Analysing the Relationship between Income Inequality and Health in China

A thesis submitted in partial fulfilment of the requirement for the degree of Master of Research

Submitted by:

Han Cheng

Supervised by:

Dr Kompal Sinha and A/Prof Tony Bryant

10 December 2019

Department of Economics

Macquarie Business School

Macquarie University

Statement of Originality

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Date: 11 October 2019

Han Cheng

Acknowledgement

I would first like to express my sincere gratitude to my principal supervisor Dr Kompal Sinha for the continuous support of my study, for her motivation, enthusiasm, immense knowledge and unlimited patience. She consistently allowed this thesis to be my own work, but steered me in the right the direction whenever she thought I needed it. I would also like to thank my associate supervisor A/Prof Tony Bryant for his continuous encouragement, inspiration, admirable guidance, and relentless support.

I would also like to thank Master of Research advisor Dr Edwin Franks for his support and inspiration throughout this program. I am grateful to the HDR Academic Communication Specialist, Dr Frank Song for his support to me through academic communication skill development. I also appreciate Macquarie Business School HDR office staff Ms Lin Bai and Ms Mel Hubbard for their valuable and efficient assistance. My sincere thanks also go to Dr Lei Si, Senior Research Fellow, Office of the Chief Scientist, the George Institute for Global Health, for the inspiration during the data collection period.

I must express my very profound gratitude to my parents, Zhiqi Cheng and Yueyun Wang, whose love and guidance are with me in whatever I pursue. Last but not least, I wish to thank my loving and supportive wife, Yuan Zhou for her love, sacrifice, and providing me with unending encouragement throughout my study and through the process of researching and writing this thesis.

Dedication

I dedicate this thesis

to the memory of my beloved

late grandmother

Xueying Niu (1934-2012)

Abstract

This study tests the association between income inequality and individual's health, an association known as the Income Inequality Hypothesis (IIH) for China using the CHNS dataset. We adopt a dynamic approach to account for lagged effects of inequality on health outcomes using objective measures of health: nurse-collected health measures and blood-based biomarkers. Using a balanced panel data across 24 years from China we employ pooled OLS and fixed effects regression models controlling for individual level variables, county/city fixed effects, and year dummies. We find current inequality level not to be associated with health outcomes as reported in a recent empirical study by Bakkeli (2016). However considering a dynamic framework, we find lagged inequality level to have a statistical significantly associate with certain health outcomes after accounting for a host of demographic factors. Our analysis identifies the association between lagged inequality and health outcomes and the need for policy makers to account for the dynamic association between inequality and health outcomes.

Contents

List of Tables

Chapter 1. Income Inequality and Health

1.1. Introduction

From a microeconomics perspective, the purpose of economic development is to maximise an individual's utility. Over the past few decades, economies around the world have experienced significant growth alongside an increasing income gap. Health is the most important and fundamental indicator of people's wellbeing. As income inequality grows, the number of health issues is also on the rise, particularly in richer countries.

The relationship between income and health suggests more income helps people become healthier, due to an increased capacity to pay for health care. The relationship between income inequality and disparities in health have been a matter of concern for researches and policy makes around the world. This relationship became popular after Wilkinson (1992) argued that severe income inequality levels may impede the average growth expectancy of the Organisation for Economic Co-operation and Development (OECD) countries. Coined as the Income Inequality Hypothesis (IIH), Wilkinson's analysis suggested that an individual's health is affected by two aspects: 1) higher income brings better health and 2) higher income inequality lowers health outcomes. Starting from this point, several hypotheses were developed that studies have tested. To better understand these issues in the context of China, this study evaluates the relationship between income inequality and health (i.e., the IIH).

We find most studies in the literature show support for the IIH (Avendano & Hessel, 2015; Pickett & Wilkinson, 2015; Rostila et al., 2012). Previous studies have attempted many ways to understand the true relationship between income inequality and health. While the hypothesis was primarily designed to study the relationship in rich nations, recent researches evaluate the validity of this hypothesis in low income and middle income nations. Although these studies used various datasets, empirical methods, techniques and health measures, there is a lack of focus on the dynamic framework on income inequality. This is due to data availability, as measuring the lifetime accumulation effect of income inequality is difficult to achieve. However, based on recently available data the accumulated effect is worth measuring (Truesdale & Jencks, 2016).

Similar to other developed nations, economic development in China has been accompanied by a growth in income inequality, especially during the economic reform periods. Since the economic reform in 1978, the Gini coefficient has increased from 0.2 to around 0.5 (Chen & Zhou, 2005). The growth in inequality puts China at the same income inequality level as the United States (US). Meanwhile, the health care system in China is not largely developed compared to the economic side. According to Tang et al. (2008), the average life expectancy in China has not risen at the same level as the economy has developed. Not enough attention is paid to healthy lifestyle, as smoking and alcohol intake are still prevalent across the country (Shi et al., 2008). Conversely, diseases caused by affluence have emerged as people's income rises, with outstanding cases including obesity and diabetes. Since 1978, the prevalence of diabetes and obesity has increased dramatically. In the 1980s, the prevalence of diabetes was around 0.67%. In 2010, the same indicator was 11.6% (Wang et al., 2017). The huge increase also applies to the prevalence of obesity, which has increased two to three times since the reform (He et al., 2014). However, we found few studies placed any focus on China. These issues make China an outstanding context to study. This study delves into the context of China and applies multiple methods to provide a better understanding of the relationship between income inequality and individual health in China.

1.2. Thesis Structure

This dissertation is organised as follows. Chapter 2 reviews the relevant theoretical literature and builds a basic model to present all possible kinds of income inequality effects, including direct effects and indirect effects. Chapter 3 provides an overview of the empirical literature, in terms of the emerging hypotheses, the methods previous studies adopted and the countries and locations where the hypotheses were tested. Chapter 4 presents the institutional settings in China and comprehensively describes the China Health and Nutrition Survey (CHNS) dataset. This is followed by the methodology this study employs, including the constructions of all variables and empirical regression models used. Chapter 5 presents the major findings of this study. This chapter is divided into eight subsections showing original results, robustness checks and interpretations of the results. Chapter 6 presents a discussion, conclusion and directions for future studies.

Chapter 2. Review of Theoretical Literature

2.1. Introduction

The motivation behind IIH is that income levels and income inequality likely affect health outcomes across a cohort. Nevertheless, most of the literature concentrates on the relationships between income inequality, average heath and health disparity at the individual level (Truesdale & Jencks, 2016).

Growing income inequality affects an individual's health in two ways: direct effects and indirect effects. Higher income inequality directly affects an individual's income, through changes in health-related expenditure. The indirect effects result from changes in the relative position of individuals in society, including differences in income of societal peers, modified economic and political institutions in a society and people's lifestyle and customs as well as ideals. These indirect effects influence an individual's health through changes in their health behaviours and motivations. The basic relationship between an individual's health and income can be expressed as:

$$Health_i = \alpha + \beta(Income_i) + \mu_i \tag{1}$$

Where $Health_i$ is the individual's health, $Income_i$ is the sum of all sources of the individual's income and μ_i is the random error term with zero mean. There are three pathways through which income can affect an individual's health: direct change in $Income_i$, indirect change through change in β capturing changes the slope and change in α that can shift the curve up or down. The change of the intercept α stands for the change in average health level.

2.2. Direct Effects

Truesdale and Jencks (2016) suggest that the relationship between income inequality and health should have an upward sloping line, because increasing purchasing power disparities are driven by a rising income inequality. However, due to extremely limited literature showing the graph as a straight line, the majority of studies prefer a concave graph (Truesdale & Jencks, 2016). The intuition behind the concavity is the concept of

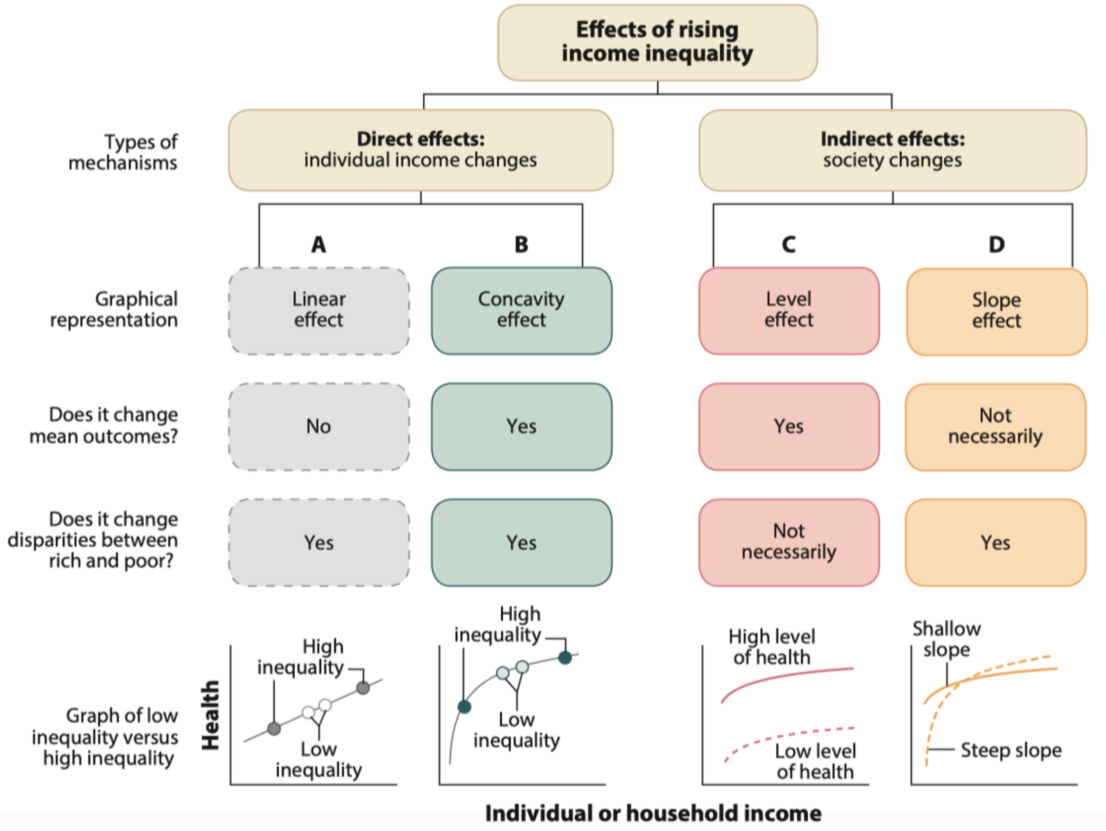

Figure 1: Summary of impacts of rising income inequality on health. Source: (Truesdale & Jencks, 2016)

marginal diminishing returns. Assuming the health returns from an extra dollar of income discount while income increases, the transfer from rich to poor will increase the average health level. Conversely, the money flow from poor to rich will lower the average health level. Previous studies named this direct effect as the concavity effects on average health or effects of individual income (Backlund et al., 1996), also known as the absolute income hypothesis (AIH) (Lynch et al., 2004). AIH is easy to understand, but is far from realistic evidence (Wagstaff & Van Doorslaer, 2000). The effects from income inequality are only significant when the income inequality level has increased significantly from the previous generation (Truesdale & Jencks, 2016). With the assumption that higher income contributes to better health, regardless of whether the connection is linear or concave, higher income inequality enlarges health disparities.

The concavity shows the marginal effects of income inequality on health, which happens at the bottom and the top. It is well understood that the transfer from the richest to the poorest is always preferred, whereas the opposite is always the worst transfer. For example, a person who is starving will benefit enormously from a little improvement in income, while the same level of improvement will do nothing for the CEO of the richest firm.

2.3. Indirect Effects

According to Equation (1), rising income inequality has indirect effects on health through two pathways. Both pathways affect individuals through the influences of social changes that are caused by the modification of the relationship between individual income and individual health. A clearer depiction would be a change in education, political institutions, consumption patterns and financial merchandise could indirectly touch people's health. The complicated social outcomes make the prediction function of the basic model inaccurate. The complication also demonstrates the pathways are not limited to two. However, in this study, we focus on the basic model with two indirect effects: level effects and slope effects.

Level effects change the average outcomes evenly for every individual at each income level. From the perspective of a graph as Column C in Figure 1 shows, level effects theoretically shift the whole curve up and down. A hypothesis proposed by Pickett and Wilkinson (2010) argued that if income inequality is higher, then every income group's average health will be reduced by the same percentage. Such a hypothesis emphasises the existence of level effects, yet the literature barely discovered an equal portion effect of income inequality on health. A refined version of level effects is well supported and demonstrates that, although the effects from changes in income inequality on all income groups' health are in the same direction, the proportion does not have to be the same. The difference between the two level effects offers a diversity of results of health disparity. While the hypothesis by Pickett and Wilkinson (2010) rules out the possibility of change in health disparities, the refined less restricted version allows alternation in health disparities.

Slope effects alter the coefficient β in Equation (1), changing the intensity of how income inequality affects individual health. Slope effects directly influence health disparities, which in column D of Figure 1 changes the slope of the income inequality-health curve. While excluding level effects, the sole slope effects do not necessarily change the mean health.

Slope effects are usually found as a motivation coming from outside (Truesdale & Jencks, 2016). The following air pollution example clearly describes the mechanism. Once a government sets strict rules in air pollution regulation, the overall better air quality benefits everyone in society. Such a reduction on air pollution not only contributes to the society's average health level, but also smooths the curve connecting income and airborne diseases. Thus, the slope of the curve is flatter after the regulation is applied. If one day there happens to be a war, the government and the regulation may disappear, leading to heavily polluted air. Under this circumstance, the slope of the curve is increased, meaning only affluent people have the capacity to purchase air purifying machines and enjoy better air. Conversely, poor people may not even consider the option of better air, due to limited budget.

Slope effects offer a different influence from another perspective. Apart from affecting people directly in health trouble, slope effects also affect health disparities after experiencing health problems. In our concavity case, the health disparity in a lower income inequality society will not increase significantly after the increase of the slope. However, when income inequality is higher the society has to navigate a tough path to manage increasing health disparities. Truesdale and Jencks (2016) noted that slope effects have barely noticeable implications on measuring how other social changes affect health disparities in terms of the change of the purchasing power of money.

2.4. Theories on Direct Effects and Indirect Effects

Although Equation (1) presents the effects of income inequality on individual health and health disparities, few studies address the theory on the part of health disparities. We found six theories offer indications and predictions on this topic. Overall, based on the groups of direct and indirect effects of income inequality, the six theories can be grouped into three, namely: direct concavity effects, indirect level effects and indirect slope effects.

Direct concavity effects suppose that increasing income inequality reduces average health and causes health disparities. Lynch et al. (2000) argued from the vision of the neo-materialist in that health is affected by both individual level and community level material resources. Regarding material resources as income, they hypothesised that higher material resources inequality leads to larger health disparities and harms average health, which is in line with concavity effects (also called marginal diminishing returns of health). The theory of scarcity is proposed by Mullainathan and Shafir (2013) with a different perspective of pathway through which income affects health. The scarcity theory assumes individual health is directly affected by individual income, different from the neo-materialist and the effects are cognitive. More specifically, the key assumption of the scarcity theory is the imposition of a ‘bandwidth tax’. Such a tax hinders an individual’s healthy behaviour and affects their long-term plans, guiding the individual to be more risk taking.

Indirect level effects postulate that upward trending income inequality lowers all income groups’ health as well as average health, although health disparities do not have to be affected. Indirect level effects are advocated by Pickett and Wilkinson (2010), Kawachi et al. (1997), Subramanian and Kawachi (2004) and Marmot (2007). The central argument of indirect level effects is the depreciation of social capital. The depreciation in detail presents that income inequality crushes the social framework and, thus, damages social trust as well as social capital. The less trustworthy society reduces public investment and increases living stress, harming the health of everyone in the society including the rich and the poor. In summary, indirect level effects may cause different effects on different income groups. The effect could be the same for everyone and it could also differ by income groups.

Indirect slope effects receive the most support, stating that increasing income inequality enlarges health disparities by enhancing health for the rich while lessening health for the poor. It could also be the case that both the health of the rich and the poor are reduced by increasing income inequality. Pickett and Wilkinson (2010) and Marmot (2007) proposed the relative deprivation theory that suggests that social comparison to better conditioned friends, colleagues and neighbourhoods within the community generates stress and health issues for people whose living condition is lower than their reference group. The major assumption of relative deprivation theory is that while the poor are living within a higher-ranked cohort who pay for premium infrastructure, the motivation effects are influenced by the depression effects for the poor causing health issues. Conversely, the theory does not offer a prediction of the effects for the rich.

Gilens (2012) and Bartels (2018) focused on the political part with the political capture theory that argues that increasing income inequality could enhance the political influence of the rich and predicts the health of the poor may be harmed if the rich's policy preference impedes investment in public goods and services on health (such as constructions of playgrounds, sanitation and education). Lastly, Link and Phelan (Glied & Lleras-Muney, 2008; Phelan et al., 2004; Phelan et al., 2010) turn to technology and innovations to suggest the theory of diffusion of innovations. The idea is that rich and educated individuals could gain access to the latest technology first, increasing health disparities. The diffusion of technology and innovations to the poor and less-educated individuals could be slower. Real world scenarios show support for the diffusion theory, such as looking for advanced medical treatments, filtered drinking water and easier ways to quit smoking.

2.5. Lagged Effects

While the apple you eat today will not contribute to your health right away, it may do tomorrow or in the future. Such a belief inspires us to consider applying the same mechanism to the focus of this study. Previous studies have touched on this field in an empirical way and we review these empirical studies in next chapter. To the best of our knowledge, the literature still lacks a systematic theory of how lagged inequality affects health outcomes. Although this study concentrates on the empirical side, we attempt to contribute to the development of a theory on the effect of lagged inequality on health outcomes.

2.6. Conclusion

From the theoretical work surveyed in this chapter, we have observed that the relationship between income inequality and health can be expressed in two ways: direct effects and indirect effects. Based on these two effects, we found six theories in three groups: direct concavity effects, indirect level effect and indirect slope effects. Following the intuition of lagged effects, we found little or no theories indicating lagged effects. Chapter 3 reviews the empirical studies that apply these theories to the real world.

Chapter 3. Review of Empirical Literature

3.1. Introduction

This chapter surveys the rich empirical literature that tests various versions of the IIH. Highlights of this literature include the evolution of the IIH, worldwide testing of the IIH, studies that focus on lag effects and methods adopted by previous studies.

3.2. Evolution of Hypotheses

The first study on the relationship between income inequality and health was delivered by Loftin (1974). Loftin investigated the relationship between homicide and inequality among the states in the US in the earliest paper on this topic. One year later, Preston (1975) who focused on the context of both "Western and non-Western areas" highlighted the importance of income redistribution, suggesting the redistribution of income would benefit the poor more and harm the rich less. Marmot et al. (1978) used a sample of 17,530 civil servants in the UK and found a significant adverse relationship between employment grade and mortality, which known as the Whitehall study. Rodgers (1979) attempted to uncover the relationship between income inequality and both infant mortality and life expectancy at age 5 in a group of 56 developed and developing countries. In the late 1980s and early 1990s, several studies showed a robust upward tendency of violence when higher income inequality applied (Kawachi & Kennedy, 1999). Wilkinson (Wilkinson, 1992; Wilkinson, 1996; Wilkinson, 1997) achieved a milestone with his set of three hypotheses on the relationship between income inequality and health namely, the AIH, the Relative Income Hypothesis (RIH) and the IIH. Wilkinson (1992) developed the AIH in his first paper on this topic. The AIH claims the existence of a positive relationship between one's health and the absolute level of income. Wilkinson then revised the income variable as the RIH. The

RIH postulates an individual's health is affected by the distribution of income within society. After attempting for a couple of years to derive the causality on health inequality, Wilkinson (1996) refined the hypothesis as the IIH. This hypothesis suggests people's health is affected by the income inequality level in the area in which they live, as well as by their own income. The IIH has inspired many studies to examine the relationship through various methodologies and in different countries around the world. Most studies conclude that health tended to be worse in societies with a higher level of income inequality (Pickett & Wilkinson, 2015). Starting in 2005, researchers shifted their focus on physical health to other detailed social outcomes such as mental health and teenage birth rates. Based on this trend, it is clear that this topic has been generalised to a broader range of health outcomes, thus, requiring a more general explanation of the IIH that focused more on physical health.

Such a requirement yielded a refinement of the hypothesis by Wilkinson and Pickett (2006). The generalised hypothesis postulates the level of income inequality may intensify the differentiation of socioeconomic status that is driven by social gradients. The evolved hypothesis inspired studies to adopt more detailed and general specifications when constructing the dependent variable of health. To the best of our knowledge, most of the studies to date confirm the hypothesis that income inequality harms health and other social outcomes.

However, there is counter-evidence that shows little or no such relationship. Mackenbach (2002) argued that supportive evidence for the hypothesis had disappeared. An interesting study by Deaton and Lubotsky (2003) counteracts IIH, with the researchers finding it was not income inequality that affected health, but the number of black residents in areas.

3.3. Debate of The Hypothesis

Studies around the world were in line with either no or an extremely weak relationship between income inequality and individual health. Studies in New Zealand (Blakely et al., 2003), Japan (Shibuya et al., 2002), England (Weich et al., 2002) and Canada (McLeod et al., 2003) find no relationship. Conversely, some studies in Canada and England found income inequality triggered better health outcomes (Craig, 2005; McLeod et al., 2003). Rostila et al. (2012) noted that income inequality could have a different consequence on different groups of people, specifically the rich and the poor. They also found that spending on social goods such as education explained the mechanism behind politically autonomous units in Sweden. The Sweden-based study concluded with limited evidence for the existing relationship between income inequality and health at the city level, whereas no relationship was found at the smaller neighbourhood-level. Such an output emphasised the important role geographic level plays when analysing with different aggregation levels. A recent literature review by Avendano and Hessel (2015) evaluated whether the IIH should be rejected. Referring to three key studies based on Switzerland (Clough-Gorr et al., 2015), Spain (Regidor et al., 2015) and a group of 43 European countries (Hu et al., 2015), Avendano and Hessel developed the argument that there was no causality between income inequality and health, although they were correlated with each other.

The debate of this topic is ongoing. Pickett and Wilkinson (2015) argued that studies with unsupportive evidence for the IIH should consider revision in the following elements: the geographic scale, the selection of lag length and the measurement of income inequality and health. If not chosen correctly, all these elements would bias the results.

3.4. Studies of Lag Effects

The Dutch famine study (Roseboom et al., 2006) argued the famine on newborn infant is a unique case of a relative deprivation in "a well-nourished population", which ultimately associates with health outcomes later. Such a lifetime potential relationship initiates the interest in investigating the lag effects of income inequality. Early studies had reviewed empirical studies concerned with how lagged income inequality affects individual health. Zheng (2012) covered 79 studies that analysed the relationship between income inequality and mortality and found 11 studies focused on the lag effects, especially the lag up to 10 years. Four studies tested through the aggregated level, while the other seven went through multi-level analysis. These studies found mixed results. The common feature of these studies is regarding the lagged income inequality as time-invariant variables. Lagged income inequalities were tested in a particular year, but not controlled as a series of past, subsequent and current income inequalities. Reviewed the link with self-rated health this study found seven significant influences of income inequality with up to 15 years of lag.

Broadening the horizon to worldwide studies, with a particular focus on panel studies and time-series studies of income inequality and health, Zheng (2012) summarised 53 additional analyses within nine studies. These studies used various measures of health including: self-rated health, mortality, infant mortality, life expectancy and under-5 survival rate. Over half of these studies advocated for a time-based longitudinal income inequality effect on health, 40% showed no such relationship and less than 10% found mixed results. Note that the issue of testing the lagged inequalities time-invariant variables still existed.

Zheng (2012) reviewed the lag effects of income inequality when preparing analysis of the way in which cross-country income inequality affects mortality in the US. Using a sample of over 700,000 individuals across 21 years, the study found income inequality started to have a harmful effect on mortality after 3–5 years, reaching its peak at 7 years and faded in 12 years. Although this study found a 12-year time range of the effects of income inequality to show up, Zheng mentioned the difficulty of theoretically confirm the time order of lag effects of income inequality. Instead, Zheng proposed a potential intuition that the lag effects may be due to the interval of medication to take effect, and the latency time range through which the health risks develop. This large-scale research set an excellent example and standard in terms of the estimated average lag time. It is worth mentioning that the average lag time in this study benefits the research of mortality, as different measures of health should be assigned for a proper lag time. Age cohort is another potential influencing factor, as responses from different age groups are expected to be different.

When we stretch the relationship of inequality and health to a long-term range, the question of causality may arise, that is, what is the exact causality between income inequality and health? Does unequally distributed income lower health performance, or does an individual's long-term sick body earns them less income? Both arguments seem reasonable on the surface, but empirical studies found unsupportive evidence for the reverse causality statement. Time-series studies (Pickett & Wilkinson, 2015) depicted that there is a considerable gap between changes in income inequality and changes in health outcomes. Further, Pickett and Wilkinson (2007) found another essential fact on the opposite side of the reverse causality argument. They found substantial evidence backing the finding that income inequality affects the health outcomes of children and even infants through low birth weight, infant mortality, child mental health and child overall wellbeing. This finding strongly proves worse health as the reason for pessimistic income and further income inequality.

Individual conditions and social gradients are also seriously considered in the lag effects analysis. Potential contributors such as education, gender, occupation and employment sector, transportation and environmental regulation do not instantly affect health outcomes. Another social element that could seriously influence health is culture. There have been many studies that argue that culture is the core reason behind the relationship of income inequality and health. We found that all culture-focused studies aimed at finding the connection that health indicator changes happen after several years of income being unequally distributed. Despite the culture difference between Japan and north European countries, both locations have a low level of income inequality and their societies appear more stable with less violence. Conversely, Russia and South Africa maintain very high income inequality levels, both lack social stability and the cultural differences between the two nations can be easily distinguished. Therefore, hypotheses on culture must consider the compatibility of different parts of the world in terms of development level and variety in culture (Pickett & Wilkinson, 2015).

3.5. Methodology

Early studies tended to analyse from a micro or macro perspective, while recent studies usually use both to develop a multi-level analysis model. The micro perspective focuses on the relationship between an individual's income and health. The macro approach switches to the aggregate level to analyse the relationship between income inequality of the area, health inequality and people's health. The selection of micro and macro approaches make a difference, as Pickett and Wilkinson (2015) indicated that geographical level is the key determinant of the direction of the relationship between income inequality and health. Most studies prefer to adopt the Gini coefficient as the measurement of income inequality. Some studies have used the Theil index (Bakkeli,

2016) and deprivation has also been used as an income-related metric (Kondo et al., 2008).

The literature has used various options for the health variables as the key dependent variable. There are three types of health measurements: subjective measures (i.e., self-rated health), objective measures (i.e., physical measurement of health markers) and health measures for an area at a higher geographic level. Self-rated health is the most widely used measure of an individual's health. According to Clarke and Ryan (2006), self-rated health is a stable measurement of health and has been confirmed by previous studies. For a more precise and objective measurement of individual health, previous studies have adopted diverse benchmarks such as mortality, mental health, body mass index (BMI), blood pressure and teenager health, among others. These detailed measurements of objective individuals' health enabled the certainty of investigation between income inequality and precise health outcomes. Those measurements also opened wider options for studies to explore. Finally, the national characteristic of health provides a measurement from a macro perspective. As an example, concentration index (CI) has been utilised in many studies to analyse the relationship between income inequality and health at a higher geographic level such as state or province and country level (Zhou et al., 2017).

3.6. Focus of the Study

Early studies of the income inequality–health relationship focused predominantly on developed countries. As theories and analytical tools were updated, Wilkinson (2011) noted that it was more important for the developing and poor countries to be mindful that people's average wellbeing was not only depend on economic growth and national income, but also the overall equality. The present study will focus on the context of China.

From the late 1970s, China has experienced miraculous economic growth and has made the largest contribution to global poverty reduction, with more than 650 million people escaping poverty. From 1981 to 2005, the poverty headcount ratio in China dropped from 44% to 2% in urban areas and from 94% to 26% in rural areas (Baeten et al., 2013). Meanwhile, the country also witnessed an upward trend of income inequality as well as disparities in individuals' health status (Tang et al., 2008). With the unique experience of development, China has become a prominent context in which to test the relationship between income inequality and individuals' health (Yu & Chiu, 2016).

To the best of our knowledge, there is an increasing number of studies that explore the relationship between income inequality and other social outcomes, while the number of analysis of the relationship between income inequality and health in China is not increasing as rapid as in developed counties. Among nine relevant reviewed studies, the majority advocated for the IIH (Baeten et al., 2013; Feng et al., 2012; Li & Zhu, 2006; Pei & Rodriguez, 2006; Sun et al., 2012; Yang & Liu, 2018; Yang & Kanavos, 2012; Zhou et al., 2017). However, there is one outstanding study conducted more recently (Bakkeli, 2016) that shows negative support for the hypothesis. Almost every study in China has adopted the Gini coefficient or a revised version of the Gini coefficient to suit certain research environments and conditions.

For health variables, these China-based studies used self-reported health as their main measurement of an individual's health. Li and Zhu (2006) applied physical functions (PF) and activities of daily living (ADL), in addition to self-rated health. Sun et al. (2012) applied multiple measurements of health including self-reported depression, perceived stress and cigarette smoking. Unlike the above two studies, Baeten et al. (2013) developed two new health variables: income-related health inequalities (IRHI) and a Gini index of health. Studies analysing higher macro levels adopt a different metric for health variable (i.e., health inequality). Zhou et al. (2017) used health-related quality of life (HRQoL) measured by the EQ-5D, which is a CI of the health inequity. Yang and Liu (2018) tested the relationship between air pollution, public health and health inequality. They adopted the slope indices of inequality (SII) and the health concentration curve to develop their health variables. Bakkeli (2016) adopted a combination of four objective measures of health including: blood pressure, waist-to-hip ratio (WHR), mid-upper arm muscle circumference (MAMC) and obesity to quantify health status.

While some studies around the world disclaimed the IIH (Clough-Gorr et al., 2015; Hu et al., 2015; Regidor et al., 2015), Bakkeli (2016) delivered the first study that cast doubt on the IIH in China. The study tested the relationship in China to find an astonishing and statistically robust conclusion that income inequality had no or little effect on health. These findings are in stark contrast to those reported in the context of China in which existing research has found the IIH to hold. According to Pickett and Wilkinson (2015), there is a strong relationship between income inequality and health and studies with unsupportive outcomes should consider testing at a higher level of geographic level. Dorling and Barford (2009) offered a similar idea that supports the hypothesis. Conversely, Avendano and Hessel (2015) tended to reject the hypothesis. There has been no study that attempts to confirm or cast doubt on Bakkeli's findings, as far as we are aware.

Pickett and Wilkinson (2015) suggested future studies consider different measures of income inequality, test specific causal pathways with different time lags for different outcomes and incorporate wealth inequality in the discussion of inequality. The present study incorporates objective measures of health into the analysis to evaluate and extend the IIH in China. Using objective measures of health allows us to avoid the measurement bias and reporting bias associated with self-reported measures of health (Sinha et al., 2018). The outcome of this study will provide insights into improving people's overall wellbeing in China and in other developing countries.

3.7. Conclusion

This chapter surveyed the empirical literature on various versions of income inequality–health relationship. Based on this survey, it appears that the situation of developing countries (particularly China), has been relatively neglected, which leads to the focus of the present study. Chapter 4 introduces the dataset and methods this study employs including processing the dataset, specification of empirical regression models and definitions of variables.

Chapter 4. Research Method

4.1. Introduction

This chapter describes the methods by which we intend to pursue our objective of making an empirical study of the income inequality–health relationship in China. Highlights include the dataset description and processing, specifications of regression models and definitions of variables.

4.2. Data

This study utilises the data from the CHNS, which is a longitudinal panel dataset collected from 1989 to 2015. This survey is an international collaborative project conducted by the Carolina Population Center at the University of North Carolina at Chapel Hill in the US and the National Institute for Nutrition and Health, formerly the National Institute of Nutrition and Food Safety, at the Chinese Center for Disease Control and Prevention (CCDC). CHNS is designed to evaluate the influence of health, nutrition and family planning policies and other social policies imposed by local governments. The primary objective of this dataset is to reveal the relationship between China's social and economic development and its population's health and nutrition status. All socioeconomic and health data for this study are accessible in the CHNS dataset.

4.2.1. Data Description

At the time this study was undertaken, the CHNS dataset contained data across nine provinces: Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi, Guizhou, Liaoning (missed 1997), Heilongjiang (added 1997), Beijing (added 2011), Shanghai (added 2011) and Chongqing (added 2011). The nine provinces are distributed in the western, eastern, middle and northern areas of China and consist of 72 cities or counties. The

CHNS sample is not designed to be representative at the country level but at the provincial level according to Popkin et al. (2009), the provinces and counties surveyed are differed substantially in geography, public resources, economic development, and health indicators. The CHNS sample are picked from randomly selected households to demonstrate a wide range of economic and demographic characters. This study utilises the CHNS dataset up to the tenth wave, including 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009, 2011 and 2015.

4.2.2. Collection of Data

CHNS is an open source project, and its dataset can be found and downloaded through their website (i.e., https://www.cpc.unc.edu/projects/china). The website offers data in each category including: biomarker, agriculture, asset, business, individual energy, childcare, constructed income, individual education, ever-married women, health care, ID, income categories, infant feeding, livestock, macronutrients, media, nutrition, physical examination, relationship, time use and urban index. All the data in these categories are compressed into zip file packages. Unzipping all files provides the original CHNS datasets in SAS format. For straightforwardness and convenience, we use the software Stat/Transfer and transfer the SAS-formatted original CHNS datasets to the DTA format, which could be used in Stata. After the confirmation of a success full transfer, we merged relevant datasets.

4.2.3. Merging Datasets

For convenience of analysis, we merged the relevant data into two datasets. The following data are merged for the analysis of nurse-collected markers: constructed income, individual education, ever-married women, physical examination, ID and urban index. The nurse-collected markers dataset covers the full time range from 1989 to 2015, hence, the merged dataset of nurse-collected measures is a longitudinal unbalanced panel dataset across 26 years.

The analysis of blood-based biomarkers is based on the dataset that consists of biomarker, constructed income, individual education, ever-married women, physical examination, ID and urban index. In contrast with nurse-collected measures, the blood-based biomarkers are only available for 2009. Thus, the analysis of blood-based biomarkers is focused on a cross-sectional dataset in 2009.

4.2.4. Cleaning Dataset

For better control and understanding of the relationship between income inequality and individual health, we set the first restriction that observations are limited to people aged 16–69. The second restriction imposed is that observations must be employed. These two restrictions remove over half the total observations. After the cleaning process, the total number of individuals is 22,819 from 1989 to 2015, resulting in 74,970 observations in total. Note that because this study adopts the same method of cleaning dataset as Bakkeli (2016), thus the representativeness is not dropped after cleaning dataset. Table 1 presents detailed descriptive statistics for all variables.

To capture the city and county level fixed effects, we generate a variable stand for the county ID based on the original administrative division in the CHNS dataset. All 72 cities and counties are labelled uniquely.

Table 1 Descriptive Statistics for the Reduced Sample

Variables	Observations (N)	Mean (SD)	Min/Max
Health (Nurse-collected Measures)			
Blood Pressure	58,341	0.46 (0.50)	0/1
WHR	51,964	0.56 (0.50)	0/1
MAMC	63,059	0.07 (0.25)	0/1
BMI	64,253	0.24 (0.43)	0/1
Waist Circumference	74,970	0.55 (0.50)	0/1
Health (Blood-based Biomarkers)			
C-Reactive Protein	5,760	0.08 (0.27)	0/1
HbA1c	5,760	0.07 (0.25)	0/1
Cholesterol Rate	6,668	0.30 (0.46)	0/1
Inequality			
Gini Coefficient*	74,970	0.45 (0.09)	0.22/0.75
Theil L*	74,970	0.46 (0.19)	0.09/1.4
Theil T*	74,970	0.41 (0.19)	0.1/2.25
Theil V*	74,970	0.76 (1.06)	0.11/5.30
Income			
Individual Income	74,970	11510 (34189)	0/4800000
County Average Income	74,970	11495 (13224)	758/78049
Individual Controls			
Age	74,970	42.83 (13.54)	16/69
Gender	74,970	1.49 (0.50)	0/1
Marital Status	74,970	0.88 (0.32)	0/1
Majority	74,970	0.88 (0.33)	0/1
Years of Education	74,970	7.76 (4.40)	0/18
Urban	74,970	0.31 (0.46)	0/1
Occupation			
Service Class	62,944	0.09 (0.28)	0/1
Non-manual Worker	62,944	0.09 (0.29)	0/1
Skilled Worker/Supervisor	62,944	0.10 (0.31)	0/1
Semi-/non-skilled Worker	62,944	0.21 (0.41)	0/1
Farmers	62,944	0.47 (0.50)	0/1
Others	62,944	0.40 (0.19)	0/1
Employment Sector			
State	62,438	0.30 (0.46)	0/1
Collective	62,438	0.37 (0.48)	0/1
Family Farming	62,438	0.17 (0.37)	0/1
Individual Enterprise	62,438	0.14 (0.35)	0/1
Private Three-cap. Enterprise	62,438	0.01 (0.09)	0/1
Others	62,438	0.02 (0.13)	0/1

Note. *Author's calculations. See Appendix A for further detail. WHR = waist-to-hip ratio. MAMC = mid–upper arm muscle circumference. BMI = body mass index. CRP = C-reactive protein.

4.3. Methods of Data Analysis Employed

Given our goals and objectives, this study adopts a quantitative approach from a large number of individuals with the application of statistical techniques to recognise overall patterns in the relations of processes. In line with findings from previous studies, geographic scale does implement a significant effect on research outputs (Chen & Gotway Crawford, 2012). According to previous studies, cities and counties in CHNS are large enough to capture the contextual information required (Pickett & Wilkinson, 2015). Therefore, this study selects city and county level as the basic geographic level to analyse.

4.3.1. Data Analysis Techniques

Inspired by Bakkeli (2016) and Sinha et al. (2018), this study uses objective measures of health to compare results from different measurements of health, then further explores whether past income inequalities affect health. To capture the time-invariant factors while controlling the possibility of bias due to unobserved heterogeneity, this study creates linear probability models with the main research method of fixed effect analysis. Each model is adjusted for individual control variables and year dummies. Although the method of fixed effect analysis is similar to the study conducted by Bakkeli (2016), this study includes more objective health metrics and lagged inequalities to make comparisons.

4.4. Modelling Strategy

To fit the set goals, we first confirm the relationship between income inequality and nurse-collected health markers, then compare the outcomes with blood-based biomarkers. Finally, we test whether lagged inequalities make any difference. We achieve this by constructing eight econometric models. All models are estimated with each health variable respectively and further adjusted for individual control variables.

Model 1 and Model 2 mainly focus on the goal of refreshing Bakkeli's study with the latest CHNS dataset and blood-based biomarkers. Model 3 and Model 4 analyse this relationship but incorporate income inequalities of the past. All models are regressed separately for males and females.

Model 1 and Model 2 test the relationship between income inequality and health using nurse-collected health markers and blood-based biomarkers. As an extension to Bakkeli's study, this group utilises the latest CHNS dataset with data up to 2015. Overall, the development of the four models follows Bakkeli's process. The expected functions of Model 1 and Model 2 would be in line with the previous study, thus, the results from these two models can be judged as reinforcing or opposing Bakkeli's no relationship conclusion.

Model 1 contains three equations from Equation (2) to Equation (4). Equation (2) includes income inequality, individual income and average income in a city or county:

$$Health_{ict} = \alpha + \beta_1 incomeinequality_{ct} + \beta_2 income_{ict} + \gamma_1 countyincome_{ct} + \eta_{\mathrm{ict}} \tag{2}$$

Equation (3) absorbs the detailed individual control variables:

$$Health_{ict} = \alpha + \beta_1 incomeinequality_{ct} + \beta_2 income_{ict} + \gamma_1 countyincome_{ct} + \gamma_2 C_{\mathrm{ict}} + \eta_{\mathrm{ict}} \tag{3}$$

Equation (4) adds years, dummies and county and city level fixed effects to Model 1:

$$Health_{ict} = \alpha + \beta_1 incomeinequality_{ct} + \beta_2 income_{ict} + \gamma_1 countyincome_{ct} + \gamma_2 C_{\mathrm{ict}} + \gamma_3 D_{\mathrm{t}} + \varepsilon_{\mathrm{c}} + \upsilon_{\mathrm{t}} + \eta_{\mathrm{ict}} \tag{4}$$

Model 2 includes Equation (5) and Equation (6). It focuses on health status as measured by blood-base biomarkers. We apply the ordinary least squares (OLS) technique in this model to test the relationship between income inequality and blood-based biomarkers. Due to the blood-based biomarkers only being available in 2009, the analysis from Model 2 is based on a cross-sectional dataset instead of a panel dataset. We use OLS in Model 2, to keep the most similarity between nurse-collected measures analysis and blood-based biomarkers. Due to the analysis of blood-based biomarkers being undertaken on a cross-sectional dataset, county average income would be omitted if we adopted the fixed effect model. Thus, we employ the cross-section regression technique of OLS for this analysis. In a similar set up to Model 1, Model 2 introduces more detailed and reasonable factors into the models to reveal the connection between income inequality and specific health issues.

Equation (5) retains the identical specifications with the exception of health variables:

$$Health_{ic} = \alpha + \beta_1 incomeinequality_{c2009} + \beta_2 income_{ic} + \gamma_1 countyincome_c + \eta_{\mathrm{ic}} \tag{5}$$

Based on Equation (5), Equation (6) adds individual control variables:

$$Health_{ic} = \alpha + \beta_1 incomeinequality_{c2009} + \beta_2 income_{ic} + \gamma_1 countyincome_c + \gamma_2 C_{\mathrm{ic}} + \gamma_3 \mathrm{D} + \varepsilon_{\mathrm{c}} + \upsilon + \eta_{\mathrm{ict}} \tag{6}$$

As reviewed in section 3.4, the money you earn today will influence your life tomorrow. In other words, income inequality in the past contributes to life today. Therefore, lag effects should be considered in the analysis. Previous researchers discuss the topic of lag effects, with some finding supportive evidence to show that lag effects are essential while others hold opposite opinions. This study utilises the method in accordance with Blakely et al. (2000). Although Blakely et al. (2000) argued 15 years of lag is appropriate, due to the data limitation we moderate the lag to nine years. From the perspective of dynamic framework, the changing amount of income inequality is credited more with resulting in people's lives. Model 3 and Model 4 are extensions of Model 1 and Model 2, in incorporating the inspiration of lagged inequalities. Model 3 extends Model 1 with three waves of lagged inequalities, while Model 4 does the same to Model 2.

The model specification for Model 3 is presented in Equation (7). Three specifications of this model are estimated with three waves up to nine years of lags of income inequality added into the analysis:

$$\begin{aligned} Health_{ict} = \alpha &+ \beta_1 incomeinequality_{ct} + \beta_2 incomeinequality_{c(t-1)} \\ &+ \beta_3 incomeinequality_{c(t-2)} + \beta_4 incomeinequality_{c(t-3)} \\ &+ \beta_5 income_{ict} + \gamma_1 countyincome_{ct} + \eta_{\mathrm{ict}} \end{aligned} \tag{7}$$

$$\begin{aligned} Health_{ict} = \alpha &+ \beta_1 incomeinequality_{ct} + \beta_2 incomeinequality_{c(t-1)} \\ &+ \beta_3 incomeinequality_{c(t-2)} + \beta_4 incomeinequality_{c(t-3)} \\ &+ \beta_5 income_{ict} + \gamma_1 countyincome_{ct} + \gamma_2 C_{\mathrm{ict}} + \eta_{\mathrm{ict}} \end{aligned} \tag{8}$$

$$\begin{aligned} Health_{ict} = \alpha &+ \beta_1 incomeinequality_{ct} + \beta_2 incomeinequality_{c(t-1)} \\ &+ \beta_3 incomeinequality_{c(t-2)} + \beta_4 incomeinequality_{c(t-3)} \\ &+ \beta_5 income_{ic} + \gamma_1 countyincome_{c} + \gamma_2 C_{\mathrm{ic}} + \gamma_3 \mathrm{D} + \varepsilon_{\mathrm{c}} + v + \eta_{\mathrm{ict}} \end{aligned} \tag{9}$$

Model 4 employs the same methods as Model 3, but differs in the set up. Model 4 could be regarded as a variant of Model 2 and it contains Equation (10) and Equation (11):

$$\begin{aligned} Health_{ic} = \alpha + \beta_1 incomeinequality_{c2009} + \beta_2 incomeinequality_{c2006} \\ + \beta_3 incomeinequality_{c2004} + \beta_4 incomeinequality_{c2000} \\ + \beta_5 income_{ic} + \gamma_1 countyincome_c + \eta_{\mathrm{ic}} \end{aligned} \tag{10}$$

$$\begin{aligned} Health_{ic} = \alpha + \beta_1 incomeinequality_{c2009} + \beta_2 incomeinequality_{c2006} \\ + \beta_3 incomeinequality_{c2004} + \beta_4 incomeinequality_{c2000} \\ + \beta_5 income_{ic} + \gamma_1 countyincome_c + \gamma_2 C_{\mathrm{ic}} + \gamma_3 D + v + \eta_{\mathrm{ic}} \end{aligned} \tag{11}$$

In all models, i indicates ID of each individual, c stands for county ID, t stands for the wave which the survey was taken, C is a vector consisting of individual control variables, D is the vector of year dummies, β is the coefficient this study will test, ε_{c} is interpreted as an unknown intercept of county units, v_{t} is the year of dummies error and η is the term of disturbance. Table 2 summarises specifications of all models.

Table 2 Summary of Models' Specification

Equation #	Lags	Data Type	Estimation Method	Individual Control
Model 1				
2	No	Panel	OLS	No
3	No	Panel	OLS	Yes
4	No	Panel	Fixed-effects	Yes
Model 2				
5	No	Cross-sectional	OLS	No
6	No	Cross-sectional	OLS	Yes
Model 3				
7	Yes	Panel	OLS	No
8	Yes	Panel	OLS	Yes
9	Yes	Panel	Fixed-effects	Yes
Model 4				
10	Yes	Cross-sectional	OLS	No
11	Yes	Cross-sectional	OLS	Yes

Following Bakkeli (2016), health variables are specified as binary variables that have two indicators. Poor health is assigned the value of health variables equal to 1, while the value of 0 indicates a healthy function. With the assumption of health indicators and common knowledge, this study hypothesises that lower income inequality reduces health issues and, conversely, higher income mitigates health risks. Therefore, in this study we expect the coefficient on income inequality to be greater than zero and the coefficient on income to be less than zero. Applying this to the models above, $\beta_1 > 0$ and $\beta_2 < 0$. To ensure the results from all models are reliable, we undertake robustness checks using diverse measures of income inequality. Specifically, this study chooses three variants of Theil indices as the replacement for Gini coefficients: Theil L, Theil T and Theil V. Theil L is sensitive to bottom level income changes and is the average logarithm deviation. Theil T is known as the Theil index and is sensitive to upper level income changes. Theil V is the variation of half the squared coefficient. Conversely, Gini coefficient is sensitive to changes in income at the middle level (Bakkeli, 2016).

The assumption of independent and identically distributed random variables (iid) in a panel dataset is indispensable. Serial correlation is a potential issue in this study, as the observations can be correlated over time and within the county or city. This could lead to the collapse of the assumption of *iid*. To address this risk, two-way cluster-robust standard errors are employed on both individual and county and city level units (Cameron et al., 2011).

4.5. Variable Definition

Following are the definitions of the variables referred to in the above equations.

4.5.1. Dependent Variable

As discussed, the dependent variables are all health markers that can be grouped into two. To capture more detail of the connection between inequality and health, this study follows Sinha et al. (2018) to use two types of health measurements: nurse-collected health measures and blood-based biomarkers.

4.5.2. Nurse-collected Markers

Model 1 and 2 use nurse-collected markers including blood pressure, WHR, MAMC and obesity. Previous studies have found blood pressure to be an accurate health indicator that can predict potential diseases such as heart disease, stroke, chronic kidney disease and coronary artery disease (Trialists'Collaboration, 2008). Systolic blood pressure and diastolic blood pressure were measured three times for every individual. To obtain the most accurate results, we skip the first reading and take the average of the second and third measures. The cut-off value of blood pressure is 120/80 mmHg (Blakely et al., 2003; MacMahon et al., 1990). Therefore, a normal blood pressure is characterised as at or below 120/80 mmHg.

WHR measures the distribution of body fat. Although WHR is usually a metric of obesity, it captures more responses compared with BMI because WHR considers body structure. WHR has been regarded as an important indicator for older people's mortality (Price et al., 2006) and general health for average people. WHR is also considered an accurate metric of cardiovascular diseases and, more importantly, is found to be a more suitable indicator for Asian populations (Wu, 2006). This study sets the bar at 0.80 for women and 0.90 for men, according to the World Health Organization (WHO (2011b), with any value over the cut-off value judged as abnormal.

MAMC is often used to assess the status of malnutrition by measuring protein storage as well as muscle mass. This measure has been utilised in the literature as an indicator of body composition (Jarvis, 2011) in developing context studies. Following WHO norms, we use the cut-off value of 20.88 for women and 22.77 for men (WHO, 2004)

Body Mass Index (BMI) is a traditional indicator of obesity. According to WHO (2004), the suggested cut-off point for the Chinese population is 25 kg/m2, meaning an observation with a BMI under 25 kg/m2 would be excluded from obesity. This study also uses waist circumference as the enhancement of the BMI to measure obesity. Waist circumference is a more informative measure of obesity comparing to BMI. Waist circumference also correlates with cardiovascular risk and is more accurate for Asian populations. The cut-off value of waist circumference differs in gender, with 80 for women and 90 for men. Any measure over the cut-off value is regarded as an indicator of obesity (Wu, 2006).

4.5.3. Blood-based Biomarkers

This study adopts three categories in blood-based biomarkers including: inflammatory, blood glucose and cholesterol ratio. Inflammation is measured by C-reactive protein (CRP), a type of protein that could reflect chronic inflammation. According to Ishii et al. (2012), a CRP value above 5 mg/L is considered a high risk infection and a value above 10 mg/L is considered a severe risk infection. HbA1c, known as glycated Haemoglobin, is the examiner of diabetes. The cut-off values of HbA1c are 48 mmol/mol for diagnosed diabetes and 42 for the predictable risk for diabetes (WHO, 2011a). The cholesterol rate is the amount of fat in the blood. The value of cholesterol rate over 4 suggests there is elevated atherosclerotic risk (Millán et al., 2009).

4.5.4. Independent Variables

The variable $income$ in the models is the aggregated income including all income sources of the individual. To obtain individual income, we calculate the sum of individual income from all sources after deducting individual expenditures. Note, this calculated individual income is not per capita income within the household. The logic is that some income categories such as household subsidies and other income could not be assigned to a specific member in the household and, thus, could not be a part of individual income.

Income inequality is mainly presented by the Gini coefficient that is calculated at the county level. We employed Stata software to calculate the Gini coefficient. The code utilises the definition equation to calculate (see Appendix A). As this study focuses on the county and city geographic level, we first create 72 location indicators for each county and city. Then, we compute the Gini coefficient for the assessed units at county and city level. To obtain the proper consequence of income inequality on health status, this study includes a dynamic method that is the time-lag in calculating the coefficient in models 3 and 4. The lagged Gini coefficients adopted are across waves up to last three waves. Note that the potential issue of individuals' heterogeneous exposure to income inequality is avoid in this study because this study captures same individuals over time by identification through individual ID, thus the computation of Gini indices is done through the same cohort of observations. Furthermore, according to Popkin et al. (2009), although there exists attrition issue, the team conducting the CHNS data collection did find a remedy to keep the dataset representative. Since the CHNS data is collected in various years, the typical lag length in Model 3 is over nine years and eight years. Model 4 focuses on the biomarker data from CHNS wave 2009. Therefore, one wave lag is the year 2006, two wave lags lead to the year 2004 and three wave lags stands for the year 2000. The lagged Gini coefficients are allocated accordingly. All Gini coefficients including the current lags are standardised to their related z-score, which means the standardised variables fit the normal distribution with zero average and one standard deviation.

For the effect of geographic scale, the analysis is further adjusted for average county and city income using the variable $county\ income$ available in CHNS. This variable assesses the average disposable income in the county or city where respondents live. The computation of this variable follows the location indicators we generated when obtaining the county and city level Gini coefficients. As we already have the location indicators, we use the average of individual income within the county or city. Although the adoption of both individual income and county average income in the models is inspired by Bakkeli (2016), this study tests the potential issue of correlation and multicollinearity, in both tests the null hypotheses are not significant, detailed test results are shown in Appendix A. For consistency, the $county\ income$ is also standardised by the corresponding z-score. All income variables and income inequality variables are analysed as continuous variables.

The individual control vector D consists of individual control variables including age, gender, marital status, racial (i.e., ethnic) majority, education, whether the individual lives in an urban area and sector of employment. Age is restricted to the range of 16–69 and analysed as a continuous variable. This study takes gender into account as well as marital status and racial majority. Gender is applied, as every model conducts male and female separately. The original data of marital status has the following category: never married, married, divorced, widowed and separated. We convert these categories into a binary variable, labelling never married as value 0 and the rest as married with value 1. Racial majority is defined with value 1 if the individual is Han Chinese, otherwise the value is 0. Variable education is measured by years of education, representing the participants' highest level of education. The original education data is more detailed and contains the following levels: primary school, lower middle school, upper middle school, technical school and college. We transform and classify the original data by combining technical school and college together to make the level of education easy to read. The level of education is examined by the following metric, with 18 years as the highest education level (PhD) and 0 as never-educated.

The data on occupation in CHNS data has 16 categories:

1. Senior professional/technical
2. Junior professional/technical
3. Administrator/executive/manager
4. Office staff
5. Farmer, fisherman, hunter
6. Skilled worker
7. Non-skilled worker
8. Army officer, police officer
9. Ordinary soldier, policeman
10. Driver
11. Service worker
12. Athlete, actor, musician
13. Other
14. Small household business
15. Homemaker
16. Student.

Following Bakkeli's study, this study groups the above 16 values into six groups:

1. The service class (1) (3) (8)
2. The non-manual worker (2) (4)
3. The skilled workers/supervisor (6) (9) (10) (12)
4. Semi-skilled and non-skilled workers (7) (11)
5. Farmers (5)
6. Other (13) (14) (15) (16).

Employment sector is also considered alongside occupation. The CHNS dataset offers nine values in the employment sector:

1. Government
2. State service/institute
3. State-owned enterprise
4. Small collective enterprise
5. Large collective enterprise
6. Family contract farming
7. Private, individual enterprise
8. Three-capital enterprise
9. Unknown.

In the analysis, we created six groups to present the employment sector:

1. The state sector (1) (2) (3)
2. The collective sector (4) (5)
3. The sector of family farming (6)
4. The individual enterprise (7)
5. Three-capital enterprise (8)
6. Others (9)

4.6. Conclusion

In this chapter we have explained the empirical methodology employed in this study. The following chapter discusses the implementation and the interpretation of the results of this methodology, as applied to the CHNS dataset.

Chapter 5. Empirical Results and Analysis

5.1. Introduction

This chapter presents the results of our empirical analysis. We begin with our empirical analysis with a replication of the recent study by Bakkeli (2016). The reason for doing this is to test whether Bakkeli's conclusions hold with the updated dataset. We then progress to our own study that extends the IIH with detailed objective health measures and lagged income inequality effects.

5.2. Replication Results

First, we update Bakkeli's study using the same set up of models and extended CHNS data up to 2015. The reference group is married Han-ethnic farmers with rural household registration. This group work in the collective sector and do not have any education. Overall, the results are consistent with Bakkeli's findings, that is, a weak connection is found between income inequality and nurse-collected health variables.

5.2.1. Empirical Results

The regression results are shown in Table 3. We find that estimating Equation (2) when only including Gini, individual's income and county average income, Gini indices make a significant contribution to a higher risk of abnormal blood pressure, higher WHR and BMI for both men and women. An increase in income for women leads to a greater improvement in their blood pressure, WHR, BMI and waist circumference, while there is no noticeable change for men. For blood pressure, WHR and BMI, the significant positive signed coefficients on county average income is consistent with Bakkeli's findings, which is still shocking. Conversely, for MAMC and waist circumference, the negative significant sign indicates local economy development improves these health performances, which follows rational logic.

Table 3 OLS and Fixed Effects Model for Nurse-Collected Health Outcomes

	Female			Male		
	(2)	(3)	(4)	(2)	(3)	(4)
Blood Pressure						
Gini	0.032***	0.014**	0.019**	0.030***	0.017***	0.016
	(0.006)	(0.006)	(0.009)	(0.006)	(0.006)	(0.010)
County Income	0.101***	0.051***	–0.012	0.110***	0.067***	0.003
	(0.009)	(0.007)	(0.020)	(0.008)	(0.008)	(0.016)
Income	–0.032***	–0.007*	–0.005	–0.000	0.003**	0.003**
	(0.007)	(0.004)	(0.004)	(0.001)	(0.001)	(0.001)
Constant	0.390***	–0.218***	0.012	0.519***	–0.003	0.274***
	(0.011)	(0.033)	(0.041)	(0.012)	(0.035)	(0.047)
No. of Obs.	29,438	22,345	22,345	28,903	24,612	24,612
No. of Groups			72			72
WHR						
Gini	0.023***	0.009	0.002	0.024***	0.024***	0.007
	(0.005)	(0.006)	(0.008)	(0.008)	(0.007)	(0.007)
County Income	0.039***	0.027***	0.002	0.092***	0.066***	–0.003
	(0.005)	(0.006)	(0.017)	(0.007)	(0.008)	(0.016)
Income	–0.021***	–0.002	–0.003	0.004	0.001	0.002
	(0.005)	(0.006)	(0.005)	(0.004)	(0.003)	(0.003)
Constant	0.740***	0.426***	0.373***	0.350***	–0.039	–0.013
	(0.008)	(0.032)	(0.035)	(0.010)	(0.032)	(0.031)
No. of Obs.	26,235	19,201	19,201	25,729	21,445	21,445
No. of Groups			72			72
MAMC						
Gini	0.000	0.002	–0.001	0.000	–0.000	0.001
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
County Income	–0.012***	–0.008**	0.021**	–0.028***	–0.018***	0.017*
	(0.003)	(0.003)	(0.008)	(0.004)	(0.003)	(0.010)
Income	0.000	–0.001	–0.001	–0.001	0.001	0.000
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.052***	0.107***	0.032	0.088***	0.205***	0.059**
	(0.005)	(0.022)	(0.025)	(0.007)	(0.038)	(0.028)
No. of Obs.	31,827	24,469	24,469	31,232	26,825	26,825
No. of Groups			72			72
BMI						
Gini	0.011**	0.004	0.006	0.011*	0.009*	0.003
	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)	(0.004)
County Income	0.073***	0.041***	–0.024*	0.101***	0.069***	–0.003
	(0.007)	(0.007)	(0.013)	(0.007)	(0.008)	(0.012)
Income	–0.029***	–0.015***	–0.011***	0.005	0.000	0.001

	Female			Male		
	(2)	(3)	(4)	(2)	(3)	(4)
	(0.005)	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)
Constant	0.241***	–0.084**	0.056	0.237***	–0.098***	0.146***
	(0.011)	(0.034)	(0.038)	(0.012)	(0.023)	(0.033)
No. of Obs.	32,340	24,961	24,961	31,913	27,483	27,483
No. of Groups			72			72
Waist Circumference						
Gini	–0.001	0.013	0.013*	–0.016	0.007	0.006
	(0.008)	(0.012)	(0.007)	(0.012)	(0.012)	(0.007)
County Income	–0.016***	–0.023***	–0.001	–0.054***	–0.030***	0.020
	(0.006)	(0.009)	(0.018)	(0.010)	(0.009)	(0.019)
Income	–0.023***	–0.006	–0.005	0.006	0.003	0.002
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)
Constant	0.604***	0.628***	0.336***	0.491***	0.684***	0.340***
	(0.011)	(0.040)	(0.037)	(0.013)	(0.039)	(0.034)
No. of Obs.	36,605	28,759	28,759	38,365	33,549	33,549
No. of Groups			72			72

Note. OLS = ordinary least squares. WHR = waist-to-hip ratio. MAMC = mid-upper arm muscle circumference. BMI = body mass index. Obs. = observations. (2) = basic OLS, (3) = OLS with individual controls and (4) = fixed effects.

When including individual controls variables, Equation (3) shows that the power of an increase in income inequality is still significantly and is associated with a higher chance of having abnormal blood pressure for both genders, WHR and BMI for men, but all magnitudes have been reduced. Similar changes can also apply to income growth, both the magnitude and significance from earning more money are diminishing for women. Interestingly for men, as income grows the power is changed both in its sign and significance, meaning men are more likely to present unusual blood pressure. Uniformly, all parameters on county average income decline after adding the individual controls.

Equation (4) reveals the response of county and year fixed effects. Overall, we find the relationship between income inequality and nurse-collected health measures is disconnected. Although the results of inequality on blood pressure and waist circumference are still significant at 10% for women, compared to Equation (2) and Equation (3), the strength of such outcomes are discounted. Individual income is also disconnected with most of the health measures for both genders, except men's blood pressure. With the exception of MAMC for both genders and BMI for women, the significance of average county income disappears for all genders and nurse-collected health measures.

Appendix B provides a table of the results from Equation (2) to Equation (4), in which we find individual controls reveal informative responses. Old age has a significant effect on health risks, which is intuitive and aligned with human body mechanism. Like age, marital status is closely connected to nurse-collected health outcomes for both genders, but interestingly the connection dichotomises. Married individuals are more likely to have normal blood pressure, waist circumference and MAMC. Conversely, people who marry face a risk of higher BMI and WHR. Men who live in urban areas have a higher risk of having abnormal blood pressure and BMI. The power of education levels is also prevalent across all health outcomes. Disregarding gender, a better-educated individual has a lower chance of abnormal blood pressure and WHR. For women, a higher education level also provides a better indicator of BMI. Whereas for men, a higher education level leads to obesity (BMI and waist circumference), but better MAMC. Occupation and employment sectors also play a role, especially for men. All six categories for men are significant, but with different signs. Working causes harm to health except for MAMC, meaning the indicator of MAMC improves as men work. The outstanding example of employment sector is women's waist circumference. We observed mixed outcomes, as women working in an office tended to have a better waist circumference, whereas women doing manual work were more likely to be exposed to obesity.

5.2.2. Robustness Check

The robustness analysis replaces the Gini with the Theil indices. As shown in Table 4, the outcome of Theil indices exhibits mostly identical trends as shown by the Gini coefficient, while none of the inequality measures show their significance. Men's individual income is statistically robust, as all three Theil indices support the result. Modest evidence shows county average income is robust for both genders' MAMC. In women's BMI, county average income and individual income are found to be statistically robust. Overall in all cases, there is little evidence to support the significance of county average income and individual income.

Table 4 Fixed Effects Models on Nurse-Collected Health with Theil's Indices

	Female			Male		
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
Blood Pressure						
Inequality	–0.010	–0.006	–0.007	0.000	0.004	0.000
	(0.007)	(0.006)	(0.005)	(0.009)	(0.007)	(0.005)
County income	–0.006	–0.003	–0.000	0.005	0.001	0.004
	(0.021)	(0.024)	(0.025)	(0.017)	(0.019)	(0.019)
Income	–0.005	–0.005	–0.005	0.003**	0.003**	0.003**
	(0.004)	(0.004)	(0.004)	(0.001)	(0.001)	(0.001)
Constant	0.027	0.019	0.016	0.289***	0.294***	0.290***
	(0.040)	(0.041)	(0.041)	(0.043)	(0.044)	(0.045)
No. of Obs.	22,345	22,345	22,345	24,612	24,612	24,612
	Female			Male		
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
WHR						
Inequality	0.004	0.011	0.007	0.003	0.005	0.001
	(0.009)	(0.009)	(0.006)	(0.007)	(0.006)	(0.005)
County Income	0.001	–0.011	–0.008	–0.003	–0.008	–0.003
	(0.018)	(0.021)	(0.021)	(0.017)	(0.018)	(0.018)
Income	–0.003	–0.003	–0.003	0.002	0.002	0.002
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
Constant	0.372***	0.367***	0.366***	–0.017	–0.019	–0.017
	(0.035)	(0.036)	(0.037)	(0.031)	(0.031)	(0.032)
No. of Obs.	19,201	19,201	19,201	21,445	21,445	21,445
	Female			Male		
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
MAMC						
Inequality	0.004	0.005	0.002	0.002	0.001	–0.002
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)

	Female			Male		
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
County Income	0.019**	0.016	0.018	0.017	0.017	0.019*
	(0.009)	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)
Income	–0.001	–0.001	–0.001	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.100***	0.097***	0.097***	0.159***	0.159***	0.160***
	(0.015)	(0.016)	(0.016)	(0.027)	(0.027)	(0.027)
No. of Obs.	24,469	24,469	24,469	26,825	26,825	26,825
	Female			Male		
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
BMI						
Inequality	0.000	0.004	0.004	0.003	0.004	0.002
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.004)
County Income	–0.023*	–0.027*	–0.028*	–0.003	–0.006	–0.004
	(0.013)	(0.015)	(0.015)	(0.013)	(0.013)	(0.014)
Income	–0.011***	–0.011***	–0.011***	0.001	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Constant	–0.107***	–0.109***	–0.110***	–0.089***	–0.091***	–0.090***
	(0.029)	(0.029)	(0.030)	(0.028)	(0.028)	(0.028)
No. of Obs.	24,961	24,961	24,961	27,483	27,483	27,483
	Female			Male		
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
Waist Circumference						
Inequality	0.002	0.004	0.000	0.004	0.007	0.004
	(0.008)	(0.006)	(0.004)	(0.009)	(0.007)	(0.004)
County Income	0.001	–0.002	0.001	0.019	0.013	0.015
	(0.018)	(0.020)	(0.022)	(0.018)	(0.020)	(0.022)
Income	–0.005	–0.005	–0.005	0.002	0.002	0.002
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Constant	0.872***	0.870***	0.872***	0.982***	0.978***	0.979***
	(0.035)	(0.035)	(0.036)	(0.032)	(0.032)	(0.033)
No. of Obs.	28,759	28,759	28,759	33,549	33,549	33,549

Note. Fixed effects number of groups: 72. WHR = waist-to-hip ratio. MAMC = mid-upper arm muscle circumference. BMI = body mass index. Obs. = observations.

5.3. Extension with Blood-based Biomarker Analysis

This study takes the advantages of blood-based biomarkers that could offer deeper and more detailed responses from income inequality. We use almost the same models as in 5.2, but with two differences. First, we substitute the nurse-collected health variable with blood-based biomarkers. Second, we use OLS regression instead of the fixed effects model. The reference group is still married Han-ethnic farmers with rural household registration working in the collective sector with no education.

5.3.1. Empirical Results

The results of income inequality from blood-based biomarkers model are not as significant as the ones from nurse-collected markers. Table 5 presents the results of the blood-based biomarkers analysis. Overall, in both models the power of income inequality is insignificant. Without individual control variables, Equation (5) only discovers that individual income is significantly associated with the total cholesterol rate for both genders. With more individual detail, Equation (6) shows that for men, all income inequality, average county income and individual income are disassociated with blood-based biomarkers. The only variable that is significant at the 10% confidence level is average county income on total cholesterol for women, indicating that development of the local economy contributes to better health.

Table 5 OLS Models for Blood-Based Biomarkers

	Female		Male	
	(5)	(6)	(5)	(6)
Total Cholesterol				
Gini	–0.019	–0.012	–0.015	–0.015
	(0.029)	(0.036)	(0.028)	(0.030)
County Income	–0.028	–0.097*	–0.013	–0.039
	(0.044)	(0.052)	(0.044)	(0.048)
Income	0.020**	0.014	0.026**	0.011
	(0.010)	(0.012)	(0.011)	(0.011)
Constant	0.309***	0.103	0.310***	0.373***
	(0.034)	(0.126)	(0.038)	(0.106)
No. of Obs.	3,285	2,364	3,383	2,851
	Female		Male	
CRP				
Gini	0.005	0.009	–0.004	–0.011
	(0.009)	(0.009)	(0.010)	(0.011)
County Income	–0.007	0.001	0.009	0.007
	(0.015)	(0.015)	(0.019)	(0.020)
Income	–0.005	–0.004	–0.000	0.002

	(0.005)	(0.005)	(0.007)	(0.008)
Constant	0.077***	0.043	0.083***	0.112*
	(0.012)	(0.065)	(0.016)	(0.058)
No. of Obs.	2,889	2,056	2,871	2,390
	Female		Male	
HbA1c				
Gini	0.002	0.003	–0.003	–0.005
	(0.010)	(0.008)	(0.010)	(0.012)
County Income	0.020	0.009	0.033	–0.007
	(0.013)	(0.013)	(0.025)	(0.023)
Income	0.011	0.015	0.004	0.002
	(0.013)	(0.015)	(0.005)	(0.006)
Constant	0.048***	–0.068	0.068***	–0.051
	(0.010)	(0.052)	(0.013)	(0.064)
No. of Obs.	2,889	2,056	2,871	2,390

Note. OLS = ordinary least squares. CRP = C-reactive protein. Obs. = observations. (5) = basic OLS and (6) = OLS with individual controls.

Tables showing the responses from individual control variables can be found in Appendix C. Compared to the results from nurse-collected health outcomes analysis, individual controls exhibit similar trends. Age is still the dominating factor causing women to worsen in all three biomarkers, whereas for men only HbA1c worsens. Being a Han Chinese for women is good for their CRP, while it is bad for men. Living in an urban area is worse for men's total cholesterol. Interestingly, a higher education level for women indicates worse total cholesterol. The role of occupation and employment sector is not as impressive as that demonstrated in 5.2.1. Overall, we find women are more sensitive to occupation and employment sector and both genders who do not do manual work perform better in health. Both genders who work in three-capital enterprises improve their HbA1c performance.

5.3.2. Robustness Check

Table 6 shows the results of replacing Gini coefficients with Theil indices. Unlike the result shown in 5.2.1, in the blood-based biomarker analysis the robustness check results do not uniformly agree with the empirical results. This is particularly the case for men's total cholesterol and women's CRP. Theil's L index is significant at the 10% confidence level and its positive sign indicates inequality reduces people's chance of having a healthier total cholesterol. Women's CRP is much stronger, as all three Theil indices are positive and significant at the 5% confidence level. Compared to the insignificant Gini coefficients, the responses from Theil indices show inequality increases people's risk of having abnormal CRP levels. While inequality shows its strength, other variables remain insignificant.

Table 6 OLS Models for Blood-Based Biomarkers with Theil Indices

	Female			Male		
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
Total Cholesterol						
Inequality	0.020	0.018	0.025	0.030*	0.019	0.000
	(0.018)	(0.014)	(0.017)	(0.016)	(0.013)	(0.005)
County Income	–0.081*	–0.098**	–0.108**	–0.013	–0.036	0.004
	(0.046)	(0.046)	(0.047)	(0.039)	(0.041)	(0.019)
Income	0.013	0.012	0.011	0.009	0.009	0.003**
	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.001)
Constant	0.072	0.063	0.058	0.333***	0.339***	0.290***
	(0.123)	(0.122)	(0.121)	(0.106)	(0.109)	(0.045)
No. of Obs.	2,364	2,364	2,364	2,851	2,851	2,851
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
CRP						
Inequality	0.009**	0.008**	0.012**	0.007	0.006	0.002
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.006)
County Income	0.001	–0.007	–0.012	0.016	0.011	0.011
	(0.014)	(0.014)	(0.014)	(0.016)	(0.016)	(0.017)
Income	–0.004	–0.004	–0.004	0.002	0.002	0.002
	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)
Constant	0.035	0.030	0.029	0.098	0.098	0.104*
	(0.065)	(0.064)	(0.064)	(0.059)	(0.059)	(0.059)
No. of Obs.	2,056	2,056	2,056	2,390	2,390	2,390
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
HbA1c						
Inequality	–0.004	–0.000	0.001	–0.007	–0.006	–0.007

	(0.005)	(0.005)	(0.008)	(0.007)	(0.007)	(0.011)
County Income	0.006	0.007	0.007	–0.008	–0.002	0.000
	(0.013)	(0.012)	(0.011)	(0.022)	(0.022)	(0.023)
Income	0.016	0.015	0.015	0.003	0.003	0.003
	(0.014)	(0.014)	(0.014)	(0.006)	(0.006)	(0.006)
Constant	–0.061	–0.066	–0.067	–0.045	–0.044	–0.045
	(0.051)	(0.052)	(0.053)	(0.065)	(0.065)	(0.065)
No. of Obs.	2,056	2,056	2,056	2,390	2,390	2,390

Note. CRP = C-reactive protein. Obs. = observations.

5.4. The Analysis of Lag Effects of Inequality on Nurse-collected Health Measures

As mentioned in Chapter 2, intuition allows this study to explore how lagged inequality affects individual health. With lagged Gini coefficients added, the other specification of models is identical to section 5.2.1. The reference group is also the same group as in section 5.2.1.

5.4.1. Empirical Results

Focusing on the key variables provided in Table 7, the overall trend is that the significance and magnitude of both the current Gini coefficient and the lagged Gini coefficients are declining, while this trend also applies to county average income and individual income. In the basic OLS model, although inequality and its lagged version show their significant explanation power almost everywhere, we find the relationship is weak for MAMC for men, BMI for men and waist circumference for women. The difference appears in the results, as several coefficients of inequality are negative meaning income inequality contributes to better health. The negative inequality coefficients are all lagged ones. The majority coefficients of individual income show a negative sign that is in line with our assumption that increases in individual income improve individual health. In contrast, 9 out of 10 coefficients on county average income are significant, but with different signs. While local economic development is good for both genders' MAMC, it negatively affects both genders' blood pressure, WHR, BMI and waist circumference. After controlling individual variables, the results corroborate previous results without individual controls, but with a few modest adjustments. When further adding the fixed effects, we find the overall significance of the Gini coefficients and the lagged Gini is reduced. Only two significant lagged coefficients are found overall to show inequality is bad for individual health. Although modest and limited evidence is found on the power of income inequality on individual health, it still reveals the importance of lag effects.

Table 7 OLS and Fixed Effects Model for Nurse-Collected Health Outcomes with Lagged Inequality

	Female			Male		
	(7)	(8)	(9)	(7)	(8)	(9)
Blood Pressure						
Gini	0.022**	0.008	0.023	0.003	0.004	0.007
	(0.011)	(0.011)	(0.017)	(0.011)	(0.012)	(0.016)
Gini L1	0.010	–0.009	0.011	–0.007	–0.027**	–0.018
	(0.013)	(0.014)	(0.017)	(0.012)	(0.012)	(0.016)
Gini L2	0.026*	0.010	0.031*	0.030**	0.019	0.017
	(0.013)	(0.017)	(0.017)	(0.015)	(0.018)	(0.019)
Gini L3	–0.010	0.012	0.003	–0.023	–0.012	0.009
	(0.017)	(0.021)	(0.020)	(0.016)	(0.019)	(0.018)
County Income	0.101***	0.052***	–0.056*	0.125***	0.079***	0.030
	(0.016)	(0.016)	(0.029)	(0.015)	(0.016)	(0.025)
Income	–0.035**	–0.007	0.002	–0.007	–0.006	–0.007
	(0.015)	(0.011)	(0.010)	(0.008)	(0.008)	(0.008)
Constant	0.435***	–0.389***	–0.217**	0.554***	0.048	0.168**
	(0.016)	(0.065)	(0.090)	(0.014)	(0.082)	(0.084)
No. of Obs.	6,490	4,998	4,998	6,744	5,764	5,764
No. of Groups			54			54
WHR						
Gini	0.014**	0.004	–0.006	0.020*	0.026**	–0.008
	(0.007)	(0.007)	(0.010)	(0.010)	(0.011)	(0.013)
Gini L1	0.017**	0.007	0.000	0.012	–0.001	–0.013
	(0.008)	(0.011)	(0.012)	(0.011)	(0.010)	(0.011)
Gini L2	0.028**	0.018	0.015	0.024**	0.026**	0.019
	(0.011)	(0.013)	(0.014)	(0.010)	(0.011)	(0.013)
Gini L3	–0.014	–0.011	–0.007	–0.012	0.002	–0.005
	(0.010)	(0.011)	(0.013)	(0.012)	(0.012)	(0.013)
County Income	0.031***	0.011	–0.015	0.085***	0.049***	–0.008
	(0.010)	(0.015)	(0.030)	(0.017)	(0.017)	(0.035)
Income	–0.009	0.010	0.011	0.024***	0.014**	0.012*

		Female			Male	
	(7)	(8)	(9)	(7)	(8)	(9)
	(0.009)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	0.790***	0.483***	0.537***	0.370***	–0.092	0.291***
	(0.009)	(0.058)	(0.084)	(0.016)	(0.067)	(0.083)
No. of Obs.	6,553	5,032	5,032	6,819	5,834	5,834
No. of Groups			54			54
MAMC						
Gini	0.003	0.001	–0.007	0.003	0.004	–0.002
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.006)
Gini L1	–0.008*	–0.008	–0.007	0.002	0.004	0.013*
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
Gini L2	–0.001	–0.001	0.003	–0.002	–0.008	–0.002
	(0.005)	(0.006)	(0.007)	(0.008)	(0.008)	(0.008)
Gini L3	0.018***	0.015**	0.010	0.005	0.000	–0.010
	(0.006)	(0.006)	(0.008)	(0.007)	(0.006)	(0.009)
County Income	–0.019***	–0.016***	0.007	–0.028***	–0.027***	0.002
	(0.006)	(0.006)	(0.011)	(0.008)	(0.008)	(0.016)
Income	0.001	–0.004	–0.002	–0.003	0.005	0.005
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Constant	0.049***	0.020	–0.023	0.087***	0.070*	–0.065
	(0.006)	(0.040)	(0.044)	(0.009)	(0.039)	(0.055)
No. of Obs.	6,612	5,077	5,077	6,871	5,878	5,878
No. of Groups			54			54
BMI						
Gini	0.005	0.002	0.008	0.013	0.018*	–0.002
	(0.008)	(0.010)	(0.012)	(0.010)	(0.010)	(0.010)
Gini L1	0.012	0.012	0.007	0.013	0.008	0.013
	(0.008)	(0.011)	(0.012)	(0.009)	(0.009)	(0.010)
Gini L2	–0.012	–0.010	–0.008	–0.009	–0.007	0.002
	(0.010)	(0.011)	(0.009)	(0.009)	(0.010)	(0.010)
Gini L3	–0.021**	–0.016	0.003	–0.005	0.001	0.008
	(0.009)	(0.011)	(0.013)	(0.011)	(0.010)	(0.012)
County Income	0.087***	0.062***	–0.051*	0.085***	0.054***	0.000
	(0.015)	(0.020)	(0.029)	(0.015)	(0.017)	(0.024)
Income	–0.022**	–0.013	–0.005	0.039***	0.018***	0.017**
	(0.009)	(0.009)	(0.010)	(0.007)	(0.006)	(0.007)
Constant	0.257***	0.113	0.254***	0.243***	–0.099*	0.129**
	(0.015)	(0.083)	(0.093)	(0.016)	(0.051)	(0.064)
No. of Obs.	6,659	5,119	5,119	6,935	5,940	5,940
No. of Groups			54			54
Waist Circumference						
Gini	0.016	0.013	0.007	0.021	0.032**	0.007
	(0.012)	(0.014)	(0.019)	(0.015)	(0.014)	(0.011)
Gini L1	0.017	0.011	0.007	–0.016*	–0.001	0.013
	(0.011)	(0.013)	(0.014)	(0.009)	(0.010)	(0.013)
Gini L2	0.006	0.008	0.008	–0.004	0.009	0.020

	Female			Male		
	(7)	(8)	(9)	(7)	(8)	(9)
	(0.016)	(0.017)	(0.015)	(0.012)	(0.013)	(0.014)
Gini L3	–0.020	–0.010	0.027	0.003	0.000	0.020
	(0.017)	(0.018)	(0.017)	(0.015)	(0.014)	(0.013)
County Income	0.083***	0.050**	–0.042	0.029	0.018	0.047*
	(0.018)	(0.024)	(0.042)	(0.018)	(0.016)	(0.027)
Income	–0.013	–0.004	0.003	0.033***	0.014**	0.014**
	(0.015)	(0.011)	(0.010)	(0.008)	(0.008)	(0.008)
Constant	0.499***	0.229***	0.365***	0.333***	0.227***	0.331***
	(0.020)	(0.083)	(0.101)	(0.020)	(0.061)	(0.083)
No. of Obs.	7,103	5,481	5,481	7,842	6,778	6,778
No. of Groups			54			54

Note. OLS = ordinary least squares. WHR = waist-to-hip ratio. MAMC = mid-upper arm muscle circumference. BMI = body mass index. Obs. = observations. (7) = basic OLS, (8) = OLS with individual controls and (9) = fixed effects.

Individual control variables are shown in detail in the tables in the Appendix D. Age is everywhere significant, especially for both genders' blood pressure, WHR and MAMC. Men's BMI and women's waist circumference are also strongly affected by age. Marital status shows conflicting force for blood pressure—married women tend to have abnormal blood pressure, whereas marriage improves blood pressure for men. When living in urban areas, both men and women gain weight, although the indicator of obesity for women is waist circumference and BMI for men. A higher education level contributes to worse WHR for men, but the consequence on obesity differs by gender. Demonstrated by both BMI and waist circumference, better education makes women thinner, but men heavier. For occupation and employment sectors, men are more sensitive overall than women. For both genders, MAMC improves as people work, while other indicators show the opposite.

5.4.2. Robustness Check

Table 8 presents the robustness check results of 5.4.1. All the significances in women's blood pressure disappeared after the replacement of Theil indices. While the second wave lagged Gini coefficient and county average income are significant, the robustness check shows the opposite direction. None of the variables is significant for men's blood pressure in the original outputs, in contrast, the third wave lagged Theil's T is significant at the 10% confidence level. WHR is similar to blood pressure, whereas men's individual income fully supports the results of the Gini coefficient with both identical value at 10% confidence level significance. The strong support from individual income continues with men's BMI and waist circumference having the same performance, in which the identical value and significance happen.

Table 8 Fixed Effects Model for Nurse-Collected Health Outcomes with Lagged Theil Indices

	Female			Male		
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
Blood Pressure						
Inequality	–0.012	–0.005	0.002	–0.015	–0.008	–0.009
	(0.013)	(0.011)	(0.009)	(0.012)	(0.011)	(0.011)
Inequality L1	–0.002	0.004	0.002	0.006	0.002	–0.008
	(0.014)	(0.013)	(0.015)	(0.013)	(0.011)	(0.015)
Inequality L2	0.021	0.022	0.012	–0.007	–0.004	–0.011
	(0.014)	(0.014)	(0.019)	(0.016)	(0.013)	(0.014)
Inequality L3	–0.004	–0.006	–0.003	–0.015	–0.019*	–0.020
	(0.017)	(0.016)	(0.018)	(0.014)	(0.011)	(0.019)
County Income	–0.028	–0.026	–0.037	0.038	0.039	0.041
	(0.032)	(0.035)	(0.035)	(0.026)	(0.028)	(0.026)
Income	0.001	0.002	0.002	–0.007	–0.007	–0.008
	(0.010)	(0.010)	(0.010)	(0.008)	(0.008)	(0.008)
Constant	0.323***	0.320***	0.307***	0.128*	0.124*	0.121*
	(0.071)	(0.069)	(0.069)	(0.070)	(0.072)	(0.072)
No. of Obs.	5,689	5,689	5,689	6,597	6,597	6,597
WHR						
Inequality	–0.002	0.003	0.010	0.002	–0.001	–0.007
	(0.011)	(0.011)	(0.009)	(0.009)	(0.008)	(0.008)
Inequality L1	–0.021*	–0.007	–0.004	–0.010	–0.012	–0.017*
	(0.012)	(0.010)	(0.012)	(0.010)	(0.008)	(0.009)
Inequality L2	–0.011	–0.015	–0.013	–0.014	–0.010	–0.018
	(0.013)	(0.013)	(0.015)	(0.011)	(0.011)	(0.014)
Inequality L3	–0.015	–0.005	0.004	0.011	0.012	0.009
	(0.013)	(0.012)	(0.014)	(0.013)	(0.013)	(0.017)
County Income	–0.031	–0.030	–0.035	–0.017	–0.012	–0.009
	(0.028)	(0.029)	(0.028)	(0.033)	(0.034)	(0.035)
Income	0.010	0.010	0.010	0.012*	0.012*	0.012*
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	0.523***	0.517***	0.517***	0.193***	0.190***	0.189***

	Female			Male		
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
	(0.072)	(0.072)	(0.071)	(0.067)	(0.069)	(0.069)
No. of Obs.	5,736	5,736	5,736	6,695	6,695	6,695
MAMC						
Inequality	0.007	0.006	0.003	–0.006	–0.004	–0.003
	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)
Inequality L1	–0.005	–0.006	–0.007	–0.006	–0.003	0.004
	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)	(0.009)
Inequality L2	–0.000	–0.003	–0.006	0.002	–0.001	0.001
	(0.005)	(0.004)	(0.004)	(0.007)	(0.006)	(0.009)
Inequality L3	0.011*	0.000	–0.011*	0.003	–0.001	–0.008
	(0.006)	(0.006)	(0.006)	(0.007)	(0.008)	(0.013)
County Income	0.010	0.003	0.004	–0.002	–0.001	–0.002
	(0.011)	(0.012)	(0.012)	(0.018)	(0.018)	(0.018)
Income	–0.001	–0.001	–0.002	0.005	0.005	0.005
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Constant	–0.018	–0.008	–0.012	–0.062	–0.060	–0.063
	(0.043)	(0.042)	(0.042)	(0.050)	(0.051)	(0.050)
No. of Obs.	5,793	5,793	5,793	6,749	6,749	6,749
BMI						
Inequality	0.005	–0.010	0.003	0.001	0.004	0.004
	(0.008)	(0.009)	(0.011)	(0.008)	(0.008)	(0.008)
Inequality L1	0.012	–0.004	–0.005	–0.001	–0.006	–0.012
	(0.008)	(0.010)	(0.009)	(0.010)	(0.009)	(0.011)
Inequality L2	–0.012	0.003	0.003	0.005	–0.003	–0.016
	(0.010)	(0.008)	(0.008)	(0.008)	(0.008)	(0.011)
Inequality L3	–0.021**	0.006	0.002	0.000	0.004	–0.002
	(0.009)	(0.010)	(0.013)	(0.008)	(0.008)	(0.013)
County Income	0.087***	–0.027	–0.045	–0.001	–0.007	–0.012
	(0.015)	(0.031)	(0.029)	(0.023)	(0.026)	(0.028)
Income	–0.022**	–0.006	–0.006	0.017**	0.017**	0.017**
	(0.009)	(0.009)	(0.009)	(0.006)	(0.006)	(0.006)
Constant	0.257***	0.205**	0.221**	0.083	0.094*	0.100*
	(0.015)	(0.086)	(0.083)	(0.051)	(0.050)	(0.052)
No. of Obs.	5,831	5,831	5,831	6,805	6,805	6,805
Waist Circumference						
Inequality	0.004	0.003	0.012	0.004	0.003	0.004
	(0.016)	(0.012)	(0.012)	(0.014)	(0.011)	(0.010)
Inequality L1	–0.014	–0.010	–0.015	–0.015	–0.025***	–0.040***
	(0.012)	(0.012)	(0.014)	(0.012)	(0.009)	(0.013)
Inequality L2	–0.006	–0.021	–0.030*	0.003	–0.004	–0.018
	(0.016)	(0.014)	(0.015)	(0.011)	(0.010)	(0.012)
Inequality L3	0.015	0.004	–0.010	0.005	0.001	–0.013
	(0.016)	(0.015)	(0.015)	(0.011)	(0.012)	(0.020)

	Female			Male		
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
County Income	–0.032	–0.037	–0.053	0.052*	0.046	0.041
	(0.036)	(0.038)	(0.038)	(0.028)	(0.029)	(0.031)
Income	0.002	0.001	0.001	0.014**	0.014**	0.014**
	(0.010)	(0.011)	(0.010)	(0.006)	(0.006)	(0.006)
Constant	0.263***	0.272***	0.281***	0.267***	0.284***	0.291***
	(0.077)	(0.078)	(0.079)	(0.074)	(0.075)	(0.075)
No. of Obs.	6,235	6,235	6,235	7,729	7,729	7,729

Note. WHR = waist-to-hip ratio. MAMC = mid-upper arm muscle circumference. BMI = body mass index. Obs. = observations.

5.5. Analysis of Lag Effects of Inequality on Blood-based Biomarkers

In this section, we apply the dynamic lag effects analysis on blood-based biomarkers. The model set up is similar to section 5.3; however, with the extension of added lag effects. We first use OLS regression with basic key variables and then add individual control variables. The reference group is consistent with previous analyses.

5.5.1. Empirical Results

The results are shown in Table 9. Equation (10) regresses the basic OLS model with the Gini coefficient and its lags, county average income and individual income. We find mixed results in that while the current inequality level does not significantly attach to blood-based biomarkers, lagged inequality has mixed outcomes on health outcomes. For both men and women, total cholesterol is significantly affected by lagged inequality measures. While the three years' lag inequality improves the total cholesterol indicator, the five years' lag Gini coefficients for both genders and nine years' lag version for men impede healthier improvement. For CRP, the only significant variable for women is the nine years' lag Gini coefficient. The outlier in the OLS model is men's CRP, as both current and lagged inequality do not affect this indicator. The case of HbA1c is similar to the case of total cholesterol, in which for women the three years' lag and five years' lag inequality are in significantly opposite directions, whereas for men only the five years' lag inequality is powerful enough to support an improvement in HbA1c.

Equation (11) adds individual control variables, although some lagged inequality levels still improve health, the overall magnitude of lagged inequality draw the power back to reducing health performances. Most of the variables in total cholesterol stay the same sign and significance, except county average income. Adding individual controls has a negative and significant power leading to the interpretation that the development of local economy contributes to residents' health. CRP for women loses its significant three years' lag inequality measure, meanwhile for men the weight of individual income becomes significant and stronger, in line with the assumption of this study. HbA1c for men retains the same status, but women's coefficients change slightly. The nine years' lag inequality shows its power to make women's HbA1c worse, while women's individual income loses its significance. Overall, we find the trend in this biomarker analysis that all current Gini coefficients are not significant. County average incomes are all in the expected direction, but not significant. Individual income is only beneficial and significant for men's CRP.

Table 9 OLS Model for Blood-Based Biomarkers with Lagged Inequality

	Female		Male	
	(10)	(11)	(10)	(11)
Total Cholesterol				
Gini	–0.007	–0.015	–0.012	–0.017
	(0.050)	(0.055)	(0.027)	(0.033)
Gini L1	–0.108**	–0.135**	–0.078**	–0.067*
	(0.047)	(0.063)	(0.035)	(0.038)
Gini L2	0.074*	0.082*	0.090***	0.097***
	(0.038)	(0.048)	(0.029)	(0.034)
Gini L3	0.018	0.016	0.065***	0.065***
	(0.026)	(0.035)	(0.022)	(0.022)
County Income	–0.091	–0.178*	–0.021	–0.053
	(0.064)	(0.092)	(0.046)	(0.075)
Income	0.021	0.029	0.035	0.009
	(0.025)	(0.027)	(0.021)	(0.019)
Constant	0.329***	0.018	0.311***	0.337*
	(0.043)	(0.590)	(0.033)	(0.201)
No. of Obs.	958	714	1,037	883

	Female		Male	
	(10)	(11)	(10)	(11)
CRP				
Gini	–0.016	–0.017	–0.019	–0.018
	–0.012	–0.012	–0.019	–0.019
Gini L1	0.023	0.040**	–0.002	–0.003
	–0.018	–0.016	–0.024	–0.023
Gini L2	–0.001	–0.001	–0.013	–0.012
	–0.016	–0.016	–0.017	–0.017
Gini L3	–0.034**	–0.043***	–0.019	–0.016
	–0.013	–0.013	–0.016	–0.015
County Income	–0.022	–0.036	–0.017	–0.034
	–0.026	–0.03	–0.036	–0.045
Income	0.007	0.015	–0.011	–0.015**
	–0.015	–0.014	–0.008	–0.007
Constant	0.072***	0.889***	0.101***	0.130
	(0.017)	(0.165)	(0.024)	(0.143)
No. of Obs.	878	657	933	787
HbA1c				
Gini	–0.006	–0.019	–0.001	–0.004
	(0.011)	(0.011)	(0.016)	(0.016)
Gini L1	–0.038*	–0.064**	–0.020	–0.006
	(0.020)	(0.025)	(0.027)	(0.029)
Gini L2	0.056***	0.073***	0.043*	0.048*
	(0.015)	(0.019)	(0.022)	(0.024)
Gini L3	0.016	0.034**	0.023	0.028
	(0.011)	(0.015)	(0.018)	(0.018)
County Income	0.005	–0.017	0.050	–0.010
	(0.021)	(0.024)	(0.031)	(0.034)
Income	–0.016*	–0.012	0.012	0.010
	(0.010)	(0.011)	(0.015)	(0.016)
Constant	0.065***	–0.234*	0.063***	0.008
	(0.015)	(0.118)	(0.019)	(0.101)
No. of Obs.	878	657	933	787

Note. OLS = ordinary least squares. CRP = C-reactive protein. Obs. =observations

Individual control variables offer more detailed feedback (see the Appendix E for tables). Age significantly drags down the HbA1c level for both genders, while women's total cholesterol and men's CRP are also significantly affected by age. The response from marital status shows married women are healthier in terms of CRP. Ethnic majority emphasises that being a Han Chinese harms both genders' HbA1c. Women living in cities tend to be worse in total cholesterol. Education levels have no association on women, but slightly impede men's total cholesterol. It is worth noting the power of occupation and employment sectors in this analysis, as all statistically significant variants for both genders in this analysis contribute to better health.

5.5.2. Robustness Check

Similar to all previous robustness checks, we employ Theil indices in the full model with individual control variables. The details are presented in Table 10. The feedback from lagged total cholesterol partially supports the original Gini-based results. For women, both Theil's T and Theil's V make three years' lag inequality robust, while the nine years' lag inequality gains the significance. Men's total cholesterol is not well supported by the Theil indices, as only Theil's L shows five years' lag is significant while others are not supported. In the case of CRP, for women we find little support, but for men the story is different. Alongside the strong and significant indicator of five years' lag and three years' lag Theil indices, individual income is statistically robust. For HbA1c, nine years' lag inequality is proved statistically robust for both genders, although other inequality measures are not backed by the Theil indices.

Table 10 OLS Model for Blood-Based Biomarkers with Lagged Theil Indices

		Female			Male	
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
Total Cholesterol						
Inequality	–0.021	–0.009	0.019	0.014	0.016	0.025
	(0.033)	(0.024)	(0.023)	(0.025)	(0.021)	(0.027)
Inequality L1	0.041	0.044*	0.067*	0.009	0.011	0.014
	(0.029)	(0.025)	(0.039)	(0.023)	(0.025)	(0.042)
Inequality L2	0.024	0.004	–0.026	0.059**	0.055	0.056
	(0.033)	(0.034)	(0.054)	(0.026)	(0.034)	(0.057)
Inequality L3	0.037	0.069**	0.172***	–0.003	–0.008	–0.020
	(0.034)	(0.029)	(0.051)	(0.027)	(0.027)	(0.054)
County Income	–0.067	–0.054	–0.087	0.034	–0.028	–0.063
	(0.086)	(0.079)	(0.073)	(0.062)	(0.061)	(0.063)
Income	0.025	0.027	0.025	0.016	0.015	0.014
	(0.028)	(0.027)	(0.026)	(0.021)	(0.021)	(0.021)
Constant	–0.009	–0.031	–0.034	0.237	0.252	0.287
	(0.568)	(0.550)	(0.530)	(0.216)	(0.216)	(0.213)

	Female			Male		
	Theil's L	Theil's T	Theil's V	Theil's L	Theil's T	Theil's V
No. of Obs.	714	714	714	883	883	883
CRP						
Inequality	0.003	–0.003	–0.003	–0.001	–0.002	0.003
	(0.013)	(0.011)	(0.013)	(0.012)	(0.010)	(0.009)
Inequality L1	–0.017**	–0.007	–0.001	–0.018**	–0.012*	–0.011
	(0.008)	(0.006)	(0.011)	(0.007)	(0.007)	(0.011)
Inequality L2	0.035***	0.023	0.012	0.045***	0.050***	0.061**
	(0.013)	(0.018)	(0.024)	(0.013)	(0.017)	(0.025)
Inequality L3	–0.018	–0.001	0.016	0.000	0.001	–0.025
	(0.011)	(0.012)	(0.019)	(0.011)	(0.013)	(0.025)
County Income	0.000	0.007	0.013	–0.001	–0.020	–0.045
	(0.032)	(0.029)	(0.029)	(0.030)	(0.031)	(0.032)
Income	0.013	0.016	0.016	–0.013*	–0.014**	–0.014**
	(0.015)	(0.015)	(0.015)	(0.007)	(0.007)	(0.007)
Constant	0.861***	0.875***	0.874***	0.100	0.089	0.109
	(0.160)	(0.166)	(0.167)	(0.142)	(0.146)	(0.143)
No. of Obs.	657	657	657	787	787	787
HbA1c						
Inequality	–0.015	–0.009	–0.008	–0.002	0.001	0.003
	(0.012)	(0.008)	(0.009)	(0.018)	(0.013)	(0.013)
Inequality L1	0.004	0.008	0.007	–0.002	–0.003	–0.007
	(0.012)	(0.012)	(0.019)	(0.011)	(0.010)	(0.016)
Inequality L2	0.010	0.003	0.006	–0.008	–0.005	0.002
	(0.014)	(0.013)	(0.021)	(0.017)	(0.021)	(0.030)
Inequality L3	0.021**	0.030**	0.052**	0.017	0.033*	0.084***
	(0.010)	(0.011)	(0.024)	(0.018)	(0.019)	(0.031)
County Income	0.030	0.042*	0.028	0.012	0.030	0.027
	(0.024)	(0.025)	(0.025)	(0.045)	(0.039)	(0.036)
Income	–0.016	–0.016	–0.016	0.011	0.012	0.012
	(0.011)	(0.012)	(0.012)	(0.017)	(0.017)	(0.017)
Constant	–0.216*	–0.211*	–0.198	–0.017	–0.045	–0.056
	(0.118)	(0.120)	(0.120)	(0.107)	(0.103)	(0.100)
No. of Obs.	657	657	657	787	787	787

Note. CRP = C-reactive protein. Obs. = observations.

5.6. Conclusion

This chapter presents our empirical analysis of income inequality–health for China. The major findings first are that Bakkeli's conclusions hold with the updated CHNS dataset. We extend the test with blood-based biomarkers and past inequalities. We then find several connections for the relationship between income inequality and health. Chapter 6 summarises and discusses our findings and considerations.

Chapter 6. Final Thoughts and Conclusions

In this study, we first extend Bakkeli's study making two significant contributions: 1) we employ the up-to-date dataset and 2) we extend the health measures with detailed information on blood-based biomarkers. In the nurse-collected health measures analysis, unlike the absolute conclusion that no association is found between income inequality and nurse-collected health measures after controlling individual conditions, county and city level fixed effects and year dummies, we find Gini coefficients significantly affect individual health in few cases. This finding does not fully comply with the conclusion by Bakkeli (2016). These results hold when we extend the study with blood-based biomarkers. Due to the availability of dataset, we replaced fixed effects regression with OLS analysis and found inequality was still disconnected with blood-based biomarkers. This result further confirms Bakkeli's findings.

To reveal more detailed responses and uncover the potential relationship between income inequality and health outcomes, this study extends the measurement of income inequality by adding past income inequalities. Then, we test how the lagged inequality levels affect both types of health outcomes. Considering evidence from previous studies (Blakely et al., 2000; Foverskov & Holm, 2016; Pickett & Wilkinson, 2015), we employ three waves that typically equal nine year lag measures of inequalities. In the analysis of nurse-collected health measures, we find little evidence of the outcome of income inequality on health for both genders. Conversely, the blood-based biomarkers analysis shows that the lagged inequality significantly affects blood-based biomarkers. We find three interesting characteristics from the lagged inequalities analysis. First, all significant Gini coefficients are lagged, which means none of the contemporary Gini coefficients is significant. Second, mixed results are found in this case. Both positive and negative signs are attached to Gini coefficients. Third, significant coefficients on individual income and county average income show a negative relationship with health, as we expected. We assume an improvement in individual income and local economy contribute to better individual health outcomes. Following this assumption, signs on both variables are expected to be negative.

This study adopts a dynamic approach to understanding income inequality using the objective measures of health and a multidimensional approach to individuals' socioeconomic position in society. We include multiple individual control variables to better understand how these details perform in the whole model. Age, marital status and years of education are each found to play important roles in all analyses. The influence of ethnic majority and location of residency are only observable in the analysis of blood-based biomarkers. Age is an outstanding individual control variable and is almost a statistically significant negative everywhere, especially for women. This confirms the logic that body function discounts as our age increases. Another influential individual control is marital status. We find people's health overall tends to improve if they marry, emphasising the importance of forming a family. Education level also shows its affect, the prevailing trend is that a higher education level provides people with better health conditions. The location where individuals live also touches people's health status. Our finding shows people living in urban areas are healthier than people living in rural areas. This is in line with better infrastructure and medical treatment conditions in the cities that offer residents a better chance of being healthy.

The feedback from occupations and employment sectors tell us first that men are more sensitive to occupations than women. Almost all occupations affect men's health, with the majority harming health. Such a finding could be interpreted as men's health would be worse as long as they keep working. Second, in both nurse-collected measures analyses, all MAMC indicators are negatively connected with current and lagged inequalities. MAMC is the indicator of malnutrition; therefore, we can conclude that when we broaden the time range people's nutrition status would improve as long as they are working, no matter what and where they work. This finding is complementary to the conclusion from Bakkeli (2016) that argues that men in rural areas are more likely to be malnourished. Lastly, in the blood-based biomarkers analysis, although mixed signs are found in the analysis without lags, after adding the lags all significant occupations and employment sectors factors are allocated with negative signs. The negative relationship shows in the long-term for both genders that working improves their performance of blood-based biomarkers.

To find the mechanism of how inequality affects health, we attempt to control every variable and obtain its response. Interestingly, we find the lag effects become stronger when we remove the year dummies. The cases of significance of inequality doubles and exhibit the trend that men are more exposed to the influence of lagged inequality. The interpretation is that when year dummies are added into models, they extract the significance and effect from each year and present these separately. The outcomes with year dummies also highlight the aggregate longitudinal effects.

Although not shown in this thesis, we attempt the option to use actual values for each health measures. By obtaining the weights of lagged inequality, we make the substitution to try to provide a better interpretation of negative coefficients on inequality (which means the controversial idea that inequality improves individual health). The outcomes from this try are overall similar to the original results and indicate there is no sufficient need to replace.

China has been experiencing fast economic growth. Following the path of already developed nations, growing income inequality alongside the country's development is currently emerging as an issue not to be ignored. Urban–rural diversity is a unique characteristic of Chinese society structure and such a diversity initiated and currently enlarges the inequality between rural and urban areas.

Previous studies (Bakkeli, 2016; Kawachi & Kennedy, 1999; Li & Zhu, 2006; Marmot, 2007) and the experiences from developed countries tell us that consumption patterns will change according to income change. Specifically, after people become affluent, the consumption of cigarettes and alcohol goes up. Based on the path of Chinese economy development, the first batch of rich people emerged after the reform of Chinese economy in late 1980s. Therefore, it is worth focusing on such a cohort to find out how income and income inequality affect this cohort's specific health outcomes related to their lifestyle change.

Due to the limited timeframe and availability of dataset, this study could not achieve some of our initial targets. First, we could not apply the biomarker analysis to the whole panel dataset. The author is aware that the biomarker data from CHNS wave 2015 will be released in late 2019. Unfortunately, until the time this study is undertaken, we could not access the latest biomarker data. Second, further extensions could be applied. Pickett and Wilkinson (2015) pointed out future studies should focus on more detailed measurements of income inequality as well as health. While we extend the measure of health with the blood-based biomarkers and include past inequalities, we are aware that the utilisation of relative deprivation and multidimensional inequality as the extension measure of income inequality could help us observe more all-round findings on this topic. The utilisation of multidimensional inequality as the measurement of inequality will make a comparison between traditional income inequality and all-round inequality. Lastly, instead of using income inequality as the key independent variable, future researches could attempt to pick wealth as the substitution.

Appendix

Appendix A. Specification of Income Inequality Variables

This study employs the Gini coefficient and Theil indices (the Generalised Entropy index) to measure income inequality level. We specifically use three types of Theil indices namely: Theil L (GE(0)), Theil T (GE(1)), and Theil V (GE(2)). These four income inequality variables are calculated through Stata code - inequal – which is developed by Whitehouse (1995).

Gini coefficient is calculated by Stata code using the following equation

$$Gini = 1 + \left(\frac{1}{N}\right) - \frac{2}{m \cdot N^2} \sum_{i=1}^{n} (N - i + 1) y_i$$

where individuals are ranked by y_i with ascending order. The Generalised Entropy index is calculated by the following equation

$$GE(a) = \frac{1}{a(1-a)} \left\{ \left[\sum_{i=1}^{n} f_i \left(\frac{y_i}{m}\right)^a \right] - 1 \right\}, a \neq 0, a \neq 1$$

then the Theil indices are given by

$$GE(0) = \sum_{i=1}^{n} f_i \log\left(\frac{y_i}{m}\right)$$

$$GE(1) = \sum_{i=1}^{n} f_i \left(\frac{y_i}{m}\right) \log\left(\frac{y_i}{m}\right)$$

$$GE(2) = -\frac{1}{2} \left\{ \left[\sum_{i=1}^{n} f_i \left(\frac{y_i}{m}\right)^2 \right] - 1 \right\}.$$

Test result of correlation

	Gini (z-score)	Mean county income (z-score)	Income (z-score)	Age	Married	Majority	Urban	Years of education
Gini (z-score)	1.0000							
Mean county income (z-score)	0.1369	1.0000						
Income (z-score)	0.0537	0.3874	1.0000					
Age	0.0915	0.2591	0.0562	1.0000				
Married	0.0517	0.1150	0.0414	0.5041	1.0000			
Majority	-0.0919	0.0791	0.0362	0.0134	0.0233	1.0000		
Urban	-0.0578	0.1917	0.1237	0.1021	0.0382	0.1044	1.0000	
Years of education	0.0322	0.3170	0.2003	-0.2414	-0.1014	0.0923	0.3681	1.0000

Test result of multicollinearity

Variable	VIF	1/VIF
Gini (z-score)	1.11	0.900936
Mean county income (z-score)	1.59	0.629242
Income (z-score)	1.16	0.862993
Age	1.68	0.593876
Married	1.35	0.742318
Majority	1.04	0.962563
Urban	1.52	0.655907
Years of education	1.88	0.531285
Mean VIF	1.42	

Appendix B. Tables for Analysis of Nurse-collected Health Measures (Chapter 5.2)

Table B.1 OLS and fixed-effects regressions with blood pressure.

Blood pressure	Female (1)	Female (2)	Female (3)	Male (1)	Male (2)	Male (3)
County variables						
Gini (z-score)	0.032***	0.014**	0.019**	0.030***	0.017***	0.016
	(0.006)	(0.006)	(0.009)	(0.006)	(0.006)	(0.010)
Mean county income	0.101***	0.051***	-0.012	0.110***	0.067***	0.003
(z-score)	(0.009)	(0.007)	(0.020)	(0.008)	(0.008)	(0.016)
Individual variables						
Income (z-score)	0.032***	-0.007*	-0.005	-0.000	0.003**	0.003**
	(0.007)	(0.004)	(0.004)	(0.001)	(0.001)	(0.001)
Age		0.013***	0.013***		0.010***	0.010***
		(0.001)	(0.000)		(0.000)	(0.000)
Married		0.060***	0.068***		-0.018	-0.027**
		(0.011)	(0.012)		(0.012)	(0.013)
Majority		0.052*	0.014		0.023	-0.014
		(0.027)	(0.014)		(0.026)	(0.015)
Urban		-0.004	0.016		0.005	0.035***
		(0.015)	(0.015)		(0.014)	(0.013)
Years of education		0.001	-0.002*		0.001	-0.002*
		(0.002)	(0.001)		(0.002)	(0.001)
Occupation (ref.=farmer)						
Service class		-0.044*	-0.052**		0.085***	0.046**
		(0.025)	(0.024)		(0.024)	(0.018)
Non-manual worker		-0.024	-0.033		0.083***	0.039*
		(0.022)	(0.020)		(0.026)	(0.021)
Skilled-workers/		0.022	0.018		0.075***	0.031*
supervisor		(0.026)	(0.023)		(0.022)	(0.016)
Semi-skilled/		0.024	0.018		0.076***	0.038**
non-skilled worker		(0.019)	(0.018)		(0.021)	(0.018)
Others		-0.013	-0.010		0.056**	0.041*
		(0.020)	(0.018)		(0.024)	(0.021)
Sector (ref.=collective)						
State		0.006	0.004		-0.004	-0.005
		(0.013)	(0.013)		(0.017)	(0.013)
Family farming		0.069***	0.034*		0.077***	-0.009
		(0.016)	(0.018)		(0.018)	(0.018)
Individual enterprise		0.024	0.003		0.035**	-0.013

		(0.017)	(0.014)		(0.017)	(0.015)
Private/three-capital		-0.051	-0.058		-0.032	-0.053
enterprises		(0.042)	(0.044)		(0.037)	(0.034)
Others		-0.024	-0.039		-0.020	-0.067*
		(0.024)	(0.025)		(0.034)	(0.035)
Wave						
1991			0.222***			-0.278***
			(0.050)			(0.051)
1993			0.199***			-0.249***
			(0.048)			(0.046)
1997			0.147***			-0.202***
			(0.046)			(0.049)
2000			0.160***			-0.173***
			(0.044)			(0.039)
2004			0.155***			-0.108***
			(0.042)			(0.035)
2006			0.159***			-0.137***
			(0.039)			(0.033)
2009			0.099***			-0.093***
			(0.029)			(0.026)
2011			-0.066**			-0.058*
			(0.029)			(0.030)
Constant	0.390***	0.218***	0.012	0.519***	-0.003	0.274***
	(0.011)	(0.033)	(0.041)	(0.012)	(0.035)	(0.047)
Observations	29,438	22,345	22,345	28,903	24,612	24,612
R-squared	0.047	0.135	0.138	0.056	0.113	0.112
Number of groups			72			72

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.2 OLS and fixed-effects regressions with waist-hip ratio.

Waist-hip ratio	Female			Male		
	(1)	(2)	(3)	(1)	(2)	(3)
County variables						
Gini (z-score)	0.023***	0.009	0.002	0.024***	0.024***	0.007
	(0.005)	(0.006)	(0.008)	(0.008)	(0.007)	(0.007)
Mean county income	0.039***	0.027***	0.002	0.092***	0.066***	-0.003
(z-score)	(0.005)	(0.006)	(0.017)	(0.007)	(0.008)	(0.016)
Individual variables						
Income (z-score)	0.021***	-0.002	-0.003	0.004	0.001	0.002
	(0.005)	(0.006)	(0.005)	(0.004)	(0.003)	(0.003)
Age		0.007***	0.008***		0.005***	0.004***

	(0.000)	(0.000)	(0.001)	(0.000)
Married	0.050***	0.049***	0.046***	0.033**
	(0.017)	(0.016)	(0.015)	(0.015)
Majority	-0.023*	-0.018	0.049**	-0.012
	(0.013)	(0.016)	(0.021)	(0.018)
Urban	0.002	-0.017	0.006	0.019
	(0.015)	(0.016)	(0.016)	(0.016)
Years of education	0.005***	0.004***	0.004**	0.000
	(0.001)	(0.002)	(0.002)	(0.001)
Occupation (ref.=farmer)				
Service class	-0.021	-0.027	0.108***	0.081***
	(0.024)	(0.022)	(0.021)	(0.020)
Non-manual worker	-0.025	-0.027	0.086***	0.052**
	(0.026)	(0.025)	(0.023)	(0.021)
Skilled-workers/	-0.024	-0.004	0.086***	0.060***
supervisor	(0.030)	(0.024)	(0.020)	(0.018)
Semi-skilled/	-0.007	0.005	0.061***	0.037**
non-skilled worker	(0.022)	(0.021)	(0.018)	(0.017)
Others	-0.018	-0.003	0.074***	0.054**
	(0.026)	(0.025)	(0.024)	(0.022)
Sector (ref.=collective)				
State	-0.037**	-0.024*	-0.002	0.001
	(0.015)	(0.014)	(0.017)	(0.013)
Family farming	0.025	-0.034	0.025*	-0.056***
	(0.017)	(0.021)	(0.014)	(0.017)
Individual enterprise	0.030	-0.005	0.014	-0.030*
	(0.019)	(0.015)	(0.018)	(0.015)
Private/three-capital	-0.053*	-0.056**	-0.006	-0.014
enterprises	(0.029)	(0.027)	(0.052)	(0.049)
Others	0.047*	0.021	0.046*	0.006
	(0.025)	(0.023)	(0.024)	(0.026)
Wave				
1993		-0.004		0.036*
		(0.024)		(0.018)
1997		0.027		0.078***
		(0.024)		(0.016)
2000		0.075**		0.145***
		(0.029)		(0.022)
2004		0.076**		0.158***
		(0.029)		(0.026)
2006		0.081*		0.188***
		(0.041)		(0.029)

2009			0.110***			0.233***
			(0.041)			(0.033)
2011			0.113*			0.289***
			(0.058)			(0.048)
Constant	0.740***	0.426***	0.373***	0.350***	-0.039	-0.013
	(0.008)	(0.032)	(0.035)	(0.010)	(0.032)	(0.031)
Observations	26,235	19,201	19,201	25,729	21,445	21,445
R-squared	0.011	0.074	0.068	0.043	0.071	0.060
Number of groups			72			72

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.3 OLS and fixed-effects regressions with MAMC.

MAMC	Female			Male		
	(1)	(2)	(3)	(1)	(2)	(3)
County variables						
Gini (z-score)	0.000	0.002	-0.001	0.000	-0.000	0.001
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Mean county income	0.012***	-0.008**	0.021**	0.028***	0.018***	0.017*
(z-score)	(0.003)	(0.003)	(0.008)	(0.004)	(0.003)	(0.010)
Individual variables						
Income (z-score)	0.000	-0.001	-0.001	-0.001	0.001	0.000
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Age		0.001**	0.001*		0.002***	0.002***
		(0.000)	(0.000)		(0.000)	(0.000)
Married		0.056***	0.051***		0.085***	0.078***
		(0.009)	(0.009)		(0.009)	(0.009)
Majority		-0.029*	-0.007		-0.066**	-0.027
		(0.016)	(0.005)		(0.027)	(0.022)
Urban		0.011	0.006		0.013	0.006
		(0.009)	(0.014)		(0.012)	(0.013)
Years of education		-0.000	0.000		0.003***	-0.002*
		(0.001)	(0.001)		(0.001)	(0.001)
Occupation (ref.=farmer)						
Service class		-0.001	-0.012		0.042***	0.038***
		(0.010)	(0.009)		(0.010)	(0.010)
Non-manual worker		-0.008	-0.012		0.042***	0.034***
		(0.008)	(0.009)		(0.012)	(0.012)
Skilled-workers/		0.001	0.000		0.041***	0.037***
supervisor		(0.011)	(0.012)		(0.009)	(0.010)
Semi-skilled/		-0.013*	-0.017**		0.036***	0.033***
non-skilled worker		(0.007)	(0.008)		(0.008)	(0.009)

Others		-0.009	-0.018		-0.026**	-0.023*
		(0.010)	(0.013)		(0.011)	(0.012)
Sector (ref.=collective)						
State		-0.001	0.002		-0.004	-0.001
		(0.007)	(0.007)		(0.011)	(0.011)
Family farming		-0.020**	-0.014*		0.053***	-0.018*
		(0.008)	(0.008)		(0.008)	(0.009)
Individual enterprise		-0.008	0.002		-0.025**	0.003
		(0.007)	(0.006)		(0.010)	(0.009)
Private/three-capital		-0.022*	-0.014		-0.002	0.019
enterprises		(0.012)	(0.012)		(0.017)	(0.015)
Others		-0.009	-0.005		0.045***	-0.024
		(0.014)	(0.013)		(0.017)	(0.017)
Wave						
1989			0.067***			0.100***
			(0.025)			(0.025)
1991			0.079***			0.119***
			(0.027)			(0.028)
1993			0.084***			0.137***
			(0.025)			(0.024)
1997			0.061**			0.112***
			(0.023)			(0.027)
2000			0.055***			0.085***
			(0.018)			(0.022)
2004			0.077***			0.073***
			(0.022)			(0.022)
2006			0.060***			0.058***
			(0.018)			(0.020)
2009			0.035**			0.044**
			(0.017)			(0.017)
2011			0.020*			0.016
			(0.011)			(0.012)
Constant	0.052***	0.107***	0.032	0.088***	0.205***	0.059**
	(0.005)	(0.022)	(0.025)	(0.007)	(0.038)	(0.028)
Observations	31,827	24,469	24,469	31,232	26,825	26,825
R-squared	0.003	0.010	0.009	0.010	0.035	0.026
Number of groups			72			72

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.4 OLS and fixed-effects regressions with BMI.

BMI		Female			Male	

	(1)	(2)	(3)	(1)	(2)	(3)
County variables						
Gini (z-score)	0.011**	0.004	0.006	0.011*	0.009*	0.003
	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)	(0.004)
Mean county income	0.073***	0.041***	-0.024*	0.101***	0.069***	-0.003
(z-score)	(0.007)	(0.007)	(0.013)	(0.007)	(0.008)	(0.012)
Individual variables						
Income (z-score)	0.029***	0.015***	0.011***	0.005	0.000	0.001
	(0.005)	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)
Age		0.004***	0.004***		0.002***	0.002***
		(0.001)	(0.001)		(0.000)	(0.000)
Married		0.065***	0.059***		0.066***	0.054***
		(0.010)	(0.010)		(0.010)	(0.010)
Majority		0.046	0.004		0.045*	-0.004
		(0.029)	(0.016)		(0.025)	(0.026)
Urban		-0.008	0.011		0.027*	0.052***
		(0.013)	(0.013)		(0.016)	(0.015)
Years of education		-0.003	0.005***		0.006***	0.002**
		(0.002)	(0.002)		(0.001)	(0.001)
Occupation (ref.=farmer)						
Service class		0.070**	0.044		0.164***	0.136***
		(0.029)	(0.027)		(0.021)	(0.016)
Non-manual worker		0.064***	0.045**		0.122***	0.094***
		(0.023)	(0.021)		(0.020)	(0.016)
Skilled-workers/		0.076***	0.070***		0.097***	0.069***
supervisor		(0.021)	(0.017)		(0.018)	(0.012)
Semi-skilled/		0.106***	0.099***		0.084***	0.059***
non-skilled worker		(0.019)	(0.017)		(0.019)	(0.012)
Others		0.099***	0.110***		0.063***	0.061***
		(0.020)	(0.017)		(0.020)	(0.016)
Sector (ref.=collective)						
State		-0.006	-0.008		0.000	-0.007
		(0.016)	(0.015)		(0.016)	(0.012)
Family farming		0.081***	0.064***		0.046***	0.036***
		(0.012)	(0.016)		(0.011)	(0.013)
Individual enterprise		-0.003	-0.006		0.032**	-0.016
		(0.015)	(0.018)		(0.015)	(0.015)
Private/three-capital		0.123***	0.102***		-0.028	-0.027
enterprises		(0.027)	(0.032)		(0.044)	(0.042)
Others		0.046	0.037		-0.000	-0.044*
		(0.029)	(0.028)		(0.026)	(0.025)
Wave						

1989			0.163***			0.235***
			(0.027)			(0.022)
1991			0.138***			0.214***
			(0.028)			(0.024)
1993			0.122***			0.200***
			(0.026)			(0.024)
1997			0.088***			0.165***
			(0.025)			(0.021)
2000			-0.054**			0.126***
			(0.022)			(0.021)
2004			0.090***			0.080***
			(0.018)			(0.017)
2006			0.089***			0.069***
			(0.015)			(0.015)
2009			0.051***			0.039***
			(0.012)			(0.009)
2011			0.058***			0.067***
			(0.019)			(0.018)
Constant	0.241***	-0.084**	0.056	0.237***	0.098***	0.146***
	(0.011)	(0.034)	(0.038)	(0.012)	(0.023)	(0.033)
Observations	32,340	24,961	24,961	31,913	27,483	27,483
R-squared	0.027	0.054	0.056	0.062	0.107	0.086
Number of groups			72			72

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.5 OLS and fixed-effects regressions with waist circumference.

Waist circumference	Female			Male		
	(1)	(2)	(3)	(1)	(2)	(3)
County variables						
Gini (z-score)	-0.001	0.013	0.013*	-0.016	0.007	0.006
	(0.008)	(0.012)	(0.007)	(0.012)	(0.012)	(0.007)
Mean county income	0.016***	0.023***	-0.001	0.054***	0.030***	0.020
(z-score)	(0.006)	(0.009)	(0.018)	(0.010)	(0.009)	(0.019)
Individual variables						
Income (z-score)	0.023***	-0.006	-0.005	0.006	0.003	0.002
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)
Age		0.001	0.005***		0.004***	-0.000
		(0.001)	(0.000)		(0.001)	(0.000)
Married		-0.031*	0.051***		0.006	-0.028**
		(0.016)	(0.014)		(0.015)	(0.011)
Majority		0.054*	-0.027*		0.077***	0.001

	(0.029)	(0.014)	(0.027)	(0.016)
Urban	0.032**	-0.016	0.050***	0.007
	(0.015)	(0.010)	(0.014)	(0.012)
Years of education	0.016***	-0.002	0.013***	0.003***
	(0.002)	(0.001)	(0.001)	(0.001)
Occupation (ref.=farmer)				
Service class	0.034	0.072***	0.070***	0.144***
	(0.028)	(0.023)	(0.023)	(0.017)
Non-manual worker	0.013	0.069***	0.026	0.091***
	(0.021)	(0.019)	(0.024)	(0.018)
Skilled-workers/	0.074***	0.107***	0.022	0.107***
supervisor	(0.022)	(0.017)	(0.020)	(0.015)
Semi-skilled/	0.070***	0.141***	0.009	0.104***
non-skilled worker	(0.016)	(0.016)	(0.018)	(0.014)
Others	0.138***	0.067***	0.075***	0.089***
	(0.022)	(0.019)	(0.027)	(0.014)
Sector (ref.=collective)				
State	0.033*	0.060***	0.065***	0.038***
	(0.017)	(0.014)	(0.015)	(0.012)
Family farming	-0.046*	0.051***	0.165***	-0.039**
	(0.024)	(0.016)	(0.027)	(0.015)
Individual enterprise	0.099***	0.049***	0.120***	0.074***
	(0.022)	(0.014)	(0.022)	(0.015)
Private/three-capital	0.157***	-0.104**	-0.088**	-0.077**
enterprises	(0.036)	(0.040)	(0.042)	(0.038)
Others	-0.059*	0.041	0.118***	-0.023
	(0.032)	(0.029)	(0.030)	(0.026)
Wave				
1989		0.537***		0.643***
		(0.037)		(0.041)
1991		0.547***		0.641***
		(0.037)		(0.041)
1993		-0.101**		-0.078*
		(0.040)		(0.042)
1997		-0.054		-0.051
		(0.038)		(0.042)
2000		-0.011		-0.003
		(0.029)		(0.029)
2004		0.076***		-0.062**
		(0.024)		(0.028)
2006		0.078***		-0.052**
		(0.025)		(0.025)

2009			-0.029*			-0.032*
			(0.017)			(0.017)
2011			0.069**			0.060**
			(0.030)			(0.026)
Constant	0.604***	0.628***	0.336***	0.491***	0.684***	0.340***
	(0.011)	(0.040)	(0.037)	(0.013)	(0.039)	(0.034)
Observations	36,605	28,759	28,759	38,365	33,549	33,549
R-squared	0.004	0.038	0.257	0.013	0.064	0.330
Number of groups			72			72

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix C. Tables for Analysis of Blood-based Biomarkers (Chapter 5.3)

Table C.1 OLS and OLS with individual controls regressions with total cholesterol.

Total cholesterol	Female (1)	Female (2)	Male (1)	Male (2)
County variables				
Gini (z-score)	-0.019	-0.012	-0.015	-0.015
	(0.029)	(0.036)	(0.028)	(0.030)
Mean county income	-0.028	-0.097*	-0.013	-0.039
(z-score)	(0.044)	(0.052)	(0.044)	(0.048)
Individual variables				
Income (z-score)	0.020**	0.014	0.026**	0.011
	(0.010)	(0.012)	(0.011)	(0.011)
Age		0.006***		-0.001
		(0.001)		(0.001)
Married		-0.075		0.033
		(0.046)		(0.039)
Majority		-0.037		-0.049
		(0.050)		(0.036)
Urban		0.027		0.050*
		(0.033)		(0.027)
Years of education		0.007*		0.001
		(0.004)		(0.004)
Occupation (ref.=farmer)				
Service class		-0.057		0.115*
		(0.082)		(0.065)
Non-manual worker		0.004		0.069
		(0.080)		(0.066)
Skilled-workers/		-0.026		0.079
supervisor		(0.078)		(0.057)
Semi-skilled/		-0.038		0.059
non-skilled worker		(0.076)		(0.056)
Others		0.002		0.059
		(0.083)		(0.061)
Sector (ref.=collective)				
State		0.054		-0.039

		(0.058)		(0.076)
Family farming		-0.019		-0.044
		(0.088)		(0.084)
Individual enterprise		0.008		-0.053
		(0.048)		(0.064)
Private/three-capital enterprises		-0.053		0.025
		(0.082)		(0.104)
Others		0.105*		-0.055
		(0.061)		(0.076)
Constant	0.309***	0.103	0.310***	0.373***
	(0.034)	(0.126)	(0.038)	(0.106)
Observations	3,285	2,364	3,383	2,851
R-squared	0.047	0.135	0.138	0.056

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C.2 OLS and OLS with individual controls regressions with C-reactive protein.

C-reactive protein	Female		Male	
	(1)	(2)	(1)	(2)
County variables				
Gini (z-score)	0.005	0.009	-0.004	-0.011
	(0.009)	(0.009)	(0.010)	(0.011)
Mean county income (z-score)	-0.007	0.001	0.009	0.007
	(0.015)	(0.015)	(0.019)	(0.020)
Individual variables				
Income (z-score)	-0.005	-0.004	-0.000	0.002
	(0.005)	(0.005)	(0.007)	(0.008)
Age		0.002***		0.001
		(0.001)		(0.001)
Married		-0.005		-0.016
		(0.021)		(0.022)
Majority		-0.047***		0.016
		(0.016)		(0.018)
Urban		0.002		0.002
		(0.017)		(0.018)
Years of education		0.001		-0.002
		(0.002)		(0.002)
Occupation (ref.=farmer)				
Service class		-0.054		-0.036
		(0.051)		(0.048)

Non-manual worker		-0.093**		-0.061
		(0.043)		(0.045)
Skilled-workers/		-0.069		-0.040
supervisor		(0.051)		(0.037)
Semi-skilled/		-0.087*		-0.021
non-skilled worker		(0.046)		(0.040)
Others		-0.069		-0.002
		(0.044)		(0.051)
Sector (ref.=collective)				
State		0.018		-0.014
		(0.019)		(0.032)
Family farming		-0.053		-0.052
		(0.045)		(0.052)
Individual enterprise		0.043**		-0.032
		(0.020)		(0.032)
Private/three-capital		-0.019		-0.042
enterprises		(0.018)		(0.061)
Others		0.053		-0.046
		(0.050)		(0.046)
Constant	0.077***	0.043	0.083***	0.112*
	(0.012)	(0.065)	(0.016)	(0.058)
Observations	2,889	2,056	2,871	2,390
R-squared	0.00	0.021	0.000	0.006

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C.3 OLS and OLS with individual controls regressions with HbA1c.

HbA1c	Female		Male	
	(1)	(2)	(1)	(2)
County variables				
Gini (z-score)	0.002	0.003	-0.003	-0.005
	(0.010)	(0.008)	(0.010)	(0.012)
Mean county income	0.020	0.009	0.033	-0.007
(z-score)	(0.013)	(0.013)	(0.025)	(0.023)
Individual variables				
Income (z-score)	0.011	0.015	0.004	0.002
	(0.013)	(0.015)	(0.005)	(0.006)
Age		0.002***		0.002***
		(0.001)		(0.001)
Married		0.005		0.000
		(0.009)		(0.019)
Majority		0.016		0.042**

		(0.010)		(0.018)
Urban		0.008		0.012
		(0.014)		(0.016)
Years of education		-0.000		0.002
		(0.002)		(0.002)
Occupation (ref.=farmer)				
Service class		0.015		0.005
		(0.050)		(0.046)
Non-manual worker		0.005		0.024
		(0.045)		(0.053)
Skilled-workers/		0.005		0.014
supervisor		(0.046)		(0.044)
Semi-skilled/		0.019		-0.001
non-skilled worker		(0.043)		(0.044)
Others		-0.006		0.022
		(0.042)		(0.052)
Sector (ref.=collective)				
State		-0.026		-0.049
		(0.027)		(0.036)
Family farming		0.008		-0.052
		(0.056)		(0.046)
Individual enterprise		-0.006		-0.047
		(0.024)		(0.036)
Private/three-capital		-0.049*		-0.109***
enterprises		(0.025)		(0.034)
Others		0.012		-0.087**
		(0.040)		(0.034)
Constant	0.048***	-0.068	0.068***	-0.051
	(0.010)	(0.052)	(0.013)	(0.064)
Observations	2889	2056	2871	2390
R-squared	0.002	0.019	0.002	0.018

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix D. Tables for Analysis of Nurse-collected Health Measures with Lagged Inequality (Chapter 5.4)

Table D.1 OLS and fixed-effects regressions with blood pressure with lagged inequality.

Blood pressure	Female (1)	Female (2)	Female (3)	Male (1)	Male (2)	Male (3)
County variables						
Gini (z-score)	0.022**	0.008	0.023	0.003	0.004	0.007
	(0.011)	(0.011)	(0.017)	(0.011)	(0.012)	(0.016)
Gini lag 1 (z-score)	0.010	-0.009	0.011	-0.007	-0.027**	-0.018
	(0.013)	(0.014)	(0.017)	(0.012)	(0.012)	(0.016)
Gini lag 2 (z-score)	0.026*	0.010	0.031*	0.030**	0.019	0.017
	(0.013)	(0.017)	(0.017)	(0.015)	(0.018)	(0.019)
Gini lag 3 (z-score)	-0.010	0.012	0.003	-0.023	-0.012	0.009
	(0.017)	(0.021)	(0.020)	(0.016)	(0.019)	(0.018)
Mean county income	0.101***	0.052***	-0.056*	0.125***	0.079***	0.030
(z-score)	(0.016)	(0.016)	(0.029)	(0.015)	(0.016)	(0.025)
Individual variables						
Income (z-score)	-0.035**	-0.007	0.002	-0.007	-0.006	-0.007
	(0.015)	(0.011)	(0.010)	(0.008)	(0.008)	(0.008)
Age		0.014***	0.014***		0.010***	0.010***
		(0.001)	(0.001)		(0.001)	(0.001)
Married		0.090***	0.068**		-0.054	-0.058*
		(0.032)	(0.030)		(0.036)	(0.033)
Majority		0.041	-0.015		0.002	-0.008
		(0.044)	(0.032)		(0.035)	(0.036)
Urban		-0.063**	-0.030		0.005	0.034
		(0.031)	(0.036)		(0.022)	(0.024)
Years of education		0.002	0.001		-0.001	-0.001
		(0.003)	(0.003)		(0.003)	(0.003)
Occupation (ref.=farmer)						
Service class		-0.057	-0.048		0.113**	0.086**
		(0.055)	(0.056)		(0.046)	(0.037)
Non-manual worker		0.015	0.014		0.158***	0.102**
		(0.039)	(0.038)		(0.054)	(0.044)
Skilled-workers/		0.029	0.010		0.082*	0.035
supervisor		(0.058)	(0.052)		(0.044)	(0.035)
Semi-skilled/		0.057*	0.041		0.110***	0.072**
non-skilled worker		(0.032)	(0.034)		(0.038)	(0.033)
Others		0.067	0.029		0.120***	0.099**
		(0.046)	(0.051)		(0.045)	(0.045)
Sector (ref.=collective)						

State		-0.017	-0.001		0.001	0.017
		(0.030)	(0.034)		(0.028)	(0.025)
Family farming		0.062**	0.099***		0.061*	0.035
		(0.029)	(0.035)		(0.036)	(0.030)
Individual enterprise		-0.014	0.023		0.004	0.004
		(0.032)	(0.030)		(0.027)	(0.027)
Private/three-capital		-0.014	-0.038		0.072	0.057
enterprises		(0.107)	(0.118)		(0.069)	(0.066)
Others		-0.002	0.037		-0.031	-0.038
		(0.068)	(0.076)		(0.049)	(0.049)
Wave						
1997			-0.139			-0.132*
			(0.093)			(0.067)
2000			-0.154*			-0.083
			(0.088)			(0.061)
2006			0.218***			-0.094**
			(0.074)			(0.045)
2009			0.166***			-0.046
			(0.055)			(0.044)
2011			-0.097*			-0.055
			(0.053)			(0.043)
Constant	0.435***	0.389***	-0.217**	0.554***	0.048	0.168**
	(0.016)	(0.065)	(0.090)	(0.014)	(0.082)	(0.084)
Observations	6,490	4,998	4,998	6,744	5,764	5,764
R-squared	0.041	0.107	0.108	0.047	0.086	0.080
Number of groups			54			54

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.2 OLS and fixed-effects regressions with waist-hip ratio with lagged inequality.

Waist-hip ratio	Female			Male		
	(1)	(2)	(3)	(1)	(2)	(3)
County variables						
Gini (z-score)	0.014**	0.004	-0.006	0.020*	0.026**	-0.008
	(0.007)	(0.007)	(0.010)	(0.010)	(0.011)	(0.013)
Gini lag 1 (z-score)	0.017**	0.007	0.000	0.012	-0.001	-0.013
	(0.008)	(0.011)	(0.012)	(0.011)	(0.010)	(0.011)
Gini lag 2 (z-score)	0.028**	0.018	0.015	0.024**	0.026**	0.019
	(0.011)	(0.013)	(0.014)	(0.010)	(0.011)	(0.013)
Gini lag 3 (z-score)	-0.014	-0.011	-0.007	-0.012	0.002	-0.005
	(0.010)	(0.011)	(0.013)	(0.012)	(0.012)	(0.013)
Mean county income	0.031***	0.011	-0.015	0.085***	0.049***	-0.008
(z-score)	(0.010)	(0.015)	(0.030)	(0.017)	(0.017)	(0.035)
Individual variables						

Income (z-score)	-0.009	0.010	0.011	0.024***	0.014**	0.012*
	(0.009)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
Age		0.007***	0.007***		0.005***	0.005***
		(0.001)	(0.001)		(0.001)	(0.001)
Married		0.033	0.040		0.029	0.011
		(0.043)	(0.042)		(0.043)	(0.040)
Majority		-0.046**	-0.029		0.085***	-0.038
		(0.019)	(0.027)		(0.028)	(0.029)
Urban		0.031	0.017		0.012	0.022
		(0.027)	(0.027)		(0.024)	(0.026)
Years of education		-0.004*	-0.004		0.007***	0.004*
		(0.003)	(0.003)		(0.002)	(0.002)
Occupation (ref.=farmer)						
Service class		-0.057	-0.070		0.092**	0.084**
		(0.048)	(0.048)		(0.039)	(0.037)
Non-manual worker		-0.052	-0.064		0.083**	0.056
		(0.042)	(0.043)		(0.036)	(0.037)
Skilled-workers/ supervisor		-0.053	-0.047		0.110***	0.082**
		(0.036)	(0.034)		(0.035)	(0.034)
Semi-skilled/ non-skilled worker		0.012	0.022		0.071**	0.046
		(0.024)	(0.023)		(0.029)	(0.030)
Others		0.011	0.030		0.087**	0.071*
		(0.038)	(0.042)		(0.041)	(0.041)
Sector (ref.=collective)						
State		-0.011	-0.004		-0.018	-0.027
		(0.024)	(0.025)		(0.026)	(0.028)
Family farming		0.026	-0.011		0.004	-0.044*
		(0.026)	(0.037)		(0.020)	(0.026)
Individual enterprise		0.029	0.007		-0.002	-0.021
		(0.033)	(0.037)		(0.032)	(0.033)
Private/three-capital enterprises		0.045	0.062		-0.085	-0.103
		(0.070)	(0.059)		(0.156)	(0.132)
Others		0.095**	0.070		0.095*	0.057
		(0.037)	(0.044)		(0.047)	(0.051)
Wave						
1997			-0.132**			-0.296***
			(0.064)			(0.076)
2000			-0.090			-0.268***
			(0.060)			(0.077)
2006			-0.035			-0.161**
			(0.046)			(0.067)
2009			-0.058*			-0.114**
			(0.031)			(0.051)
2011			-0.038			-0.106***
			(0.030)			(0.035)

Constant	0.790***	0.483***	0.537***	0.370***	-0.092	0.291***
	(0.009)	(0.058)	(0.084)	(0.016)	(0.067)	(0.083)
Observations	6,553	5,032	5,032	6,819	5,834	5,834
R-squared	0.014	0.052	0.048	0.041	0.066	0.060
Number of groups			54			54

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.3 OLS and fixed-effects regressions with MAMC with lagged inequality.

MAMC	Female			Male		
	(1)	(2)	(3)	(1)	(2)	(3)
County variables						
Gini (z-score)	0.003	0.001	-0.007	0.003	0.004	-0.002
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.006)
Gini lag 1 (z-score)	-0.008*	-0.008	-0.007	0.002	0.004	0.013*
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
Gini lag 2 (z-score)	-0.001	-0.001	0.003	-0.002	-0.008	-0.002
	(0.005)	(0.006)	(0.007)	(0.008)	(0.008)	(0.008)
Gini lag 3 (z-score)	0.018***	0.015**	0.010	0.005	0.000	-0.010
	(0.006)	(0.006)	(0.008)	(0.007)	(0.006)	(0.009)
Mean county income	0.019***	0.016***	0.007	0.028***	0.027***	0.002
(z-score)	(0.006)	(0.006)	(0.011)	(0.008)	(0.008)	(0.016)
Individual variables						
Income (z-score)	0.001	-0.004	-0.002	-0.003	0.005	0.005
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Age		0.002***	0.002***		0.003***	0.003***
		(0.000)	(0.001)		(0.001)	(0.001)
Married		-0.034	-0.037		-0.019	-0.009
		(0.027)	(0.027)		(0.019)	(0.020)
Majority		-0.025*	0.004		-0.067**	-0.022
		(0.015)	(0.013)		(0.025)	(0.036)
Urban		0.007	0.002		0.002	-0.001
		(0.014)	(0.016)		(0.017)	(0.018)
Years of education		-0.000	-0.000		-0.004**	-0.002
		(0.001)	(0.001)		(0.002)	(0.002)
Occupation (ref.=farmer)						
Service class		0.021	-0.002		0.047***	-0.044**
		(0.022)	(0.019)		(0.017)	(0.019)
Non-manual worker		0.005	-0.010		-0.051**	-0.041
		(0.017)	(0.015)		(0.023)	(0.025)
Skilled-workers/		-0.023*	-0.030**		-0.033*	-0.028
supervisor		(0.014)	(0.015)		(0.018)	(0.020)
Semi-skilled/		-0.019*	-0.023*		-0.034**	-0.028*

non-skilled worker		(0.011)	(0.012)		(0.015)	(0.015)
Others		-0.017	-0.034		-0.015	-0.007
		(0.018)	(0.021)		(0.022)	(0.022)
Sector (ref.=collective)						
State		0.003	0.009		0.021	0.022
		(0.013)	(0.013)		(0.017)	(0.016)
Family farming		0.023***	0.039***		-0.042**	-0.008
		(0.008)	(0.014)		(0.017)	(0.022)
Individual enterprise		0.001	-0.001		-0.024*	0.002
		(0.012)	(0.014)		(0.012)	(0.014)
Private/three-capital		0.055	0.057		0.051	0.082
enterprises		(0.041)	(0.042)		(0.072)	(0.058)
Others		-0.018	-0.026		0.069***	0.059***
		(0.024)	(0.025)		(0.016)	(0.016)
Wave						
1997			0.039			0.096**
			(0.033)			(0.038)
2000			0.033			0.070*
			(0.027)			(0.039)
2006			0.064***			0.031
			(0.023)			(0.029)
2009			0.029			0.015
			(0.019)			(0.022)
2011			0.006			0.000
			(0.014)			(0.020)
Constant	0.049***	0.020	-0.023	0.087***	0.070*	-0.065
	(0.006)	(0.040)	(0.044)	(0.009)	(0.039)	(0.055)
Observations	6,612	5,077	5,077	6,871	5,878	5,878
R-squared	0.008	0.021	0.017	0.007	0.041	0.032
Number of groups			54			54

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.4 OLS and fixed-effects regressions with BMI with lagged inequality.

BMI	Female			Male		
	(1)	(2)	(3)	(1)	(2)	(3)
County variables						
Gini (z-score)	0.005	0.002	0.008	0.013	0.018*	-0.002
	(0.008)	(0.010)	(0.012)	(0.010)	(0.010)	(0.010)
Gini lag 1 (z-score)	0.012	0.012	0.007	0.013	0.008	0.013
	(0.008)	(0.011)	(0.012)	(0.009)	(0.009)	(0.010)
Gini lag 2 (z-score)	-0.012	-0.010	-0.008	-0.009	-0.007	0.002
	(0.010)	(0.011)	(0.009)	(0.009)	(0.010)	(0.010)
Gini lag 3 (z-score)	-0.021**	-0.016	0.003	-0.005	0.001	0.008
	(0.009)	(0.011)	(0.013)	(0.011)	(0.010)	(0.012)

Mean county income	0.087***	0.062***	-0.051*	0.085***	0.054***	0.000
(z-score)	(0.015)	(0.020)	(0.029)	(0.015)	(0.017)	(0.024)
Individual variables						
Income (z-score)	-0.022**	-0.013	-0.005	0.039***	0.018***	0.017**
	(0.009)	(0.009)	(0.010)	(0.007)	(0.006)	(0.007)
Age		0.001	0.002		0.002**	0.002**
		(0.001)	(0.001)		(0.001)	(0.001)
Married		0.043	0.023		0.041	0.019
		(0.043)	(0.044)		(0.031)	(0.027)
Majority		0.046	-0.032		0.055*	-0.036
		(0.044)	(0.044)		(0.029)	(0.041)
Urban		0.019	0.043		0.029	0.042*
		(0.031)	(0.034)		(0.026)	(0.024)
Years of education		-0.007**	-0.005**		0.010***	0.008***
		(0.003)	(0.002)		(0.003)	(0.002)
Occupation (ref.=farmer)						
Service class		0.115*	0.075		0.145***	0.146***
		(0.061)	(0.059)		(0.035)	(0.034)
Non-manual worker		0.103*	0.086		0.137***	0.119***
		(0.052)	(0.052)		(0.039)	(0.037)
Skilled-workers/		0.121**	0.106**		0.126***	0.112***
supervisor		(0.052)	(0.049)		(0.026)	(0.025)
Semi-skilled/		0.184***	0.163***		0.075**	0.064**
non-skilled worker		(0.039)	(0.039)		(0.031)	(0.026)
Others		0.224***	0.228***		0.036	0.046
		(0.048)	(0.047)		(0.041)	(0.039)
Sector (ref.=collective)						
State		-0.056	-0.064		0.024	0.009
		(0.037)	(0.041)		(0.023)	(0.023)
Family farming		0.070***	0.077**		0.042*	0.011
		(0.025)	(0.035)		(0.021)	(0.022)
Individual enterprise		-0.035	-0.009		0.033	0.021
		(0.033)	(0.036)		(0.030)	(0.030)
Private/three-capital		-0.141*	-0.118		0.020	0.037
enterprises		(0.082)	(0.094)		(0.148)	(0.143)
Others		0.051	0.082		-0.044	-0.056
		(0.064)	(0.056)		(0.044)	(0.038)
Wave						
1997			-0.186***			-0.158***
			(0.061)			(0.051)
2000			-0.142**			-0.127**
			(0.059)			(0.050)
2006			-0.194***			-0.098*
			(0.046)			(0.050)
2009			-0.086**			-0.078*

			(0.033)			(0.040)
2011			-0.041			-0.063*
			(0.026)			(0.034)
Constant	0.257***	0.113	0.254***	0.243***	-0.099*	0.129**
	(0.015)	(0.083)	(0.093)	(0.016)	(0.051)	(0.064)
Observations	6,659	5,119	5,119	6,935	5,940	5,940
R-squared	0.022	0.040	0.039	0.043	0.094	0.070
Number of groups			54			54

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.5 OLS and fixed-effects regressions with waist circumference with lagged inequality.

Waist circumference	Female			Male		
	(1)	(2)	(3)	(1)	(2)	(3)
County variables						
Gini (z-score)	0.016	0.013	0.007	0.021	0.032**	0.007
	(0.012)	(0.014)	(0.019)	(0.015)	(0.014)	(0.011)
Gini lag 1 (z-score)	0.017	0.011	0.007	-0.016*	-0.001	0.013
	(0.011)	(0.013)	(0.014)	(0.009)	(0.010)	(0.013)
Gini lag 2 (z-score)	0.006	0.008	0.008	-0.004	0.009	0.020
	(0.016)	(0.017)	(0.015)	(0.012)	(0.013)	(0.014)
Gini lag 3 (z-score)	-0.020	-0.010	0.027	0.003	0.000	0.020
	(0.017)	(0.018)	(0.017)	(0.015)	(0.014)	(0.013)
Mean county income	0.083***	0.050**	-0.042	0.029	0.018	0.047*
(z-score)	(0.018)	(0.024)	(0.042)	(0.018)	(0.016)	(0.027)
Individual variables						
Income (z-score)	-0.013	-0.004	0.003	0.033***	0.014**	0.014**
	(0.011)	(0.011)	(0.010)	(0.006)	(0.006)	(0.007)
Age		0.004***	0.006***		-0.001	-0.001
		(0.001)	(0.001)		(0.001)	(0.001)
Married		-0.039	-0.055		-0.038	-0.047
		(0.049)	(0.048)		(0.039)	(0.037)
Majority		0.084*	-0.038		0.126***	-0.040
		(0.043)	(0.052)		(0.034)	(0.048)
Urban		0.073**	0.067**		0.021	0.023
		(0.034)	(0.028)		(0.028)	(0.028)
Years of education		-0.007**	-0.004*		0.008***	0.008***
		(0.004)	(0.003)		(0.003)	(0.002)
Occupation (ref.=farmer)						
Service class		0.114*	0.093		0.140***	0.173***
		(0.058)	(0.056)		(0.048)	(0.039)
Non-manual worker		0.121**	0.138***		0.091*	0.115**
		(0.046)	(0.044)		(0.051)	(0.043)

Skilled-workers/		0.063	0.082*		0.121***	0.143***
supervisor		(0.052)	(0.048)		(0.031)	(0.026)
Semi-skilled/		0.161***	0.167***		0.105***	0.134***
non-skilled worker		(0.031)	(0.031)		(0.036)	(0.029)
Others		0.181***	0.203***		0.062	0.085*
		(0.037)	(0.036)		(0.051)	(0.044)
Sector (ref.=collective)						
State		0.102***	0.119***		-0.035	-0.023
		(0.034)	(0.034)		(0.030)	(0.027)
Family farming		0.053	0.067**		-0.058**	0.003
		(0.036)	(0.028)		(0.026)	(0.029)
Individual enterprise		-0.050	-0.029		-0.069**	-0.045
		(0.042)	(0.037)		(0.033)	(0.036)
Private/three-capital		-0.026	0.004		-0.224*	-0.140
enterprises		(0.089)	(0.093)		(0.120)	(0.127)
Others		0.009	0.021		-0.077*	-0.042
		(0.065)	(0.058)		(0.044)	(0.046)
Wave						
1997			-0.176*			0.043
			(0.094)			(0.092)
2000			-0.099			0.077
			(0.082)			(0.079)
2006			-0.158**			0.001
			(0.077)			(0.070)
2009			-0.082			-0.018
			(0.061)			(0.047)
2011			-0.070			-0.046
			(0.055)			(0.040)
Constant	0.499***	0.229***	0.365***	0.333***	0.227***	0.331***
	(0.020)	(0.083)	(0.101)	(0.020)	(0.061)	(0.083)
Observations	7,103	5,481	5,481	7,842	6,778	6,778
R-squared	0.024	0.040	0.045	0.012	0.054	0.043
Number of groups			54			54

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix E. Tables for Analysis of Blood-based Biomarkers with Lagged Inequality (Chapter 5.5)

Table E.1 OLS and OLS with individual controls regressions with total cholesterol with lagged inequality.

Total cholesterol	Female		Male	
	(1)	(2)	(1)	(2)
County variables				
Gini (z-score)	-0.007	-0.015	-0.012	-0.017
	(0.050)	(0.055)	(0.027)	(0.033)
Gini lag 1 (z-score)	-0.108**	-0.135**	-0.078**	-0.067*
	(0.047)	(0.063)	(0.035)	(0.038)
Gini lag 2 (z-score)	0.074*	0.082*	0.090***	0.097***
	(0.038)	(0.048)	(0.029)	(0.034)
Gini lag 3 (z-score)	0.018	0.016	0.065***	0.065***
	(0.026)	(0.035)	(0.022)	(0.022)
Mean county income	-0.091	-0.178*	-0.021	-0.053
(z-score)	(0.064)	(0.092)	(0.046)	(0.075)
Individual variables				
Income (z-score)	0.021	0.029	0.035	0.009
	(0.025)	(0.027)	(0.021)	(0.019)
Age		0.011***		-0.000
		(0.003)		(0.002)
Married		-0.268		-0.021
		(0.566)		(0.114)
Majority		0.011		0.026
		(0.078)		(0.042)
Urban		0.174*		0.048
		(0.093)		(0.042)
Years of education		0.008		0.010*
		(0.007)		(0.006)
Occupation (ref.=farmer)				
Service class		-0.292**		-0.027
		(0.134)		(0.134)
Non-manual worker		-0.265**		-0.102
		(0.117)		(0.140)
Skilled-workers/		-0.002		-0.015
supervisor		(0.146)		(0.145)
Semi-skilled/		-0.088		-0.082
non-skilled worker		(0.103)		(0.126)
Others		-0.045		-0.163
		(0.146)		(0.135)
Sector (ref.=collective)				

State		0.115		-0.061
		(0.121)		(0.099)
Family farming		-0.027		-0.091
		(0.151)		(0.151)
Individual enterprise		0.008		-0.029
		(0.125)		(0.088)
Private/three-capital		0.047		-0.096
enterprises		(0.095)		(0.216)
Others		0.189		0.050
		(0.185)		(0.151)
Constant	0.329***	0.018	0.311***	0.337*
	(0.043)	(0.590)	(0.033)	(0.201)
Observations	958	714	1,037	883
R-squared	0.012	0.081	0.024	0.036

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table E.2 OLS and OLS with individual controls regressions with C-reactive protein with lagged inequality.

C-reactive protein	Female		Male	
	(1)	(2)	(1)	(2)
County variables				
Gini (z-score)	-0.016	-0.017	-0.019	-0.018
	(0.012)	(0.012)	(0.019)	(0.019)
Gini lag 1 (z-score)	0.023	0.040**	-0.002	-0.003
	(0.018)	(0.016)	(0.024)	(0.023)
Gini lag 2 (z-score)	-0.001	-0.001	-0.013	-0.012
	(0.016)	(0.016)	(0.017)	(0.017)
Gini lag 3 (z-score)	-0.034**	-0.043***	-0.019	-0.016
	(0.013)	(0.013)	(0.016)	(0.015)
Mean county income	-0.022	-0.036	-0.017	-0.034
(z-score)	(0.026)	(0.030)	(0.036)	(0.045)
Individual variables				
Income (z-score)	0.007	0.015	-0.011	-0.015**
	(0.015)	(0.014)	(0.008)	(0.007)
Age		0.001		0.002**
		(0.002)		(0.001)
Married		-0.841***		-0.037
		(0.120)		(0.099)
Majority		-0.023		0.025
		(0.022)		(0.026)
Urban		-0.018		0.009
		(0.037)		(0.028)
Years of education		0.002		0.001
		(0.002)		(0.004)

Occupation (ref.=farmer)				
Service class		-0.082		-0.069
		(0.081)		(0.069)
Non-manual worker		-0.106		-0.092
		(0.071)		(0.063)
Skilled-workers/		-0.131*		-0.021
supervisor		(0.073)		(0.067)
Semi-skilled/		-0.101		-0.060
non-skilled worker		(0.075)		(0.062)
Others		-0.032		-0.085
		(0.087)		(0.069)
Sector (ref.=collective)				
State		0.016		-0.067
		(0.029)		(0.058)
Family farming		-0.049		-0.140*
		(0.069)		(0.082)
Individual enterprise		0.034		-0.078
		(0.024)		(0.049)
Private/three-capital		0.006		-0.106**
enterprises		(0.022)		(0.052)
Others		0.028		-0.091
		(0.137)		(0.070)
Constant	0.072***	0.889***	0.101***	0.130
	(0.017)	(0.165)	(0.024)	(0.143)
Observations	878	657	933	787
R-squared	0.007	0.051	0.007	0.022

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table E.3 OLS and OLS with individual controls regressions with HbA1c with lagged inequality.

HbA1c	Female		Male	
	(1)	(2)	(1)	(2)
County variables				
Gini (z-score)	-0.006	-0.019	-0.001	-0.004
	(0.011)	(0.011)	(0.016)	(0.016)
Gini lag 1 (z-score)	-0.038*	-0.064**	-0.020	-0.006
	(0.020)	(0.025)	(0.027)	(0.029)
Gini lag 2 (z-score)	0.056***	0.073***	0.043*	0.048*
	(0.015)	(0.019)	(0.022)	(0.024)
Gini lag 3 (z-score)	0.016	0.034**	0.023	0.028
	(0.011)	(0.015)	(0.018)	(0.018)
Mean county income	0.005	-0.017	0.050	-0.010
(z-score)	(0.021)	(0.024)	(0.031)	(0.034)

Individual variables				
Income (z-score)	-0.016*	-0.012	0.012	0.010
	(0.010)	(0.011)	(0.015)	(0.016)
Age		0.003**		0.003**
		(0.001)		(0.001)
Married		0.042		-0.061
		(0.083)		(0.085)
Majority		0.058***		0.051*
		(0.017)		(0.028)
Urban		0.056		0.026
		(0.046)		(0.033)
Years of education		0.001		0.003
		(0.003)		(0.003)
Occupation (ref.=farmer)				
Service class		0.022		-0.039
		(0.085)		(0.070)
Non-manual worker		0.031		0.033
		(0.079)		(0.079)
Skilled-workers/		0.015		0.010
supervisor		(0.073)		(0.061)
Semi-skilled/		0.058		0.002
non-skilled worker		(0.081)		(0.062)
Others		-0.000		-0.045
		(0.077)		(0.074)
Sector (ref.=collective)				
State		-0.010		-0.058
		(0.070)		(0.044)
Family farming		0.078		-0.081
		(0.097)		(0.068)
Individual enterprise		0.051		-0.042
		(0.075)		(0.053)
Private/three-capital		-0.003		-0.137***
enterprises		(0.085)		(0.050)
Others		0.061		-0.129***
		(0.107)		(0.043)
Constant	0.065***	-0.234*	0.063***	0.008
	(0.015)	(0.118)	(0.019)	(0.101)
Observations	878	657	933	787
R-squared	0.018	0.047	0.018	0.045

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

References

Avendano, M., & Hessel, P. (2015). The income inequality hypothesis rejected? *Official journal of the European Epidemiological Federation associated to the International Epidemiological Association, 30*(8), 595-598. doi:10.1007/s10654-015-0084-8

Backlund, E., Sorlie, P. D., & Johnson, N. J. (1996). The shape of the relationship between income and mortality in the United States: evidence from the National Longitudinal Mortality Study. *Annals of epidemiology, 6*(1), 12-20.

Baeten, S., Van Ourti, T., & van Doorslaer, E. (2013). Rising inequalities in income and health in China: Who is left behind? *Journal of Health Economics, 32*(6), 1214-1229. doi:10.1016/j.jhealeco.2013.10.002

Bakkeli, N. Z. (2016). Income inequality and health in China: A panel data analysis. *Social Science & Medicine, 157*, 39-47. doi:10.1016/j.socscimed.2016.03.041

Bartels, L. M. (2018). *Unequal democracy: The political economy of the new gilded age*: Princeton University Press.

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. doi:10.1136/jech.57.4.279

Blakely, T. A., Kennedy, B. P., Glass, R., & Kawachi, I. (2000). What is the lag time between income inequality and health status? *Journal of Epidemiology & Community Health, 54*(4), 318-319.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics, 29*(2), 238-249.

Chen, Z., & Gotway Crawford, C. A. (2012). The role of geographic scale in testing the income inequality hypothesis as an explanation of health disparities. *Social Science & Medicine, 75*(6), 1022-1031. doi:10.1016/j.socscimed.2012.04.032

Clarke, P. M., & Ryan, C. (2006). Self‐reported health: reliability and consequences for health inequality measurement. *Health Economics, 15*(6), 645-652. doi:10.1002/hec.1089

Clough-Gorr, K., Egger, M., & Spoerri, A. (2015). A Swiss paradox? Higher income inequality of municipalities is associated with lower mortality in Switzerland. *Official journal of the European Epidemiological Federation associated to the International Epidemiological Association, 30*(8), 627-636. doi:10.1007/s10654-015-9987-7

Craig, N. (2005). Exploring the generalisability of the association between income inequality and self-assessed health. *Social Science & Medicine, 60*(11), 2477-2488. doi:10.1016/j.socscimed.2004.11.018

Deaton, A., & Lubotsky, D. (2003). Mortality, inequality and race in American cities and states. *Social Science & Medicine, 56*(6), 1139-1153. doi:10.1016/S0277-9536(02)00115-6

Dorling, D., & Barford, A. (2009). The inequality hypothesis: Thesis, antithesis, and a synthesis? *Health and Place, 15*(4), 1166-1169. doi:10.1016/j.healthplace.2009.06.005

Feng, Z., Wang, W. W., Jones, K., & Li, Y. (2012). An exploratory multilevel analysis of income, income inequality and self-rated health of the elderly in China. *Social Science & Medicine, 75*(12), 2481-2492. doi:10.1016/j.socscimed.2012.09.028

Foverskov, E., & Holm, A. (2016). Socioeconomic inequality in health in the British household panel: tests of the social causation, health selection and the indirect selection hypothesis using dynamic fixed effects panel models. *Social Science & Medicine, 150*, 172-183.

Gilens, M. (2012). *Affluence and influence: Economic inequality and political power in America*: Princeton University Press.

Glied, S., & Lleras-Muney, A. (2008). Technological innovation and inequality in health. *Demography, 45*(3), 741-761.

Hu, Y., Lenthe, F., & Mackenbach, J. (2015). Income inequality, life expectancy and cause-specific mortality in 43 European countries, 1987–2008: a fixed effects study. *Official journal of the European Epidemiological Federation associated to the International Epidemiological Association, 30*(8), 615-625. doi:10.1007/s10654-015-0066-x

Ishii, S., Karlamangla, A. S., Bote, M., Irwin, M. R., Jacobs Jr, D. R., Cho, H. J., & Seeman, T. E. (2012). Gender, obesity and repeated elevation of C-reactive protein: data from the CARDIA cohort. *PloS one, 7*(4), e36062.

Jarvis, C. (2011). *Physical Examination and Health Assessment-E-Book*: Elsevier Health Sciences.

Kawachi, I., & Kennedy, B. P. (1999). Income inequality and health: pathways and mechanisms.(A Memorial to Alice Hersh). *Health Services Research, 34*(1), 215.

Kawachi, I., Kennedy, B. P., Lochner, K., & Prothrow-Stith, D. (1997). Social capital, income inequality, and mortality. *American journal of public health, 87*(9), 1491-1498.

Kondo, N., Kawachi, I., Subramanian, S. V., Takeda, Y., & Yamagata, Z. (2008). Do social comparisons explain the association between income inequality and health?: Relative deprivation and perceived health among male and female Japanese individuals. *Social Science & Medicine, 67*(6), 982-987. doi:10.1016/j.socscimed.2008.06.002

Li, H., & Zhu, Y. (2006). Income, income inequality, and health: Evidence from China. *Journal of Comparative Economics, 34*(4), 668-693. doi:10.1016/j.jce.2006.08.005

Loftin, C. (1974). Regional subculture and homicide an examination of the Gastil-Hackney thesis. *American sociological review : ASR : official journal of the American Sociological Association, 39*(5), 714-724. doi:10.2307/2094316

Lynch, J., Smith, G. D., Harper, S. A., Hillemeier, M., Ross, N., Kaplan, G. A., & Wolfson, M. (2004). Is income inequality a determinant of population health? Part 1. A systematic review. *The Milbank Quarterly, 82*(1), 5-99.

Lynch, J. W., Smith, G. D., Kaplan, G. A., & House, J. S. (2000). Income inequality and mortality: importance to health of individual income, psychosocial environment, or material conditions. *BMJ, 320*(7243), 1200-1204.

Mackenbach, J. P. (2002). Income inequality and population health: evidence favouring a negative correlation between income inequality and life expectancy has disappeared. (Editorials).(Editorial). *British Medical Journal, 324*(7328), 1.

MacMahon, S., Peto, R., Collins, R., Godwin, J., Cutler, J., Sorlie, P., . . . Stamler, J. (1990). Blood pressure, stroke, and coronary heart disease: part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *The Lancet, 335*(8692), 765-774.

Marmot, M. (2007). *The Status Syndrome: How Social Standing Affects Our Health and Longevity*: Macmillan.

Marmot, M. G., Rose, G., Shipley, M., & Hamilton, P. J. (1978). Employment grade and coronary heart disease in British civil servants. *Journal of Epidemiology & Community Health, 32*(4), 244-249.

McLeod, C. B., Lavis, J. N., Mustard, C. A., & Stoddart, G. L. (2003). Income inequality, household income, and health status in Canada: a prospective cohort study.(Author

Abstract). *The American Journal of Public Health, 93*(8), 1287. doi:10.2105/AJPH.93.8.1287

Millán, J., Pintó, X., Muñoz, A., Zúñiga, M., Rubiés-Prat, J., Pallardo, L. F., . . . González-Santos, P. (2009). Lipoprotein ratios: physiological significance and clinical usefulness in cardiovascular prevention. *Vascular health and risk management, 5*, 757.

Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why having too little means so much*: Macmillan.

Pei, X., & Rodriguez, E. (2006). Provincial income inequality and self-reported health status in China during 1991–7. *60*(12), 1065-1069. doi:10.1136/jech.2005.043539

Phelan, J. C., Link, B. G., Diez-Roux, A., Kawachi, I., & Levin, B. (2004). “Fundamental causes” of social inequalities in mortality: a test of the theory. *Journal of health and social behavior, 45*(3), 265-285.

Phelan, J. C., Link, B. G., & Tehranifar, P. (2010). Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. *Journal of health and social behavior, 51*(1_suppl), S28-S40.

Pickett, K., & Wilkinson, R. (2010). *The spirit level: Why equality is better for everyone*: Penguin UK.

Pickett, K. E., & Wilkinson, R. G. (2007). Child wellbeing and income inequality in rich societies: ecological cross sectional study. *BMJ, 335*(7629), 1080.

Pickett, K. E., & Wilkinson, R. G. (2015). Income inequality and health: A causal review. *Social Science & Medicine, 128*(C), 316-326. doi:10.1016/j.socscimed.2014.12.031

Popkin, B. M., Du, S., Zhai, F., & Zhang, B. (2009). Cohort Profile: The China Health and Nutrition Survey—monitoring and understanding socio-economic and health change in China, 1989–2011. *International Journal of Epidemiology, 39*(6), 1435-1440.

Preston, S. H. (1975). The Changing Relation between Mortality and level of Economic Development. *Population Studies, 29*(2), 231-248. doi:10.1080/00324728.1975.10410201

Price, G. M., Uauy, R., Breeze, E., Bulpitt, C. J., & Fletcher, A. E. (2006). Weight, shape, and mortality risk in older persons: elevated waist-hip ratio, not high body mass index, is associated with a greater risk of death. *The American Journal of Clinical Nutrition, 84*(2), 449-460. Retrieved from https://doi.org/10.1093/ajcn/84.2.449. doi:10.1093/ajcn/84.2.449

Regidor, E., Vallejo, F., Giráldez-García, C., Ortega, P., Santos, J., Astasio, P., & Fuente, L. (2015). Low mortality in the poorest areas of Spain: adults residing in provinces with lower per capita income have the lowest mortality. *Official journal of the European Epidemiological Federation associated to the International Epidemiological Association, 30*(8), 637-648. doi:10.1007/s10654-015-0013-x

Rodgers, G. B. (1979). Income and inequality as determinants of mortality: An international cross-section analysis. *Population Studies, 33*(2), 343-351. doi:10.1080/00324728.1979.10410449

Roseboom, T., de Rooij, S., & Painter, R. (2006). The Dutch famine and its long-term consequences for adult health. *Early human development, 82*(8), 485-491.

Rostila, M., Kölegård, M. L., & Fritzell, J. (2012). Income inequality and self-rated health in Stockholm, Sweden: A test of the 'income inequality hypothesis' on two levels of aggregation. *Social Science & Medicine, 74*(7), 1091-1098. doi:10.1016/j.socscimed.2011.11.027

Shibuya, K., Hashimoto, H., & Yano, E. (2002). Individual income, income distribution, and self rated health in Japan: cross sectional analysis of nationally representative sample. *BMJ, 324*(7328), 16. doi:10.1136/bmj.324.7328.16

Sinha, K., Davillas, A., Jones, A. M., & Sharma, A. (2018). *Distributional analysis of the role of breadth and persistence of multiple deprivation in the health gradient measured by biomarkers*. Retrieved from

Subramanian, S. V., & Kawachi, I. (2004). Income inequality and health: what have we learned so far? *Epidemiologic reviews, 26*(1), 78-91.

Sun, P., Unger, J. B., Palmer, P., Ma, H., Xie, B., Sussman, S., & Johnson, C. A. (2012). Relative income inequality and selected health outcomes in urban Chinese youth. *Social Science & Medicine, 74*(1), 84-91. doi:10.1016/j.socscimed.2011.10.010

Tang, S., Meng, Q., Chen, L., Bekedam, H., Evans, T., & Whitehead, M. (2008). Tackling the challenges to health equity in China. *The Lancet, 372*(9648), 1493-1501. doi:10.1016/S0140-6736(08)61364-1

Trialists'Collaboration, B. P. L. T. (2008). Effects of different regimens to lower blood pressure on major cardiovascular events in older and younger adults: meta-analysis of randomised trials. *BMJ, 336*(7653), 1121-1123.

Truesdale, B. C., & Jencks, C. (2016). The Health Effects of Income Inequality: Averages and Disparities. *Annu. Rev. Public Health, 37*(1), 413-430. doi:10.1146/annurev-publhealth-032315-021606

Wagstaff, A., & Van Doorslaer, E. (2000). Income inequality and health: what does the literature tell us? *Annual review of public health, 21*(1), 543-567.

Weich, S., Lewis, G., & Jenkins, S. P. (2002). Income inequality and self rated health in Britain. *Journal of Epidemiology and Community Health, 56*(6), 436. doi:10.1136/jech.56.6.436

WHO. (2004). Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet (London, England), 363*(9403), 157.

WHO. (2011a). *Use of glycated haemoglobin (HbA1c) in diagnosis of diabetes mellitus: abbreviated report of a WHO consultation*. Retrieved from

WHO. (2011b). Waist circumference and waist-hip ratio: report of a WHO expert consultation, Geneva, 8-11 December 2008.

Wilkinson, R. G. (1992). Income distribution and life expectancy. *British Medical Journal, 304*(6820), 165. doi:10.1136/bmj.304.6820.165

Wilkinson, R. G. (1996). *Unhealthy Societies: The Afflictions of Inequality*: Routledge.

Wilkinson, R. G. (1997). Comment: income, inequality, and social cohesion. *The American Journal of Public Health, 87*(9), 1504. doi:10.2105/AJPH.87.9.1504

Wilkinson, R. G. (2011). How economic inequality harms societies. In.

Wilkinson, R. G., & Pickett, K. E. (2006). Income inequality and population health: A review and explanation of the evidence. *Social Science & Medicine, 62*(7), 1768-1784. doi:10.1016/j.socscimed.2005.08.036

Wu, Y. (2006). Overweight and obesity in China. *BMJ, 333*(7564), 362-363. Retrieved from https://www.bmj.com/content/bmj/333/7564/362.full.pdf. doi:10.1136/bmj.333.7564.362

Yang, T., & Liu, W. (2018). Does air pollution affect public health and health inequality? Empirical evidence from China. *Journal of Cleaner Production, 203*, 43-52. doi:10.1016/j.jclepro.2018.08.242

Yang, W., & Kanavos, P. (2012). The less healthy urban population: income-related health inequality in China. *BMC Public Health, 12*, 804. doi:10.1186/1471-2458-12-804

Yu, W.-h., & Chiu, C.-T. J. S. I. R. (2016). Growing Pains: Changes in Psychological Well-Being in Urban China. *129*(3), 1349-1382. Retrieved from https://doi.org/10.1007/s11205-015-1166-4

Zheng, H. (2012). Do people die from income inequality of a decade ago? *Social Science & Medicine, 75*(1), 36-45.

Zhou, Z., Fang, Y., Zhou, Z., Li, D., Wang, D., Li, Y., . . . Chen, G. (2017). Assessing Income-Related Health Inequality and Horizontal Inequity in China.(Report)(Author abstract). *Social Indicators Research, 132*(1), 241. doi:10.1007/s11205-015-1221-1