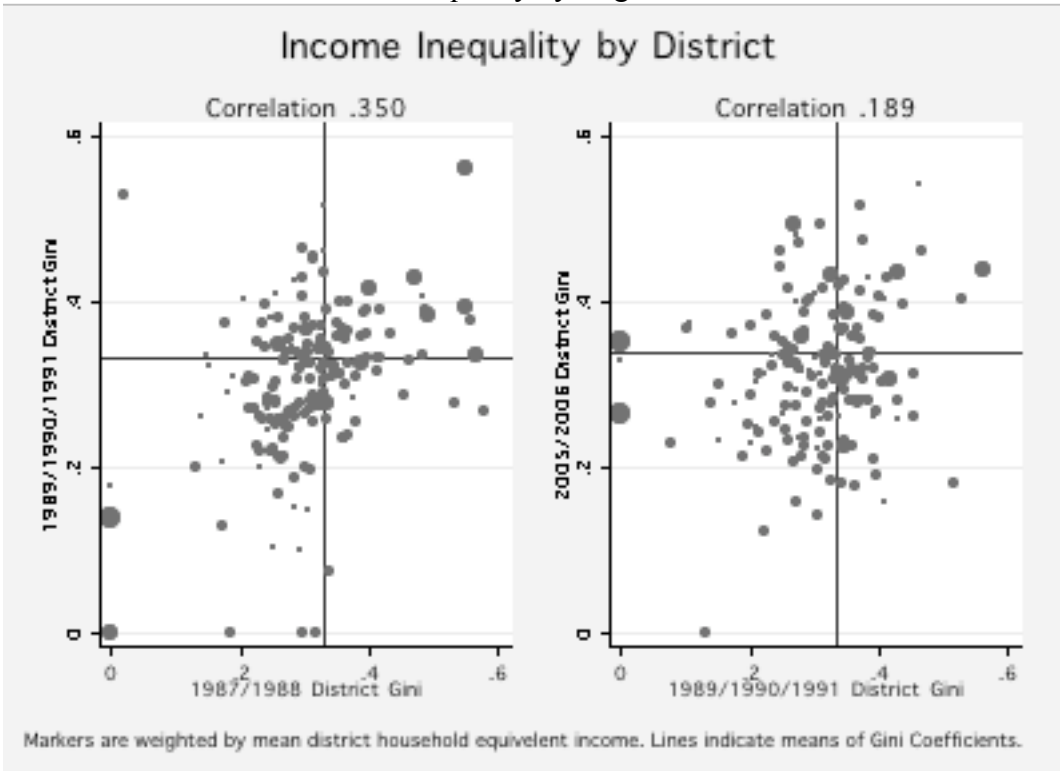
Source: Author's calculations from Costa Rican Household Survey for Multiple Purposes. For 4 time points a 5-year growth rate of the Gini Coefficient is estimated using a least squares method. Magnitudes of the growth rates are mapped. Not all growth rates are statistically significant. Panel A uses all the data while Panel B only uses cantons with more than 250 observations in any time period. The two are similar overall.

Figure 2.10: Map of estimated 5-year growth rates of Unemployment, In-Migration, Household Equivalent Monthly Income (HHEMI), and Extreme Poverty from 1989-2005



Source: Author's calculations from Costa Rican Household Survey for Multiple Purposes. Growth rates of the variables are estimated using a least squares method. Magnitudes of the growth rates are mapped. Not all growth rates are statistically significant, but cut-offs were chosen to reflect those changes that were often significant.

Figure 2.11: Scatter Plot of Income Inequality by Region and Year



Source: Author's calculation based on Costa Rican Household Survey for Multiple Purposes.

Chapter 3: Ecological Relationship Between Income Inequality and Mortality

3.1 Introduction

For over 30 years social scientist have noted the association between area-level inequality and health outcomes, especially mortality. Studies have examined this relationship across nations (Asafu-Adjaye, 2004; R. G. Wilkinson, 1992), states (Kaplan, Pamuk, Lynch, Cohen, & Balfour, 1996; Kennedy et al., 1996; J. Lynch et al., 2004b; S. V. Subramanian & Kawachi, 2003), and even counties (J. W. Lynch, Kaplan, Pamuk, Cohen, Heck, Balfour et al., 1998). Early studies often explored the cross-sectional association, while more recent studies have looked at the longitudinal relationship between income inequality and all-cause mortality to account for area-level differences. Yet, distinguishing between causes of death is particularly important as income inequality may exert its harmful effects through specific pathways that affect only certain causes of death. Few studies have carefully looked at cause-specific mortality at the sub-national level over time. One previous study did examine changes in inequality over time and its relation to cause-specific mortality (J. Lynch et al., 2004b). However that study was based in the United States and did not use multivariate methods in a setting where race, access to care, and other contextual variables are thought to confound the association.

Given this gap in the literature, this chapter will explore the area-level longitudinal association of income inequality and mortality rates for all-cause mortality, 4 grouped mortality categories and 18 cause-specific mortality rates in a middle-income country with near universal health insurance. Using the longitudinal dataset constructed in Chapter 2, the following analyses will focus on the association between changes in mortality rates and changes in income inequality in Costa Rica from 1989 to 2005. The cause of death analysis may be particularly useful for differentiating between the mechanisms presented in Chapter 1. The contribution of income inequality to different causes of death are explored, as they are expected to vary more for conditions that may respond more quickly to changes in economic conditions (such as cardiovascular diseases) than for conditions that are thought to be primarily genetic (such as cancer). This analysis also serves as a reference point for the final study in this dissertation.

3.2 Data

The data for these analyses come from three sources. The number of deaths came from the Vital Statistics Registry of Costa Rica. These data include canton of residence at time of death, date of death, age, gender, and the International Classification of Disease (ICD) codes for the cause of death. Accordingly, these data can be used to calculate death rates by canton, year of death, age, gender, and cause of death. The denominator for each canton, year, age, and gender cell comes from intercensal interpolations of population for years 1987-2006 from the 1984 and 2000 censuses. The area-level covariates were constructed from the July Household surveys as described in Chapter 2. The population projection data, Vital Statistics Mortality data, as well as the Household Surveys were all obtained from the Central American Population Center (CCP), which is a research center at the University of Costa Rica.

To construct the dependent variable, counts of deaths from the Vital Statistics Registry were organized into cells by year, canton, five-year age, gender, and cause of

death. The CCP classified cause of death into 24 categories by grouping ICD8, ICD9 & ICD10 codes together based on underlying cause of death. The cause of death groupings constructed by the CCP will be used throughout this chapter.[49] [50] Next I link these counts of deaths with the population projection for each canton-year-age-gender cells. In order to merge these files correctly, I recode age groups for those 75 and over so that they match across the two files. In particular, all the counts of death are combined for those aged 75 and over because the population data's last age category is 75 and over. All age groups below 15 years old were dropped and only adult mortality was assessed to enhance comparability with the individual analyses presented in Chapter 4. Next, I create canton-year-age-sex mortality rates for each cause of death by dividing the total number of deaths in each group by the population in that group. Finally, I further organize each cause of death into 4 broad categories: (1) infectious disease, (2) cancers, (3) lifestyle diseases, and (4) external causes. This categorization allows for the creation of 4 causes of death categories, in which the deaths in each category are aggregated. I also aggregate all deaths across all-causes to create a canton-year-age-sex all-cause mortality rate. These mortality rates are multiplied by one thousand to get a mortality rate per thousand. These values are subjected to a logarithmic transformation to create the mortality rate variable used in the regressions. Five causes of death categories are not included because these categories have either too few deaths (malnutrition, HIV and congenital disease), or they are associated with childbearing (perinatal mortality and maternal mortality), or their underlying cause is unspecified (residual mortality).[51] The other 18 causes of death account for 86% percent of all deaths for those aged 15-59 and 89% of deaths for those aged 60 and over in the years considered in the analysis. [52]

The area-level independent variables of interest are the same as those examined in Chapter 2: income inequality, mean household equivalent income, percent in extreme poverty, in-migration, and unemployment rate. These area-level measures are merged into the mortality data by canton and year. In order to obtain similar spacing between time periods and because of data limitations, I use only the 1989, 1995, 2001, and 2005 measures.[53] This allows for an easier interpretation of the subsequent results. Any association with these measures can be interpreted in the context of approximately 5-year change in each variable. I also estimate a longer-difference model with a 15-year time gap, in order to determine whether or not short-run changes in income inequality have different effects than long-term ones.

3.3 Methods

The main goal of this chapter is to test the strong form of the income inequality hypothesis. This hypothesis will be assessed separately for all-cause mortality, the four

[49] The CCP classifications were compared with published ones from the United States Center for Disease Control and they were very similar.

[50] A few deaths without medical certificates are weighted to account for under reporting. For example, of the over 15,000 deaths registered in 2005, only 85 are weighted and their weight is only 1.3.

[51] These causes are included in the all-cause mortality measure, but excluded from the 4 causes of death categories.

[52] Analyzing these data separately for those aged 60 and over allows us to compare the results with another nationally representative sample of Costa Ricans aged 60+.

[53] These years are also the years for which the necessary geocodes were available to construct the income inequality measures. Refer to Chapter 1 for detailed discussion.

cause categories, and for each specific cause of death by estimating the following model (Model 1):

$$M_{ctag} = \beta_0 + \beta_1 (I_{ct}) + \beta_2 (W_{ct}) + \boldsymbol{\beta_{3Z}} (\mathbf{Z}_{ct}) + \beta_4 \mathbf{Y_t} + \beta_5 \mathbf{A_a} + \beta_6 \mathbf{G_g} + \beta_7 \mathbf{A_a Y_t} + \boldsymbol{\beta_8 \gamma_c} + \varepsilon_{ctag}$$

where M_{ctag} is the mortality rate in canton *c* at year *t* for each gender *g* and 5-year age group *a*; I_{ct} is the income inequality in canton *c* at time *t*, W_{ct} is the mean income in canton *c* at time *t*; $\mathbf{Z}_{ct}$ is a vector of area-level controls (canton unemployment rate, a measure of the percent of households in extreme poverty, and percent migration into the area); $\mathbf{Y_t}$ is a series of dummies for the year; $\mathbf{A_a}$ is a series of dummies for each age group; $\mathbf{G_g}$ is a gender dummy; $\mathbf{A_a Y_t}$ is a series of interactions of all age and year dummies to account for time varying improvements to mortality rates for age-specific groups that vary beyond the general time and age trends (Lee, 2000), and $\boldsymbol{\gamma_c}$ is a vector of canton specific fixed effect.[54] Finally, the ε_{ctag} accounts for the remaining idiosyncratic error. The $\mathbf{Y_t}$ will account for time trends in area-level mortality or any specific yearly shock to mortality. The $\boldsymbol{\gamma_c}$ will account for any time-invariant unobserved area-level factors, such as the quality of the health clinic in the area, which could be correlated with both mortality rates and any of the other independent variables. The addition of these fixed effects will limit confounding as long as there are *no systematic* time invariant shocks. However, including these fixed effects substantially limits the variation used to estimate the inequality and health relationship, thus the model above are also examined without the fixed effects for comparison.

The canton-level controls, $\mathbf{Z}_{ct}$, are added to capture potential time varying shocks and canton-level characteristics that are correlated with both inequality and mortality. For example, areas with rapid economic growth may experience both increasing inequality and in-migration of healthy workers; thus, even if there were a real negative health impact from the growing inequality it may be suppressed by a positive effect of the influx of healthy people.[55] As a result, an in-migration variable, which measures the proportion of the population that has moved into a canton in the last two years, is included in Model 1. This variable may also serves as a proxy for economic activity in area.[56] Likewise, the unemployment rate variable is included in Model 1 because unemployment may be related to both income inequality and health outcomes (or health behaviors). For example, increases in unemployment may result in greater income inequality.[57] Meanwhile, those that are unemployed my also experience more depressed affect (Dooley et al., 1994) or adopt worse health behaviors (Strully, 2009) which could in turn affect area-level mortality. Finally, a poverty measure will be included to account for a non-linear relationship between income and health outcomes. For instance, if the income and health relationship is concave, then it may be that poverty is positively related to

[54] Any variable that is bolded is meant to be a vector and the corresponding β is also a vector.
[55] This is an important consideration in the period of this study because in Chapter 2 there was evidence that income inequality was associated with economic development.
[56] Increases in-migration are positively and significantly correlated with increases in mean income across these years. See Table 3.3 for details.
[57] In Costa Rica increases in unemployment rates variable are weakly but positively associated with increases in income inequality (see Table 3.3 for details).

mortality even after controlling for mean income. However, the poverty measure could also be acting as an additional inequality control, and so it will be both included and excluded from the models.

These controls are by no means exhaustive. Other time-varying area-level factors that are related to both income inequality and mortality, such as health services supply, transportation routes, or public health infrastructure, might still cause biases in the results. For example, if the concentration of health centers increased (decreased) in areas with increasing income inequality, then the results would have a negative (positive) bias since health services tend to be negatively related to mortality. In fact, healthcare supply did improve in terms of distance to the nearest health center due to health sector reforms from 1995 forward (Luis Rosero-Bixby, 2004), but as long as these time-varying improvements are not systematically correlated with changes in income inequality, the results will remain unbiased.

The outcome variable in Model 1, M_{ctag}, will be a log transformation of the raw mortality rates. To validate this functional form assumption, Figure 3.1 and 3.2 present the scatter plots between mortality rates and the Gini coefficients. Figure 3.1 shows the population weighted cross-sectional relationship between the logged mortality measure and the Gini coefficient for the 2005 data. Similarly, Figure 3.2 shows the population weighted relationship between changes in this logged mortality measure and changes in the Gini coefficient. These graphs indicate that an approximate linear relationship exists (if any relationship exists) between these logged mortality and the Gini coefficient within a canton. These graphs also preview the overall results for all-cause mortality.

Model 1 will be estimated separately for adults and the elderly as well as for all-cause and each cause-specific mortality rate. Standard errors will be clustered at the canton level to correct for arbitrary serial correlation.[58] For five regressions, the standard errors could not be clustered. The large number of fixed effects and the small magnitude of the rates caused the covariance matrices to be non-invertible for these regressions, and as a result only Huber-White robust standard errors are presented.[59] Each observation in these data is weighted by the total population in the canton in each year. This ensures generalizability of the results to the entire country and reduces heteroskedasticity. Finally, since the value of β_2 gives the estimated effect of a one-unit change in area income inequality on area-level mortality, this value is not directly interpretable because the Gini coefficient ranges from 0-1. Consequently, in the text the reported coefficient will be transformed in terms of an increase in the Gini coefficient by one-standard deviation, a 0.06 unit increase.

---

[58] I later test for auto-regressive structures in the residuals and find strong evidence for serial correlation. Bertrand, Duflo, and Mullainathan (Bertrand, Duflo, & Mullainathan, 2004) and Anderson (Anderson, 2008) suggest that clustered standard errors perform better than AR(1) type corrections especially if the number of cantons is much larger than the number of time periods and residuals are correlated within a given canton, but are independent across different cantons at given time period.

[59] For five regressions, the standard errors could not be clustered because the large number of fixed effects and the small rates caused the covariance matrices to be non-invertible. These five regressions are the ones with diarrhea, respiratory, cervical, breast and prostate cancer mortality rates for the adult sample as the dependent variable. For these five regressions, robust standard errors will be shown in the tables.

3.4 Results

Table 3.1 presents the means of mortality rates by year and cause per 1,000 persons in each canton-year-age-sex cell. Doing so makes the table interpretable and shows the average number of deaths per canton-year-age-sex cell (assuming similar sized cells). Panel A presents mortality rates for adults aged 15-59 and Panel B presents mortality rates for those aged 60 and over. Table 3.1 Panel A shows that all-cause mortality is stable during the study period, ranging between 3.9 to 3.2 deaths per 1,000 (in 1995 the death rate increases because of increases in external causes, lifestyle diseases and infectious diseases). For those aged 15-59, the most common causes of death are cardiovascular disease, other cancers, digestive cancer, motor accidents, and alcohol related liver disease. Table 3.1 Panel B also shows that all-cause mortality was decreasing during this time period for those aged 60 and over, from 105 to 78 deaths/1,000. In this age group mortality rates are obviously larger. The most common causes of death are cardiovascular disease, chronic respiratory disease, digestive cancers, other cancers, other accidents, diabetes, and prostate cancer. Note that the mortality rates for cardiovascular disease and digestive cancer are decreasing over this time period for both age groups. These are both causes of death that are affected by increased screening or preventative measures (Wallerstein, 2001) implying that primary health services may have been improving during this period.

Table 3.2 presents the summary statistics for the dependent variables. As in Chapter 2, the Gini is increasing slightly from 0.34 to 0.36 on average across the cantons. Meanwhile, poverty rates are decreasing substantially from 9.6 % to 6.2%. Mean area-level household equivalent monthly income (HHEMI) is also increasing with slightly increasing standard deviations (these measures are inflation corrected values). The other independent variables such as unemployment and in-migration are relatively stable over this period. These data cover 80 or 79 of the 81 cantons in Costa Rica.

To interpret how the independent variables are related to each other, the correlation coefficients between the area-level variables are examined. Table 3.3 Panel A presents the correlations of these variables at baseline in 1987. The Gini coefficient in 1987 is strongly related to the income level at the 90^{th} percentile of the income distribution. This suggests again that the variation at the top of the income distribution is correlated with inequality. Additionally, the results indicate that migrants were moving to areas that were more affluent.

Next the correlation coefficients between the within canton *changes* are examined since the specification in Model 1 uses the variation in changes in income inequality within cantons. Table 3.3 Panel B presents the correlation coefficients between the growth rates (or changes) in each of the time varying variables presented. This table shows that increases in the Gini coefficient are positively and significantly correlated with increases in HHEMI, but they are only weakly correlated to the other area-level variables. HHEMI is also positively and significantly related to in-migration, implying that areas experiencing economic growth are attracting new migrates. As expected, HHEMI is also negatively correlated with extreme poverty. Otherwise, the correlations between these *changes* are small.

3.4.1 All-Cause Mortality

The association between income inequality and all-cause mortality have the

broadest implications for population health. If income inequality is protective for some diseases and hazardous for others, all-cause mortality will weigh the relative importance of each of these relationships. Table 3.4 presents the association between all-cause mortality rate and income inequality. Columns 1 and 2 present the regressions without canton-level fixed effects while columns 3 and 4 present the regressions with fixed effects. In column 1 there is an imprecisely estimated small positive association between income inequality and all-cause mortality for those aged 15-59 without the fixed effects. The magnitude implies that a one standard deviation increase in income inequality is associated with a 1.04% increase in the all-cause mortality rate with confidence intervals ranging from -1.55% to 3.64%. Column 2 presents the same model for those aged 60 and over. The estimated magnitude suggests that a one standard deviation increase in income inequality is associated with a 0.45% decrease in the all-cause mortality rate with confidence intervals ranging from 2.97% to -3.88%. With the inclusion of canton-fixed effects, the estimated relationship is again small but negative for both age groups. The estimated magnitude for those aged 15-59 in models with the canton fixed effects suggests that that a one standard deviation increase in income inequality is associated with a -0.14 % decrease in all-cause mortality with confidence intervals ranging from 2.05% to -2.33%. The magnitude in column 4 for those aged 60 and over further suggests that a one standard deviation increase in income inequality is associated with a -0.40% decrease in the all-cause mortality rate with confidence intervals ranging from 3.3% to -4.15%. These essentially null results confirm the lack of relationship observed in Figure 3.2 where changes in canton level Gini coefficients were graphed against changes in log mortality rates.

In comparing the results for the other independent variable across models, the estimates indicate that in the models without the canton fixed effects, the area-level mean income, unemployment, migration and extreme poverty variables were significantly associated with mortality in both age groups. With the inclusion of canton-level fixed effects, only mean income nears significance in the hypothesized negative direction for the older sample (aged 60 and over). The inclusion of the fixed effects substantially reduces the point estimate for the unemployment rate and in-migration variables which suggests the there are other canton-level factors, such as economic development, that are associated with these labor force characteristics and health. Furthermore, while income inequality was not significantly related to all-cause mortality even in the cross-sectional case, mean are income was *positively* associated with mortality rate (opposite to the hypothesize direction).

All-cause mortality encapsulates many heterogeneous causes of death and so the lack of results may be masking a more subtle relationship between income inequality and other more specific causes of death. In order to better understand pathways through which inequality could affect health, the relationship between income inequality and 18 specific causes of death are explored. Table 3.5-3.8 presents these results.

3.4.2 Infectious Disease

The relationship between causes of death related to infections including Diarrhea, Tuberculosis, acute respiratory infections (Pneumonias, Influenzas), all other infections, and all infections (the previous 4 causes grouped together) are presented first. Taken together these causes account for 3.3% of all deaths for adults aged 15-59 and 5.4% for

those aged 60 and over. These diseases are generally easily treatable and so they may be related more directly to health services access as opposed to the psychosocial, resource allocation or neomaterial mechanism presented in Figure 1.1.

As hypothesized, income inequality was not related to these infectious causes of death in either age group. However, Table 3.5 does demonstrate a consistent association between unemployment rates and infectious disease mortality for both adults and the elderly. A one standard deviation increase in unemployment (0.02 unit change) is associated with a 0.15% decrease in Diarrhea mortality rates and a 0.28% decrease in Tuberculosis mortality rates for those aged 15-59. For the older sample, unemployment rates are associated with a 2.63% *decrease* in acute respiratory infection related mortalities. While these point estimates are small and two are only marginally significant, they are consistent with previous studies that area-level recessions are associated with short-term reductions in area-level mortality (D. Miller, Page, Stevens, & Filipski, 2009; Ruhm, 2000). Increasing unemployment may cause more people to stay at home and provide more care for other family members (especially the elderly), and also cause people to interact less and which limits the opportunity for infectious diseases to spread, thus reducing mortality from infectious diseases.

3.4.3 Cancer Mortality

Costa Rica has an exceptionally high rate of cancer mortality compared to other Central American countries, and its cancer rates are on par with those in the United States. Cancer related mortality accounts for 23% of all deaths for adults aged 15-59 and 24% of deaths for those 60 and over. The estimates in Table 3.6 show that income inequality is not significantly associated with adult cancer mortality, but is associated with breast cancer for the elderly. A one standard deviation increase in income inequality is associated with a 3.06 % *decrease* in breast cancer mortality rates for those aged 60 and over. Also a 15,000 colones increase (the average increase) in mean area income is associated with a 9.9% increased in breast cancer mortality. While a similar reverse gradient where women of higher SES have higher breast cancer rates has also been observed in the United States, these results based on 5 year changes in income are surprising, which suggests that there may be other factors, such as prevention measures, that are driving these results. Moreover, mortality from all other cancers is positively associated with percent in extreme poverty for those aged 60 and over, indicating that income plays a nuanced role in cancer mortality.

3.4.4 Lifestyle or Behavior Related

Next, the association between income inequality and causes of death often linked with behavioral choices are considered. This "Lifestyle" related mortality category includes chronic respiratory causes (linked with smoking), cardiovascular disease (linked with diet, exercise, and smoking), diabetes (linked with diet), liver disease (linked with alcohol consumption), and all these lifestyle diseases. These causes account for the bulk of all deaths, 30% of adult mortality (aged 15-59) and 56% of elderly mortality (aged 60+). These cause are also thought to be particularly sensitive to stress. To the extent that income inequality increases stress through invidious comparisons, leading directly to poorer health or health behavior as hypothesized in the biological pathway laid out in Figure 1.1, then a positive relationship between income inequality and mortality for these

cause of death is predicted.

Table 3.8 presents the relationship between income inequality and these lifestyle related causes. The analysis reveals a negative association between income inequality and cardiovascular mortality for adults aged 15-60. A one-standard deviation increase in income inequality is associated with a 1.69% *decrease* in cardiovascular mortality—the opposite of the hypothesized direction. In addition, an inverse SES gradient, where a 15,000 colones increase in mean area-level income is related to a 0.51% increase in the cardiovascular mortality rate was found. Likewise, some evidence for a healthy worker effect was also found where a one standard deviation increase in-migration is associated with a 1.23% decrease cardiovascular mortality rate.

Furthermore, for those aged 60 and over, mortality from all these lifestyle diseases together are negatively associated with increase in inequality. Looking at individual causes, only diabetes and chronic respiratory causes have a negative, though insignificant, coefficient. For cardiovascular mortality, there results demonstrate some support for the absolute income hypothesis, where a 15,000 colones increase in monthly income is marginally associated with 14.9 % decrease in cardiovascular mortality. In addition, a one standard deviation increase in in-migration is associated with a 6.9% decrease in diabetes mortality and areas with slower decreases in extreme poverty have lower diabetes mortality!

The results for lifestyle disease are indeed exactly the opposite of what was expected *a priori*. Yet, they are inline with another study in Costa Rica, where richer Costa Ricans are adopting poorer nutritional habits (Luis Rosero-Bixby & Dow, 2009b). Nevertheless, these outcomes deserve further scrutiny at the individual-level to ensure that they are not due to ecological fallacy.

3.4.4 External Causes

The final mortality category is external causes. The number of deaths in this category accounts for 30% of deaths for adult (aged 15-59) and 5% of deaths for the elderly (aged 60 and over). These injury-related causes are thought to be less associated with our biology, but strongly associated with our social condition and thus may help distinguish between pathways. Table 3.9 shows that suicide mortality is related to area income inequality. A one standard deviation increase in income inequality is associated with a 0.84% increase in suicide mortality rates for adults. For the elderly, increases in the unemployment rate and in-migration are associated with lower mortality from motor vehicles accidents. Here again lower economic activity (characterized by more unemployment and less migration) may lead to fewer people driving and thus fewer accidents.

3.5 Sensitivity Analysis

3.5.1 Pooled OLS Models and Longer-Term Changes Models

The models above show the effects of a "5-year change" in income inequality on mortality rates using only the limited variation in mortality rates and inequality within cantons. Thus, the above estimates could err on the conservative side, as these short-term change models may not capture the time dynamics of how income inequality affects

mortality correctly. Therefore, longer duration change models may be informative.

Accordingly, both the cross-sectional association of income inequality on mortality and a "15-year change" in income inequality and mortality are examined. The cross-sectional association examines the relationship using a pooled OLS models similar to Model 1 but the *without* canton-level fixed effects. This specification exploits both within and across canton variation in inequality. The "15-year change" models parallel Model 1 *with* canton fixed effects but include only data from 1989 and 2005. The beta coefficients on the Gini coefficient and the standard errors for the pooled OLS models, the "15-year change" models, and the "5-year change" models are presented in Table 3.10. Panel A displays the results for those aged 15-59. No significant association between income inequality and any of the cause specific mortality rates were found for the pooled models. For the 15 -year change models, cardiovascular disease and suicide mortality rates were associated with changes in income inequality. These results mirror the ones found for the 5-year change models. Here the magnitudes in absolute terms are larger for the longer -term change models as would be expected if the results were indeed valid and the significance for the suicide mortality result increases.

Table 3.10 Panel B presents these same three versions of Model 1, the pooled OLS, 15-year change" and "5-year change" models for those aged 60 and over. In the pooled OLS analyses, there is a positive association between income inequality and other infections and liver disease mortality. For the longer-term change models, income inequality is associated with all other infections again and chronic respiratory cause. Note, that neither of these associations persist in the 5-year change fixed effects regressions.

The estimated coefficients from the pooled OLS models are likely to be more biased than the longitudinal ones. In examining the estimated coefficients from the pooled models, both in-migration and unemployment rates were significantly related to mortality, but once the fixed effects were included these strong association often disappeared, as was also shown in the all–cause mortality model in Table 3.4 (the same as the pooled OLS model). This suggests that the inclusion of the fixed effects account for area-level characteristics associated with the broader labor market factors, implying that these controls may be are necessary to limit confounding. For example, if increases in income inequality are associated with economic development (as was suggested in Chapter 2) and economic development is negatively associated with suicide rates (as is suggested by the negative coefficient on mean HHEMI and suicide rates), then the longitudinal model are likely to be negatively biased for this cause.

3.5.2 Model Specification

In estimating the above models, several assumptions about functional form were made that merit further inspection. To build confidence in the models and subsequent results, the specification and functional form assumptions where examined to determine if the overall results changed with respect to income inequality.

3.5.2.1 Exclusion of Extreme Poverty

Some have suggested that it is not sufficient to control for just mean income in an

area when examining the income inequality hypothesis and that it is also necessary to control for other non-linearities in income (A. Deaton, 2003). In the analysis above, a control variable for percentage in an area in extreme poverty is included as the non-linear income control. However, this variable may be unnecessary since the model already controls for mean area income. Moreover, its inclusion might be biasing the results or causing the income inequality effect to disappear. In the Costa Rican setting, the inclusion of this poverty measure is also unlikely to change the results because the upper end of the income distribution is driving the change in the Gini coefficients. Nonetheless, this is examined empirically.

Accordingly, the "5-year" change regressions are estimated again now excluding the control for extreme poverty (results not shown). In doing so, only one of the inequality results changed. For those aged 15-59, the Gini coefficient becomes significantly *protective* for respiratory cancer. Here the standard errors do not change across the two models, but the magnitude of the estimate becomes more negative.

3.5.2.2 Poisson and Negative Binomial Models

Next, the functional form assumptions are considered. Since the underlying data used to construct the mortality rate measure are based on counts of deaths, Poisson and Negative Binomial models are employed. While the results are not shown, but they are largely consistent with what was found above. With the Poisson model, the number of deaths from respiratory cancer [60] and suicide[61] was associated with changes in income inequality for those under 60. For those over 60, breast cancer mortality was again associated with changes in the Gini coefficient.[62] These regressions were also estimated using a negative binomial model, but many of the models did not converge. Where they did, the suicide and breast cancer results still appeared significant and magnitudes were the same as in the Poisson models.

Table 3.11 presents the transformed parameter estimates for each of the analyses using the 5-year model. This allows us to interpret the transformed coefficients as the change in mortality rate bases on a one-standard deviation change in the Gini coefficient (the raw estimates are multiplied by 0.06). This table also summarized the overall results and notes significance across the models without poverty and using a Poisson distribution. The most robust results are for suicide, cardiovascular disease mortality, and respiratory cancer for those aged 15-59 and breast cancer mortality for those aged 60 and over.

3.5.3 Measurement Error

To better understand whether measurement error is affecting the results, the models are again estimated with random error added to the inequality measures. As discussed in Chapter 2, the number of observations over which the inequality measures were created caused some concern for measurement error in the inequality measures.

---

[60] A one standard deviation increase is income inequality associated with 22% decrease in the *number* of respiratory deaths and this is significant at the 5% level.

[61] A one standard deviation increase in income inequality is associated with 9% increase in the *number* of suicide deaths and this is significant at the 10% level.

[62] A one standard deviation increase in income inequality is associated with 20% decrease in the *number* of breast cancer deaths and this is significant at the 10% level.

Using the distributions in Table 2.7, random noise was added to the inequality measures bases on the estimated standard error of the measures according to the number of observations used to create the Gini measures. In this procedure, a normal random sample is drawn with mean zero and standard error according to Table 2.7 (for cantons with fewer than 100 observations the mean =0 and std=0.037, for cantons with fewer than 150 the mean= 0 and the std=0.029, for cantons with fewer than 200 the mean=0 and std=0.024, for cantons with fewer than 250 the mean= 0 and std=0.019, and for cantons with fewer than 300 the mean= 0 and std=0.018). From these distributions, noise is added to each inequality measure in cantons where fewer than 300 observations were used in constructing the Gini measures, and the models for suicide and cardio vascular mortality are estimated 50 times for those aged 15-59. The average of the estimated coefficient on the income inequality variable is then obtained. This procedure provides estimated bounds for the attenuation bias.

The results of this procedure indicate that attenuation bias is likely to be small. Comparing the estimates, the average coefficient on the income inequality variable for suicide mortality was 0.134 (SE=0.0126) with the addition of noise versus 0.140 that was estimated before. The average coefficient on the income inequality variable for cardiovascular mortality was -0.270 (SE=0.025) with noise versus -0.280 that was estimated before. Thus, the attenuation bias due to measurement of income inequality due to small samples is likely to be small (on the order of 4-5%).[63]

3.6 Discussion

Generally, the results provide limited evidence for the strong form of the income inequality hypothesis. Even so, several results stand out and deserve further attention.

First, regardless of the model specification, no association between income inequality and all-cause mortality was observed.[64] In light of previous studies in China and Chile that find a relationship between income inequality and self-reported health, the null results with respect to all-cause mortality are especially interesting. One reason for this difference is that income inequality appears more closely related to perceptions of health than actualizations of health (Deaton, 2003). Next, important differences in these settings may account for the differences in results. In the case of Chile, income distribution may affect health through political channels. Chile's development strategy from 1970-1990 focused on privatization and decentralization of public services (Mesa-Lago, 2000), which may explain the differences. With regard to China, exceptionally fast growth on the order of 10% per year for nearly 15 years accompanied many other social changes, which may be driving the income inequality and health relationship there.

Secondly, in the Costa Rican context the analyses indicate that for those over 60 years old breast cancer mortality is negatively associated with income inequality. An increase of one standard deviation in income inequality was associated with a 3%

[63] Note that these are weighted regressions and so cantons with smaller populations (and observations) are given less weight.

[64] In fact, two other specifications where also examined for all-cause mortality (not presented, but available). Following Li &Zhu (2006), a quadratic inequality measure is included in one specification and a 5-year lagged Gini measure is examined in another. In the first specification, neither the level nor the quadratic term was statistically significant. In the second model, the Gini lagged term was not statistically significant.

decrease in breast cancer mortality. While this result has the largest magnitude of all the associations found with income inequality, it may be biased. In another nationally representative sample of Costa Ricans aged 60 and over, a clear income gradient was found for those obtaining mammograms within the previous 3 years (Brenes-Camacho & Rosero-Bixby, 2009). If richer Costa Ricans living in areas that are more unequal seek more mammograms, and if breast cancer survival rates are strongly associated with early detection, then this result may be biased. It may not be income inequality *per se,* rather poorer detection in more equal areas could explain the association found.

Third, higher suicide mortality was observed in areas of increased income inequality. This finding is limited to those aged 15-59, and as such it is particularly interesting. The most obvious mechanism is one of mental distress similar to that hypothesized to function through the negative psychological effects of invidious comparisons. These effects may be further enhanced by the breakdown of social capital buffering the emotional states of those who are already unstable. This points to a much more nuanced definition of social vulnerability than has been previously used in this literature. Here, vulnerability is not distinguished by age (Muramatsu, 2003), but rather by sensitivity to inequality or by previous psychiatric problems. This result is somewhat consistent with the income inequality hypothesis and the psychosocial mechanism. Furthermore, similar results have previously been found in the United States for younger adults (J. R. Miller, Piper, Ahern, Tracy, & Tardiff, 2005) and more generally (J. Lynch et al., 2004b). While this is not a particularly novel finding, it merits further exploration in the context of Costa Rica. This cause of death has much more to do with psychology than physiology, and thus it points to behavioral changes at the extremes rather than changes in underlying biological processes.

Fourth, the results of the analyses imply that income inequality is protective for cardiovascular disease, which is quite counter-intuitive. While in similar analyses in the United States results have been found suggesting that income inequality is protective for all-cause mortality (Angus Deaton & Lubotsky, 2003), all these results counter our initial assumptions. If the results of the Deaton & Lubotsky study were confounded by access to health services or by differences in health behaviors across races, then the results in the Costa Rican setting are more illuminating. However, the current results may also be confounded by the nutritional transition in which the richer are adopting the poor eating habits of the United States. The results of the analysis in this chapter will be further considered in the individual-level analyses in Chapter 4.

Finally, when extreme poverty was omitted from the model, areas with increasing income inequality were found to have lower mortality from respiratory cancers. Again, since respiratory cancers are often caused by smoking, which we would have expected to increase in high income inequality areas *a priori*, this result again counters our initial hypothesis. Some of the changes could be explained if smoking is decreasing faster in areas that have higher inequality. For example, if richer Costa Ricans (who live in more unequal areas) were reducing their smoking rates, this would be consistent with these results. Yet, this is purely a conjecture.

3.7 Limitations

While this chapter of the this thesis has the advantage of differentiating between multiple causes of death, using the Vital Statistics Registry of Costa Rica, and less prone

to confounding, its ecological design is somewhat concerning. Since the relationship between individual income and health status is thought to be concave such that each additional income brings about a decreasing amount of health, the aggregate-level relation between income distribution and health may be the result of this functional form. In the analysis, controls for extreme poverty are included and excluded to capture the functional relationship between income and health. While this does not change the main results of the analysis, the relationship may be much more complicated and thus neither approach is capturing the true underlying relationship. Moreover, if other community-level time varying confounders are omitted from the model, aggregation may lead to biases in the regression coefficients that reflect both omitted variables and compositional effects.

Another limitation is that the weak form of the income inequality hypothesis could not be tested with these data. If indeed certain groups (such as the rich) thrive in unequal areas while another group (the poor) suffer then these two effects could be counter acting each other and thus the association with income inequality at this aggregate level would be zero where as the underlying relationship may be negative or positive depending on the group of interest. Many speculate that the strong form of the income inequality hypothesis is unlikely to be valid, but that certain vulnerable groups suffer more and to that extent, it is more interesting to look at the weak form in detail.

Another key problem is that there may be some measurement error in income inequality measures. To the extent that measurement error was related to the number of observations used to create these measures per canton, error was simulated and added to the inequality measures. The results of these simulations suggest that the attenuation bias is small.

Another problem in interpreting the standard errors in these regressions is related to serial correlation. While there was evidence of serial correlation in the errors across time periods, it does not make the coefficient estimates inconsistent, but it does affect the standard errors, making inference difficult. As mentioned above the standard errors are corrected for arbitrary correlation using the STATA cluster function. To the extent that observations within a given canton are correlated over time, this method should correct the standard errors. However, if there is dependence between cantons within a given year or if there were some sort of geographical clustering (beyond canton), then the standard errors may still be wrong. In this case, the possibility of geographical clustering over time is more concerning than the dependence across cantons within a given year.

Finally, cause of death coding is always a problem in these types of analyses. While the ICD8, ICD9 and ICD10 codes were correctly matched to the cause categorized by CCP, the problem may be even deeper as the underlying causes of death are notoriously poorly enumerated on death certificates. Mortality rates are based on underlying cause of death as entered on a death certificate by a physician. Incorrect coding of the cause of death may still occur since some individuals may have had competing causes of death and often autopsies are not done. When the physician enters multiple causes of death, ICD provides standard rules to ensure consistent reporting but some misclassification might still occur. Considering that Costa Rica's mortality registry is thought to be one of the best in the developing world and that this is the best available

data, this problem is difficult to address.[65]

[65] Costa Rica's Death Registry has been rated complete, the highest rating possible, by United Nations since 1961.

Table 3.1: Weighted Mean Mortality Rate by Year and Cause

Panel A: Aged 15-59				
	1989	**1995**	**2001**	**2005**
All-Cause Mortality	3.48	3.90	3.46	3.20
Infectious Disease	0.115	0.174	0.100	0.111
Diarrhea	0.004	0.014	0.009	0.012
Tuberculosis	0.015	0.026	0.016	0.020
Acute Respiratory (Infection)	0.029	0.065	0.039	0.047
All other Infections	0.067	0.069	0.036	0.031
Cancers	0.736	0.711	0.679	0.695
Cancer of Digestive System	0.229	0.214	0.189	0.081
Cancer of Respiratory System	0.034	0.035	0.040	0.052
Cervical Cancer	0.072	0.079	0.061	0.044
Breast Cancer	0.062	0.081	0.061	0.062
Prostate Cancer	0.003	0.004	0.004	0.005
Other Cancers	0.337	0.298	0.324	0.451
Lifestyle Disease	0.961	1.021	0.898	0.817
Chronic Respiratory Causes	0.064	0.099	0.074	0.067
Cardio Vascular Disease	0.609	0.606	0.542	0.475
Diabetes	0.073	0.077	0.072	0.095
Alcohol/Liver Disease	0.215	0.239	0.210	0.180
External Cause	1.20	1.32	1.22	1.07
Motor Vehicle Accident	0.443	0.454	0.537	0.335
Other Accidents	0.333	0.266	0.210	0.204
Suicide	0.180	0.236	0.160	0.192
Homicide	0.243	0.360	0.313	0.343

Table 3.1 cont.: Mean Mortality Rate by Year and Cause- Aged 60 and over

Panel B: Aged 60 and over	**1989**	**1995**	**2001**	**2005**
All-Cause Mortality	104.8	102.5	88.9	78.2
Infectious Disease	6.03	8.60	4.68	4.21
Diarrhea	0.604	1.193	0.593	0.638
Tuberculosis	0.626	0.458	0.233	0.180
Acute Respiratory (Infection)	3.80	5.89	3.34	2.91
All other Infections	0.992	1.054	0.520	0.484
Cancers	23.79	23.39	20.82	18.85
Cancer of Digestive System	11.31	10.08	8.72	2.92
Cancer of Respiratory System	2.81	2.11	1.88	1.71
Cervical Cancer	0.839	1.040	0.925	0.853
Breast Cancer	1.11	1.21	1.36	0.98
Prostate Cancer	1.90	2.05	2.14	1.87
Other Cancers	5.81	6.92	5.80	10.51
Lifestyle Disease	58.67	54.98	50.13	41.14
Chronic Respiratory Causes	9.93	8.35	7.72	6.05
Cardio Vascular Disease	42.61	40.58	36.69	28.99
Diabetes	4.02	3.84	3.82	3.84
Alcohol/Liver Disease	2.11	2.20	1.90	2.26
External Cause	5.15	5.06	3.68	3.64
Motor Vehicle Accident	1.115	1.099	0.692	0.518
Other Accidents	3.51	3.34	2.54	2.50
Suicide	0.173	0.172	0.169	0.159
Homicide	0.356	0.447	0.279	0.458

Notes: Author's calculations based on Vital Statistics Mortality data. Panel A present rates for those aged 15-59 and Panel B presents rates for those aged 60 and over. Including all deaths over all age groups generates results very similar to published death rates.

Table 3.2: Summary Statistics of Independent Variables

	Year	Obs	Weighted Mean	Std. Dev.
Gini Coefficient	1989	80	0.354	0.049
Gini Coefficient	1995	80	0.362	0.047
Gini Coefficient	2001	79	0.401	0.056
Gini Coefficient	2005	79	0.368	0.045
(Mean HHEMI)/1000 in 2006 Colones	1989	80	189.11	54.30
(Mean HHEMI)/1000 in 2006 Colones	1995	80	169.14	55.43
(Mean HHEMI)/1000 in 2006 Colones	2001	79	209.27	74.93
(Mean HHEMI)/1000 in 2006 Colones	2005	79	201.60	71.20
Extreme Poverty	1989	80	0.096	0.088
Extreme Poverty	1995	80	0.065	0.060
Extreme Poverty	2001	79	0.063	0.066
Extreme Poverty	2005	79	0.051	0.051
Migration	1989	80	0.042	0.023
Migration	1995	80	0.047	0.025
Migration	2001	79	0.043	0.026
Migration	2005	79	0.051	0.026
Unemployment Rate	1989	80	0.046	0.018
Unemployment Rate	1995	80	0.057	0.022
Unemployment Rate	2001	79	0.066	0.027
Unemployment Rate	2005	79	0.062	0.019

Notes: Means weighted by canton population. All statistics calculated for over those aged 12 and over as has been done in other studies of Costa Rica.

Table 3.3: Correlation Coefficients of Main Independent Variables

Panel A: Correlation Coefficients in 1987	Gini Coefficient	Household Equivalent Mean Income (HHEIM)	Household Equivalent Income (90th percentile)	Extreme Poverty	Migration
Gini Coefficient	1				
HHEIM	0.1827	1			
	[0.1047]				
Household Equivalent Income (90th percentile)	**0.2969**	**0.9524**	1		
	[0.0075]**	**[0.000]****			
Extreme Poverty	0.1488	**-0.6273**	**-0.5091**	1	
	[0.1876]	**[0.000]****	**[0.000]****		
Migration	-0.0904	**0.5635**	**0.543**	**-0.3631**	1
	[0.4253]	**[0.000]****	**[0.000]****	**[0.0009]****	
Unemployment	0.1043	0.026	0.0624	-0.0206	0.0573
	[0.3572]	[0.8186]	[0.5827]	[0.856]	0.6137

Table 3.3 cont.: Correlation Coefficients of Main Independent Variables

Panel B: Correlation Coefficients of Growth Rates	Gini Coefficient	Household Equivalent Mean Income (HHEIM)	Extreme Poverty	Migration	Unemployment
Gini Coefficient					
HHEMI	**0.394**				
	[0.0004]**				
Extreme	0.169				
Poverty		**-0.314**	1		
	[0.148]	**[0.0062]****			
Migration	-0.040	**0.271**	-0.042	1	
	[0.728]	**[0.017]***	[0.722]		
Unemployment	0.0398	0.037	0.029	0.080	1
	[0.731]	[0.750]	[0.807]	[0.492]	

Note: Growth rates for each time varying variable are calculated using a least-squares method to get at secular changes by canton using 1989,1996, 2001, 2005 data. Correlation coefficient are presented along with the corresponding p-values. * significant at 5%; ** significant at 1%

Table 3.4: Association of Inequality and All-cause Mortality Rates

	Ln(All-Cause Mortality)			
	Without Canton Fixed effects		With Canton Fixed effects	
	Aged 15-59	Aged 60 and over	Aged 15-59	Aged 60 and over
Gini Coefficient	0.174	-0.076	-0.024	-0.066
	[0.217]	[0.286]	[0.183]	[0.313]
(Mean HHEMI)/100,000	0.315	0.779	0.001	-0.525
	[0.232]	[0.325]**	[0.031]	[0.388]
Unemployment Rate	2.071	2.533	0.298	-0.089
	[0.624]***	[0.711]***	[0.45]	[0.699]
Migration	-1.265	-3.063	-0.360	-0.454
	[0.519]**	[0.69]***	[0.443]	[0.622]
Extreme Poverty	-0.554	-1.562	-0.336	-0.163
	[0.197]***	[0.309]***	[0.254]	[0.3]
Constant	1.354	4.149	0.215	2.378
	[0.103]***	[0.113]***	[0.088]**	[0.126]***
Observations	5724	2544	5724	2544
R-squared	0.6	0.72	0.63	0.75

Robust and clustered standard errors in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: Controls for age (5 year brackets), gender, year fixed effects and age and year indicator interactions are included. Observations are weighted by canton population.

The unit of observation is canton, year, gender, 5-year age group. For those aged 15-59, there are 9 age group, 2 gender groups, 4 years and either 79 or 80 cantons. 2(79*2*9+80*2*9)=5724. For those aged 60 and over, there are 4 age groups, 2 gender groups, 4 year groups and either 79 or 80 cantons. 2(79*2*9+80*2*9)=2544

Table 3.5: Association of Inequality and Infectious Disease Mortality Rates

	Diarrhea	**Tuberculosis**	**Acute Respiratory (Infection)**	**All other Infections**	**All Infections**
	Panel A: Aged 15-59				
Gini Coefficient	0.027	-0.012	0.045	0.020	0.006
	[0.022]	[0.04]	[0.048]	[0.045]	[0.008]
(Mean HHEMI)/100,000	-0.001	-0.002	0.0001	**0.009**	0.074
	[0.002]	[0.006]	[0.006]	**[0.005]***	[0.065]
Unemployment Rate	**-0.076**	**-0.141**	-0.013	-0.076	**-0.274**
	[0.043]*	**[0.052]*****	[0.07]	[0.15]	**[0.163]***
Migration	-0.002	-0.033	-0.030	-0.062	-0.123
	[0.035]	[0.074]	[0.082]	[0.095]	[0.138]
Extreme Poverty	**0.032**	-0.024	-0.030	-0.020	-0.042
	[0.016]**	[0.022]	[0.03]	[0.044]	[0.056]
Constant	-0.006	0.015	-0.010	-0.017	-0.018
	[0.006]	[0.014]	[0.018]	[0.02]	[0.028]
Observations	5724	5724	5724	5724	5724
R-squared	0.04	0.06	0.08	0.05	0.12
	&&				
	Panel B: Aged 60+				
Gini Coefficient	0.025	0.189	0.022	0.156	-0.058
	[0.187]	[0.162]	[0.326]	[0.208]	[0.049]
(Mean HHEMI)/100,000	-0.248	0.406	-0.652	-0.384	0.245
	[0.413]	[0.248]	[0.434]	[0.293]	[0.332]
Unemployment Rate	0.028	-0.096	**-1.419**	-0.444	**-1.313**
	[0.328]	[0.377]	**[0.682]****	[0.395]	**[0.748]***
Migration	-0.144	0.145	0.170	-0.172	0.091
	[0.29]	[0.321]	[0.64]	[0.36]	[0.646]
Extreme Poverty	0.066	0.142	-0.231	-0.148	-0.143
	[0.147]	[0.158]	[0.269]	[0.147]	[0.316]
Constant	0.069	-0.101	0.225	0.088	0.249
	[0.078]	[0.066]	[0.125]*	[0.072]	[0.121]**
Observations	2544	2544	2544	2544	2544
R-squared	0.18	0.12	0.53	0.17	0.52

Note: Robust and clustered standard errors in brackets. Controlling for age (5-year brackets), sex, canton fixed effects, year fixed effects, canton fixed effects, and age*time interactions. Observations are weighted by canton population.
* significant at 10%; ** significant at 5%; *** significant at 1%
&& indicates that standard errors are NOT clustered and only robust standard errors are reported

Table 3.6: Association of Inequality and Cancer Related Mortality Rates

	Cancer of Digestive System	Cancer of Respiratory System	Cervical Cancer	Breast Cancer	Prostate Cancer	Other Cancers	All Cancers
	Panel A: Aged 15-59						
Gini Coefficient	-0.073	-0.085	-0.008	0.017	0.017	0.077	0.015
	[0.102]	[0.056]	[0.058]	[0.067]	[0.025]	[0.101]	[0.023]
(Mean HHEMI)/100,000	0.008	-0.008	0.007	-0.006	0.003	0.015	-0.083
	[0.014]	[0.009]	[0.009]	[0.01]	[0.003]	[0.018]	[0.133]
Unemployment Rate	0.295	0.148	-0.017	-0.029	-0.032	0.078	0.288
	[0.253]	[0.123]	[0.132]	[0.144]	[0.054]	[0.26]	[0.305]
Migration	0.119	0.126	-0.041	-0.028	0.045	-0.327	-0.040
	[0.193]	[0.126]	[0.111]	[0.127]	[0.056]	[0.23]	[0.265]
Extreme Poverty	-0.090	**-0.089**	-0.039	-0.048	-0.023	0.032	-0.165
	[0.116]	**[0.048]***	[0.049]	[0.059]	[0.019]	[0.103]	[0.141]
Constant	-0.022	0.031	0.034	0.056	-0.016	-0.011	0.074
	[0.042]	[0.023]	[0.022]	[0.026]**	[0.009]*	[0.049]	[0.057]
Observations	5724	5724	5724	5724	5724	5724	5724
R-squared	0.28	0.12	0.14	0.17	0.07	0.22	0.44
		&&	&&	&&	&&		
	Panel B: Aged 60+						
Gini Coefficient	-0.121	-0.433	0.141	**-0.510**	-0.042	-0.319	0.002
	[0.415]	[0.42]	[0.246]	**[0.216]****	[0.273]	[0.398]	[0.047]
(Mean HHEMI)/100,000	0.548	0.353	-0.049	**0.656**	0.107	-0.307	-0.362
	[0.505]	[0.59]	[0.252]	**[0.312]****	[0.396]	[0.498]	[0.459]
Unemployment Rate	0.606	0.668	-0.436	-0.614	**0.971**	-0.467	-0.456
	[1.067]	[0.748]	[0.514]	[0.436]	**[0.528]***	[0.882]	[0.897]
Migration	-1.601	0.303	-0.555	-0.277	-0.332	-0.670	-0.747
	[0.996]	[0.705]	[0.484]	[0.381]	[0.521]	[0.761]	[0.804]
Extreme Poverty	-0.070	-0.364	-0.009	0.235	-0.173	**0.959**	0.500
	[0.381]	[0.38]	[0.174]	[0.174]	[0.228]	**[0.376]****	[0.448]
Constant	0.529	0.197	0.236	0.275	-0.312	0.586	1.225
	[0.184]***	[0.148]	[0.092]**	[0.077]***	[0.108]***	[0.159]***	[0.195]***
Observations	2544	2544	2544	2544	2544	2544	2544
R-squared	0.45	0.23	0.22	0.27	0.44	0.38	0.49

Note: Robust and clustered standard errors in brackets. Controlling for age (5-year brackets), sex, canton fixed effects, year fixed effects, canton fixed effects, and age*time interactions. Observations are weighted by canton population.
* significant at 10%; ** significant at 5%; *** significant at 1%
&& indicates that standard errors are NOT clustered. Robust standard errors are reported.

Table 3.7: Association of Inequality and Lifestyle Related Mortality Rate

	Chronic Respiratory Causes	Cardio Vascular Disease	Diabetes	Alcohol/Liver Disease	All Lifestyle
	Panel A: Aged 15-59				
Gini Coefficient	-0.037	**-0.281**	-0.044	0.098	-0.007
	[0.067]	**[0.126]****	[0.069]	[0.088]	[0.02]
(Mean HHEMI)/100,000	0.004	**0.035**	**-0.022**	-0.021	-0.141
	[0.012]	**[0.017]****	**[0.012]***	[0.014]	[0.149]
Unemployment Rate	0.169	-0.249	-0.033	-0.129	-0.212
	[0.124]	[0.355]	[0.15]	[0.168]	[0.308]
Migration	-0.052	**-0.492**	0.186	-0.206	-0.437
	[0.148]	**[0.289]***	[0.139]	[0.193]	[0.334]
Extreme Poverty	0.064	0.087	-0.014	-0.086	-0.034
	[0.064]	[0.128]	[0.076]	[0.079]	[0.147]
Constant	-0.004	0.017	0.057	-0.016	0.044
	[0.029]	[0.048]	[0.027]**	[0.038]	[0.053]
Observations	5724	5724	5724	5724	5724
R-squared	0.13	0.49	0.18	0.25	0.59
	Panel B: Aged 60+				
Gini Coefficient	-0.007	0.115	-0.130	0.264	**-0.102**
	[0.466]	[0.367]	[0.367]	[0.335]	**[0.049]****
(Mean HHEMI)/100,000	-1.058	**-0.992**	0.247	-0.329	-0.072
	[0.659]	**[0.536]***	[0.669]	[0.384]	[0.322]
Unemployment Rate	-0.578	-0.221	0.247	-0.791	0.066
	[0.884]	[0.891]	[1.035]	[0.726]	[0.842]
Migration	0.464	-0.538	**-2.779**	-0.483	-0.852
	[0.723]	[0.744]	**[0.827]*****	[0.588]	[0.72]
Extreme Poverty	**-0.655**	-0.372	**-0.867**	-0.003	-0.567
	[0.372]*	[0.342]	**[0.375]****	[0.231]	[0.341]
Constant	0.603	1.349	0.629	0.248	1.861
	[0.15]***	[0.166]***	[0.142]***	[0.11]**	[0.135]***
Observations	2544	2544	2544	2544	2544
R-squared	0.57	0.66	0.29	0.15	0.68

Note: Robust and clustered standard errors in brackets. Controlling for age (5-year brackets), sex, canton fixed effects, year fixed effects, canton fixed effects, and age*time interactions. Observations are weighted by canton population.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.8: Association of Inequality and External Related Mortality Rates

	Motor Vehicle Accident	Other Accidents	Suicide	Homicide	All External
	Panel A: Aged 15-59				
Gini Coefficient	0.033	-0.025	**0.140**	-0.092	-0.005
	[0.101]	[0.109]	**[0.082]***	[0.074]	[0.029]
(Mean HHEMI)/100,000	-0.012	0.010	-0.018	0.002	0.034
	[0.019]	[0.017]	[0.012]	[0.012]	[0.169]
Unemployment Rate	0.052	0.006	0.156	0.105	0.227
	[0.245]	[0.237]	[0.138]	[0.164]	[0.422]
Migration	0.361	-0.053	-0.101	-0.032	0.081
	[0.245]	[0.185]	[0.147]	[0.112]	[0.359]
Extreme Poverty	-0.108	0.046	-0.085	0.088	0.016
	[0.088]	[0.102]	[0.073]	[0.088]	[0.169]
Constant	-0.006	0.004	-0.018	-0.007	-0.014
	[0.047]	[0.05]	[0.031]	[0.03]	[0.086]
Observations	5724	5724	5724	5724	5724
R-squared	0.21	0.14	0.09	0.17	0.36
	Panel B: Aged 60+				
Gini Coefficient	-0.156	-0.232	-0.094	0.044	0.025
	[0.261]	[0.406]	[0.143]	[0.178]	[0.054]
(Mean HHEMI)/100,000	0.441	-0.313	0.105	-0.208	-0.336
	[0.43]	[0.462]	[0.231]	[0.261]	[0.427]
Unemployment Rate	**-1.476**	0.095	**0.437**	-0.223	-0.322
	[0.592]**	[0.711]	**[0.249]***	[0.352]	[0.928]
Migration	-1.011	0.372	-0.212	0.219	-0.782
	[0.579]*	[0.7]	[0.345]	[0.333]	[0.868]
Extreme Poverty	-0.194	-0.557	-0.054	-0.053	**-0.863**
	[0.228]	[0.27]**	[0.1]	[0.119]	**[0.319]*****
Constant	0.114	0.299	0.021	0.071	0.402
	[0.093]	[0.136]**	[0.051]	[0.072]	[0.156]**
Observations	2544	2544	2544	2544	2544
R-squared	0.17	0.36	0.08	0.09	0.34

Note: Robust and clustered standard errors in brackets. Controlling for age (5-year brackets), sex, canton fixed effects, year fixed effects, canton fixed effects, and age*time interactions. Regressions are weighted by canton population.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.9: Comparison of Pooled OLS, 15-year Change and 5-year Change Models

	Panel A: Aged 15-59						Panel B: Aged 60+					
	Pooled OLS		15 year changes		5 year changes		Pooled OLS		15 year changes		5 year changes	
	Beta	SE	Beta	SE	Beta	SE	Beta	SE	Beta	SE	Beta	SE
All-Cause Mortality	0.174	[0.217]	-0.095	[0.322]	-0.024	[0.183]	-0.076	[0.287]	-0.118	[0.504]	-0.066	[0.314]
Diarrhea	0.030	[0.019]	*0.029*	*[0.032]*	*0.027*	*[0.022]*	-0.046	[0.166]	0.218	[0.299]	0.025	[0.187]
Tuberculosis	-0.005	[0.029]	-0.050	[0.042]	-0.012	[0.041]	0.189	[0.148]	0.220	[0.248]	0.189	[0.163]
Acute Respiratory (Infection)	0.054	[0.041]	0.032	[0.074]	0.045	[0.049]	0.087	[0.339]	0.424	[0.469]	0.022	[0.326]
All other Infections	0.021	[0.036]	-0.107	[0.071]	0.020	[0.045]	**0.321**	**[0.162]***	**0.721**	**[0.246]*****	0.156	[0.208]
Cancer of Digestive System	0.058	[0.092]	-0.021	[0.128]	-0.073	[0.103]	0.110	[0.498]	-0.550	[0.625]	-0.121	[0.416]
Cancer of Respiratory System	-0.019	[0.043]	*-0.093*	*[0.085]*	*-0.085*	*[0.057]*	0.328	[0.262]	-0.777	[0.69]	-0.433	[0.42]
Cervical Cancer	-0.010	[0.039]	*-0.089*	*[0.082]*	*-0.008*	*[0.059]*	0.157	[0.193]	0.258	[0.284]	0.141	[0.246]
Breast Cancer	-0.047	[0.046]	*0.135*	*[0.111]*	*0.017*	*[0.068]*	-0.129	[0.227]	0.239	[0.314]	**-0.510**	**[0.216]****
Prostate Cancer	0.018	[0.016]	*-0.016*	*[0.03]*	*0.017*	*[0.026]*	-0.125	[0.242]	-0.160	[0.274]	-0.042	[0.274]
Other Cancers	0.046	[0.091]	0.057	[0.17]	0.077	[0.102]	-0.229	[0.453]	0.336	[0.616]	-0.319	[0.399]
Chronic Respiratory Causes	-0.004	[0.069]	0.173	[0.127]	-0.037	[0.068]	0.611	[0.428]	**-1.094**	**[0.581]***	-0.007	[0.466]
Cardio Vascular Disease	-0.069	[0.133]	**-0.408**	**[0.202]****	**-0.281**	**[0.126]****	0.229	[0.405]	-0.460	[0.651]	0.115	[0.368]
Diabetes	-0.036	[0.068]	-0.137	[0.109]	-0.044	[0.069]	0.126	[0.34]	-0.426	[0.625]	-0.130	[0.368]
Alcohol/Liver Disease	0.039	[0.104]	0.089	[0.139]	0.098	[0.088]	**0.524**	**[0.31]***	0.365	[0.489]	0.264	[0.335]
Motor Vehicle Accident	0.127	[0.108]	-0.101	[0.181]	0.033	[0.102]	0.163	[0.225]	-0.373	[0.368]	-0.156	[0.262]
Other Accidents	0.066	[0.068]	0.053	[0.139]	-0.025	[0.109]	0.461	[0.346]	0.293	[0.609]	-0.232	[0.406]
Suicide	0.066	[0.066]	**0.302**	**[0.118]****	**0.140**	**[0.083]***	-0.094	[0.102]	-0.376	[0.244]	-0.094	[0.143]
Homicide	-0.059	[0.091]	-0.111	[0.107]	-0.092	[0.074]	0.146	[0.158]	0.089	[0.237]	0.044	[0.179]

* significant at 10%; ** significant at 5%; *** significant at 1%. Note: Controls for Age (5 year brackets), gender, year fixed effects, age and year indicator interactions, canton-level mean household monthly equivalent income, canton-level unemployment rate, canton-level in-migration, and canton-level extreme poverty are included in all models. 15-year change and 5-year change models also include canton fixed effects. Italics indicate where Standard Errors are not clustered.

Table 3.10: Summary of Sensitivity Analysis

	Panel A: Aged 15-59				Panel B: Aged 60+			
	Transformed Beta (%)	95% Confidence Intervals (%)		Significance Across Models	Transformed Beta (%)	95% Confidence Intervals (%)		Significance Across Models
All-Cause Mortality	-0.143	2.042	-2.328		-0.396	3.354	-4.145	
Diarrhea	0.164	0.426	-0.099		0.152	2.385	-2.081	
Tuberculosis	-0.069	0.420	-0.559		1.133	3.079	-0.813	
Acute Respiratory (Infection)	0.270	0.855	-0.315		0.132	4.024	-3.761	
All other Infections	0.122	0.660	-0.415		0.933	3.417	-1.550	
Cancer of Digestive System	-0.439	0.791	-1.669		-0.728	4.239	-5.695	
Cancer of Respiratory System	-0.511	0.170	-1.191	&&, ++	-2.598	2.417	-7.613	
Cervical Cancer	-0.048	0.656	-0.753		0.845	3.783	-2.092	
Breast Cancer	0.099	0.911	-0.713		-3.059	-0.480	-5.638	**,&&,++
Prostate Cancer	0.101	0.412	-0.209		-0.249	3.023	-3.521	
Other Cancers	0.463	1.681	-0.755		-1.915	2.849	-6.679	
Chronic Respiratory Causes	-0.219	0.593	-1.031		-0.045	5.519	-5.609	
Cardio Vascular Disease	-1.689	-0.184	-3.193	**,++	0.690	5.084	-3.704	
Diabetes	-0.264	0.560	-1.088		-0.780	3.519	-5.078	
Alcohol/Liver Disease	0.586	1.636	-0.465		1.583	5.523	-2.357	
Motor Vehicle Accident	0.195	1.413	-1.023		-0.933	2.195	-4.061	
Other Accidents	-0.149	1.152	-1.450		-1.391	3.457	-6.238	
Suicide	0.842	1.833	-0.149	*, +	-0.565	1.143	-2.272	
Homicide	-0.550	0.334	-1.433		0.264	2.402	-1.873	

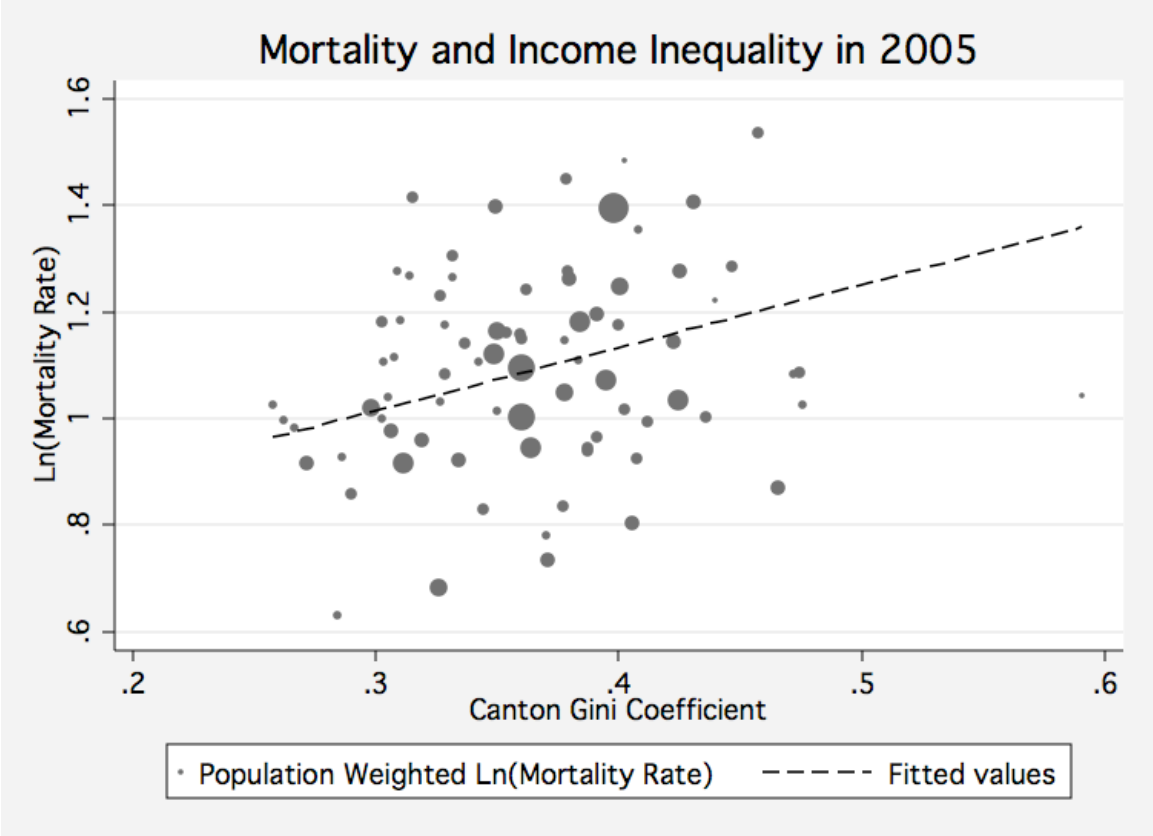
Note: Betas have been transformed and can be interpreted as an X percentage change in age-sex-cause specific mortality rate for a one standard deviation increase in income inequality. The betas reported reflect the main estimates from the clustered weighted semi-log least squares model including controls for age (5 year brackets), gender, year, age and year indicator interactions, canton fixed effects, canton-level mean household monthly equivalent income, canton-level unemployment rate, canton-level in-migration, and canton-level extreme poverty.

* significant at 10%; ** significant at 5%; *** significant at 1% in semi-log least squares "5-year change" models including extreme poverty.

& significant at 10%; && significant at 5%; &&& significant at 1% in semi-log least squares "5-year change" models including extreme poverty.

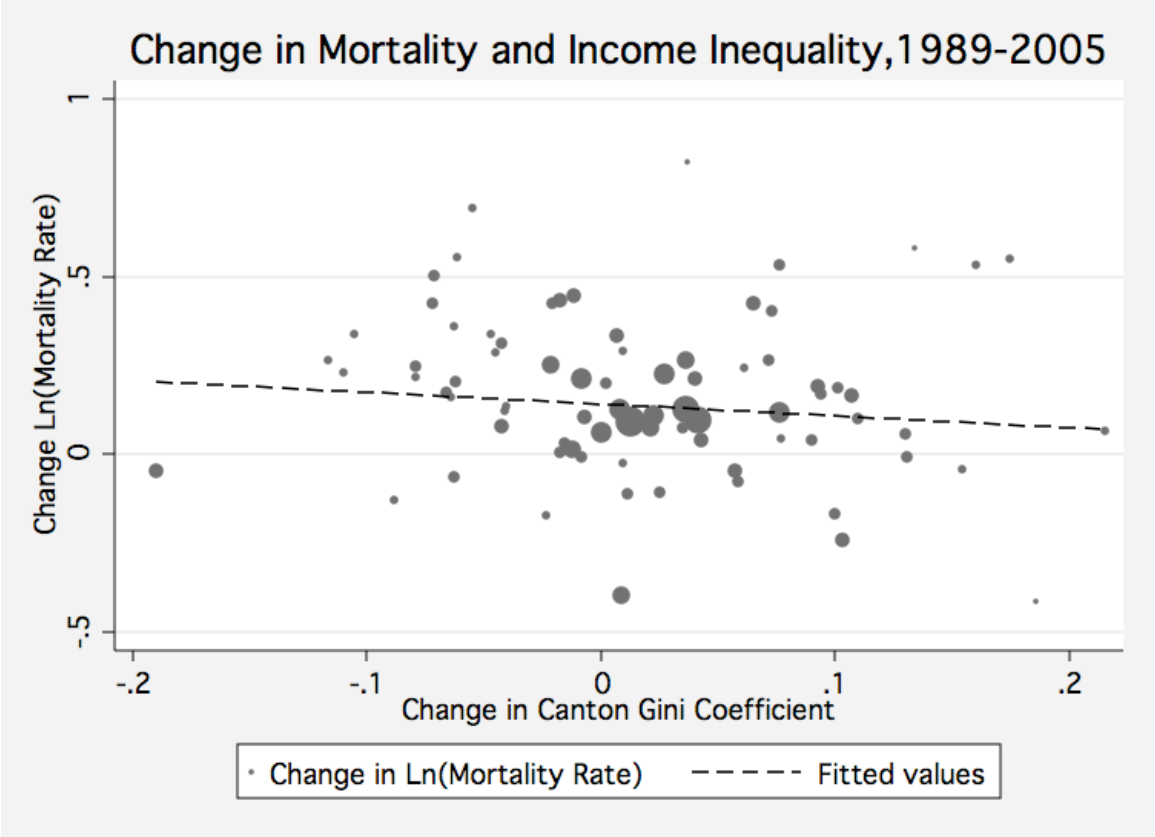
+ significant at 10%; ++ significant at 5%; +++ significant at 1% in Poisson Models (and Negative Binomial Models)

Figure 3.1: Scatter Plot of Canton Income Inequality and Mortality Rate in 2005



Source: Author's calculations from Costa Rican Household Survey for Multiple Purposes. The size of the markers reflects the canton population size.

Figure 3.2: Scatter Plot of Change in Mortality and Income Inequality between 1989-2005



Source: Author's calculations from Costa Rican Household Survey for Multiple Purposes. The size of the markers reflects the canton population size.

Chapter 4: Individual-level Association of Income Inequality and Mortality

While the previous study assessed the ecological association of income inequality and cause-specific mortality using canton fixed effects to control for time-invariant canton characteristics, those analyses could not address certain fundamental questions. First, ecological studies cannot satisfactorily distinguish inequality effects from non-linear absolute income effects (Wagstaff & van Doorslaer, 2000). Second, the heterogeneity of effects for different individuals within the cantons may be masked in the previous analyses. Accordingly, this study assesses the association of income inequality at the individual-level with the intent of examining whether or not poor individuals are more adversely affected by income inequality as predicted by the weak form of the income inequality hypothesis. To do so, this study will analyze a nationally representative prospective cohort to determine whether residence in cantons with increasing income inequality is independently associated with individual-level mortality risk over a 20-year follow-up. Survival will be assessed controlling for individual-level socioeconomic and demographic covariates. Thus, this chapter makes the following two contributions: (1) an oft-cited source of confounding, individual-level resources/wealth can be explicitly controlled for, and (2) the weak form of the income inequality hypothesis, can be tested explicitly. Moreover, this is the first such study that I know of from a non-industrialized country to examine these relationships in terms of long-term survival.

4.1 Data Collection

This analysis takes advantage of a new data source that links 1984 census records to mortality data in the Costa Rican Vital Statistics registries until December 31, 2007. These data have the distinct advantage of allowing us to follow a nationally representative prospective cohort of approximately 16,000 Costa Ricans adults for over 20 years.

This census-mortality dataset was collected in three phases. First, a random stratified sample of approximately 20,000 individuals was selected from the 1984 census. Next, the original census files were consulted in order to retrieve the name of the selected individuals, and then names were linked with the Civil Registry to obtain these individuals' "*cedula*" (a unique national identification number similar to a social security number). These matches are performed with an algorithm that employs alphabetic "triads" and observable demographic characteristics to link names from the census questionnaires with the Civil Registry names. Finally, the *cedula* is used to link these records with voting registries and, more importantly, with the Vital Statistics National Death Registry through 2007.

In order to obtain the desired sample size, 302 *segmento*s (census tracts) were randomly selected. From the 1984 census microdata files, all adults aged 30 and above in these 302 *segmento*s were selected into the sample.[66] Enumerating all those who met the

[66] The underlying population of interest for this data collection was Costa Ricans (by birth or naturalized) registered in the 1984 census. This population numbered at 753, 363 and the desired sample size was nearly 20,000, thus a 2.8 % random sample was required. 302 census tracts account for 2.8% of census tracts in Costa Rica.

sample criteria in the *segmentos* led to a sample size of 21, 930. Of these observations, 4% were discarded because they represented foreigners, leaving an initial sample size of 21,161.

Next this sample was sent to the Instituto Nacional de Estadisticos y Censos (INEC), which copied the full names from the census questionnaires of the selected individuals, and linked them with the Civil Registry, in order to obtain the individuals' "*cedulas*" (unique national identification number). To obtain names and assure accuracy, a computer program was developed to help the enumerators copy the names off the census ballots and perform several consistency checks. For 2.2% of the sample, there were problems in obtaining names from the census ballots due to ballot loss, missing names, or illegible names.

The final step required the names obtained from the census ballots to be matched to names and *cedulas* in the Civil Registry. To do this, a matching algorithm divided the names into triads and matched the rare triad. For each match found across the two datasets, a weight was assigned according to: (a) the similarity of the names (70%); (b) the agreement of the provinces of birth (10%); and (c) the similarity between the year of birth registered in the Civil Registry and the year of birth according to the reported census age (20%). Based on these criteria, when there were multiple matches to a single name, the one with the higher weight was selected as the correct match. This method matched 18,258 of the census names to *cedulas*. Only those observations with a high probability of a high-quality match are used in this study leading to a sample size of 16,316.[67] Table 4.1 summarizes the data creation process. In this sample of high-quality matches, there were 4,346 deaths by December 31, 2007. Of these deaths, 3,952 could be linked to the Vital Statistics Records, which further indicates the cause of death information.

In addition to the linkage between the census and mortality data, the *cedula* links observations to Costa Rica's voting registry data. Costa Ricans are required to vote, and consequently they are required to register every four years. Thus the canton where they register to vote can also be tracked. This also allows for some corrections to the data. For example, for the 112 deceased individuals whose date of death is unknown, the date of death was assigned between the date they were last observed in the voting registries and December 31, 2007. Since exclusion of such deaths would lead to a downward bias, the above corrections are preferred to dropping these observations (Luis Rosero-Bixby & Antich, 2009).

To further assess the potential for the under-registration of deaths, the CCP carried out several consistency checks. First, of the 151 persons aged 85+ in 1984, in the data none were deemed alive by December 2007 and only 3 were lost to follow-up. For

[67] Since census names could match multiple names in the Civil Registry (e.g. Maria Lopez), an automated process was used to rate the likelihood that the link indeed identified the same person across the two datasets. In this study, I include individuals that either had the best match or the second best match. A match was considered to be a best match if its weight was over 80, the similarity of the year of birth in was no more than + / - 2 years, and the next closest match was at least 10 points or more lower. The second best match were cases where the weight exceeded 70 and had a difference of 15 points or more to the next closest match. Details of matching can be found in Spanish in Luis Rosero-Bixby & Antich, 2009 (Luis Rosero-Bixby & Antich, 2009).

those aged 75-84 in 1984, only 16 were alive by December 2007, and several of these individuals were confirmed to be alive and were actually contacted. [68]

4.2 Measures

Throughout all the models, demographic characteristics (gender, age, age squared, age cubed), insurance status, and whether the individual is in a consensual union in 1984 are included as controls. Likewise, the information in the 1984 census is also employed to create several non-time varying socio-economic measures.

4.2.1 Individual-Level Covariates

4.2.1.1 Education

Since the sample consists of those aged 30 years or older, most individuals have already completed their education. Consequently, the education variable accounts for the highest level of education completed. It has four categories: (1) No formal education or some primary education; (2) completed primary or completed primary with some secondary education; (3) completed secondary education or completed secondary education with some college education; and (4) received a college or graduate degree.

4.2.1.2 Wealth

Since there is no income measure in the 1984 census, a wealth measure is included to control for individual-level resources. The wealth measure is a summary of index based on owning certain assets (telephone, hot water heater, refrigerator, radio, television, and car) or having certain amenities in the household (access to electricity, piped in water, sewage, and non-dirt floors) in 1984. This measure ranges from 0 to 10 and is divided into three wealth categories: (a) low wealth, defined as those with a wealth measure of less than 5 (33% of the sample); (b) average wealth, defined as those with a wealth measure greater than 5 but less than 9 (48% of the sample) and; (c) high wealth, defined as those with wealth measure greater than 9 (19% of the sample). It should also be noted that there were clear regional disparities in this wealth measure. In the province of San José, 29% of the population had the maximum possible value for the wealth measure, while in Limón province only 3% had the maximum possible. Figure 4.1 presents the box plot of the created wealth index by Province.

4.2.1.3 Place of Residence

Figure 4.1 shows that four provinces have consistently wealthier residents. These four provinces, San José, Alajuela, Cartago, and Heredia, also contain parts of the San José Metro area, which reflects the higher wealth in the capital and surrounding central valley. A control variable for living in the San José Metro area is included in models without area fixed effects (discussed below). This may be an important proxy for SES in the models without fixed effects. Moreover, a control variable for whether or not one lives in a rural or urban area is also added to the models. Urbanicity varies within cantons, and as such, it is included as a control even in models with fixed effects.

[68] In general, Costa Rica's Vital Statistics Registry is ranked as complete by WHO and considered of high quality similar to that of the United States, Japan etc. (Mahapatra, Shibuya, Lopez, Coullare, Notzon, Rao et al., 2007)

4.2.1.4 Relative Deprivation

In some analyses an alternate version of the relative income hypothesis similar to the weak form of the income inequality hypothesis will be examined using a measure of relative deprivation. The relative deprivation measure will be calculated over the canton of residence and an individual's age (based on two age groups 30-44 and 45+). Following Eibner and Evans (2005), relative deprivation (RD) will be defined as the sum of the differences in the wealth index between person i and all others who have more wealth than person i in their canton and age group, j. This measure is calculated as follows

$RD_i=1/N*\sum(y_j-y_i)\ \forall\ y_j> y_i$ or $RD_i=[E(y|y > y_i)- y_i]* Pr(y> y_i)$

This measure can be interpreted as the expected difference in wealth between person i and all other individuals with greater wealth than that individual in his/her community and age group.

Table 4.2 presents the means of all the individual-level variables from the 1984 census.

4.2.2 Time Varying Variables

The time varying variables are the same as those created in Chapter 2, namely, income inequality, mean household equivalent monthly income (HHEMI), the unemployment rate, the proportion of in-migration in the last two years, and the proportion of the canton households living in extreme poverty. Each observation is split into six periods to coincide with the registration dates in the voting registries. Each of these periods is then matched to the closest year for which canton-level measures are available. Table 4.3 summarizes the timing and the mapping of the area-level measures onto the individual-level data.

For each individual, the voting registry data has the canton where the individual registered to vote in the November prior to the previous election (November 1989, November 1993, November 1997, November 2001, November 2005). If the most recent voting registration cantons had been used in the analyses, this information would have allowed for additional variation because those who had moved could have had their area-level covariates updated to reflect their most recent canton of residence. However, there may be selection problems in using this information, because many people move when approaching death. Those who are near death are more likely to move to obtain better health services (for example people with terminal illnesses may move to San José to obtain better tertiary medical care) or to live with family members. If such is the case, using the updated information may lead to biases in the survival analyses. This case is explicitly tested, and the analysis suggests that those who move are more likely to die sooner. Consequently, people who move are not followed in the voting registry data. Therefore, canton-level measures are assigned based on canton of residence in 1984.

4.2.3 Outcomes

The outcome of interest is time to death. Survival between January 1, 1987, and December 31, 2007, is examined for a potential 21-year follow up.[69] In the analysis the

[69] For the relative deprivation model in the final analysis, the survival time begins from September 9, 1984, leading to a 24-year follow-up period.

follow-up period begins on January 1, 1987, because 1987 is the first year for which consistent inequality measures are available, even though the follow-up in the initial data begins on September 9, 1984. Examining the data from 1987 onward, 4,085 deaths are recorded, and 3,725 of them have an associated cause of death. The causes of death based on ICD 9 and 10 codes are further classified in these analyses. Table 4.4 details the classification for the five causes of death thought to be most susceptible to a negative income inequality effect (these are suicide and the lifestyle diseases discussed in Chapter 3). Table 4.5 displays the number of deaths observed in these data for these causes. Of the 3,725 deaths with known causes, 1,052 of those deaths were due to heart disease, 219 due to chronic lower respiratory disease, 187 deaths due to diabetes, 65 due to liver disease, and only 2 were due to suicide. Heart disease is the only cause of death for which there is a sufficient numbers of deaths to examine in a separate sub-analysis.

Figure 4.2 and 4.3 present Kaplan-Meier estimate of the survival function for these two outcomes by gender. Separate survival curves are drawn for those living in a canton with increasing or decreasing income inequality.[70] These descriptive graphs do not show any significant difference in survival rate according to changes in income inequality. In order to understand these relationships further, I examine them in a multivariate context.

4.3 Analysis Methods

This study employs maximum likelihood survival regressions with a Gompertz hazard distribution. The Gompertz distribution has been widely used to study mortality, and these models are known to fit these data (L. Rosero-Bixby & Dow, 2009a). The key assumption in the Gompertz model is that the underlying rate of mortality is monotonically increasing (or decreasing) with time of exposure, or in our case, age.[71]

The Gompertz hazard equations is

$$H(t)=\lambda e^{\gamma t}$$

where γ is fixed for all individuals and represents the external "force" of mortality, while λ is estimated based on a log-linear equation. The λ is a function of the covariates, $\lambda=\exp(\boldsymbol{\beta X})$. The first two sets of models will examine the income inequality hypothesis in its various formulations. Two groups of models will be estimated, without and with canton fixed effects for the income inequality hypothesis.

Group 1
$$\boldsymbol{\beta X} = \beta_0 + \beta_1 (I_{ct}) + \beta_2 (\mathbf{A}_{ct}) + \boldsymbol{\beta}_{3Z} (\mathbf{Z}_i) + \boldsymbol{\varepsilon}$$
$$\boldsymbol{\beta X} = \beta_0 + \beta_1 (I_{ct}) + \beta_2 (\mathbf{A}_{ct}) + \boldsymbol{\beta}_{3Z} (\mathbf{Z}_i) + \boldsymbol{\beta}_4 I_{ct} * LW_i + \boldsymbol{\varepsilon}$$

Group 2
$$\boldsymbol{\beta X} = \beta_0 + \beta_1 (I_{ct}) + \beta_2 (\mathbf{A}_{ct}) + \boldsymbol{\beta}_{3Z} (\mathbf{Z}_i) + \boldsymbol{\beta}_5 y_c + \boldsymbol{\varepsilon}$$
$$\boldsymbol{\beta X} = \beta_0 + \beta_1 (I_{ct}) + \beta_2 (\mathbf{A}_{ct}) + \boldsymbol{\beta}_{3Z} (\mathbf{Z}_i) + \boldsymbol{\beta}_4 I_{ct} * LW_i + \boldsymbol{\beta}_5 y_c + \boldsymbol{\varepsilon}$$

[70] Increasing or decreasing inequality is defined based on growth rate regressions outlined in Chapter 2. Cantons with growth rates >0.5% are categorized as increasing and those with growth rates <-.5% are categorized as decreasing. See Figure 2.9.

[71] If a hazard is known to depend strongly on age but weakly on time, such as death, then age is the best way to define the time scale (Allison, 1984).

In the first group of models, I_{ct} is the income inequality in canton *c* at time *t*, A_{ct} is a vector of the canton-level controls at time *t*, $\mathbf{Z}_i$ is a vector of individual controls measured in 1984, and I_{ct} ***LW**$_i$ is an indicator variable for having low wealth interacted with the canton Gini coefficient.[72] The estimate for β_1 is meant to capture the association between income inequality and survival risk. The interaction is included to examine if income inequality has a different impact on the poor as is assumed in the weak form of the income inequality hypothesis.

For the second group of models, similar models are estimated but canton level fixed effects, $\mathbf{y}_c$, are also included. These fixed effects are meant to account for time invariatant canton characteristics that are corrolated with both changes income inequality and mortalitiy risk. As discussed in Chapter 3, the addition of these fixed effects are meant to limit confounding as long as there are *no systematic* time invariant shocks that are correlated with changes in income inequality.

A third group of models will assess the relative deprivation version of the relative income hypotheses (see discussion of in Chapter 1). These models will use wealth as the underlying metric to create the relative deprivation measure and will not have any time varying variables. Consequently, they are not directly comparable to the models above, but will rather serve as complementary analysis of the weak form of the income inequality hypothesis.

Group 3 $\quad \boldsymbol{\beta X} = \beta_0 + \boldsymbol{\beta_{3Z}}(\mathbf{Z}_i) + \beta_6 RD_i + \boldsymbol{\varepsilon}$

$\quad\quad\quad\quad \boldsymbol{\beta X} = \beta_0 + \boldsymbol{\beta_{3Z}}(\mathbf{Z}_i) + \beta_6 RD_i + \boldsymbol{\beta_7 y_c} + \boldsymbol{\varepsilon}$

Here $\mathbf{Z}_i$ is a vector of individual controls measured in 1984 (gender, age, age squared, age cubed, insurance status, consensual union, wealth, education, and place of residence); RD_i is the individual's relative deprivation as compared to those living within their canton and in their age group; and $\mathbf{y}_c$ is a vector of area-level fixed effects. Here the canton fixed effects will capture common unobserved factors shared among those living in the same canton, such as access to health services that may be related to both the variation in deprivation and mortality.[73] The RD variable varies at the individual-level, so area-level controls are not considered.

All models are weighted by the probability that the observation had a successful match and are clustered at the canton level.

[72] Note that I_{ct} and A_{ct} are time varying while $\mathbf{Z}_i$ and LW_i are not.
[73] Eibner and Evans (2005) also include these types of fixed effects in their models.

4.4 Results

4.4.1 Income Inequality Hypotheses

4.4.1.1 All-Cause Mortality

Table 4.7 presents the results for all-cause mortality using the Gompertz models described above. The first three models (columns 1-3) present the relationship between income inequality and survival time without fixed effects, whereas the last three models include canton fixed effects (columns 4-6). Columns 1 and 4 present the basic model without including the canton-level percent in extreme poverty. Columns 2 and 5 include this extra poverty control. Finally columns 3 and 6 include an interaction term between income inequality and having low wealth as a test for the weak form of the income inequality hypothesis.

Across all models there are similar patterns for the individual-level time-invariant charactersitics. As expected, women have a 30% lower hazard of mortality. Those who complete college have a 24% lower mortality hazard (HR= .76 CI .65-.87), however, they account for only 10 percent of the sample and these tend to be concentrated in the younger generations. Those who live in rural area also have 9-13% lower hazard, but this coefficient becomes only marginally significant upon the inclusion of fixed effects. Moreover, there is also a survival advantage for those partnered in 1984. Those with high wealth have a 10% lower mortality hazard (HR= .90 CI .84-.94) in the models without the fixed effects. Finally, those without insurance in 1984 also live longer providing evidence that there was adverse selection into the public insurance program at that time. Also, these models show that area-level mean income is not associated with survival time when controls for individual wealth are included.

The canton-level variable estimates differ in the models with and without the fixed effects. The first group of models shows a protective effect of income inequality, i.e. contrary to expectation those living in areas with increasing inequality have better survival. Adding poverty, which in this case acts as a secondary control for income distribution, reduces the magnitude on the inequality variable, but the poverty measure becomes signifcant. The hazard ratio is only slightly protective at 6.9% lower for a one standard deviation increase in area-level income inequality (HR=.931, CI .91-.95). Adding an interaction term for income inequality and low wealth does not yeild a significantly different inequality effect for those with low wealth. The magnitude on this interaction term is positive indicating that the poor are harmed by income inequality, but the coeffiecnt is not significant. In the fixed effects model, the income inequality term is no longer significant, and niether is the area-level poverty measure.[74] Note that in the fixed effects models the interaction of poverty (having low wealth) and canton-level income inequality is not significant either. Moreover, unemployment is related to survival. Areas with a one-standard deviation increase in unemployment have an 8% lower hazard rate (HR= .92, CI .90-94). This is again consistent with Ruhm's argument that mortality and unempoyment maybe procyclical (Ruhm, 2000), especially for certain causes.

[74] Log-likelihood ratio tests indicate that taken together all fixed effects are not statistically significant, but individual ones are.

4.4.1.2 Heart Disease Mortality

Table 4.8 presents the results for heart disease mortality. Here again the first three models (columns 1-3) present the relationship between income inequality and survival time without the canton fixed effects whereas the last three models include canton fixed effects. Columns 2 and 5 include the extra poverty control and columns 3 and 6 include the interaction term (albeit without the extra poverty control this time).

Fewer individual-level characteristics were found to be determinants of heart disease mortality. Women have a lower hazard of cardiovascular mortality, as do those with college education. The parameter estimate for the decrease in hazard for those with a college education is quite large, even in the fixed effect model (HR= .53, CI-.47-.61).[75] Also, the wealth gradient is not present for heart disease. The results also show that those with no insurance in 1984 have reduced heart disease mortality, again pointing to adverse selection into the public insurance program.

The results for the canton-level variables are unexpected. Areas with increases in mean income have a *higher* hazard for heart disease mortality, whereas areas with increases in income inequality have a *lower* hazard of mortality. In these models, when the extreme poverty variable is added, the coefficient on income inequality becomes more negative while the coefficient on the poverty measure is positive. This might be due to the redundancy of including this measure in addition to controls for mean income, which may be measuring a similar construct. In the final specification with the interaction, this measure is excluded. In doing so, the results show that although the interaction is not significant, the main effect of living in an area with a one standard deviation increase in income inequality decreases the hazard of cardiovascular mortality by 18% (HR=0.82, CI 0.75-0.88). The hazard further decreases when the area-level fixed effects are included (HR=0.63, CI 0.52-0.76). Moreover, in areas with increases in unemployment the mortality hazard is also quite large and positive. An increase of one standard deviation in the unemployed rate is associated with a 38% increase in the hazard of mortality (HR=1.38, CI 1.23-1.54) in the fixed effects model. Comparing the contradictory effect of unemployment from Table 4.7 to those in Table 4.8 underscores how unemployment may affect mortality through a multifaceted pathway.[76]

4.4.1.3 Comparison of Parametric and Semi-Parametric Models

Given the peculiarities in the results above, several different parametric models are compared to the Gompertz models to ensure that the distributional assumptions are not driving the results. Indeed, in the models above, the gamma for the baseline hazard in the Gompertz model is negative for cardiovascular mortality suggesting that the hazard of cardiovascular mortality is decreasing with age, which is contrary to common sense (and the Kaplan-Meier graphs presented). Upon closer inspection, the estimated gamma becomes negative only after adding control variables beyond gender; thus this is less worrisome. The Gompertz models were then compared with Cox Proportional Hazard models, which have less stringent parametric assumptions. Table 4.9 presents these regressions. Here the coefficients across these two sets of models are very similar.

[75] This may be because of health behaviors. There is some evidence from another Costa Rican dataset that college graduates smoke and drink less. These are two main contributors to cardiovascular disease.
[76] In the ecological models in Chapter 3, unemployment was associated with lower infectious disease and motor vehicle accidents, but it was not associated with cardiovascular mortality.

Furthermore, a variety of parametric models are compared using the Akaike's information criterion (AIC) (Akaike, 1981). Of the tested parametric models, the Gompertz had the lowest AIC score for both all-cause mortality and cardiovascular disease mortality (Table 4.10)---confirming that the Gompertz model is the best parametric survival model for these data.

4.4.2 Relative Deprivation Hypothesis

As mentioned in Chapter 1, a parallel hypothesis about inequality and health argues that those who are more deprived than their peers in a community will have poorer health outcomes. In effect, a poor individual within an unequal community is also more relatively deprived. Thus, this relative deprivation formulation is similar to the weak form of the income inequality hypothesis. While the purpose of this supplemental analysis is not to directly compare across measure, it might be useful to look at this formulation given our null results in the previous analyses.

Accordingly, a measure of Relative Deprivation was created based on the 1984 wealth variable. Unlike income inequality, this variable varies at the individual-level, is based on wealth instead of income, and is not time varying and thus is not directly comparable to the Gini coefficient uses in the previous analyses. In addition, the start of the observation time is changed to September 9, 1984 since these data do not need to be mapped to another dataset.

Table 4.11 presents the results from these analyses. Columns 1-3 present the results for all-cause mortality and columns 4-6 present the results for heart disease mortality. In column 1, the baseline model without the RD measures are presented. As before women, those with college degree, those living in a rural area, those with high wealth, and those with no insurance in 1984 have higher all-cause survival. In columns 2 and 3 the RD measure is added. Each additional unit of wealth deprivation, i.e. having one fewer item than those who are wealthier, is associated with a 6% higher risk of death (HR=1.06, CI 1.08-1.04), even with the fixed effects. Also noteworthy, the RD measures eliminate the significance of the protective effects of having high wealth. This suggests the RD is capturing much of the same variation as those who are wealth are by definition less deprived.

For heart-disease mortality, the analyses confirm that as before women, those with a college education, and those with no health insurance in 1984 have better survival in the baseline model in column 5. In columns 6 and 7, the results indicate that each additional unit of deprivation is associated with a 7% increase in the risk of death from heart disease, but this is coefficient is not significant with the inclusion of canton fixed effects.

4.5 Discussion

The primary aim of this individual-level analysis was to test the weak form of the income inequality hypothesis (whether inequality affects the poor differently). For all-cause mortality the results fail to show that the poor had a significantly higher hazard of mortality in more unequal areas, but the magnitude was in this hypothesized direction. For heart disease mortality, the results again failed to find support for a supplemental inequality effect on the poor, and the magnitude of the estimated effect was in the

opposite direction. Furthermore, the results also indicate that contrary to expectations income inequality was associated with better survival for heart disease mortality.

For all-cause mortality, the results were generally consistent with the ecological analyses in Chapter 3. While the analyses reveal some evidence for an unexpected protective effect of income inequality in the models without fixed effects, no correlation between income inequality and survival was found once controls for canton fixed effects were included. The results indicate that those with the highest education and wealth have the best survival, but only at the extremes and not in a graded fashion. The analyses also suggest that residents in areas with increasing unemployment have better survival even with fixed effects and controls for living in a rural area. This is similar to results in Chapter 3 where increasing unemployment was associated with lower rates of infectious disease and motor accident mortality, which may be influencing the results here.

In the case of cardiovascular mortality, several unexpected, but not unprecedented associations were found. First, those who reside in cantons with increasing inequality actually exhibited longer survival. Based on the initial hypotheses, the opposite results were predicted. This result then begs the question of how income inequality or other sources of heterogeneity can be protective. A standard argument proposes that in the short-term there are tradeoffs between economic growth and equality. Thus, in cantons with high growth, there may be short-term increases in income inequality (Kuznets, 1955), but benefits from this growth outweigh any potential negative effects of income inequality. This may especially be true in settings with abundant safety nets and where inequality is not associated with the presence of a historical underclass. Another argument could be based on so-called demonstration effects. To the extent that specific subgroups exhibit better health behaviors, then those living in an environment with greater heterogeneity may benefit from witnessing the good behaviors of those outside their group (Galea & Ahern, 2005). For example if the rich smoke less and the poor eat more healthy diets (there is some evidence of both in Costa Rica), and if these behaviors are observable then, potentially, they may be more likely to be adopted in more heterogeneous settings. Whether or not these mechanisms are driving our results, we fail to find evidence for the detrimental effects of income inequality as expected.[77]

In addition to the surprising inequality finding, the results for heart disease also suggest a notable relationship with unemployment. Unemployment is associated with a greater risk of mortality. The association in this case is the opposite of that observed for all-cause mortality. Given that heart disease accounts for almost 25% of mortality in this sample, these results further suggest that the protective effects of area-level unemployment on other health causes are even larger than the estimate for all-cause mortality. These paradoxical results suggest that the effects of unemployment on mortality may be quite nuanced and different depending on the cause. In Chapter 3, the results show that unemployment was related to lower mortality for specific causes that may be related to the positive externalities of unemployment for some members of the family (more time to take care of family members and less time interacting with other

[77] Another notable difference between the all-cause mortality results and the heart disease mortality results highlights the different ways in which income effects mortality. For all-cause mortality individual level wealth was related to mortality, but area-level income was not. For heart disease mortality the opposite was found. These results suggest that heart disease may somehow be more sensitive to ecological changes in SES.

people may lead to lower mortality from infectious diseases, and less driving may lower motor vehicle accidents when unemployment is high). Yet unemployment may still have detrimental effects on the individual's health. Bezruchka makes this kind of argument when she attempts to reconcile different studies of the effects of unemployment on health (Bezruchka, 2009).

Again, with regard to cardiovascular mortality, the results demonstrate a survival advantage for those with a college education, but none for those with wealth. Those who complete college may be less likely to drink or smoke, which may be driving the results because smoking is the largest risk factor for heart disease. Contrary to expectation, there is no survival advantage for the wealthy. As suggested above, it may be that there is a mix of health behaviors, some concentrated among the poor and some concentrated among the rich, which are effectively canceling each other out. This has indeed been found in an older population of Costa Ricans (Luis Rosero-Bixby & Dow, 2009b).

Finally, there was some evidence that those who are more deprived compared to their peers within a community had lower survival rates in these analyses. This relative deprivation formulation was meant as an alternate test to the weak form of the income inequality hypothesis. However, this test was based on different underlying data and was essentially a cross-sectional test, so the results are not directly comparable to the fixed effects results discussed above. Additionally, these results are not directly comparable to models that do not include fixed effects, but interestingly they show an relationship opposite to those exhibited in columns 1-3 in Table 4.7. The results from the RD metric are most closely comparable to the analysis present in Figure 3.1, which also shows a slight positive association between inequality and all-cause mortality, albeit in 2005. In examining the RD metric at the canton-level, a weak association between average relative deprivation based on wealth measures and income inequality was found (correlation coefficient =0.18 p-value=0.101). While the results cannot be compared with the previous analysis, the inconsistency between the *positive* association of relative deprivation and survival versus the *negative* relationship of changes in income inequality and survival demands that further scrutiny.

4.6 Limitations

While this study has several methodological advantages over the previous ecological design, there are some noteworthy limitations. First, the time intervals at which the canton-level covariates are measured are not continuous whereas the time intervals for the outcome variables are. The way that canton-level covariates were matched is somewhat problematic since the changes in income inequality were mapped to a specific point in time. The change is considered instantaneous, where as in reality it is not so. Allison suggests that when time of the outcome event is measured more precisely than the intervals at which the explanatory variables are measured, matching the closest time to the event is a valid a approximation (Allison, 1984). A more nuanced method would be to linearly interpolate the changes. This may be done in the future.

Second, as mentioned above, while there is available information on the actual canton of residence during the November before each election year, the new area-level measures are not used for those who move. This is because those who are moved were found to be more likely to die. This relationship may because some individuals move to obtain better services or to live with family members when approaching death. While

certainly some move for other reasons and there are potential efficiency gains from using this variation, this variation was not used for fear of introducing new bias into the results.

Third, some of the problems that were present in Chapter 3 are worth mentioning again. As discussed in the previous chapter, using canton as a reference group maybe problematic if individuals do not compare themselves to those living nearby, but rather compare themselves to the nation as a whole. This is an inherent problem in much of this literature that requires further clarification. Another potential problem is measurement error. To estimate the extent of the potential attenuation bias, the Gompertz models in Table 4.9, column 2, were estimated 50 times, each time adding error to the Gini Measure as described in Chapter 3. The attenuation bias on the inequality estimate was about 11%. This attenuation bias is not likely to be large enough to have changed the overall results significantly. In addition, the cause of death coding in the Vital Statistic Registry may still be problematic. While these are the best data available, they may still include some misclassifications.

Table 4.1: Census Mortality Linkage Sample Creation

Sample Creation	N
*Stratified Random Sample**	
Met selection criteria in 1984 census (30+, No Foreigners)	21,161
*Names Obtained From Census Ballots**	
Names Found	19,954
*Names Linked to Cedula Number**	
All Matches	18,258
Good Matches (used in analysis)	16,316
Deaths	
Dead by December 31, 2007	4,316
Cause of Death Known	3,925

*Details found in Spanish in Luis Rosero-Bixby & Antich, 2009

Table 4.2: Means of Individual-level Variables

	Weighted Mean	Linearized SD	Observations
Age			
	46.777	0.262	16316
Gender			
Female (%)	49.73%	0.80%	16316
Education			
No Formal Education Completed (%)	51.56%	3.12%	16316
Completed Primary (%)	31.89%	1.51%	16316
Completed Secondary (%)	6.39%	0.80%	16316
Completed College (%)	10.15%	1.33%	16316
Geographical Area			
Live in San José Great Metro Area in 1984 (%)	55.55%	8.38%	16316
Live in Rural Area (%)	50.04%	7.09%	16316
Wealth Group			
Wealth Index [Range 0-10]	6.256	0.265	16203
Low Wealth (<=5) (%)	33.72%	4.40%	16203
Middle Wealth (>5, <9) (%)	47.44%	2.30%	16203
High Wealth (>9) (%)	18.84%	3.16%	16203
Insurance Status			
No insurance in 1984 (%)	22.80%	0.93%	16316
In Union			
Union (%)	74.85%	1.29%	16316
Relative Deprivation Measure			
Relative Deprivation [Range 0-7.9]	0.980	0.030	16316

Weighted by probability of match and SD clustered by canton.

Table 4.3: Mapping Time Varying Variables

Split in Census-Mortality Data	Group	Year Matched	Mean Income 2006 inflation correction multiplier**
9-Sep-84-1-Nov-89*	1	1987, 1988	12.7
1-Nov-89-1-Nov-93	2	1989, 1990, 1991	8.18
1-Nov-93-1-Nov-97	3	1992, 1993	5.04
1-Nov-97-1-Nov-01	4	1995, 1996	3.2
1-Nov-01-1-Nov-05	5	2001, 2002	1.63
1-Nov-05-31-Dec-07	6	2003, 2004	1.35

*Analysis time begins in Jan 1, 1987
** CPI deflators found at Banco Central De Costa Rica (http://indicadoreseconomicos.bccr.fi.cr). These are averages across years as multiple years are included.

Table 4.4: Categorization by ICD-9 & ICD-10 Cause of Death

Cause of Death	ICD-10 Codes	ICD-9 Codes
Intentional self-harm (suicide)	X60–X84,Y87.0	E950–E959
Diseases of heart	I00–I09,I11,I13,I20–I51	390–398,402,404,410–429
Chronic lower respiratory diseases	J40–J47	490–494,496
Diabetes mellitus	E10–E14	250
Chronic liver disease and cirrhosis	K70,K73–K74	571

Source: Anderson, R., Minino, A., Hoyert, D., & Rosenberg, H. (2001). Comparability of cause of death between ICD-9 and ICD-10: preliminary estimates. *National vital statistics reports,* 49(2), 1-32.

Table 4.5: Number of Deaths by Cause in Sample

Cause of Death	Number of failures after 1/1/87*	Number of failures after 9/9/84**
All-cause	4085	4,319
Diseases of heart	1,005	1052
Intentional self-harm (suicide)	2	2
Diabetes mellitus	174	185
Chronic lower respiratory diseases	208	219
Chronic liver disease and cirrhosis	58	65

* Used in analysis with income inequality measures
** Used in analysis with relative deprivation measures

Table 4.6: Summary Statistics of Weighted Canton-Level Measures

Variable	Year	Obs	Population Weighted Mean	Std. Dev.
Gini Coefficient	1987	80	0.355	0.057
	1989	80	0.354	0.049
	1992	80	0.350	0.043
	1995	80	0.362	0.047
	2001	79	0.401	0.056
	2003	79	0.375	0.047
(Mean HHEMI)/1000 in 2006 Colones*	1987	80	181.79	62.71
	1989	80	189.10	54.30
	1992	80	169.54	51.01
	1995	80	169.14	55.43
	2001	79	209.27	74.93
	2003	79	217.35	81.13
Unemployment Rate	1987	80	0.055	0.025
	1989	80	0.046	0.018
	1992	80	0.040	0.020
	1995	80	0.057	0.022
	2001	79	0.066	0.027
	2003	79	0.066	0.020
In-Migration	1987	80	0.057	0.038
	1989	80	0.042	0.023
	1992	80	0.037	0.030
	1995	80	0.047	0.025
	2001	79	0.043	0.026
	2003	79	0.049	0.024
Extreme Poverty	1987	80	0.092	0.084
	1989	80	0.096	0.088
	1992	80	0.079	0.066
	1995	80	0.065	0.060
	2001	79	0.063	0.066
	2003	79	0.058	0.060
Weighted by canton population				
* Includes only salaried workers				

Table 4.7: Gompertz Models of Risk of Mortality and Income Inequality

	All-Cause Mortality					
Canton Level Area Covariates from July Household Surveys- Time varying						
Gini of Monthly Household Equivalent Salary	-1.179	-0.769	-0.718	-0.805	-0.731	-0.541
	[0.37]***	[0.407]*	[0.503]	[0.521]	[0.556]	[0.655]
Mean of the Monthly Household Equivalent Salary	0.00046	-0.00001	-0.00001	-0.0004	-0.00047	-0.0005
	[0.00042]	[0.00041]	[0.00041]	[0.00053]	[0.00056]	[0.00057]
Interaction of Wealth & Area Level Gini						
Gini*Low Wealth			-0.108			-0.425
			[0.647]			[0.67]
Percent Unemployed	-1.986	-2.016	-2.021	-3.994	-3.914	-3.937
	[0.854]**	[0.826]**	[0.827]**	[0.827]***	[0.857]***	[0.857]***
Percent In-Migration	-0.571	-0.595	-0.599	-0.088	-0.130	-0.134
	[0.735]	[0.753]	[0.756]	[0.888]	[0.901]	[0.908]
Percent in Extreme Poverty		-0.851	-0.848		-0.307	-0.295
		[0.341]**	[0.34]**		[0.464]	[0.468]
Individual Level Covariates from 1984 Census						
Female	-0.350	-0.350	-0.350	-0.360	-0.360	-0.360
	[0.038]***	[0.038]***	[0.038]***	[0.038]***	[0.038]***	[0.038]***
Education Catagories (None/some Primary omitted)						
Completed Primary	-0.068	-0.064	-0.064	-0.063	-0.063	-0.063
	[0.045]	[0.046]	[0.046]	[0.046]	[0.046]	[0.046]
Completed Secondary	-0.007	0.000	-0.001	0.003	0.003	0.002
	[0.062]	[0.062]	[0.061]	[0.064]	[0.064]	[0.064]
Completed College	-0.302	-0.293	-0.293	-0.280	-0.280	-0.281
	[0.077]***	[0.076]***	[0.076]***	[0.076]***	[0.076]***	[0.075]***
Geographical Area						
Live in San Jose Metro Area	0.041	0.023	0.024			
	[0.056]	[0.057]	[0.057]			
Live in Rural Area	-0.143	-0.141	-0.141	-0.092	-0.092	-0.091
	[0.043]***	[0.044]***	[0.044]***	[0.056]	[0.056]	[0.056]
Wealth Group(Middle Wealth omitted)						
Low Wealth	0.055	0.067	0.069	0.071	0.072	0.077
	[0.049]	[0.05]	[0.052]	[0.049]	[0.049]	[0.05]
High Wealth	-0.116	-0.114	-0.114	-0.082	-0.082	-0.082
	[0.053]**	[0.053]**	[0.053]**	[0.06]	[0.06]	[0.06]
Insurance Status						
No insurance in 1984	-0.090	-0.090	-0.090	-0.080	-0.080	-0.080
	[0.033]***	[0.034]***	[0.034]***	[0.035]**	[0.035]**	[0.035]**
In Union						
Union	-0.193	-0.192	-0.192	-0.198	-0.198	-0.198
	[0.043]***	[0.043]***	[0.043]***	[0.043]***	[0.043]***	[0.043]***
Constant	-9.984	-9.881	-9.882	-9.963	-9.935	-9.942
	[0.188]***	[0.202]***	[0.203]***	[0.193]***	[0.201]***	[0.203]***
Canton Fixed Effects	NO	NO	NO	YES	YES	YES
Log-Likelihood	-831.16	-828.11	-828.09	-780.05	-779.88	-779.66
Gamma	0.09	0.09	0.09	0.1	0.1	0.1
Observations	84503	84503	84503	84503	84503	84503
* significant at 10%; ** significant at 5%; *** significant at 1%						
Robust clustered standard errors in brackets						
Controls for age, age squared and age cubed included						

Table 4.8: Gompertz Models of the Risk of Heart Disease Mortality and Income Inequality

	Heart Disease Mortality					
Canton Level Area Covariates from July Household Surveys- Time varying						
Gini of Monthly Household Equivalent Salary	-3.448	-5.320	-3.37934	-8.345	-9.799	-7.708
	[1.118]***	[1.526]***	[1.313]**	[2.988]***	[3.408]***	[3.126]**
Mean of the Monthly Household Equivalent Salary	0.007	0.009	0.00683	0.021	0.022	0.021
	[0.001]***	[0.002]***	[0.001]***	[0.004]***	[0.004]***	[0.004]***
Interaction of Wealth & Area Level Gini						
Gini*Low Wealth			-0.16356			-1.688
			[1.549]			[1.847]
Percent Unemployed	7.489	8.312	7.49004	16.106	16.156	16.106
	[2.387]***	[2.636]***	[2.384]***	[5.677]***	[5.739]***	[5.644]***
Percent In-Migration	2.960	3.295	2.95346	6.363	6.151	6.411
	[1.641]*	[1.785]*	[1.657]*	[4.841]	[5.037]	[4.817]
Percent in Extreme Poverty		3.125			5.569	
		[1.226]**			[2.648]**	
Individual Level Covariates from 1984 Census						
Female	-0.320	-0.324	-0.31978	-0.311	-0.313	-0.310
	[0.068]***	[0.067]***	[0.067]***	[0.067]***	[0.067]***	[0.067]***
Education Catagories (None/some Primary omitted)						
Completed Primary	-0.193	-0.207	-0.19298	-0.158	-0.158	-0.15839
	[0.105]*	[0.105]**	[0.106]*	[0.105]	[0.105]	[0.105]
Completed Secondary	-0.079	-0.108	-0.07958	0.003	0.004	0.00113
	[0.155]	[0.156]	[0.156]	[0.152]	[0.152]	[0.153]
Completed College	-0.668	-0.702	-0.66871	-0.618	-0.621	-0.620
	[0.124]***	[0.125]***	[0.125]***	[0.124]***	[0.125]***	[0.125]***
Geographical Area						
Live in San Jose Metro Area	-0.465	-0.401	-0.46452			
	[0.139]***	[0.145]***	[0.14]***			
Live in Rural Area	0.035	0.023	0.03498	-0.021	-0.020	-0.022
	[0.091]	[0.093]	[0.091]	[0.104]	[0.104]	[0.103]
Wealth Group(Middle Wealth omitted)						
Low Wealth	-0.084	-0.125	-0.08413	-0.087	-0.087	-0.087
	[0.093]	[0.099]	[0.092]	[0.088]	[0.088]	[0.085]
High Wealth	0.018	-0.001	0.01787	0.063	0.063	0.063
	[0.092]	[0.092]	[0.091]	[0.093]	[0.093]	[0.093]
Insurance Status						
No insurance in 1984	-0.174	-0.178	-0.17449	-0.183	-0.185	-0.183
	[0.076]**	[0.078]**	[0.076]**	[0.076]**	[0.077]**	[0.076]**
In Union						
Union	-0.066	-0.066	-0.06585	-0.063	-0.064	-0.062
	[0.078]	[0.078]	[0.078]	[0.077]	[0.077]	[0.077]
Constant	-7.141	-7.558	-7.14136	-9.285	-9.704	-9.312
	[0.325]***	[0.421]***	[0.324]***	[0.519]***	[0.61]***	[0.526]***
Canton Fixed Effects	NO	NO	NO	YES	YES	YES
Log-Likelihood	-1147.45	-1135.98	-1147.45	-1038.91	-1027.51	-1038.3
Gamma	-0.17	-0.16	-0.17	-0.16	-0.15	-0.16
Observations	82187	82187	82187	82187	82187	82187
* significant at 10%; ** significant at 5%; *** significant at 1%						
Robust clustered standard errors in brackets						
Controls for age, age squared and age cubed included						

Table 4.9: Comparison of Gompertz and Cox Proportional Hazard Models

All Cause Mortality	Cox PH	Gompertz	Cox PH	Gompertz
Canton Level Area Covariates from July Household Surveys- time varying				
Gini of Monthly Household Equivalent Salary	-1.155	-1.206	-1.133	-1.179
	[0.36]***	[0.364]***	[0.365]***	[0.37]***
Mean of the Monthly Household Equivalent Salary	0.00036	0.00035	0.00046	0.00046
	[0.00034]	[0.00034]	[0.00042]	[0.00042]
Percent Unemployed			-1.913	-1.986
			[0.849]**	[0.854]**
Percent In-Migration			-0.607	-0.571
			[0.751]	[0.735]
Individual Level Covariates from 1984 Census- Non time varying				
Female	-0.351	-0.349	-0.352	-0.350
	[0.037]***	[0.038]***	[0.037]***	[0.038]***
Education Catagories (None/some Primary omitted)				
Completed Primary	-0.073	-0.068	-0.072	-0.068
	[0.045]	[0.045]	[0.046]	[0.045]
Completed Secondary	-0.012	-0.006	-0.013	-0.007
	[0.062]	[0.062]	[0.062]	[0.062]
Completed College	-0.302	-0.297	-0.306	-0.302
	[0.076]***	[0.076]***	[0.076]***	[0.077]***
Geographical Area				
Live in San Jose Metro Area	0.035	0.039	0.036	0.041
	[0.051]	[0.05]	[0.057]	[0.056]
Live in Rural Area	-0.122	-0.128	-0.136	-0.143
	[0.044]***	[0.042]***	[0.045]***	[0.043]***
Wealth Group(Middle Wealth omitted)				
Low Wealth	0.046	0.057	0.045	0.055
	[0.049]	[0.049]	[0.049]	[0.049]
High Wealth	-0.112	-0.112	-0.115	-0.116
	[0.053]**	[0.053]**	[0.053]**	[0.053]**
Insurance Status				
No insurance in 1984	-0.080	-0.091	-0.079	-0.090
	[0.031]**	[0.033]***	[0.032]**	[0.033]***
In Union				
Union	-0.188	-0.193	-0.188	-0.193
	[0.042]***	[0.042]***	[0.042]***	[0.043]***
Constant		-9.971		-9.984
		[0.18191]***		[0.18779]***
Log-Likelihood	-31357.72	-836.35	-31352.85	-831.16
Gamma		0.09		0.09
Observations	84503	84503	84503	84503
Robust clustered standard errors in brackets				
* significant at 10%; ** significant at 5%; *** significant at 1%				
Controls for age, age squared and age cubed included				

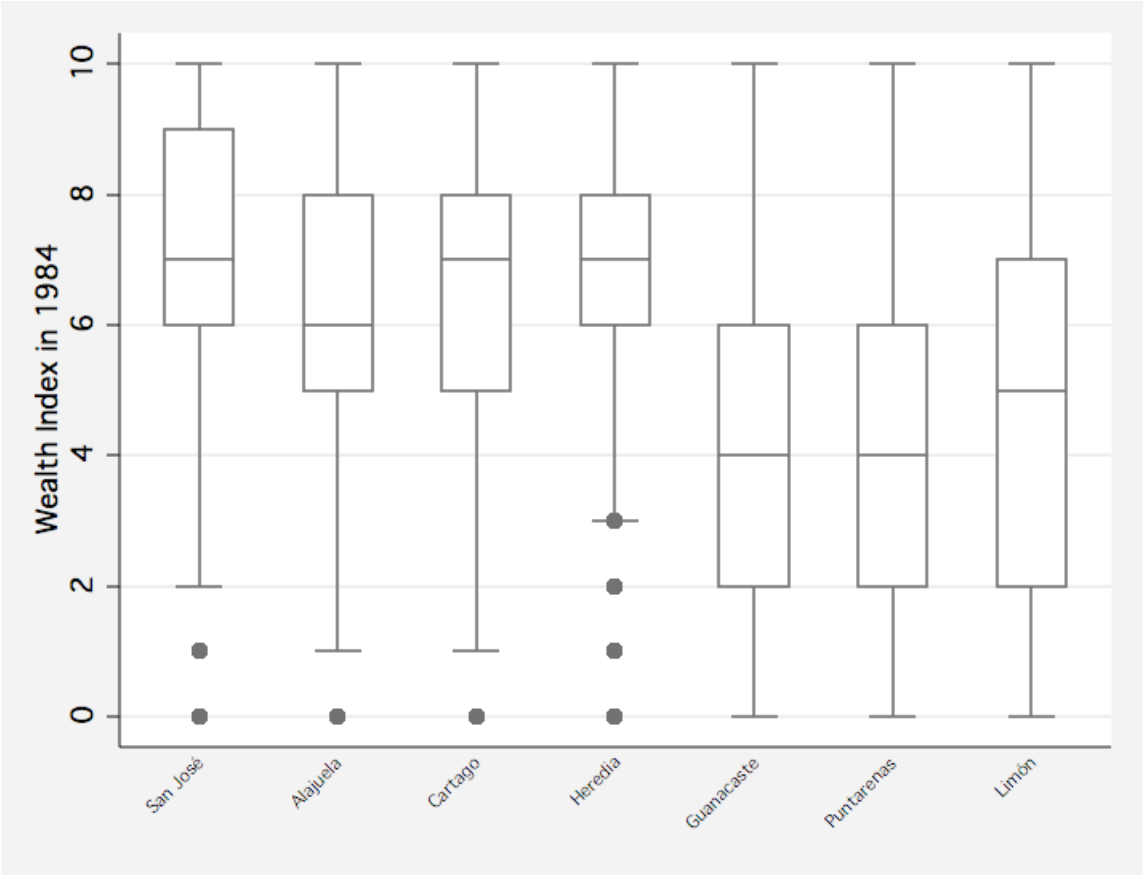
Table 4.10: Comparison of Different Parametric Models

Panel A						
Base Model Comparison of Parametric Survival Models- All-Cause Mortality	Obs	ll(null)	ll(model)	df	AIC	BIC
Exponential	85638	-4103.883	-1453.579	13	2933.159	3054.811
Gompertz	85638	-1008.774	-935.6253	14	1899.251	2030.261
Loglogistic	85638	-1291.046	-1103.207	14	2234.414	2365.424
Weibull	85638	-1050.4	-953.9503	14	1935.901	2066.911
Lognormal	85638	-1496.007	-1143.304	14	2314.609	2445.619
Gamma	85638	-1019.364	-946.9749	15	1923.95	2064.318
Panel B						
Base Model Comparison of Parametric Survival Models- Heart Disease Mortality	Obs	ll(null)	ll(model)	df	AIC	BIC
Exponential	83279	-2482.829	-1634.707	13	3295.413	3416.703
Gompertz	83279	-2148.723	-1246.609	14	2521.218	2651.837
Loglogistic	83279	-2149.654	-1672.64	14	3373.28	3503.899
Weibull	83279	Does Not Converge				
Lognormal	83279	Does Not Converge				
Gamma	83279	Does Not Converge				

Table 4.11: Gompertz Models of the Risk of Mortality and Individual-Level Relative Deprivation

	All- Cause Mortality			Heart Disease Mortality		
Individual Level Covariates from 1984 Census						
Relative Deprivation Measure		0.065	0.061		0.071	0.057
		[0.018]***	[0.02]***		[0.036]*	[0.047]
Female	-0.346	-0.343	-0.351	-0.374	-0.370	-0.377
	[0.039]***	[0.038]***	[0.038]***	[0.06]***	[0.06]***	[0.062]***
Education Catagories (None/some Primary omitted)						
Completed Primary	-0.079	-0.072	-0.058	-0.179	-0.172	-0.162
	[0.048]	[0.049]	[0.05]	[0.103]*	[0.102]*	[0.102]
Completed Secondary	-0.010	-0.002	0.020	0.050	0.056	0.051
	[0.06]	[0.06]	[0.061]	[0.154]	[0.153]	[0.15]
Completed College	-0.323	-0.314	-0.281	-0.590	-0.582	-0.579
	[0.072]***	[0.073]***	[0.072]***	[0.139]***	[0.138]***	[0.129]***
Geographical Area						
Live in San Jose Metro Area	0.091	0.050		0.100	0.058	
	[0.037]**	[0.04]		[0.112]	[0.118]	
Live in Rural Area	-0.108	-0.102	-0.090	-0.086	-0.080	-0.075
	[0.04]***	[0.041]**	[0.055]*	[0.082]	[0.083]	[0.109]
Wealth Group(Middle Wealth omitted)						
Low Wealth	0.034	-0.073	-0.054	-0.096	-0.210	-0.124
	[0.048]	[0.056]	[0.058]	[0.11]	[0.14]	[0.165]
High Wealth	-0.100	-0.056	-0.021	0.041	0.088	0.069
	[0.05]**	[0.051]	[0.059]	[0.105]	[0.109]	[0.119]
Insurance Status						
No insurance in 1984	-0.099	-0.106	-0.097	-0.228	-0.236	-0.226
	[0.032]***	[0.032]***	[0.035]***	[0.074]***	[0.073]***	[0.076]***
In Union						
Union	-0.201	-0.197	-0.203	-0.121	-0.117	-0.106
	[0.039]***	[0.039]***	[0.039]***	[0.081]	[0.081]	[0.084]
Constant	-9.932	-9.972	-10.060	-11.774	-11.819	-11.965
	[0.202]***	[0.202]***	[0.218]***	[0.407]***	[0.407]***	[0.389]***
Canton Fixed Effects	NO	NO	YES	NO	NO	YES
Log-Likelihood	-1080.15	-1074.64	-1027.58	-1531.34	-1529.88	-1483.2
Gamma	0.09	0.09	0.09	0.1	0.1	0.1
Observations	16173	16173	16173	16173	16173	16173
Robust clustered standard errors in brackets						
* significant at 10%; ** significant at 5%; *** significant at 1%						
Controls for age, age squared and age cubed included						

Figure 4.1: Wealth Index Box Plots by Province



Note: Wealth index calculated for Costa Ricans over age 30. Author's calculations based on 1984 census-10% micro sample from IPUMS

Figure 4.2: Kaplan-Meier Survival Curves for All-Cause Mortality

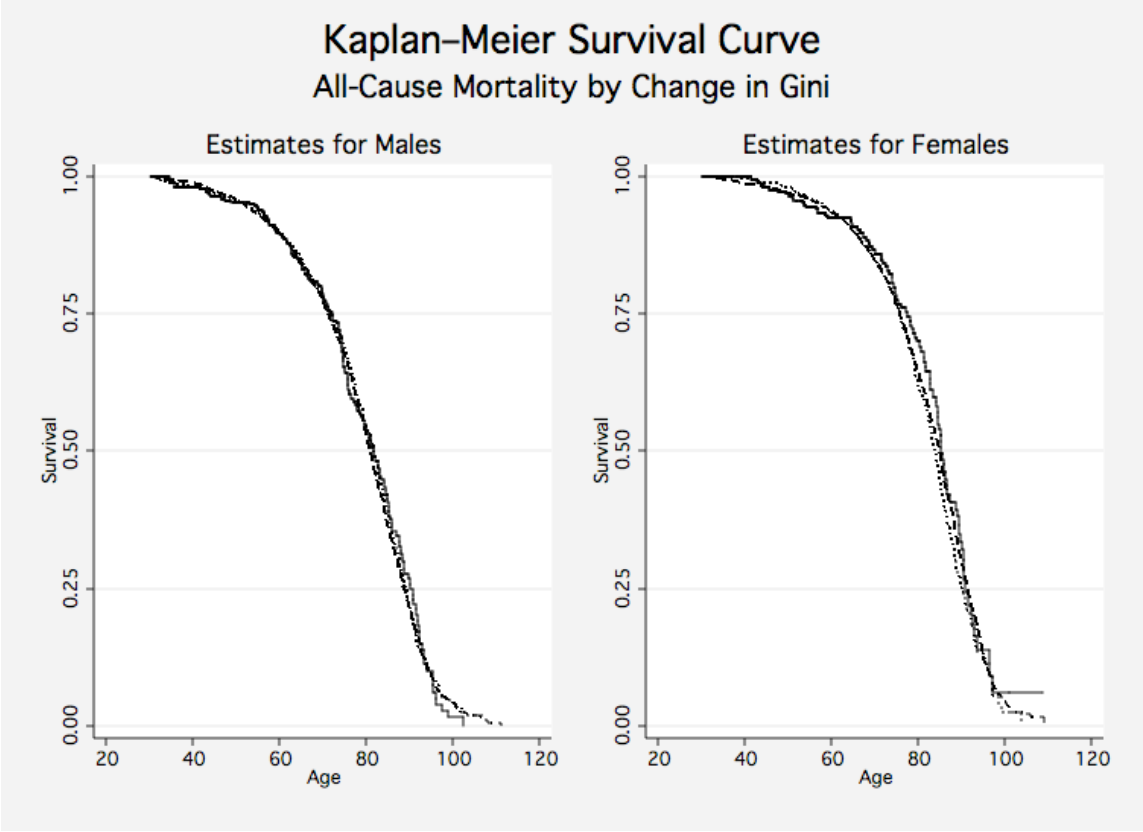
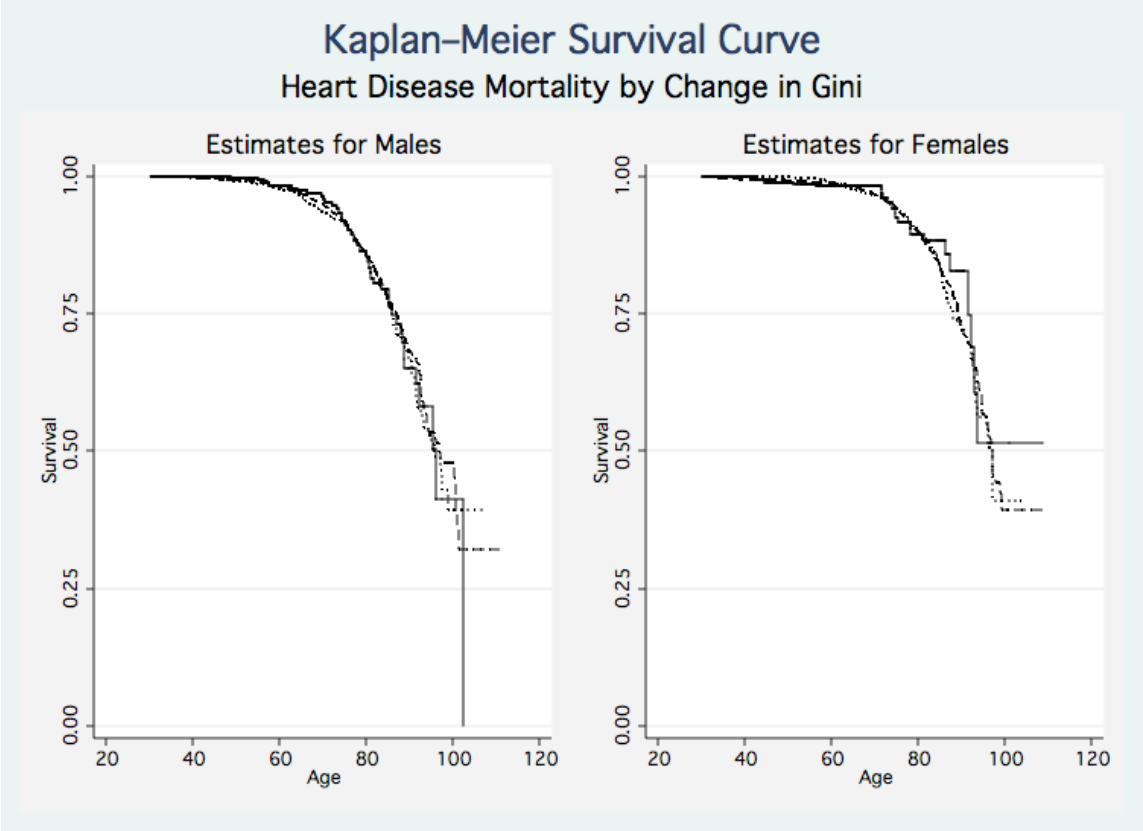


Figure 4.3: Kaplan-Meier Survival Curves for Heart Disease Mortality



Chapter 5: Conclusions

The purpose of this study has been to test the income inequality hypothesis, which argues that income inequality, as a measure of social hierarchy within a group, leads to poor health. Costa Rica is a middle-income country with limited ethnic diversity and a highly centralized political system. As such, it provides a context where the indirect effects of income inequality on health, such as through historical discrimination or local resource allocation channels, are less likely to confound the income inequality and health relationship. The study assesses the relationship between income inequality and health at both the ecological and individual-levels using longitudinal income inequality measures constructed from income data in repeated cross-sections of household surveys. In general, these analyses provide little support for either the strong versions of the income inequality hypotheses, which argue that all individuals are affected by income inequality, or for the weak version, which argues that only those lower on the hierarchy are affected by income inequality. In contrast, the results summarized in Table 5.1 suggest that in this setting residents of areas with increasing inequality experienced reduced mortality, even among diseases often associated with stress, such as cardiovascular disease. These findings imply that income inequality does not have a negative impact on health in this middle-income country where the existence of a strong safety net and political centralization may buffer the relationship.

5.1 Summary of Results

Chapter 2 described the measures of income inequality used throughout this study. Rising income inequality was determined to be affected by the increased monthly earning of individuals at the top of the income distribution. Next to be considered was the association between increasing income inequality and the degradation of social trust. Social trust is thought to be an important determinant of health that is also affected by increasing income inequality. Consequently, the study investigated whether or not areas with rising inequality experienced diminishing social trust (H1). Using an average area-level measure of interpersonal trust as a proxy for cognitive social capital, as was done previously (Kawachi, Kim, Coutts, & Subramanian, 2004), this study did not find any association between income inequality and the measure of social trust used. These results might be expected because income inequality in this case was not due to increasing poverty, but rather to the rising affluence of the rich. However, changes in cognitive social capital were examined only from 2002-2006, which might be too short of a time frame to observe such changes. Furthermore, the confidence intervals in these regressions are quite large indicating that it is not possible to rule out the existence of a relationship.

Chapter 3 first presents the ecological association between inequality and canton level all-cause mortality. Across both models without and with canton fixed-effects, no relationship was established between income inequality and all-cause mortality for either those aged 15-59 or for those aged 60 and over. These findings are especially interesting because they contrast the findings of Li & Zhu (2006) and Subramanian et al. (2003) who do find that adults in China and Chile are more likely to report poor health in areas with higher inequality.

In order to understand potential pathways through which income inequality affects health, Chapter 3 next presents the ecological association of income inequality

with 18 distinct causes of death. While the results differed by specific cause, the area-level income inequality effects, either positive or negative, were concentrated in those aged 15-60 as opposed to those aged 60+. Increases in area-level income inequality were associated with increases in adult suicide rates, which supports the main hypothesis of this dissertation (H2), but this was the only cause where income inequality had a negative impact. For other causes such as breast cancer (ages 60+), cardiovascular disease (ages 15-60), and respiratory cancers (ages 15-60), increases in income inequality are associated with lower mortality, which is contrary to the main hypothesis (H2) of this research project. Accordingly, this study found no support for its third hypothesis (H3), that income inequality is positively associated with stress-related diseases such as cardiovascular disease; and in fact there was some evidence for the opposite.

The final analytical chapter examines the individual-level association between residence in an area with increasing income inequality between all-cause and heart-disease mortality. Again, once area-level fixed effects were included, no association was established for all-cause mortality (no support for H2). Instead, a statistically significant lower hazard rate for heart-disease mortality among those living in areas with increasing inequality was observed (contradiction of H3). Additionally, the analyses of the weak version of the income inequality hypothesis with regard to poverty did not provide any support for this version of the hypothesis (H4). Finally, those who are more deprived compared to their community exhibited lower survival in the cross section.

5.2 Pathways and Mechanism

The results above do not indicate a clear or simple conclusion. Yet, they offer a number of insights into pathways and suggest avenues for further investigation.

5.2.1 Psychosocial Pathway

The psychosocial pathway discussed above has several facets. In dissecting this pathway we need to consider how repeated negative thoughts or feelings and their subsequent biological responses are exacerbated by poor health behavior as well as buffered by social support. Momentarily ignoring the moderators and focusing on a purely biological pathway resulting from a repeated negative thoughts or feelings, we expect the negative impact of income inequality to be universal across human populations. In order to support such a purely biological mechanism, we would expect to find evidence for it in all settings, regardless of the existence of social safety nets. Nevertheless, this study does not find strong support for this piece of the pathway. Yet, it does find some evidence for mechanisms through which income inequality "gets into your head" through negative thoughts, but it finds no evidence to support the theory that such inequality affects a broad set of health outcomes. It may very well be that there is indeed a psychobiological response to income inequality, but in the Costa Rican setting there appear to be ample buffers to limit excess mortality due to this response. A wide array of social support measures and biological data not considered in these analyses would be necessary to distinguish between buffering mechanisms that mediate the pathway.

The study found some evidence of behavioral changes consistent with a psychosocial pathway, but only in extreme circumstances. Higher suicide rates in areas with higher income inequality are consistent with a mechanism that acknowledges mental

distress is accompanied by a negative behavioral response. However, this may only account for the most susceptible/vulnerable portions of the population, and thus it suggests that income inequality may exacerbate pre-existing mental health conditions for some; but it has limited effects on population health in general.

The final component of the psychosocial pathway points to the breakdown of social capital and social support in areas with increasing income inequality. To test this hypothesis, the association between social capital and income inequality was examined as a mediating pathway. However, the specific measures of social capital were not statistically significantly associated with income inequality. This is likely due to the large standard errors in the relationship when fixed-effects were added, but the directionality of the magnitude of the relationship was as expected. Moreover, in previous analyses in a nationally representative sample of elderly Costa Ricans, this measure of cognitive social capital was not associated with area-level income inequality. This may be due to the measure, which is meant to capture only the cognitive component of social capital. Although this measure is the one that theory predicts to be most susceptible to changes in income inequality, other measures, such as the density of social organizations, deserve scrutiny. Better measurement with larger samples must be employed to confirm the findings here.

5.2.2 Resource Distribution Pathway

Another important mechanism previously discussed concerned the distribution of resources within a community. The importance of this pathway in Costa Rica where resources are not distributed by local municipalities is assumed to be relatively small. While this may be the case, there may be some selective sorting by physicians, which could suppress the negative health effects of income inequality, and thus warrant further investigation. Because areas in Costa Rica with higher income inequality are associated with greater wealth, there might be selection among doctors, even public general practitioners, who may prefer to work in more affluent areas.

An analysis of physician outcomes by area would be a necessary first step to examine whether differences in the provision of health services exist in more unequal areas. If systematic differences are observed, a secondary step will be necessary to distinguish between physician practice styles and patient behavior/preferences in more unequal areas, as more affluent patients may seek more preventative or better quality health services (Brenes-Camacho & Rosero-Bixby, 2009).[78] If it is found that it is indeed the difference in practices of physicians rather than patient preferences, then these differences may account for the some of the observed results with respect to heart disease mortality in the individual level analyses.[79]

[78] The results from the Brenes-Camacho & Rosero-Bixby study highlight the difficulty in addressing whether institutional characteristic versus individual preferences determine health utilization. They find that if an individual had a home visits in the past year (an institutional area-level characteristic), then these individuals are more likely to seek preventative services. Moreover, they find lower SES individuals are more likely to be visited. Yet they also find that higher SES individuals are more likely to seek preventative services. These types of details need to be considered in the secondary analysis suggested.

[79] There is a large and long-standing body of evidence pointing to different practice styles across areas in the United States (Wennberg & Gittlesohn, 1973).

5.2.3 Neomaterial Pathway

The final pathway considered in Chapter 1 posited that other changes in society drive the relationship between income inequality and health that has been observed previously. As suggested above, this mechanism requires that we distinguish between the causes of income inequality that also drive health outcomes and the effect of income inequality *itself*. It is important to note that while Costa Rica chose to engage in trade liberalization gradually, the study considered a period during a time of globalization and privatization. Consequently, wider societal changes must be considered. The evidence from these analyses indicates that Costa Rica globalized slowly enough to avoid any apparent negative consequences with the exception of suicide. However, it may well be that gains from globalization (such as increased wealth, greater industrial capacity, and economic development) are driving the observed positive results. Likewise to the extent that income inequality is associated with economic development and health behaviors in Costa Rica, it might lead to better health outcomes. For example if areas with increasing inequality are also areas with decreasing smoking rates, then such practices might explain the reduced mortality rates observed for cardiovascular disease and lung cancer, both of which are caused by smoking. Thus, it is possible that other unidentified behaviors may explain some of the health benefits found. This may be true with the exception of violence; the literature on violence does point to specific mechanisms through which income inequality *itself* increases violence within a community (J. Lynch et al., 2004a).

5.3 Interpretations

In the previous section, the empirical results are discussed within the framework of the income inequality hypothesis and the conceptual model described in Chapter 1. Yet given the limited support and mixed results found for the income inequality hypothesis in these analyses, this hypothesis requires further clarification.

Most of the support for the income inequality hypothesis has come from settings where one expects to find an association. For example, studies set in the United States, Brazil, or Chile are less informative because these settings also have deep societal divisions that have been propagated for generations and that are encapsulated in measures like income inequality. If we had found evidence for the income inequality hypothesis in a place like Costa Rica, which does not have the same historical experiences, then the results would have been very meaningful. Given that we did not, what does this really mean for the hypothesis?

Amartya Sen's idea of capability set comes to mind as a reconciliatory notion. Unlike metrics that evaluate well-being based on increases in income, Sen's capability set includes the elimination of oppression and discrimination as well as the provision of basic education, health care, and social safety nets (Sen, 1995). A country may have large income inequality (as Costa Rica does), but if it also attempts to fulfill individuals' capability sets, then income inequality or heterogeneity may not be deleterious to society. In places where income inequality is a marker for missing components of the capability set, particularly social mobility, then income inequality might be capturing the lack of those capabilities. While Costa Rica is relatively poor, especially compared to the United

States, it has made efforts to ensure a basic capability set,[80] and consequently in this case we must reinterpret what inequality or any measure of heterogeneity is capturing.

In this specific context, we must also consider how heterogeneity can be protective as is suggested in the results. Sorting based on income could affect the distribution of goods and services as well as social networks. Heterogeneity may promote health to the extent that it increases the availability of goods and services that encourage healthy behavior or increase interactions with individuals who exhibit healthy behavior (Galea & Ahern, 2005). For example, if close proximity to a wealthy person who may be more likely to undergo preventative health exams also makes a poor person more likely to do the same, then heterogeneity would be beneficial (Banerjee, Cohen-Cole, & Zanella, 2007). In this way, social or neighborhood interactions in unequal areas can result in improved health outcomes.

5.4 Contributions

There are three main contributions of this study. Methodologically, these analyses employed nationally representative data and a longitudinal design to assess the income inequality hypothesis in a middle-income country. The data used, especially the census-mortality linkage, allowed for a 20-year follow-up, which is rare in this literature. The longitudinal nature of the study addresses causality better than cross-sectional studies.[81] Also, results were compared and validated across different levels of analysis because both ecological and individual-level analyses were undertaken. Given that the results across levels largely agreed for all-cause and heart disease mortality, the other results found in the ecological analyses are further supported.

Second, Costa Rica's particular features create an ideal milieu for testing the income inequality hypothesis. As a setting characterized by low ethnic diversity, high political centralization, a long-standing democracy, and high levels of gender equality, it is the exception in many ways. Yet, it is precisely the combination of these characteristics that limit the scope for confounding from other economy-wide socioeconomic processes. Hence, if the income distribution even in this setting were found to be related to health, the propositions in the income inequality hypothesis would have been strengthened considerably. Instead, the null findings suggest that studies in other settings may be biased because inequality may be a marker for economy-wide social divisions or because income inequality is acting through the allocation of local resources. Therefore, regardless of the results it was important to investigate this setting.

This study is particularly relevant with respect to other studies that examine the income inequality hypothesis in middle-income countries. Compared to studies that find a negative impact of income inequality on health in countries such as China or Chile, this dissertation does not. First, this study examines mortality while the previous studies examine self-reported health. Yet Fiscella and Franks find that in the United States even when income inequality is found to be associated with self-reported health, it is not necessarily associated with mortality (Fiscella & Franks, 2000). They argue that while income inequality may affect perceptions of health through mental health, it does not

[80] This commitment to redistribution is apparent in the tax to GDP ratio presented in Table 1.2. Costa Rica has higher tax revenues in terms of GDP than its neighbors or the United States.

[81] However, this since one chooses one's location of residence, this study cannot address causality completely.

affect mortality. This may also be the case in middle-income countries. Secondly, while these studies find a negative association between income inequality and health, most of their analyses are cross-sectional and may be more prone to bias. Finally, the examples of China and Chile may also be particular because in the case of Chile political channels may be related to both income distribution and health, whereas rapid economic growth in China may be causing many societal changes simultaneously.

Substantively, this study challenges certain aspects of the income inequality hypothesis while confirming others. In its current formulation, some studies suggest that the income inequality hypothesis inappropriately attributes poor health outcomes to inequality, where in fact correlates of income inequality in certain settings may cause the observed poor health. This study also finds evidence for this notion. Nonetheless, there is consensus that the hypothesis may be appropriate for certain violent behaviors. In finding evidence of a higher suicide rate in areas with increasing income inequality, this study supports previous findings on suicide for which there has been some consensus in the developed world (J. Lynch et al., 2004a; J. R. Miller et al., 2005; R. G. Wilkinson & Pickett, 2006). The novel setting further of this dissertation requires us to consider how developing countries, even ones with safety nets, must address mental health issues associated with shifts in income distribution.

5.5 Limitations

As with other studies in this literature, this study used a constructed reference group. It assumed that a canton is a reasonable reference group, but this geographical interpretation may not capture individuals' perceived income inequality status. For example, an individual's conceptualization of community need not span a geographical area, and even if it does, the size of that area may vary from person to person. In a recent study, where individuals identified themselves on a scale compared to their own social network, there was some evidence for a relative deprivation effect (not income inequality) but only for the extremely deprived and the extremely wealthy (Pham-Kanter, 2009). The study finds a relationship between relative position and health status only for those who identify themselves as occupying the very top or the very bottom of their social group. It also indirectly supports a biological mechanism underlying the relative deprivation model. Ideally, future studies will attempt to consider the reference group more carefully and collect data on perceived social standing.

Another limitation of this study is that it only examines the income inequality hypothesis but not related hypotheses that consider social hierarchy generally. More recent research has focused on status while other research has focused on relative deprivation. In the final chapter some limited analyses of the relative deprivation hypothesis were performed, but a full exploration is beyond the scope of this thesis. As the literature moves forward, these types of hypotheses, which may be more fruitful, should be pursued.

Correctly specifying or determining the timing over which income inequality affects health is another concern. While the longitudinal analyses employed limited confounding, the method may not have captured the dynamics through which the timing of income inequality changes affects health. The identification strategy exploited the variation of changes in contemporaneous income inequality within cantons over five-year periods. Yet, income inequality could exert its noxious effects over a longer time period,

or the timing of the effects could vary by condition and the cause of death. Future studies should inspect whether 10, 15 or 20 year lagged inequality measures relate more consistently to health outcomes.

The only outcome measure considered was mortality, whereas health is a multi-dimensional construct. While this outcome has the benefits of including the universe of registered deaths and being an objective measure of health, it may be too coarse a measure of health. We saw that once mortality was broken down by cause, there was significant variation in its association with income inequality. Hence we would expect to see even more variation if we had considered measures of morbidity. Measures of morbidity and mental health as well as biological measures that have clinical significance may help clarify the complex web of pathways advanced above.

5.6 Further Research

The consistent findings that income inequality was associated with higher suicide rates and lower cardiovascular mortality while not associated with all-cause mortality raise a number of questions for further theoretical development of inequality theories as well as empirical testing of such abstract social concepts.

Given the results found for suicide, further exploration of violent or external causes of death may be fruitful. The first step may be looking beyond the inequality in one's own canton. A measure of changes in income inequality in surrounding cantons could be created for each canton, and the association between this new measure of inequality and cause specific mortality rates could be explored. If such an analysis found a relationship between inequality in the surrounding areas for the violent causes of death but none for the other causes, then the results would further the argument made above that violent behavior is associated with income inequality *itself*, but other health behaviors like diet, smoking, exercise, and drinking may not be.

Considering the surprising results with respect to cardiovascular mortality, in-depth investigation of health behaviors in areas with growing inequality is necessary. In an on-going study of elderly Costa Rican, Rosero-Bixby and Dow find that wealthy Costa Ricans are adopting worse nutritional behaviors. They suggest that Costa Rica is undergoing a nutritional transition from healthy traditional food to unhealthy processed food. Additionally, more affluent individuals may be adopting poor eating habits which may result in increasing cardiovascular risk (Luis Rosero-Bixby & Dow, 2009b). As mentioned above, areas experiencing economic development were also the areas with increasing income inequality. As such this nutritional transition hypothesis could be examined with respect to income inequality.

Another set of analyses to understand the lower cardiovascular mortality in more unequal areas would require an examination of the quality of doctors and the organization of health clinics in more unequal areas. Currently there is a study of the health system in 60 of the 100 health areas in Costa Rica to assess the quality of health clinics, to measure how the number of home visits, and to measure the number of treatments conducted by the local health teams. An examination of these data could uncover systematic differences in areas exhibiting changes in income inequality. To the extent that there are systematic differences in physician practice style, these differences may be related to the results found above.

Finally to further explore the null results on all-cause mortality, another set of analyses might explore the relative deprivation hypothesis using subjective social standing as does the Pham-Kanter study (Pham-Kanter, 2009). There is currently an on-going survey of the elderly in Costa Rica that measures subjective social standing using a ladder ranking. This same study collects a variety of objective, subjective, and mental health outcomes, as well as mortality. Thus Pham-Kanter's analysis could be conducted in the Costa Rican context, and such a study could further clarify the biological pathway.

Although the results of this study generally do not support the association between income inequality and health outcomes, several important avenues remain to be explored. We have only begun to understand Costa Rica's exceptional health outcomes and their relations to income inequality, or to social hierarchy more broadly; and this topic merits continued scrutiny in future work.

Table 5.1: Summary of Hypotheses

	Hypothesis 1	Hypothesis 2	Hypothesis 3	Hypothesis 4
Ecological- Level				
Social Trust	**Not Confirmed**			
All-Cause Mortality		**Not Confirmed**		
Cause-Specific Mortality		**Mixed Results By Cause** Income inequality is associated with higher suicide mortality rates and *lower* breast cancer, cardiovascular disease, and lung cancer mortality rates	**Opposite** Income inequality is associated with *lower* cardiovascular disease mortality rates	
Individual-Level				
All-Cause Mortality		**Not Confirmed**		**Not Confirmed**
Heart Disease Mortality			**Opposite** Income inequality is associated with lower heart disease mortality	**Not Confirmed**

H1: Income inequality is negatively associated with measures of social capital.
H2: Higher income inequality is associated with higher all-cause mortality.
H3: Higher income inequality is associated with higher mortality for stress related causes of death (such as cardiovascular disease).
H4: The individual-level association between area-level income inequality and health may be stronger for more vulnerable populations such as the poor.

References:

A Rodríguez-Clare, M Sáenz, & Trejos, A. (2002). Economic Growth in Costa Rica: 1950-2000. (p. 97). Washington, DC,: United States: Inter-American Development Bank, Research Department

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., & Folkman, S. (1994). Socioeconomic status and health: the challenge of the gradient. *American Psychologist,* 49(1), 15-24.

Ahern, J., & Galea, S. (2006). Social context and depression after a disaster: the role of income inequality. *Journal of Epidemiology and Community Health,* 60(9), 766-770.

Akaike, H. (1981). Likelihood of a model and information criteria. *Journal of Econometrics,* 16(1), 3-14.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth,* 8(2), 155-194.

Alesina, A., & Perotti, R. (1994). The Political Economy of Growth: A Critical Survey of the Recent Literature. *World Bank Econ Rev,* 8(3), 351-371.

Allison, P. (1984). Event History Analysis. Regression for Longitudinal Event Data. Newbury Park: SAGE Publications.

Anderson, M. (2008). Statistical Inference: Part 1, Panel Data and Clustering. Berkeley, CA.

Anderson, M., & Marmot, M. (2007). The Effects of Social Status on Heart Disease: Evidence from Whitehall.

Asafu-Adjaye, J. (2004). Income Inequality and Health: A Multi-Country Analysis. Brisbane, Australia: The University of Queensland.

Banerjee, R., Cohen-Cole, E., & Zanella, G. (2007). Demonstration Effects in Preventive Care. FRB of Boston Quantitative Analysis Unit Working Paper No. QAU7-06: SSRN.

Baum, A., Garofalo, J. P., & Yali, A. M. (1999). Socioeconomic Status and Chronic Stress: Does Stress Account for SES Effects on Health? *Annals of the New York Academy of Sciences,* 896(1), 131-144.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics,* 119(1), 249.

Bezruchka, S. (2009). The effect of economic recession on population health. *Canadian Medical Association Journal,* 181(5), 281-285.

Blakely, T., Atkinson, J., & O'Dea, D. (2003). No association of income inequality with adult mortality within New Zealand: a multi-level study of 1.4 million 25-64 year olds. *Journal of Epidemiology and Community Health,* 57(4), 279-284.

Brenes-Camacho, G., & Rosero-Bixby, L. (2009). Differentials by Socioeconomic Status and Institutional Characteristics in Preventive Service Utilization by Older Persons in Costa Rica. *J Aging Health,* 21(5), 730-758.

Chong, A., & Gradstein, M. (2006). Inequality and Institutions. *Review of Economics and Statistics,* 89(3), 454-465.

Daly, M., Wilson, M., & Vasdev, S. (2001). Income inequality and homicide rates in Canada and the United States. *Canadian Journal of Criminology,* 43(2), 219-236.
Deaton, A. (2003). Health, Inequality, and Economic Development. *Journal of Economic Literature,* 41, 113–158.
Deaton, A., & Lubotsky, D. (2003). Mortality, inequality and race in American cities and states. *Social Science & Medicine,* 56(6), 1139-1153.
Deaton, A., & Paxson, C. (1998). Aging and Inequality in Income and Health. *American Economic Review,* 88(2), 248-253.
Dickerson, S., & Kemeny, M. (2004). Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research. *Psychological Bulletin,* 130(3), 355-391.
Dooley, D., Catalano, R., & Wilson, G. (1994). Depression and unemployment: Panel findings from the Epidemiologic Catchment Area study. *American Journal of Community Psychology,* 22(6), 745-765.
Drukker, M., Feron, F. J. M., & van Os, J. (2004). Income inequality at neighbourhood level and quality of life. *Social Psychiatry and Psychiatric Epidemiology,* 39, 457-463.
Eduardo, F. (2009). Information and Values in Popular Protests: Costa Rica in 2000. *Bulletin of Latin American Research,* 28(1), 44-62.
Eibner, C. E., & Evans, W. N. (2005). Relative Deprivation, Poor Health Habits and Mortality. *Journal of Human Resources,* 40(3), 591–620.
Fiscella, K., & Franks, P. (2000). Individual income, income inequality, health, and mortality: what are the relationships? *Health Serv Res,* 35(1), 307–318.
Foster, M. (2009). What are Equivalence Scales?
Franzini, L., & Spears, W. (2003). Contributions of social context to inequalities in years of life lost to heart disease in Texas, USA. *Social Science & Medicine,* 57(10), 1847-1861.
Galea, S., & Ahern, J. (2005). Distribution of Education and Population Health: An Ecological Analysis of New York City Neighborhoods. *Am J Public Health,* 95(12), 2198-2205.
Galea, S., Ahern, J., Vlahov, D., Coffin, P. O., Fuller, C., Leon, A. C., et al. (2003). Income distribution and risk of fatal drug overdose in New York City neighborhoods. *Drug and Alcohol Dependence,* 70(2), 139-148.
Gindling, T. H. (2009). South-South Migration: The Impact of Nicaraguan Immigrants on Earnings, Inequality and Poverty in Costa Rica. *World Development,* 37(1), 116-126.
Gindling, T. H., & Trejos, J. D. (2005). Accounting for Changing Earnings Inequality in Costa Rica, 1980-99. *Journal of Development Studies,* 41(5), 898 - 926.
Gingrich, C. D., & Garber., J. D. (2009). Trade Liberalization's Impact on Agriculture in Low Income Countries: A Comparison of El Salvador and Costa Rica. *The Journal of Developing Areas,* 43(2), 1-17.
Hou, F., & Myles, J. (2005). Neighbourhood inequality, neighbourhood affluence and population health. *Social Science & Medicine,* 60(7), 1557-1569.
Iaies, G. F. (2005). Education System Evaluation Program. Inter-American Development Bank.

Kaplan, G. A., Pamuk, E. R., Lynch, J. W., Cohen, R. D., & Balfour, J. L. (1996). Inequality in income and mortality in the United States: analysis of mortality and potential pathways. *BMJ,* 312(7037), 999-1003.

Kawachi, I., & Kennedy, B. P. (1999a). Income inequality and health: pathways and mechanisms. *Health Services Research,* 34(1 Pt 2), 215-227.

Kawachi, I., & Kennedy, B. P. (1999b). Income inequality and health: pathways and mechanisms. *Health Serv Res*.

Kawachi, I., Kennedy, B. P., Lochner, K., & Prothrow-Stith, D. (1997). Social capital, income inequality, and mortality. *Am J Public Health,* 87(9), 1491-1498.

Kawachi, I., Kim, D., Coutts, A., & Subramanian, S. V. (2004). Commentary: Reconciling the three accounts of social capital. *International Journal of Epidemiology,* 33(4), 682–690.

Kennedy, B. P., Kawachi, I., & Prothrow-Stith, D. (1996). Income distribution and mortality: cross sectional ecological study of the Robin Hood index in the United States. *BMJ,* 312(7037), 1004-1007.

Krugman, P. (1996). The Spiral of Inequality. Mother Jones pp. 44-49). San Francisco

Kuznets, S. (1955). Economic Growth and Income Inequality. *The American Economic Review,* 45(1), 1-28.

Larrea, C., & Kawachi, I. (2005a). Does economic inequality affect child malnutrition? The case of Ecuador. *Social Science & Medicine*.

Larrea, C., & Kawachi, I. (2005b). Does economic inequality affect child malnutrition? The case of Ecuador. *Social Science & Medicine,* 60(1), 165-178.

Lee, R. (2000). The Lee-Carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal,* 4(1), 80-93.

Li, H., & Zhu, Y. (2006). Income, income inequality, and health: Evidence from China. *Analyzing the Socioeconomic Consequences of Rising Inequality in China,* 34(4), 668-693.

Lynch, J., Smith, G. D., Harper, S. A. M., & Hillemeier, M. (2004a). Is Income Inequality a Determinant of Population Health? Part 1. A Systematic Review. *The Milbank Quarterly,* 82(1), 5-99.

Lynch, J., Smith, G. D., Harper, S. A. M., & Hillemeier, M. (2004b). Is Income Inequality a Determinant of Population Health? Part 2. US National and Regional Trends in Income Inequality and Age-and Cause-Specific Mortality. *The Milbank Quarterly,* 82(2), 355-400.

Lynch, J. W., Kaplan, G. A., Pamuk, E. R., Cohen, R. D., Heck, K. E., Balfour, J. L., et al. (1998). Income inequality and mortality in metropolitan areas of the United States. *Am J Public Health,* 88(7), 1074-1080.

Mahapatra, P., Shibuya, K., Lopez, A. D., Coullare, F., Notzon, F. C., Rao, C., et al. (2007). Civil registration systems and vital statistics: successes and missed opportunities. *The Lancet,* 370(9599), 1653-1663.

McEwen, B. S. (2004). Protection and Damage from Acute and Chronic Stress: Allostasis and Allostatic Overload and Relevance to the Pathophysiology of Psychiatric Disorders. *Annals of the New York Academy of Sciences,* 1032(1), 1-7.

McEwen, B. S., & Seeman, T. (1999). Protective and Damaging Effects of Mediators of Stress: Elaborating and Testing the Concepts of Allostasis and Allostatic Load. *Annals of the New York Academy of Sciences,* 896, 30-47.

Mellor, J. M., & Milyo, J. (2002). Income inequality and health status in the United States: evidence from the Current Population Survey. *Journal of Human Resources,* 37(3), 29.
Mesa-Lago, C. (2000). Market, Socialist, and Mixed Economies: Comparative Policy and Performance --Chile, Cuba and Costa Rica. Maryland: The John's Hopkins University Press.
Mesa-Lago, C. (2004). Social Development with Limited Resources. In S. Palmer & I. Molina (Eds.), The Costa Rica Reader: History, Culture, Politics. Durham, North Carolina: Duke University Press.
Miller, D., Page, M., Stevens, A., & Filipski, M. (2009). Why are recessions good for your health? *American Economic Review,* 99(2), 122-127.
Miller, J. R., Piper, T. M., Ahern, J., Tracy, M., & Tardiff, K. J. (2005). Income Inequality and Risk of Suicide in New York City Neighborhoods: A Multilevel Case-Control Study. *Suicide and Life-Threatening Behavior,* 35(4), 448-459.
Milyo, J., & Mellor, J. (2003). On the importance of age-adjustment methods in ecological studies of social determinants of mortality. *Health Services Research,* 38(6 Pt 2), 1781.
Muramatsu, N. (2003). County-Level Income Inequality and Depression among Older Americans. *Health Services Research,* 38(6), 1863-1884.
Palmer, S., & Molina, I. (2004a). Cabildo of Cartago: Free Blacks, Mulattoes, and Mestizos Seek Legitimacy. In S. Palmer & I. Molina (Eds.), The Costa Rica Reader: History, Culture, Politics. Durham, North Carolina: Duke University Press.
Palmer, S., & Molina, I. (2004b). Introduction. In S. Palmer & I. Molina (Eds.), The Costa Rica Reader: History, Culture, Politics. Durham, North Carolina: Duke University Press.
Palmer, S., & Molina, I. (2004c). Student Protests. In S. Palmer & I. Molina (Eds.), The Costa Rica Reader: History, Culture, Politics. Durham, North Carolina: Duke University Press.
Pattussi, M. P., Marcenes, W., Croucher, R., & Sheiham, A. (2001). Social deprivation, income inequality, social cohesion and dental caries in Brazilian school children. *Social Science & Medicine,* 53(7), 915-925.
Pham-Kanter, G. (2009). Social comparisons and health: Can having richer friends and neighbors make you sick? *Social Science & Medicine,* 69(3), 335-344.
Preston, S. H. (1975). The Changing Relation between Mortality and Level of Economic Development. *Population Studies,* 29(2), 231-248.
Profile of the Health Services System of Costa Rica. (2002). Health Systems and Services: Pan American Health Organization
Putnam, R. D. (2000). Bowling Alone: The Collapse and Revival of American Community. New York, NY Simon & Schuster.
Quiros, C. (2004). Hunting Indians. In S. Palmer & I. Molina (Eds.), The Costa Rica Reader: History, Culture, Politics. Durham, North Carolina: Duke University Press.
Redelmeier, D. A., & Singh, S. M. (2001a). Longevity of screenwriters who win an academy award: longitudinal study. *BMJ,* 323, 1491-1496.

Redelmeier, D. A., & Singh, S. M. (2001b). Survival in Academy Award-Winning Actors and Actresses. *Annals of Internal Medicine,* 134(10), 955-962.
Rodgers, G. B. (1979). Income and Inequality as Determinants of Mortality: An International Cross-Section Analysis. *Population Studies,* 33(2), 343-351.
Rosero-Bixby, L. (2004). Spatial access to health care in Costa Rica and its equity: a GIS-based study. *Social Science & Medicine,* 58(7), 1271-1284.
Rosero-Bixby, L. (2006). Social Capital, Urban Settings and Demographic Behavior in Latin America. *Population Review,* Volume 45(2), 24-43.
Rosero-Bixby, L., & Antich, D. (2009). Descripción de una muestra de adultos del censo de 1984 enlazada con el Registro Civil para el estudio de determinantes de la mortalidad de adultos costarricenses. San Jose, Costa Rica: Centro Centroamericano de Población de la Universidad de Costa Rica.
Rosero-Bixby, L., & Dow, W. (2009a). SES gradients in adult mortality in a middle income country: the Costa Rican Longitudinal Mortality Study 1984-2007. Berkeley: UC Berkeley.
Rosero-Bixby, L., & Dow, W. H. (2009b). Surprising SES Gradients in Mortality, Health, and Biomarkers in a Latin American Population of Adults. *J GERONTOL B PSYCHOL SCI SOC SCI,* 64, 105-117.
Ruhm, C. J. (2000). Are Recessions Good for Your Health? *Quarterly Journal of Economics,* 115(2), 617-650.
Sapolsky, R. M. (2004). Social Statud and Health in Humans and Other Animals *Annual Review of Anthropology,* 33(1), 393-418.
Scheffler, R., Brown, T., Syme, L., Kawachi, I., Tolstykh, I., & Iribarren, C. (2008a). Community-level social capital and recurrence of acute coronary syndrome. *Social Science & Medicine,* 66(7), 1603-1613.
Scheffler, R. M., Brown, T. T., Syme, L., Kawachi, I., Tolstykh, I., & Iribarren, C. (2008b). Community-level social capital and recurrence of acute coronary syndrome. *Social Science & Medicine,* 66(7), 1603-1613.
Sen, A. (1995). Inequality Reexamined. Oxford: Oxford University Press.
Sewastynowicz, J. (1983). Community Power Brokers and National Political Parties in Rural Costa Rica. *Anthropological Quarterly,* 56(3), 107-115.
Shi, L. (2005). Primary Care, Social Inequalities, and All-Cause, Heart Disease, and Cancer Mortality in US Counties, 1990. *American Journal of Public Health,* 95(4), 674-680.
Shively, C. A., & Clarkson, T. B. (1994). Social status and coronary artery atherosclerosis in female monkeys. *Arteriosclerosis,* 14, 721-726.
Skalli, A. L. I., Theodossiou, I., & Etienne, J. M. Do Economic Inequalities Harm Health? Evidence from Europe. papers.ssrn.com.
Strully, K. W. (2009). Job Loss and Health in the U.S. Labor Market. *Demography,* 46(2), 26.
Subramanian, S. (2004). Income Inequality and Health: What Have We Learned So Far? *Epidemiologic Reviews,* 26(1), 78-91.
Subramanian, S. V., Delgado, I., Jadue, L., & Vega, J. (2003). Income inequality and health: multilevel analysis of Chilean communities. *British Medical Journal,* 57, 844-848.

Subramanian, S. V., & Kawachi, I. (2003). The association between state income inequality and worse health is not confounded by race. *International Journal of Epidemiology,* 32, 1022–1028.
Subramanian, S. V., & Kawachi, I. (2004). Income Inequality and Health: What Have We Learned So Far? *Epidemiologic Reviews,* 26(1), 78-91.
Subramanian, S. V., & Kawachi, I. (2006). Whose health is affected by income inequality? A multilevel interaction analysis of contemporaneous and lagged effects of state income inequality on individual self-rated health in the United States. *Health and Place,* 12, 141–156.
Szwarcwald, C. L. (1999). Income inequality and homicide rates in Rio de Janeiro, Brazil. *American Journal of Public Health,* 89(6), 845-850.
Wagstaff, A., & van Doorslaer, E. (2000). Income Inequality and Health: What Does the Literature Tell Us? *Annual Review of Public Health,* 21, 543–567.
Wallerstein, C. (2001). Costa Rica tackles high rates of stomach cancer. *BMJ,* 323.
Wennberg, J. E., & Gittlesohn, A. (1973). Small Area Variations in Health Care Delivery. *Science,* 182, 1102-1108.
Wilkinson, R. G. (1992). Income distribution and life expectancy. *BMJ,* 304(6820), 165-168.
Wilkinson, R. G. (1996). Unhealthy Societies: The Afflictions of Inequality. London: Routledge.
Wilkinson, R. G., & Pickett, K. E. (2006). Income inequality and population health: A review and explanation of the evidence. Social Science & Medicine, 62, 1768–1784.