ESSAYS ON INCOME, INEQUALITY, AND SOCIAL INSURANCE IN DEVELOPING COUNTRIES

A Dissertation

Submitted to the Faculty of the National Graduate Institute for Policy Studies (GRIPS)

in Partial Fulfillment of the Requirement for the Degree of

Ph.D. in Advanced Policy Studies

by

Trisukon Sawatrukkiat

September 2020

Abstract

Informal employment is the main source of employment in many developing countries, where more than 60% of workers work in this informal sector. These workers are mainly self-employed or working for small or family businesses who are not protected by social protection and regulation of the state. Their economic activities are generally not monitored, and the state cannot tax their associated revenues or incomes. This feature largely deviates from developed economies and is typically not considered in economic analyses.

This dissertation investigates three main issues in economic research by focusing on developing countries with a large informal sector: earnings profiles and uncertainties, inequality, and consumption and savings behaviors with policy implications. For our analysis, we select Thailand as a representative country because approximately 55% of workers are in the informal sector. Moreover, the country has developed social insurance, including social welfare programs, social security pension, and universal health insurance

We first investigate the age-earnings profiles and uncertainties over the life cycle between the formal and informal sectors. We examine the sources of earnings growth by considering different sectors of employment. Our results show that sector effect is statistically significant. We further investigate the difference of earnings shocks between the two sectors. We model the earnings shock process allowing for fixed effect, persistent shocks, and transitory shocks. Our findings show that earnings shocks in the informal sector are greater than in the formal sector; however, the shocks are more persistent in the formal sector. We also allow earnings shocks to vary between age groups in each sector. We find that earnings shocks increase with age in both sectors. Our results also show that

persistence declines with age in the formal sector while slightly increases with age in the informal sector.

Then, we examine the impact of earnings uncertainties from different sectors on earnings and consumption inequality over the life cycle. We develop a life cycle model with uninsurable earnings shocks from different sectors of employment. The model is calibrated to match Thai data. We find that key factors accounting for earnings and consumption inequality are the idiosyncratic earnings shocks in different sectors, working during old age, and social security. We also conduct policy experiments to investigate the impact of social insurance policy on inequality and individuals' consumption and saving behaviors by proposing two policy reforms: the extension of social welfare and extension of the social security system. Our results indicate that a high level of social welfare cannot reduce consumption inequality except very young workers and very old persons. However, the extension of social security to cover informal workers can reduce consumption inequality for middle-aged workers and retired persons.

Finally, we study the relationship between employment-based / private health insurance (EPHI) coverage and private saving in Thailand. Without controlling for endogeneity and using the standard Tobit model, our results show that the EPHI insured households tend to have positive savings, which is in line with the literature. After controlling for endogeneity by using the Instrumental Variables (IV) and the Full Model Maximum Likelihood (FMML) approaches, we find that households with EPHI coverage have significantly lower savings than those without coverage. Our finding provides evidence for the crowding-out effect of health insurance coverage on saving and supports precautionary saving theory.

To my beloved family

Acknowledgement

First and foremost, I would like to express my gratitude and considerable appreciation to my supervisor, Prof. Minchung Hsu, who made this dissertation possible. I am deeply indebted for his continuous guidance, support, and encouragement in writing this dissertation. He is always available when I need his help. My appreciations also goes to my dissertation committee members, Prof. Minchung Hsu, Prof. Ponpoje Porapakkarm, Prof. Junichi Fujimoto, Prof. Yoko Kijima, and Prof. Keisuke Otsu for their valuable time, suggestions, and comments. I also would like to thank Prof. Tetsushi Sonobe for his support and guidance to me as a G-cube student. I would like to express my thanks to all professors who taught me.

I would like to thank the Japanese government for granting me the G-cube Fellowship and the JICA SDGs Global Leader Fellowship Program to support my study and research, and the Bank of Thailand for allowing me to pursue doctoral studies. I also would like to thank the National Graduate Institute for Policy Studies (GRIPS), and its administrative staff, G-cube coordinator, AST, IPT, GAT, and Student Office who supported me both my academic life and daily life at GRIPS. I would like to thank all of my friends and everybody who supported me.

Finally, the most important, I would like to dedicate this work to my forever beloved family: my father, Somchai Sawatrukkiat, my mother, Lumyai Sawatrukkiat, and my younger brother, Woraphong Sawatrukkiat, who always give unconditionally love and support to me. Thank you very much.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

1.1 Motivation of the Study

This dissertation analyzes the age-earnings profiles, earnings uncertainties, inequality, and consumption and saving behaviors with their policy implications of developing countries. The life-cycle earnings profiles and earnings uncertainties are important sources of income or wealth heterogeneity in macroeconomic models to study consumption and saving behaviors, inequality, and social insurance. However, the existence of a large informal sector in developing countries that makes them different from developed countries is an important factor and cannot be ignored. By taking into account sectoral differences, it will provide a better understanding of the earnings dynamics in developing countries.,

The study of economic inequality has been associated with income, wealth, and consumption. An understanding of the sources of this type of inequality is significant in the sense that it relates to the study of consumption and saving behaviors. This also explains the ways individuals self-insure themselves against uncertainties. It also provides the scope for government policies and interventions, such as social insurance, in order to ensure consumption efficiency and maintain people's living standards. Most of the existing literature regarding these issues focuses on developed countries. However, the implications of the informal sector on earnings and consumption inequality have not been extensively studied in literature.

We have chosen Thailand for our research study for two main reasons. First, Thailand is a large developing economy, where around 55% of the country's workforce is employed in the informal sector. Second, Thailand has a number of developed social insurance policies, including social welfare programs, social-security pension, and universal health insurance.

1.2 Background Information

1.2.1 Informal Sector

The International Labour Organization (ILO) defines informal employment as those self-employed persons who work in less-organized small or unregistered businesses as well as wage workers who work without employment benefits and social security. Generally, the people working in the informal sector lack legal recognition and protection and are not supervised by a regulatory body.

Despite various attempts to formalize the informality of the economy, informal employment remains the primary source of income for people in many developing countries. A recent report from ILO (2018) states that, globally, more than 60% of the total employed population work in the informal sector. The number of workers employed in the informal sector varies across countries. Figure 1.1 displays the ratio of informal workers to the total in the non-agriculture sector of some selected developing countries. The percentage for each country exists between 30% and 80% and could be even higher if the workers in the agricultural sector were to be included.

1.2.2 Social Insurance

Social insurance is a type of support provided by the government for ensuring that its people can meet the basic needs. The social insurance provisions of developed countries are generous, providing a wide range of coverage. However, for developing countries, this type of support is still inadequate. This dissertation broadly focuses on three social insurance programs.

The first program is the social welfare program. Many countries have developed programs to fit the needs of their people. They usually consist of several programs targeting different groups of people with various requirements. Therefore, it is difficult to quantify all the benefits of social welfare of a country and compare it with the others. However, Hubbard et al. (2005) attempt to estimate the amount of subsidy in order to guarantee the minimum consumption floor in the US by considering Aid to Families with Dependent Children (AFDC), food stamps, and housing subsidies for people aged under 65. Based the study of a representative female-headed family with two dependent kids, the estimated amount of subsidy approximately corresponded to 14% of the GDP per capita in 1984. However, for developing countries, the amount of subsidies can be relatively much smaller. Thailand is a developing country that has recently introduced social welfare programs, including Child Support Grant (CSG) and a welfare card program. Following the same criteria as in case of the US, the amount of subsidy provided in Thailand was merely 5.5% of the GDP per capita in 2017.

The second is social security. The social-security pension is a crucial tool for ensuring the standard of living of the elderly after retirement and throughout their remaining lifetime. Many developing countries have developed social security pension

schemes. However, since most of the people in developing countries work in the informal sector, they are excluded from such schemes. While 90% of the older persons in developed countries receive pension, only less than 50% of those in developing countries receive pension (ILO, 2014). Moreover, it should be noted that although some developing countries provide a universal old-age benefit to support the elderly, the amount is generally small. For instance, in Thailand, the old-age benefit for old people aged 60–69 years in Thailand is 600 Thai baht a month.[1]

The third is public health insurance. Many international organizations have encouraged countries to avail universal health insurance covering the entire population, including those working in the formal as well as informal sectors (WHO, 2011). The public health insurance can help alleviate the burden of overwhelming health expenditures and, consequently, impact private savings and insurance decisions. Despite the coverage of universal public health insurance, the low-quality and limited access to public health services induces people to save more or have private health insurance for unexpected health expenditures from using private medical services (Guariglia and Rossi, 2004). Although Thailand provides universal health insurance that covers both the formal and informal sectors, workers in the formal sector receive better coverage owing to their employment-based health insurance. Therefore, the informal workers need to save more for unexpected medical expenditures that are not covered by public health insurance.

1.3 Research Strategy and Main Findings

This dissertation aims to answer the following three questions:

[1] Exchange rate = 31.33 Baht/US Dollar as of February, 2020

(1) What are the patterns of earnings profile and earnings shocks over the life cycle in developing countries?

First, we investigate the sources of earnings growth by estimating the age-earnings profiles of the formal and informal sectors controlling for an individual's specific characteristics, cohort effects, and location effects. We find that workers in the formal sector have higher earnings levels and earnings growth than those in the informal sector throughout their working lives. Following this, we identify the idiosyncratic components of labor earnings shocks, including fixed-effects, persistent shocks, and transitory shocks. We allow persistent shocks and transitory shocks to vary across sectors and age groups. We find that although the earnings shocks are more persistent in the formal sector, their magnitude is greater in the informal sector. In terms of age group, earnings shocks are more persistent for those aged between 30–50 and working in the formal sector, while the shocks are moderately persistent for the workers of other age groups in the formal sector and at all age groups for workers in the informal sector.

(2) How can earnings uncertainties explain the change of inequality in consumption and earnings over the life cycle in developing countries? And how social policy reforms affect consumption inequality and consumption and saving behavior?

We develop a life-cycle model with heterogenous agents facing mortality and earnings and employment uncertainties. During the agents' working age, they face different earnings shocks depending on the sector they work in and are allowed to change their employment sector in the next period. After retirement, all agents receive old-age transfer from the government. Our results show that earnings and consumption inequality increase with age. However, our model's consumption inequality increases faster than the data, and ends up 17.7% higher at retirement. We extend our model so as to incorporate

the mechanisms of insurance provisions available in Thailand— old-age employment and social security. Agents can continue to work after retirement age with a probability to retire at every age, and if they work in the formal sector before retirement, they qualify for receiving pension. We find that consumption inequality reduces with the presence of these two insurance mechanisms, and the model excels further in explaining consumption inequality in the data.

Next, we conduct policy experiments on two selected policies: social welfare program and social security pension extension. For the social welfare program, we use the amount of subsidies provided by the government as a measure to guarantee the minimum consumption floor. We find that a higher level of minimum consumption floor cannot reduce consumption inequality significantly, except for very young workers and very old persons, and savings among young workers is also reduced. For the extension of social security pension, we allow all agents to become eligible for pension after retirement, finding that consumption inequality is lower for middle-aged and old-aged persons compared to that in the pre-reform scenario.

(3) What is the relationship between insurance and household saving behavior in Thailand?

We follow the methodology used in Guaiglia and Rossi (2004) to investigate the association between employment-based / private health insurance (EPHI) and saving in Thailand. Firstly, we use the Tobit estimation controlling for an individual's specific characteristics and find a positive correlation between health insurance coverage and saving, it is indicated that those with coverage from the EPHI tend to save more. Following this, we apply the Instrumental Variables (IV) technique and a Full Model Maximum Likelihood (FMML) approach to control endogeneity bias. As a result, we find

a negative relationship between EPHI coverage and saving, and this effect is considerably significant in the FMML approach. Our results evidence the substitution effect between saving and insurance.

1.4 Contribution to the Literature

First, this dissertation contributes to the literature regarding life-cycle earnings profiles by incorporating the informal sector which is a unique feature of developing countries into the analysis. We have extended the specification from Kambourov and Manovskii (2009) in order to control the sector effect for investigating the differences in the earnings level and earnings growth of people with the same age but different employment sectors.

Second, this dissertation contributes to the literature pertaining to the persistence of earnings uncertainties over the life cycle. We characterize a process for identifying the idiosyncratic components of earnings shocks by allowing sectoral differences in our model. It has been highlighted that the existing literature focuses on developed countries (Meghir and Pistaferri (2004); Storesletten et al., (2004); Karahan and Ozkan (2013)). This study has made an attempt to fill the gaps in the existing literature by estimating the components of earnings shocks in the formal and informal sectors of developing countries and also allowing them to vary across different age groups.

The third contribution of this dissertation is that it provides quantitative analyses of the impacts of earnings risks on earnings and consumption inequality in developing countries by considering the special feature of such countries— the existence of the informal sector. This study is the first one to incorporate different earnings risk components of both the formal and informal sectors in a life-cycle model. We have also

incorporated the mechanisms of insurance arrangements available in Thailand— old-age employment and social security— into our analysis. Moreover, we have considered implementing some policy reforms to investigate their impacts on consumption inequality and saving behaviors.

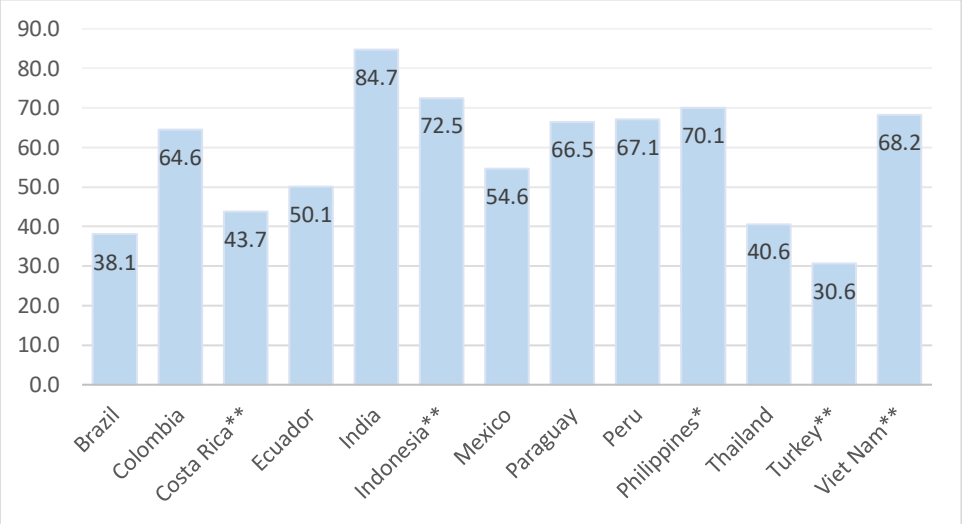
Lastly, this study offers empirical evidence to support the precautionary savings theory. We have investigated the association between health insurance and saving in Thailand. Starr-McCluer (1996) uses the US micro-data in order to examine the relationship between the demand for private medical insurance and wealth accumulation. Guariglia and Rossi (2004) examine the relationship between insurance and saving in the UK.

1.5 Organization of the Study

The dissertation proceeds as follows. In chapter 2, we estimate the age-earnings profiles of the formal and the informal sector. We also characterize an earnings shocks process for idiosyncratic components of labor earnings shocks in different employment sectors. In chapter 3, we quantitatively investigate the impact of earnings shocks from two different sectors on earnings and consumption inequality. We also investigate the effect of social policy reforms on inequality and consumption and saving behavior. In chapter 4, we investigate the effect of private health insurance on private saving in Thailand. In chapter 5, we summarize our findings and provide some policy implications.

1.A Figure

Figure 1.1 Informal Employment Share (Non-Agriculture, 2012)



Note: *data of 2008, **data of 2009. Source: International Labour Organization

CHAPTER 2

Life Cycle Earnings Profiles and Uncertainties in Developing Economies

2.1 Introduction

Understanding how earnings profiles and uncertainties change over the life cycle is essential as they directly affect consumption and saving behaviors, abilities of self-insurance, and wealth inequality. Moreover, it is a crucial source of income or wealth heterogeneity in macroeconomic models to study consumption, saving, risk-sharing, inequality, and social insurance. While most existing work on earnings profiles and earnings risks has focused on developed countries, this field of study in developing countries has not been thoroughly studied in the literature.

Why do we need to focus on developing countries? First, a remarkable feature of developing countries is the existence of informal economy, where economic activities are not legally regulated, and the employment relationships are not socially protected. Informal employment has often been associated with inferior earnings, lack of legal protection, and precarious working conditions. Consequently, they are more vulnerable to life cycle uncertainties; for example, earnings or health shocks, which can lead them to poverty. Second, the majority of the population in developing countries is engaged in the informal sector. More than 60% of the labor force in developing economies are employed in the informal sector, contrary to developed countries where less than 10% of

workers are informally employed (Bacchetta et al., 2009; ILO, 2018). As the majority of the population in developing countries is engaged in the informal sector, they cannot be neglected in the economic analysis.

Therefore, this study focuses on two main issues. First, we aim to examine whether being employed in two different sectors has a significant impact on earnings growth. We investigate the sources of earnings growth by considering different sectors of employment as well as controlling for other factors such as cohort and education. Second, we investigate whether earnings shocks in the two sectors share a similar pattern. We model the earnings shock process to identify the difference between the formal and informal sectors. Our findings show that earnings growth and shocks in developing countries have a pattern different from those in developed countries.

In the literature, most existing work on earnings profiles and persistence of earnings shocks over the life cycle has concentrated on developed countries. Kambourov and Manovskii (2009) document a significant flattening of life cycle earnings profile and increasing inequality for recent cohorts in the US. Lagakos et al. (2018) investigate the life cycle wage growth across countries and find that it is significantly higher in developed countries than in developing countries. Meghir and Pistaferri (2004) study an income process for individual earnings in the US and argue that permanent and transitory shocks are important components of income shocks. Storesletten, Telmer, and Yaron (2004) study why earnings and consumption inequality increase with age and find that highly persistent earnings shocks can explain the pattern. Karahan and Ozkan (2013) study how the persistence of earnings changes over the life cycle and find that it rises with age until midway through life, exhibiting a hump-shaped profile over the life cycle.

Studies on formal/informal earnings differential reveal informal workers receive earnings lower than those working in the formal sector with same characteristics (Dasgupta et al., 2015; Gong & Van Soest, 2002; Tansel & Kan, 2012). However, none of them study the life cycle pattern of earnings between the two sectors.

We begin our analysis by estimating age-earning profiles of the formal and informal sectors from the Household Socio-Economic Survey (HSES) conducted by the National Statistical Office (NSO) of Thailand. Our finding, higher earnings for workers in the formal sector, is in line with the literature related to formal/informal earnings differential. Moreover, we find that formal workers, given the same education level, experience higher earnings growth than the informal sector over working years.

Then, we specify an earnings shocks process to identify the idiosyncratic components of labor earnings risk. We use panel data from the Townsend Thai Household Annual Resurvey, allowing for fixed effects, persistent shocks, and transitory shocks. In this study, we apply two earning shock specifications, age-invariant and age-dependent, to estimate the autocorrelation coefficient of persistent shocks for workers in different sectors. We find that the coefficients of persistent shocks are less than unity for both sectors and higher for the formal sector. However, variances of persistent and transitory shocks are higher in the informal sector.

The remainder of this chapter is organized as follows. Section 2 investigates the life cycle earnings profiles in different employment sectors. Section 3 explains the income shocks estimation. Section 4 concludes the chapter.

2.2 Life-Cycle Earnings Profile

2.2.1 Data

In this study, we select Thailand as a representative country because informal employment dominates its labor market, with approximately 55% of workers employed informally (NSO, 2019). For this analysis, we use the data from the Household Socio-Economic Survey (HSES) conducted by the NSO from 2002 to 2017 with a two-year gap.[2] Each survey interviews around 43,000 households to collect information on household and individual characteristics such as their income, expenditure, participation in the labor market, education, and socio-economic status.

The data is restricted to male heads of households, aged 30–59, who are in the labor force.[3] We select only male workers because, in developing countries, men are more likely to engage in a full-time job than women. This study is not restricted to only wage workers. This is an important feature of this study as the majority of informal workers in developing countries are non-wage workers such as self-employed workers. Focusing only on wage workers would lead to selection bias of informal workers.

2.2.2 Identification of Formality/Informality

The definition of informality differs between countries due to their diverse social and economic conditions. Therefore, it is important to point out how this study defines the sector of employment for each individual in the sample. We consider two aspects to distinguish between formality and informality.

[2] Prior to 2006, the survey was conducted once every two years. After 2006, it is conducted annually; however, the household income data is reported once every two years. Thus, HSES data used in this study is from 2002, 2004, 2006, 2007, 2009, 2011, 2013, 2015, and 2017.
[3] We choose starting age as 30 to be consistent with earnings shock estimation in section 3.

First is social protection. According to the ILO (2003), informal workers are employees whose employment relationship is not subject to national labor legislation, income taxation, social protection, or entitlement to certain employment benefits.

Second is employment status. Informal employment covers a wide range of occupational groups, including street vendors, small farmers, workers in small unregistered factories or workshops, etc. Despite the variety, informal workers can be classified into two broad groups by their employment status: self-employed workers who operating small unregistered businesses; wage workers who work in insecure and unprotected jobs (Chen, 2005).

According to the definition by the NSO, informal employment refers to employed persons who are not covered by social security and receive no protection and benefits from their employers. Following this definition, we use medical service welfare from employers as an indicator of formality. [4] The respondent who reports receiving government/state enterprise's welfare or social security's medical card is classified as a formal worker, while that receiving other types of medical welfare is classified as an informal worker.

2.2.3 Estimation of Earnings

The focus of this study is to generate an age-earnings profile of workers in different employment sectors: formal and informal. All sources of earnings from work, both wage and non-wage incomes, are used to estimate workers' average annual earnings.[5] The data

[4] We use question 18 in Part 1 of HSES Questionnaire of Household Members and Expenditures. Household members who receive either government/state enterprise's welfare or social security's medical card are considered as workers in the formal sector.

[5] The incomes include labor income, both wage and non-wage, but not include non-labor income, such as property income, and transfer payment.

reported by the HSES is in Thai baht and nominal terms. Consumer price indices (CPI) based on the year 2015 published by the Thai Bureau of Trade and Economics are used to convert nominal earnings into real values. The earnings of each worker are calculated based on their work status as described below:

Employees: It covers government employees and private employees. Earnings of wage workers are directly determined by their total annual income, including wages/salaries, overtime, bonuses, and welfare received from employers.[6]

Self-employed workers: It covers employers and own-account workers. There are two sources of earnings for this type of workers: income from farm and non-farm businesses. Earnings are determined based on the occupation reported as a primary job. If the occupation is reported as skilled agricultural, forestry, and fishery, income from farm business is used to calculate annual earnings. For other types of occupations, income from non-farm business is used. Annual earnings are defined as gross income from business minus expenditure.[7] Earnings of self-employed workers are difficult to separate capital income and labor income. Thus, our estimation of earnings for self-employed workers could be overestimated. However, even we may include capital income, their income level are still much lower than formal workers. So, our estimate could be the upper bound for informal workers.

[6] Total earnings of wage workers = IW12+IW13+IW14

[7] Annual earnings for farm business are defined as gross receipts (IA08+IA09+IA11+(IA15-IA19)+IA20) minus expenditure (IA25+IA26+IA13+IA17+IA22), according to the NSO calculation. Annual earnings for non-farm business are gross money receipts from business (IB08) minus total operating cost (IB09), then multiplied by percent share of net profit from the business (IB11).

Contributing family workers: We use a strategy similar to Dasgupta et al. (2015) to impute earnings for self-employed workers. Total annual earnings of household business with contributing family workers are equally distributed between contributing family workers and the owners.[8]

We drop workers who are members of cooperative group and select only samples with positive earnings. We exclude the highest one percent quantile of annual earnings to remove an exceptional outlier from the sample resulting 147,619 observations.

2.2.4 Descriptive Statistics

Descriptive statistics of all samples and of samples in each sector are summarized in Tables 2.1. Approximately 71% of all samples are employed in the informal sector. Regarding the educational level, the majority of informal workers are high school dropouts (84.68%), while formal workers are at least high school graduates (55.23%). In terms of employment status, most of the informal workers are self-employed (73%), while formal workers are mostly employees (89.63%).

An alternative measure to define formality is employment status. Informal workers include small business owners, own-account workers, and contributing family workers. All others are classified as formal workers. We drop members of cooperative group from our sample as they account for only 0.01% of the total sample. Top 5% earners of employers are excluded because they might be employers in the formal sector whose earnings are usually exceptionally high. We also drop the bottom 40% of private employees who are likely employed by informal small businesses from our sample.

[8] We assume that all workers contribute to the family business equally.

Characteristics of workers by sector classified by employment status are presented in Table 2.2. The results show the similar pattern as in Table 2.1 that almost 70% of workers are employed in the informal sector. In terms of educational level, 84.14% of informal workers are high school dropouts while 52.7% of formal workers are at least high school graduates.

Figure 2.1 presents a comparison of average real annual earnings between formal and informal workers over the life cycle. It plots the average earnings within age bins of workers in each sector. This preliminary pattern of life cycle earnings profile reveals while average real earnings of workers in the formal sector increase significantly with age, earnings of informal workers show a flatter pattern along the life cycle.

Individuals with different completed education degrees earn differently. Previous studies show that workers with a higher education level would have higher earnings than those with a lower education level (Mincer, 1974). In this study, individuals are classified into either high school dropouts or high school graduates according to their latest education attainment level, as reported in the HSES. Figure 2.2 plots the average real annual earnings of workers in the formal and informal sectors by the level of education attained. For both formal and informal sectors, workers who, at least, are high school graduates have higher earnings than those without a high school degree. However, given the same level of education, formally employed workers have higher earnings than their counterparts in the informal sector.

Figure 2.3 shows the patterns of the age-earnings profile by sector from different cohorts. A cohort is denoted by the birth year of each individual. Then, they are grouped into a 10-year cohort bin (for example, individuals born during 1950–1959 belong to 1950s bin). We assume that workers within the same cohort bin are at the same age and

enter the labor market at the same time. Life cycle earnings profiles of the informal workers are flatter relative to the formal workers within the same cohort bin. The average earnings has increased for more recent cohorts for both types of workers.

2.2.5 Empirical Analysis

From the cross-sectional evidence on life cycle earnings profiles presented earlier, we ask whether the difference in the sector of employment can explain the difference in earnings patterns. We develop the model by using the following regression model:

$$ln\,(E_{i,t}) = \beta_0 + \beta_1 h_{i,t} + \beta_2 h_{i,t}^2 + \beta_3 s_{i,t} + \beta_4 s_{i,t} h_{it} + \beta_5 edu_{i,t} + \beta_6 edu_{i,t} h_{i,t} + \beta_7 edu_{i,t} s_{i,t} + \beta_8 M_t + X_{i,t}\gamma + u_{i,t}, \quad (2.1)$$

where E_{it} is real annual earnings of individual *i*, who conducted the survey at time *t*, $h_{i,t}$ is the current age in years at the time when the survey was conducted, $s_{i,t}$ is a sector dummy which equal to one if individual works in the formal sector and zero otherwise, $edu_{i,t}$ is an education dummy which equal to one if the individual has an education level of high school or above and zero otherwise, $X_{i,t}$ comprises other covariates including cohort (birth year of the interviewed individual) and location, and $u_{i,t}$ is the residual of the logarithm of earnings. M_t is the unemployment rate to capture the macroeconomic effects in a specific survey year. The quadratic in age h_{it}^2 allows for the effect of age on earnings at a decreasing rate. The interaction terms of the linear age allow different sectors and education levels to have different slopes of the earning profiles indicating the growth of earning over the life cycle.

Pooled ordinary least squares regression is performed by using the sample from the HSES described in section 2.2.1. For simplicity, we normalize initial age and cohort

to zero. Figures 2.4 and 2.5 present the age-earnings profiles from the estimated values of the logarithm of annual earnings by using coefficients from the regression equation (2.1) compared to the data. Table 2.3 summarizes estimation results. We find informal workers at a disadvantage. Coefficients on the formal sector dummy are significantly positive, indicating that earnings level of formal workers is higher than informal workers. This result is consistent with the literature on earnings differential between formal and informal employment. Our results show that earnings level of higher educated formal workers can be approximately 67% larger than that of lower educated informal workers. Moreover, the coefficient on the interaction of linear age and formal sector dummy is statistically positive, explaining higher earnings growth in the formal sector.

Moreover, on exploring effects of education on earnings growth we find that workers with higher education have higher earnings, both in terms of level and growth than workers with lower education. The coefficient of the interaction between education and sector is also significantly positive, which creates a large earnings gap between higher-educated workers in different sectors.

Our results show that higher-educated formal workers have the highest average earnings growth over 30 years of working at 4.1%, whereas informal workers with low education have the lowest average earnings growth at 0.8%. Even less-educated formal workers have higher average earnings growth than informal workers with higher education—2.9% and 1.9% growth, respectively. Given the highest average earnings growth of 4.1%, this figure is much lower than average earnings growth in the US, approximately 9.2%, as presented by Kambourov and Manovskii (2009).

2.3 Earnings Uncertainties

This section investigates uninsured idiosyncratic shocks to earnings in the two employment sectors. Although the age-earning profile provides insights on how earnings change with age, a residual deviates earnings from the earning profiles. This residual is the earnings shock or uncertainty that individuals face over time.

In this study, we assume that the earnings shock follows a functional form and constitutes three components: an individual-specific fixed effect, a persistent component, which follows an AR(1) process, and a transitory component. There are two main features in this study. First, all three components of the shocks are different between the two sectors. We assume that individuals remain in the same sector during their working life. Second, there are two specifications for the earnings process: age-invariant and age-dependent processes. In an age-invariant earnings process, the coefficient of the persistent component and variance of transitory and persistent shocks are constant over the life cycle, which is often used in the literature. However, according to Karahan and Ozkan (2013), workers at different ages face different earnings shocks. Thus, we also allow the persistent parameter and variance of shocks to vary by age group following an age-dependent earnings process.

2.3.1 Data Selection and Construction

To study how earnings vary with age due to uncertainties, panel data is required to construct the variance-covariance matrix between different ages of residuals earnings to estimate earnings shocks. Although the HSES data provides rich detail of workers' characteristics and their earnings, it is cross-sectional. Therefore, we use the panel data

from Townsend Thai Household Annual Resurvey, between 2008 and 2015 which provide information on both urban and rural areas, for this analysis. The survey is conducted on an annual basis and contains information on the earnings of households as well as their socio-economic characteristics.

The sample is restricted to males aged 30–59, who are in the labor force, have positive earnings, and are present in at least four consecutive years of the survey. Estimation of average annual earnings follows the same procedure as the HSES data. Earnings of each worker are calculated based on his work status.[10] Total earnings are converted to real earnings by using the CPI deflator based on the year 2015. The top 10% of employers and the bottom 10% of private employees are dropped from the sample. Following these criteria leaves an unbalanced panel of 477 and 1,118 individuals in the formal and informal sectors, respectively. Therefore, the total sample consists of 1,595 individuals and 9,734 observations.

To define formality, it is difficult to classify by social protection as Townsend's panel data does not provide such information. However, each household member reported their employment status. Therefore, we use this definition to indicate the formality of each laborer. This study assumes that workers who report their work status as the owner of a business, contributing family worker, and employee earning piece rate or daily wage are classified as workers in the informal sector, whereas employees with monthly wages and government workers are defined as formal workers. By applying this assumption, the

[10] We use daily wage variable (oc1c) to calculate earnings for wage earners and as a proxy to identify the number of self-employed workers (owners of business, unpaid family members, and employees who earn piece rate) in the household.
Total earnings of monthly wage worker = oc1c*12
Total earnings of daily wage worker = oc1c*12*20
Total income of informal workers = net household business income (in11)/total self-employed worker in household (oc1c=.c)

proportions of formal and informal workers in the data set are approximately 21.5% and 78.5% of total employment, respectively.[11] Table 2.4 presents descriptive statistics of all samples and by sector. The majority of formal workers are at least high school graduates (65.8%), while the education level of informal workers is mostly lower than the high school level (79.35%).

2.3.2 Age-Invariant Earnings Process

We begin with the age-invariant earnings process, in which all shock parameters are constant over time. We extract the residual earnings from the Townsend panel data by following the estimation in the previous section. The residual earnings are decomposed into a fixed effect, an AR(1) component, and a transitory component. We follow the method used in Storesletten et al. (2004) for an age-invariant earnings process. A time-series process for u_{ih} is specified as follows:

$$u_{i,h}^{s} = \alpha_i + z_{i,h}^{s} + \varepsilon_{i,h}^{s}, \tag{2.2}$$

$$z_{i,h}^{s} = \rho_s z_{i,h-1}^{s} + \eta_{i,h}^{s}, \tag{2.3}$$

where $\alpha_i \sim N(0, \sigma_\alpha^2)$, $\varepsilon_{i,h}^{s} \sim N(0, \sigma_{\varepsilon,s}^2)$, $\eta_{i,h}^{s} \sim N(0, \sigma_{\eta,s}^2)$, $z_{i,0}^{s} = 0$ and $E\left(u_{i,h}^{s}\right) = 0$ in the cross-section for all ages. α_i is an individual-specific fixed effect, determined at the beginning of the life cycle and is the same for all individuals in both sectors. The terms $z_{i,h}^{s}$ and $\varepsilon_{i,h}^{s}$ are persistent and transitory shocks of earnings, respectively, at age h of individual *i* employed in sector *s*, realized at each period over the life cycle. The transitory shock includes measurement error and temporary changes in earnings such as bonuses

[11] The possibility of informal workers being over-represented in the data set is a limitation of using Townsend's panel data.

and overtime payment. The persistent shock captures lasting changes in earnings, such as promotions and health status. Each individual faces a persistent shock of size $\eta_{i,h}^{s}$, governed by the variance $\sigma_{\eta,s}^{2}$. The persistent coefficient ρ_s determines the extent to which the shock lasts.

To estimate all parameters, we employ a minimum distance estimator method that minimizes the distance between the theoretical variance-covariance structure of residual earnings and their empirical counterparts. The variance-covariance matrix between different ages of the residual earnings is constructed from Townsend's panel data. We restrict to individuals who appear in the survey for at least four consecutive survey years and have never switched their sector of employment.[12] The theoretical variance-covariance structure of the idiosyncratic shocks is given by:

Variances:

$$var(u_0^s) = \sigma_\alpha^2 + \sigma_{\varepsilon,s}^2 + \sigma_{\eta,s}^2 \tag{2.4}$$

$$var(u_j^s) = \sigma_\alpha^2 + \sigma_{\varepsilon,s}^2 + var(z_j^s) \tag{2.5}$$

$$var(u_j^s) = \sigma_\alpha^2 + \sigma_{\varepsilon,s}^2 + \left(\sum_{k=0}^{j} \rho_s^{2k}\right)\sigma_{\eta,s}^2 \tag{2.6}$$

Covariances:

$$\begin{aligned} cov(u_0^s, u_j^s) &= \sigma_\alpha^2 + cov(z_0^s, z_j^s) \\ &= \sigma_\alpha^2 + \rho_s^j var(z_0^s) \\ &= \sigma_\alpha^2 + \rho_s^j \sigma_{\eta,s}^2 \end{aligned} \tag{2.7}$$

[12] According to Townsend's panel data, the sector transition matrix shows that 96.38% of workers in the informal sector and 89.2% of workers in the formal sector remain in the same sector of employment (see Appendix).

$$cov\left(u_n^s, u_j^s\right) = \sigma_\alpha^2 + cov\left(z_n^s, z_j^s\right)$$

$$= \sigma_\alpha^2 + \rho_s^{j-n} var(z_n^s)$$

$$= \sigma_\alpha^2 + \rho_s^{j-n}\left(\sum_{k=0}^{n} \rho_s^{2k}\right)\sigma_{\eta,s}^2 \qquad (2.8)$$

The estimation results are presented in panel (A) of Table 2.5. The first column is the variance of a fixed effect, 0.0429, which is the same for both sectors. The left panel reports all parameters of the formal sector, and the right panel represents the informal sector. We find that the persistent coefficients for the formal and informal sectors are 0.9842 and 0.7789, respectively. The estimation results indicate that earnings shocks in the formal sector are more persistent than in the informal sector. Formal workers tend to take a longer time to recover from an unexpected shock. However, the variances of persistence and transitory shocks are higher in the informal sector than the formal sector, indicating that the size of the shock is greater in the informal sector. The variances of persistence and transitory shocks are 0.0105 and 0.0421 for the formal sector, and 0.0947 and 0.1579 for the informal sector, respectively.

2.3.3 Age-Dependent Earnings Process

For an age-dependent earnings process, we follow the method used in Karahan and Ozkan (2013). However, our sample size is too small to estimate the age-profile of these parameters precisely. Therefore, we use age groups instead of age for our analysis. Each individual is classified into three age groups (g): 30–40, 40–50, and 50–59. We follow the same variance-covariance specifications as the age-invariant process and allow the

variance of persistent shock, the variance of transitory shock as well as the persistent coefficient to be different between different age groups.

Panel (B) of Table 2.5 shows results from the age-dependent earnings process. The variance of the fixed effect is 0.0571. The persistent coefficient shows a different pattern with age group between the formal and informal sectors. For the first two age groups, persistent coefficients for both sectors are quite flat and slightly decrease for both sectors. Although the persistent coefficient for the last age group continues to decrease for the formal sector, it increases for the informal sector. As the age group increases, persistent shocks for the formal sector decrease from 0.9643 to 0.9286 and 0.6429, while the persistent coefficients for the informal sector are 0.6071, 0.5814, and 0.7500. Workers in the formal sector experience higher persistent earnings shock early in working life, and it becomes moderately persistent after age 50. On the other hand, shocks in the informal sector are moderately persistent over time.

Moreover, variances of persistent shocks are higher in the informal sector except for the last age groups. Variances for age groups 30–40, 40–50, and 50–59, are 0.0102, 0.0207, and 0.1814 for the formal sector and 0.2047, 0.2389, and 0.0861 for the informal sector, respectively.

The variance of transitory shocks increases with age for the informal sector, but the size is more significant in the informal sector than the formal sector. For the formal sector, it increases from 0.0278 to 0.1107 and decreases to 0.0547. For the informal sector, it increases from 0.1791 to 0.1911 and 0.3122 as the age group increases.

2.3.4 Assessment of the Estimation Results

We compare theoretical variances of residual earnings from the two earning specifications with their empirical counterparts. Equation (2.6) presents theoretical variances. For $\rho < 1$, the residual inequality will converge to a well-defined limit in a concave fashion as the summation term converges to $\sigma_\eta^2/(1-\rho^2)$. If $\rho = 1$, the variance profile will be linear. Empirical variances of residual earnings in Figure 2.6 are from Townsend's panel data, implying that targeting these moments would result in a different ρ between the two sectors. The formal sector should have a higher ρ as the degree of concavity of its variance profile is less than the informal sector.

Figure 2.6 shows the profile of the variance of residual earnings from the age-invariant process by sector of employment. The variances increase steadily for the formal sector. On the contrary, the variances for the informal sector increase until age 35 and become stable after that.

Figure 2.6 also presents the variance profile from the age-dependent earnings process. Although the model can capture the patterns of the variance profile for each sector, the estimates of the variance from the model are higher than their empirical counterparts.

Our results from age-invariant and age-dependent earnings processes can replicate the pattern of age-profiles of residual inequality for both sectors. The age-invariant process performs better in capturing both the pattern and level of variance profiles. There are two limitations to the age-dependent specification. First, the sample size is very small, especially when we group them into age groups. Second, we cannot select parameters without constraints to only eligible values due to computational limitation while the parameters for the age-invariant are able to be chosen freely.

2.4 Conclusion

This chapter examines earnings profiles and uncertainties over the life cycle in developing economies with the existence of the informal sector. First, our results show that the impact of the sectors is statistically significant. Earnings of formal workers are found significantly greater than informal workers, in line with the literature on formal/informal earnings differential. Moreover, our study contributes to the literature on life cycle earnings, considering sector effects on earnings. We find that the growth rate of earnings is higher in the formal sector. In terms of level of education, formal workers are better-off than informal workers. Higher-educated formal workers have the highest average earnings growth. Even formal workers with lower education can have higher average earnings growth than informal workers with higher education.

Second, this study contributes to the literature on estimating the idiosyncratic earnings shocks for two aspects: sectors and age groups. We find persistent coefficients in both sectors less than unity and the shocks more persistent in the formal sector. However, their variances are higher in the informal sector, indicating that the magnitude of the shocks is more substantial in the informal sector than in the formal sector. When we allow different age groups to have different shocks, our results show that persistence declines with age in the formal sector while slightly increases with age in the informal sector. This pattern is different from Karahan and Ozkan (2013), where the persistent profile exhibits a humped-shape.

2.A Tables

Table 2.1 Characteristics of Workers by Sector

No. of observation = 147,619

	All	Formal (29.04%)	Informal (70.96%)
Education (%)			
Lower than high school	73.09	44.77	84.68
High school graduate	11.09	18.35	8.12
Vocational education	4.78	11.00	2.24
College and above	11.04	25.88	4.96
Work Status (%)			
Employer	9.14	2.43	11.88
Own-account worker	44.24	7.66	59.21
Contributing family worker	1.22	0.28	1.61
Government employee	13.03	36.39	3.48
Private employee	32.37	53.24	23.82
Earnings (in 2015 Thai Baht)			
P25	47,321	118,989	34,014
P50	95,905	198,833	68,384
P75	186,449	327,592	118,328

Source: Author's estimation is based on the HSES data from 2002 to 2017, NSO.

Table 2.2 Characteristics of Workers by Sector (Employment Measure)

No. of observation = 128,655

	Formal (31.5%)	Informal (68.5%)
Education (%)		
Lower than high school	47.3	84.14
High school graduate	17.99	8.28
Vocational education	9.83	2.66
College and above	24.88	4.92
Earnings (in 2015 Thai Baht)		
P25	120,564	28,701
P50	189,206	64,921
P75	319,835	129,311

Source: Author's estimation is based on the HSES data from 2002 to 2017 which the top 5% earners of employers and the bottom 40% of private employees are excluded, NSO.

Table 2.3 Estimation Results of Age-earnings Profile

Variable	**ln Earnings**	
Age ($h_{i,t}$)	0.0349***	(0.003)
Age Squared ($h_{i,t}^2$)	-0.0009***	(0.000)
Formal sector dummy ($s_{i,t}$)	0.1556***	(0.015)
Formal*Age ($s_{i,t} * h_{i,t}$)	0.0205***	(0.001)
High education dummy ($edu_{i,t}$)	0.2677***	(0.017)
High education*Age ($edu_{i,t} * h_{i,t}$)	0.0090***	(0.001)
High education*Formal	0.2423***	(0.015)
Cohort	0.0265***	(0.002)
Region		
Central (Excluding Bangkok Metropolis)	0.0000	(.)
Bangkok Metropolis	0.3899***	(0.009)
North	-0.5089***	(0.009)
Northeast	-0.7718***	(0.010)
South	-0.0497***	(0.009)
Public Sector	0.4964***	(0.011)
Unemployment rate	-0.0827***	(0.022)
Intercept	10.4811***	(0.092)
N	147619	

Source: Author's estimation is based on the HSES data from 2002 to 2017, NSO.
Note**:** Standard errors are reported in parentheses. *, **, *** indicate significance at the 90%, 95%, and 99% levels, respectively. Initial age and cohort are normalized to zero.

Table 2.4 Descriptive Statistics from Townsend's Panel Data (cross-sectional)

No. of observations = 9,734

	All	**Formal (21.5%)**	**Informal (78.5%)**
Education (%)			
Lower than high school	69.87	35.2	79.35
High school graduate	12.05	19.32	10.06
Vocational education	8.89	17.65	6.49
College and above	9.19	27.83	4.1
Earnings (in Thai Baht)			
P25	54,707	108,487	49,931
P50	87,364	153,061	72,705
P75	147,843	237,836	120,049

Source: Author's estimation is based on Townsend's panel data from 2008 to 2015 which the top 10% earners of employers and the bottom 10% of private employees are excluded.

Table 2.5 Estimation results – Earnings shock parameters

Specification		**Formal**			**Informal**		
	σ_α^2	ρ	σ_ε^2	σ_η^2	ρ	σ_ε^2	σ_η^2
(A) Age-invariant	0.0429	0.9842	0.0421	0.0105	0.7789	0.1579	0.0947
(B) Age Group							
30–40	0.0571	0.9643	0.0278	0.0102	0.6071	0.1791	0.2047
40–50	0.0571	0.9286	0.1107	0.0207	0.5714	0.1911	0.2389
50–59	0.0571	0.6429	0.0547	0.1814	0.7500	0.3122	0.0861

Source: Author's estimation is based on Townsend's panel data from 2008 to 2015.

2.B Figures

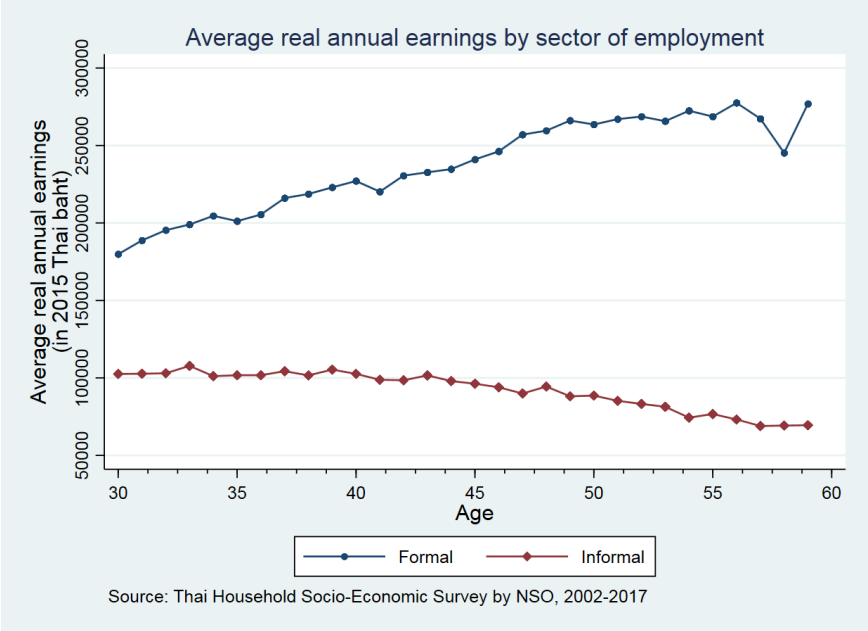


Figure 2.1 Average Real Annual Earnings of Workers by Sector, Raw data

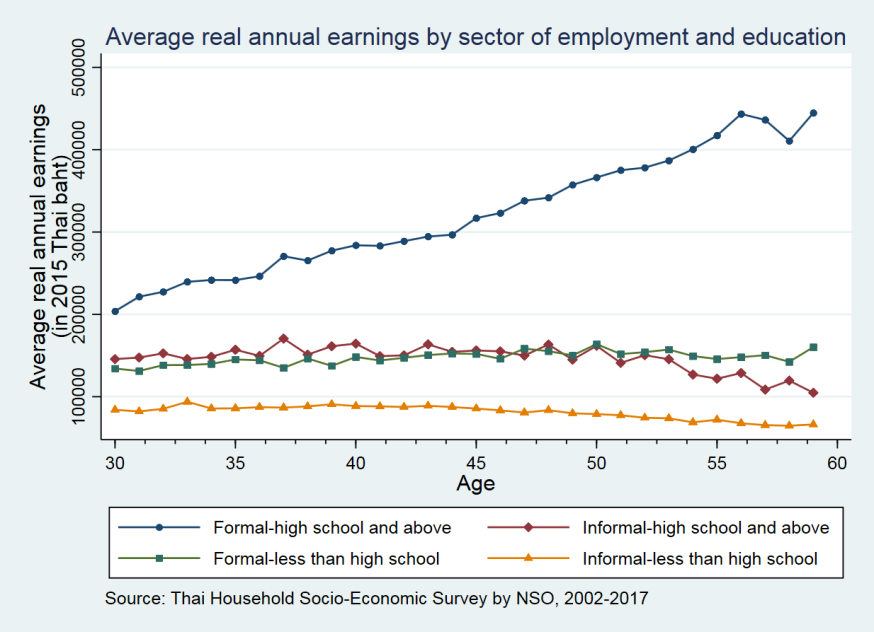


Figure 2.2 Average Real Annual Earnings of Workers by Sector and Education, Raw Data

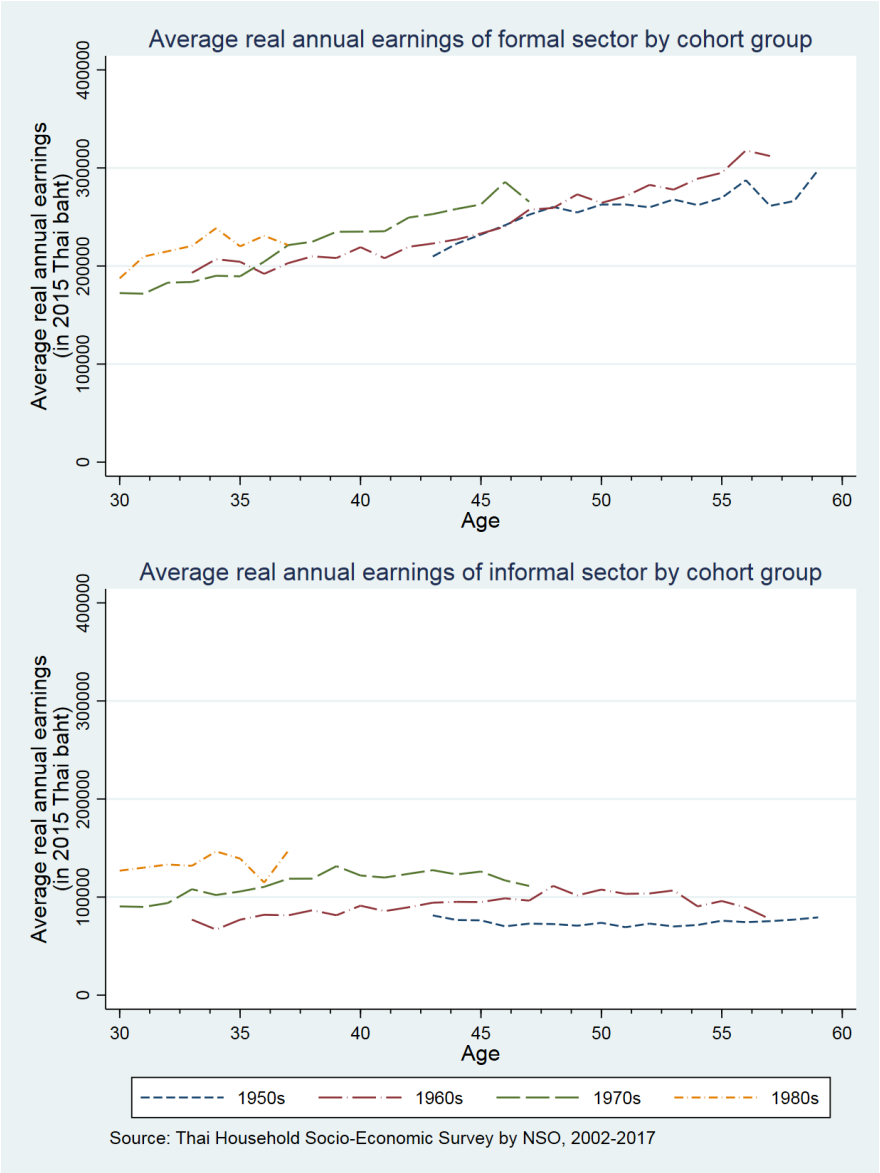


Figure 2.3 Average Real Annual Earnings of Workers by Cohort, Raw Data

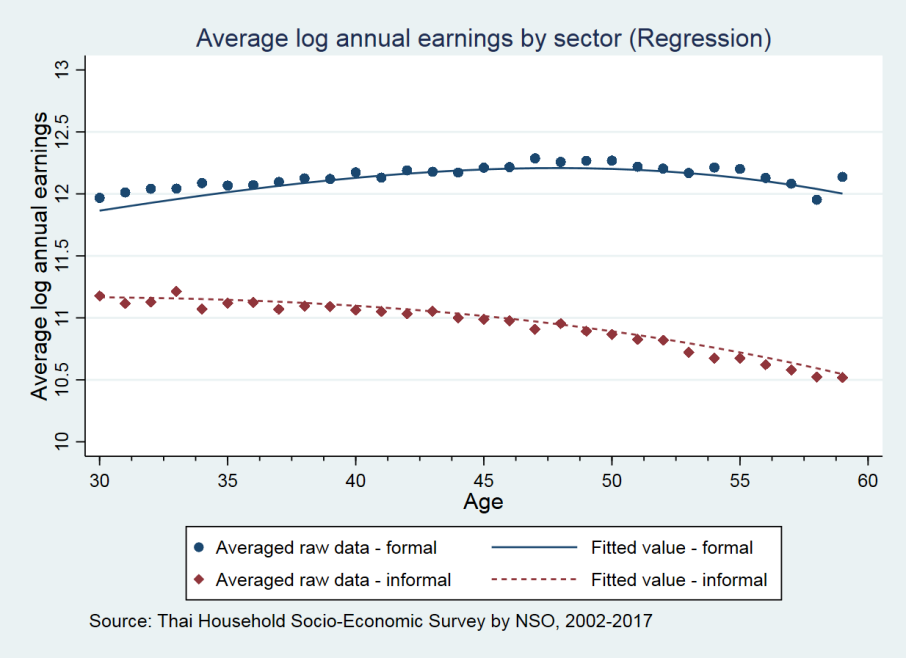


Figure 2.4 Logarithm of the Annual Earnings of Workers by Sector, Regression

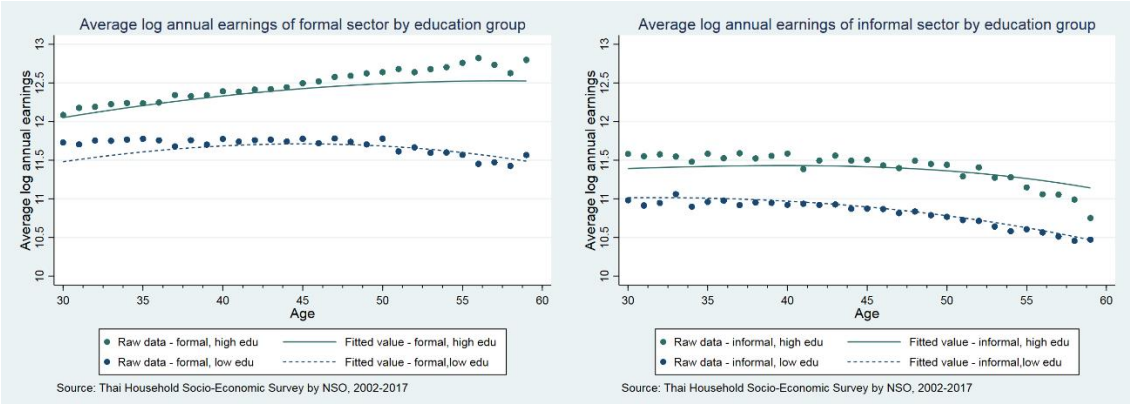


Figure 2.5 Logarithm of the Annual Earnings of Workers by Sector and Education, Regression

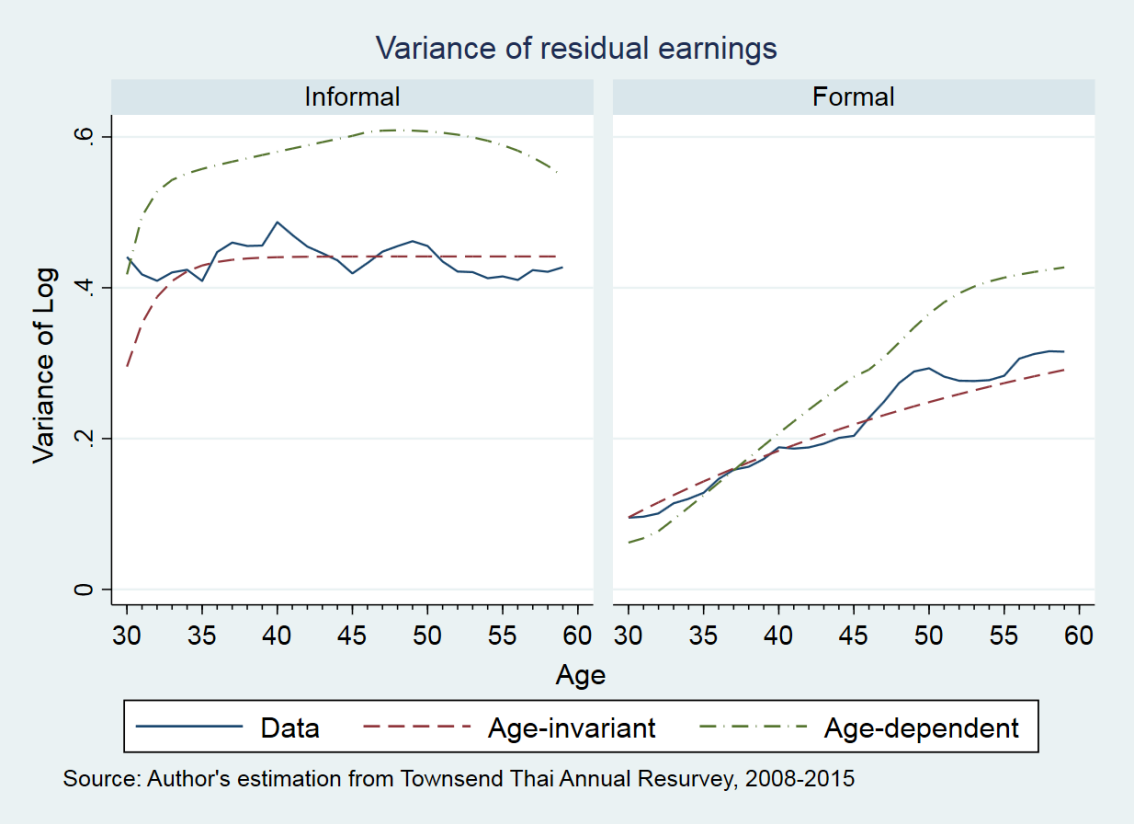


Figure 2.6 Variance of Logarithm of the Residual Earnings by Sector

Appendix

Sector Switching Transition Matrix by Age Group

Table A2.1 Transition Matrix by Age Group

t/t+1	Share (%)	
	Informal	**Formal**
Informal		
30-39	93.34	6.66
40-49	96.44	3.56
50-59	97.85	2.15
All	96.38	3.62
Formal		
30-39	10.46	89.54
40-49	12.35	87.65
50-59	9.76	90.24
All	10.80	89.20

Source: Townsend's panel data from 2008 to 2015

Robustness check

In this section, we try to estimate the labor earnings by considering the alternative setting of education. In the original setting, the education dummy equals one if an individual has an education level of high school or above. However, this 'high education' group in the formal sector could include a greater fraction of college graduates than in the informal sector. Therefore, we redefine education dummy (edu) to edu1 and edu2 where edu1 equals to one if the individual has an education level of high school or vocational and zero otherwise, and edu2 equals to one if the individual has an education level of college and above and zero otherwise.

The results of the estimate are presented in Table A2.2. The coefficient on edu1 is higher than that of edu2, indicating that the level of earnings of high school or vocational graduates is considerably higher than college graduates. However, the coefficient on the interaction terms edu2*Formal is significantly higher than edu1*Formal, meaning that the earnings of workers with a college degree are significantly higher than high school and vocational graduates if they are employed in the formal sector. Our results can be implied that education levels have less impact on earnings than sector of employment.

In terms of average earnings growth over 30 years of working, the average growth rates of workers in the formal sector are higher than workers in the informal sector at all education levels. The average growth rates for high school dropouts, high school and vocational graduates, and college graduates are 3.3%, 3.9%, and 3.9% in the informal sector, and 5.1%, 4%, and 3.9% in formal sector, respectively.

Moreover, we also run separate regressions by education levels (less than high school, high school and vocational, and college and above) according to the following regression equation:

$$\ln(E_{i,t}^{e}) = \beta_0^e + \beta_1^e h_{i,t}^e + \beta_2^e (h_{i,t}^e)^2 + \beta_3^e s_{i,t}^e + \beta_4^e s_{i,t}^e h_{i,t}^e + \beta_5^e M_t^e + X_{i,t}^e \gamma^e + u_{i,t}^e \quad (2.9)$$

The results are shown in Table A2.3. The coefficients on sector dummy are significantly positive for all educational levels and are higher for higher education levels. This result is consistent with our previous result that workers with higher education level who are employed in the formal sector have higher earnings than in the informal sector.

In terms of average earnings growth over 30 years of working, formal workers who are college graduates have the highest average earnings growth, 5.6%, whereas the average earnings growth of formal workers who are high school dropouts and high school or vocational graduates are 2.6% and 2.7%, respectively. For workers in the informal sector, their average earnings growth is much smaller than their counterparts in the formal sector. The average earnings growths for high school dropouts, high school or vocational graduates, and college graduates are 1.4%, 0.8%, and 0.6%, respectively.

Table A2.2 Estimation Results of Age-earnings Profile (Redefined Education Dummy)

Variable	**ln Earnings**	
Age	0.0346***	(0.003)
Age Squared	-0.0009***	(0.000)
Formal sector dummy	0.1877***	(0.015)
Formal*Age	0.0193***	(0.001)
High school and vocational dummy (edu1)	0.3850***	(0.019)
High school and vocational dummy*Age	0.0066***	(0.001)
High school and vocational dummy *Formal	-0.0466**	(0.016)
College and above dummy (edu2)	0.0748**	(0.028)
College and above dummy*Age	0.0100***	(0.001)
College and above dummy *Formal	0.6818***	(0.025)
Cohort	0.0261***	(0.002)
Region		
Central	0.0000	(.)
Bangkok Metropolis	0.3752***	(0.009)
North	-0.5092***	(0.009)
Northeast	-0.7825***	(0.010)
South	-0.0565***	(0.009)
Public Sector	0.4307***	(0.011)
Unemployment rate	-0.0881***	(0.022)
Intercept	10.5032***	(0.091)
N	147619	

Source: Author's estimation is based on the HSES data from 2002 to 2017, NSO.
Note**:** Standard errors are reported in parentheses. *, **, *** indicate significance at the 90%, 95%, and 99% levels, respectively. Intial age and cohort are normalized to zero.

Table A2.3 Estimation Results of Age-earnings Profile by Education Level

Variable	**Less than High school**		**High school and vocational**		**College and above**	
Age	0.0402***	(0.003)	0.0260***	(0.005)	0.0319***	(0.007)
Age Squared	-0.0009***	(0.000)	-0.0006***	(0.000)	-0.0009***	(0.000)
Formal sector dummy	0.2399***	(0.017)	0.2624***	(0.026)	0.3781***	(0.041)
Formal*Age	0.0129***	(0.001)	0.0191***	(0.002)	0.0519***	(0.003)
Cohort	0.0300***	(0.002)	0.0152***	(0.003)	0.0350***	(0.005)
Region						
Central	0.0000	(.)	0.0000	(.)	0.0000	(.)
Bangkok Metropolis	0.3899***	(0.009)	0.2267***	(0.017)	0.3373***	(0.024)
North	-0.5089***	(0.009)	-0.3949***	(0.021)	-0.6832***	(0.027)
Northeast	-0.7718***	(0.010)	-0.5415***	(0.022)	-0.2550***	(0.028)
South	-0.0497***	(0.009)	-0.1590***	(0.018)	-0.1152***	(0.025)
Public Sector	0.4964***	(0.011)	0.2082***	(0.017)	0.2990***	(0.024)
Unemployment rate	-0.0827***	(0.022)	0.0369	(0.040)	0.2468***	(0.052)
Intercept	10.4811***	(0.092)	11.1320***	(0.164)	10.2507***	(0.222)
N	99873		26595		21151	

Source: Author's estimation is based on the HSES data from 2002 to 2017, NSO.
Note**:** Standard errors are reported in parentheses. *, **, *** indicate significance at the 90%, 95%, and 99% levels, respectively. Intial age and cohort are normalized to zero.

CHAPTER 3

Uncertainties and Inequality in Developing Economies: A Structural Analysis and Policy Implications

3.1 Introduction

Understanding the dynamic of economic inequality and its determinants is important for various research studies in economics, such as education policy and economic growth. In the literature, inequality has been associated with income, wealth, consumption, and a variety of individual-specific characteristics such as educational level. However, most of the existing work is concentrated in developed countries, and the studies in developing countries are still very limited.

The fact from literature shows that inequality increases as individuals age. Deaton and Paxson (1994) investigate the consumption and income inequality within the same cohort across three different countries, namely, the US, UK, and Taiwan. They find that the within-cohort inequality increases with age. More recent studies show that the increase in inequality is due to persistent idiosyncratic earnings shocks. For example, Storesletten, Telmer, and Yaron (2004) examines why earnings and consumption inequality increase with age and find that individual-specific earnings shocks can explain the increasing pattern. Blundell, Pistaferri, and Preston (2008) propose that changes in the persistence of income shocks can explain income inequality in the US.

Therefore, this chapter aims to fill the gap in the literature by focusing on how inequality in consumption and earnings changes with age in developing countries and investigate if the change can be explained by the existence of uninsurable idiosyncratic earnings shocks. We use Thailand as a representative for calibration and quantitative analysis. Thailand is a large economy, with around 55% of workers employed informally. The social security system already exists, but its coverage is limited to workers in the formal sector.

For this purpose, we first develop a life cycle model with heterogeneous agents facing mortality, earnings, and employment uncertainties. The key innovation in our model is that there are two sectors of employment in the economy, that is, the formal and informal sectors. During working age, workers in different sectors face different earnings shocks and can move between the two sectors. After retirement, all agents receive an age-specific old-age allowance from the government. Our results show that the model can capture the increasing pattern of earnings and consumption inequality, but the level of consumption inequality is approximately 42% too high, relative to the data.

Storesletten et al. (2004) find that the divergence of consumption inequality between theory and data is due to the availability of insurance arrangements to actual agents that are not present in the model. By incorporating social security, the model can reduce consumption inequality and provide a better account of the data. Blundell, Graber, and Mogstad (2015) study income dynamics and the impact of insurance mechanisms to labor income shocks in Norway. They suggest that the tax-transfer system can reduce both permanent and transitory income shocks encountered by individuals. De Nardi, Fella, and Paz-Pardo (2019) study earnings dynamics and examine the extent to which earnings shocks can be self-insured and their implications for social insurance.

Following the literature, we extend our model to incorporate two mechanisms of insurance arrangement to earnings shocks, namely, old-age employment and social security system. In developing countries, older people are more vulnerable to poverty. Without social insurance, they have to depend on their labor earnings for daily life. Furthermore, the social security system is a formal insurance arrangement for old age, in the form of pensions. Therefore, we include these two features to our model. In an old-age employment setting, we allow elderly agents to work after their retirement age within the same sector but face retirement risk at every age. Further, we introduce a social security into the old-age employment model according to which only formal workers can receive a pension after retirement. Our finding shows that incorporating these two mechanisms can reduce consumption inequality. The model with social security provides a good account of the data and we use this extended model as a benchmark for policy experiments.

We further examine the effect of different social insurance programs on inequality as well as consumption and saving behaviors. We use the model with old-age employment and social security to conduct two policy experiments. First, we investigate the effect of a social welfare system by varying the amount of subsidy provided by the government to guarantee the minimum consumption floor. We find that a high level of social welfare cannot reduce consumption inequality significantly, except those very young and very old. Second, we extend the social security system to cover the informal sector. In this experiment, all agents are qualified for a pension after the retirement age. Our results show that the extension of pension can reduce consumption inequality significantly during old age.

The remaining contents of this chapter include: Section 2 describes the construction of earnings and consumption inequality profiles; Section 3 outlines and parameterizes the basic model; Section 4 analyzes the extension of the model; Section 5 discusses policy experiments, and Section 6 provides the conclusion.

3.2 Inequality over the Life Cycle

We use the data from the Household Socio-Economic Survey (HSES) to estimate earnings and consumption inequality profiles.[13] We define consumption by using the household average monthly consumption expenditure, which includes expenditures on nondurable goods (food, clothes, and utilities) as well as durable goods (housing and vehicles). Both consumption and earnings are converted into real values using consumer price indices (CPI) of the year 2015 published by the Thai Bureau of Trade and Economics.

To construct earnings and consumption inequality profiles, we compute the variances of the logarithm of earnings and consumption, which are decomposed into two parts. The first part is the variance from the difference in observable individual characteristics that is obtained from a regression on age, sector of employment, and educational level that rule out cohort and location effects. The second part is the variance from shocks that cannot be explained by the aforementioned observable characteristics, which is the residual from the regression.

Figure 3.1 reports the cross-sectional variance of the logarithm of earnings and consumption by age from the HSES data. There are two important features of these

[13] We use the earnings data that we constructed in Chapter 2.

findings. First, both earnings and consumption inequality increase over the working period. Second, earnings inequality increases faster with age than consumption inequality. Figure 3.2 exhibits the inequality profiles of the formal and informal sectors. The lifetime profiles of inequality are different between the formal and informal sector. Although earnings and consumption inequality in both sectors increase with age, the increase is faster in the formal sector.

3.3 Basic Model

We develop a stochastic life cycle model of workers in developing countries with the existence of a large informal sector to study the impact of earnings shocks on consumption and saving behaviors. Each individual lives for a finite period and makes optimal decisions on consumption (c) and savings (a'). In every period, individuals face employment uncertainties that determine their sector of employment and earnings. In addition, they have to face life-expectancy shocks, which define the length of their lives.

3.3.1 Demographics

The economy is populated by ex-ante heterogeneous individuals in terms of education who live up to a maximum age of *H*. Agents engage in labor market activity before they reach age $h = H^r$, where H^r is retirement age. An individual with age h can survive to the next age $h + 1$ with a probability ψ_h with $\psi_H = 0$.

3.3.2 Employment and Income Uncertainties

Individuals enter the labor market at age 25 in a specific employment sector, either formal or informal sector, and retire at age 60. At every stage along the life cycle, they face employment and income uncertainties.

At the beginning of every period, agents receive an employment shock that determines their employment status, $s \in \{1, 2, 3\}$. An agent with s = 1 or 2 can work in the formal sector or the informal sector, whereas an agent with s = 3 does not work. Agents are allowed to change their employment status in the next period. The probability of being employed in the next period depends on age, educational level, and current employment status. The level of educational attainment, which can be either low or high, $edu \in \{l, h\}$, is exogenously determined before the start of the economy and remains throughout the life cycle. Thus, the age-dependent transition probability of employment is given by the 3 × 3 matrix $P_s^{h,edu}(s'|s)$.

Before retirement, individual labor earnings are determined using the following process:

$$ln\,(E_{i,t}) = f(h, s, edu) + \alpha_i + z_{i,h}^{s} + \varepsilon_{i,h}^{s}, \tag{3.1}$$

$$f(h, s, edu) = \beta_0 + \beta_1 h_{i,h} + \beta_2 h_{i,h}^2 + \beta_3 s_{i,h} + \beta_4 s_{i,h} h_{i,h} + \beta_5 edu_i + \beta_6 edu_i h_{i,h} + \beta_7 edu_i s_{i,h}, \tag{3.2}$$

where $f(h, s, edu)$ is the average age-earnings profile conditioning on age (h), sector of employment (s), and educational attainment (edu). The coefficients $\beta_0 - \beta_7$ are the estimates in the previous chapter. $\alpha_i \sim N(0, \sigma_\alpha^2)$ is the fixed effect that is determined at

birth. $\varepsilon_{i,h}^{s} \sim N(0, \sigma_{\varepsilon,s}^{2})$ is the transitory shock that is received every period. $z_{i,h}^{s}$ is the persistent shock that is also received every period and follows an AR(1) process:

$$z_{i,h}^{s} = \rho_s z_{i,h-1}^{s} + \eta_{i,h}^{s}; \qquad \eta_{i,h}^{s} \sim N(0, \sigma_{\eta,s}^{2}), z_{i,0} = 0. \tag{3.3}$$

Workers who remain in the same sector receive earnings shocks from the AR(1) process. For those who move to a different sector, we assume that they have zero persistence. Thus, their earnings shocks do not follow the AR(1) process. After retirement, all agents receive an old-age transfer from the government according to their age and face no risk.

3.3.3 Preferences

Agents make a decision on consumption and savings to maximize their expected utility. The lifetime utility can be expressed as follows:

$$E \sum_{h=1}^{H} \beta^{h-1} \left(\prod_{t=1}^{j-1} \psi_t \right) U(c_{i,h}), \tag{3.4}$$

where β is the discount rate and $c_{i,h}$ represent an individual's consumption at age h. We assume the utility function is a constant relative risk aversion (CRRA) utility function:

$$U(c_{i,h}) = \frac{c^{(1-\mu)}}{1-\mu}. \tag{3.5}$$

3.3.4 Individual's Problem

Let V_h denote the value function of an agent at age h. The state of each individual in age h is given by $x = (a, \alpha, z, \varepsilon, s, edu, h)$. The agent's recursive problem can be presented as follows:

$$V(x) = \max_{c,\, a'}\{U(c_{i,h}) + \beta\psi_h EV(x')\} \tag{3.6}$$

subject to the following constraints:

$$c + a' = (1 + r)a + E + tr_{h \geq H^r} + T, \tag{3.7}$$

$$E = \begin{cases} e^{f(h,\, s, edu) + \alpha + z + \varepsilon} & if\ h < H^r \\ 0 & if\ h \geq H^r \end{cases}, \tag{3.8}$$

$$z = \rho_s z_{-1} + \eta; \qquad z_0 = 0, \tag{3.9}$$

$$\alpha \sim N(0, \sigma_\alpha^2),\ \ \varepsilon \sim N(0, \sigma_{\varepsilon,s}^2),\ \eta \sim N(0, \sigma_{\eta,s}^2), \tag{3.10}$$

$$T = \max\{0,\ \underline{c} - [(1 + r)a + E + tr_{h \geq H^r}]\}, \tag{3.11}$$

$$a \geq 0; c \geq 0, \tag{3.12}$$

where $EV(x')$ is the expected value from the next period with respect to employment and earnings uncertainties, $a_{i,h}$ denotes asset holdings, $f(h, s, edu)$ is the earnings profile, α is the fixed effect, z is the persistent shock, ε is the transitory shock, $tr_{h \geq H^R}$ is the age-specific old-age transfer for retired agents, $\underline{c}$ is the minimum consumption level, and T is the transfer to ensure a minimum consumption.

3.3.5 Data, Estimation, and Calibration

Table 3.1 summarizes all parameters used in this section. The model period is equivalent to one year of calendar time. Agents enter the economy at age 25, retire at age 60, and are dead at age 80. The survival probabilities ψ_h are taken from the United Nations life tables. The rate-of-return on capital (r) is set to 1.58% according to the weighted average deposit interest rate of commercial banks from 2000–2017.[14] The social welfare system is

[14] The rate-of-return on capital is the average of the weighted average deposit interest rate of commercial banks from 2000–2017. The interest rate data is retrieved from: https://www.bot.or.th/App/BTWS_STAT/statistics/BOTWEBSTAT.aspx?reportID=208&language=th

available to ensure the minimum consumption level that is denoted by $\underline{c}$. We set the amount of $\underline{c}$ to equal 5% of average earnings.[15]

The risk aversion coefficient, μ, and discount factor, β, are chosen at the same time to match the wealth-to-income ratio (aggregate and by education level). The wealth-to-income ratio is calculated from the HSES data. We select only male household heads aged 30–59 who are working and have a positive income and exclude the higher one percent wealth quantile.[16] The aggregate wealth-to-income ratio is 3.64. We also match the wealth-to-income ratio by education levels, which are 3.10 for high education and 4.19 for low education. We constrain the range of μ to (1–9) and β to (0–1), and set $\mu = 3$ and $\beta = 0.960$, which is within the range commonly used in the literature.

The transition probabilities of employment status are determined by using the panel data from Townsend's household survey (2008–2015). We select only male household members aged 30–59.[17] The sample is categorized into three age groups, that is, 30–40, 40–50, and 50–59. Further, we estimate the probability of continuation or changes in the employment status in the next year based on the educational level within that age group. The transition probabilities of employment status are presented in Table 3.2.[18]

[15] The amount of per capita subsidy from the social welfare programs provided by the Thai government is around 12,000 Thai baht per year, which is equivalent to 5.5% of the GDP per capita in the year 2017. Therefore, we use 5% of the average earnings in our model that is roughly equivalent to the real economy. (Details of the social welfare in Thailand are in appendix)

[16] Here, we use the HSES data from the years 2013, 2015, and 2017. This is because they report the actual value of financial assets. Wealth is defined as the total net worth that is calculated from the total value of household assets (including house, land, and building) minus total liability. Income is the current average monthly household income.

[17] We do not restrict the study only to male household heads. This is to prevent obtaining insufficient samples for each sector of employment in each age group.

[18] For ages 25–29 in our simulation, we use the transition of ages 30–40.

In our model, earnings are exogenous. During working age, the earnings of workers consist of a deterministic part that is the same for everyone at the same age and characteristics, and an idiosyncratic component that captures an individual's earnings risk. The deterministic component of earnings for both formal and informal workers is estimated using the equation 3.2.

For the idiosyncratic shock components of earnings, we calibrate the shock parameters to match consumption inequality from the HSES data. We set $\sigma_{\alpha}^{2} = 0.115$ for both sectors, $\rho = 0.970$, $\sigma_{\varepsilon}^{2} = 0.195$ and $\sigma_{\eta}^{2} = 0.023$ for the formal sector, and $\rho = 0.825$, $\sigma_{\varepsilon}^{2} = 0.240$ and $\sigma_{\eta}^{2} = 0.190$ for the informal sector. As we allow for the transition between sectors to take place, agents who change the sector will have zero persistence $(\rho = 0)$ while σ_{ε}^{2} and σ_{η}^{2} are set according to the new sector. All three components of earnings shocks in each sector are discretized using five grid points. Then, the transition probabilities are chosen following the procedure described in Tauchen (1986).

After retirement, everyone receives an old-age transfer from the government. There is an old-age allowance system in Thailand to provide monthly cash transfers to the elderly, aged 60 and above. The amount transferred depends on the age of recipients. In this system, 60–69-year-olds receive 600 Thai baht, 70–79-year-olds receive 700 Thai baht, 80–89-year-olds receive 800 Thai baht, and those who are 90 years of age and older receive 1000 Thai baht (Suwanrada, 2013).[20] Thus, $tr_{h \geq H^{R}}$ is set accordingly.

[20] Exchange rate = 31.33 Baht/US Dollar as of February 2020.

3.3.6 Results

We begin with an economy that we call the "basic economy" in which all agents must retire after they reach the retirement age. We match the basic model to the aggregate wealth-to-income ratio and consumption inequality profile. The model outcomes and data are presented in Table 3.3. Figures 3.3, 3.4, and 3.5 display the results from the basic model, namely, the age-profile of earnings and consumption inequality. The model is successful at capturing the main feature of the data that both earnings and consumption inequality increases with age.

Earnings inequality from the HSES data in Figure 3.3 grows slightly faster than in our model. Quantitatively, our model's earnings inequality is 20% higher at age 30 and ends up being 26% lower at retirement. Our model cannot perfectly match the earnings inequality from the data because we calibrate earnings shocks to match consumption inequality, which is our main target. However, the model can exhibit an increasing pattern of earnings inequality profile.

Consumption inequality in Figure 3.4 is consistent with the data until age 50 and starts to grow faster as the age increases, being 17.7% higher at retirement. Figure 3.5 illustrates consumption inequality profiles by sector. Consumption inequality diverge from the data after age 40 for the formal sector and around age 50 for the informal sector. A possible reason for the divergence between the basic model and data can be the existence of some insurance arrangements to actual agents that are not available in the model. Based on this assumption, we extend our model in the next section to include some features available in the Thai economy to prove whether they can improve the performance of our model in terms of explaining consumption inequality.

3.4 Extension of the Model

In the basic model, there is an extensive increase in consumption inequality starting from age 50. Storesletten et al. (2004) point out the absence of risk-sharing features such as the choice of labor hour, retirement decision, and social security can explain the deviation of the model from the data. Thus, we extend our model to allow for old-age employment and further introduce the social security system.

3.4.1 Old-Age Employment

For developing countries, old-age employment is one of the main income sources for older people. For example, in Thailand, 31% of the elderly report income from work as their main source of income.[21] Thus, we now allow old agents to continue to work until they face a retirement shock.

After reaching the retirement age, all agents face a retirement shock every period. If they are hit by the shock, they must retire. If not, they can continue to work in the same sector of employment.[22] For elderly workers, their earnings depend on their age, educational level, and sector of employment. Their earnings profile is estimated using the HSES data by selecting male heads of the household aged 60–80, who are working and have positive earnings. Table 3.4 reports the estimation results and we use these coefficients to determine the earnings for the elderly in our model. The budget constraints for the elderly are then as follows:

[21] The three main sources of income for elderly persons are transfers from child(ren) (34.7%), working (31%), and old-age allowance (20%) (NSO, 2017).
[22] The data shows that the transition between sectors after the age 60 is 12% for formal workers and only one percent for informal workers (Townsend, 2000–2017).

$$c + a' = (1 + r)a + I_r \cdot E + tr_{h \geq H^r} + T, \quad (3.13)$$

$$ln(E) = f(h, s, edu) \quad if\ h \geq H^r, \quad (3.14)$$

$$T = \max\{0, \underline{c} - [(1 + r)a_{i,h} + I_r \cdot E + tr_{h \geq H^r}]\}, \quad (3.15)$$

$$a \geq 0; c \geq 0. \quad (3.16)$$

Here, $I_r = 1$ if an agent continues to work and 0 otherwise. $f(h, s, edu)$ is an old-age earnings profile. The model is recalibrated, where $\mu = 3$ and $\beta = 0.987$.

Figures 3.6 and 3.7 present the results for the old-age employment model. The consumption dispersion after age 50 decreases and becomes close to the data. Old-age employment appears to provide a better explanation of the data. Working during old age appears to serve as an insurance arrangement after the retirement age. This is intuitive. In our basic model, the insurance arrangement for old age is very small. When agents are hit by bad earnings shocks, they must sacrifice their consumption more to guarantee that they will have enough savings for old age. As a result, it leads to an increase in consumption inequality. Similarly, if the insurance is available, unlucky agents can save less than in the basic model to ensure their consumption at old age. Therefore, the increase in inequality becomes smaller.

3.4.2 Social Security

We incorporate the social security into the old-age employment model and examine how it affects consumption inequality. The conditions for social security are based on the Thai social security system. Only agents who work in the formal sector for at least 15 years can receive a pension when they reach the retirement age. The pension benefit is calculated based on the old-age pension of the Thai social security system that pays 20%

of the average earnings before retirement, with maximum monthly earnings used to calculate at 15,000 Thai baht. The new budget constraints for old agents are then as follows:

$$c + a' = (1 + r)a + I_r \cdot E + tr_{h \geq H^r} + I_s \cdot p + T, \quad (3.17)$$

$$ln(E) = f(h, s, edu) \quad if\ h \geq H^r, \quad (3.18)$$

$$T = max\{0, \underline{c} - [(1 + r)a + I_r \cdot E + tr_{h \geq H^r} + I_s \cdot p]\}, \quad (3.19)$$

$$p = \min(\hat{p}, ss \cdot E_{H^r - 1}), \quad (3.20)$$

$$a \geq 0; c \geq 0, \quad (3.21)$$

where p is the amount of a pension received after retirement age.[23] $I_s = 1$ if the sector of employment at age $H^r - 1$ is the formal sector. *ss* is the replacement rate and $\hat{p}$ is the maximum amount of pension each individual can receive. We recalibrate the model where $\mu = 4$ and $\beta = 0.952$.

Figures 3.6 and 3.7 also plot the consumption inequality profile from the model with the social security. The pattern of the consumption inequality profile is closer to the data at all ages compared to the basic model and the old-age employment model. Figure 3.8 exhibits the average asset holdings and consumption profiles between the formal and informal sectors. On average, formal workers have higher assets and consumption levels at every age than informal workers.

In general, the old-age employment and social security can explain the unobserved risk-sharing arrangement that is absent in the basic model. We find that consumption inequality can be reduced from the basic model by 27.6% in the old-age employment

[23] We simplify our calculation for pension by using the earnings at age $H^r - 1$ instead of average monthly wage in the last 60 months before retirement, according to the Thai social security system. This is because the transition probability between sectors is lower after the age 50 (see Table 3.2).

model and 29% in the model with a combination of old-age employment and social security. Moreover, the extended model with social security can match consumption inequality better than the other two models. Therefore, we will use this model as a benchmark economy to conduct policy experiments in the next section.

3.5 Policy Experiments

In this section, we investigate the impacts of the social insurance policy on consumption and saving behaviors, as well as inequality. We consider two policy reforms, namely, the extension of the social welfare system and extension of the pension scheme. The extended model with the social security is used as a benchmark economy for all policy experiments.

3.5.1 Extension of Social Welfare

The social welfare system provides benefits or support to the people who are in need. Many countries have developed their own systems to accommodate the needs of their people that usually consist of numerous programs with different requirements. Here, we use the consumption floor ($\underline{c}$) that the government guarantees as a measure. Generally, developed countries provide good social welfare to their people. Hubbard et al. (1995) calculate the size of social welfare in the US by using the guaranteed consumption floor and suggest the amount of subsidy was approximately 15% of the GDP per capita in 1984. For developing economies, the size of the social welfare can be much smaller than that in developed countries. Thus, we conduct experiments to examine the effect of social welfare on consumption and saving behaviors.

For our analysis, we examine the impact of social welfare by varying the level of minimum consumption floor (i.e., values of $\underline{c}$). In the benchmark model, we set $\underline{c}$ to five percent of the average earnings. We conduct two policy experiments. The first experiment is on the economy that has a very small sized social welfare program with $\underline{c} = 1\%$ of average earnings, which we will refer to as "low social welfare" hereafter. The second experiment is on the economy that has a relatively large sized social welfare program with $\underline{c} = 15\%$ of the average earnings, which we will refer to as "high social welfare" hereafter.

Figure 3.9 exhibits the results of our policy experiments. We find that a larger size of the social welfare program does not have a significant impact on reducing consumption inequality, except those very young (age 30-40) and very old (age 70-80). At young ages, there are more people who affect from minimum consumption floor since their income level is still low and they do not have lots of asset holdings. For middle-aged people, it does not have an impact on reducing inequality because the minimum consumption floor affects only those who face bad earnings shocks. The high level of minimum consumption floor discourages savings of the poor people but does not affect high-income people. As a result, low-income people have fewer asset holdings after retirement age than high-income people, which increases inequality after retirement.

Table 3.5 reports the average non-negative saving rate by young agents from each experiment. The saving rate increases from 33.9% in the benchmark to 35.6% in the low social welfare economy. On the other hand, the saving rate decreases to 30.9% in the high social welfare economy. These results indicate the saving rate among the young can be reduced when the value of the minimum consumption floor is higher.

3.5.2 Pension Extension to Informal Workers

In developing countries, the development of the social security system is not comprehensive. Typically, informal workers are not eligible for the pension in old age. For this policy experiment, we aim to investigate the impact of extending the pension to cover informal workers.

In the benchmark economy, only formal workers are qualified to receive a pension after retirement. In our analysis, we remove this restriction and allow all agents to be eligible to receive a post-retirement pension. The pension is calculated using the same method as that used for formal workers.

Figure 3.10 presents the consumption inequality from the model with pension extension compared to the benchmark model. For old age, consumption inequality is reduced because everyone is eligible to receive a pension. Moreover, it also reduces inequality for young people. This is because individuals have to save when they are in the working-age for their retirement. If they are hit by bad earnings shocks, their earnings become lower, but they have to maintain the level of savings. As a result, they have to sacrifice their consumption for savings, which increases inequality. However, if a pension is available to everyone, they rely less on savings. Although they are hit by bad earnings shocks, they do not have to lower their consumption to maintain their savings for retirement. Thus, inequality is reduced.

Table 3.6 reports the average non-negative saving rate by young agents from our experiment. The saving rate decreases from 33.9% in the benchmark to 33.3% in the model with social security system.

3.6 Conclusion

This study investigates if the increase in earnings and consumption inequality with age can be explained by the earnings shocks that an individual faces over the life cycle. By developing a life cycle model with uninsured earnings shocks, we find that key factors to explain consumption inequality are the idiosyncratic earnings shocks, working during old age, and social security.

We investigate the impact of earnings shocks on consumption and saving behaviors by developing a stochastic life cycle model in developing countries with the existence of the informal sector. Agents face life expectancy, employment, and earnings shocks. They are allowed to change their employment sector throughout their working life and must retire when they reach age 60. Our model can capture the increasing pattern of earnings inequality. In terms of consumption inequality, it is 17% higher at retirement than the data. Thus, we extend our model to allow for old-age employment and introduce the social security system. We find that consumption inequality is reduced by 27.6% and 29% relative to the benchmark model in the old-age employment and social security models, respectively. The consumption inequality from the model with the social security system is well-matched with the data.

We use the model with a combination of old-age employment and social security to examine consumption and savings behaviors. We examine two policy reforms, namely, the extension of social welfare and extension of the social security system.

First, we discuss the impact of the social welfare system by varying the amount of the minimum consumption floor guaranteed by the government. Our finding indicates that a high level of social welfare can reduce consumption inequality only for very young

workers and very old persons. We find that consumption inequality is lowered than the benchmark model by around 19.8% at age 30 and 9.6% at age 80, and the saving rate decreases by four percentage points, relative to a low social welfare level.

Second, we extend the social security to cover the informal sector such that all workers can receive a pension after the retirement age. We find that the extension of social security to cover the informal sector can reduce consumption inequality significantly.

3.A Tables

Table 3.1 Parameters

Parameters	Interpretation	Value	Source/Target
h	Model period	1	Equivalent to one year and start at age 25
H^r	Retirement age	36	Equivalent to age 60
H	Maximum years after labor force entry	56	Equivalent to age 80
β	Discount factor	0.960	Match wealth-to-income ratio
μ	Risk aversion	3	Match wealth-to-income ratio
ψ_h	Survival probability at age h		(1-probability of dying between ages h and h + n) *
$P_s^{h,\,edu}(s'\|s)$	Transition probability of employment		Estimated from Townsend's panel data
r	Rate-of-return on capital	0.0158	Weighted average deposit interest rate
σ_α^2	Fixed effect	0.115	Calibration
$\{\rho_f, \rho_{inf}\}$	Persistent parameter	{0.970, 0.825}	
$\{\sigma_{\varepsilon,f}^2, \sigma_{\varepsilon,inf}^2\}$	Transitory shock	{0.195, 0.240}	
$\{\sigma_{\eta,f}^2,\ \sigma_{\eta,inf}^2\}$	Persistent shock	{0.023, 0.190}	

Note: * Life Tables by Country: Thailand, WHO.

Table 3.2 Transition Probabilities of Employment Status

Age Group	*t/t + 1*	High Education			Low Education		
		Formal	Informal	Not working	Formal	Informal	Not working
30–40	Formal	0.911	0.082	0.007	0.743	0.243	0.014
	Informal	0.083	0.904	0.013	0.036	0.954	0.010
	Not working	0.160	0.480	0.360	0.000	0.403	0.597
40–50	Formal	0.922	0.074	0.004	0.739	0.259	0.003
	Informal	0.053	0.939	0.008	0.023	0.970	0.007
	Not working	0.091	0.455	0.455	0.029	0.304	0.667
50–59	Formal	0.928	0.064	0.008	0.795	0.187	0.018
	Informal	0.043	0.944	0.013	0.018	0.973	0.010
	Not working	0.000	0.417	0.583	0.043	0.414	0.543

Source: Author's estimation based on Townsend's panel data in 2008–2015.

Table 3.3 Model Outcomes

Target variable	Data	Basic	Old-age employment	Social security
Wealth-to-income ratio				
Aggregate	3.64	3.64	3.64	3.64
High education	3.10	3.25	3.30	3.11
Low education	4.20	4.00	3.90	4.11

Table 3.4 Estimation Results of the Age–Earnings Profile for Old-Age Employment

Variable	**ln Earnings**
Age (h_{it})	-0.0336
	(0.029)
Age squared (h_{it}^2)	0.0002
	(0.000)
Formal sector dummy (s_i)	0.6019**
	(0.216)
Formal*Age ($s_i * h_{it}$)	-0.0103
	(0.006)
High education dummy (edu_i)	0.4711*
	(0.239)
High education*Age ($edu_i * h_{it}$)	-0.0127
	(0.007)
High education*Formal ($edu_i s_i$)	0.5105***
	(0.072)
Constant	11.9970***
	(0.576)
R-squared	0.265
N	30761

Source: Author's estimation based on SES data in 2002–2017, NSO.
Note: Standard errors are reported in parentheses. *, **, *** indicate significance at the 90%, 95%, and 99% levels, respectively.

Table 3.5 Saving Rates of Young Agents, Extension of Social Welfare

	1% $\bar{y}$	5% $\bar{y}$	15% $\bar{y}$
Saving rate			
Aggregate	35.6%	33.9%	30.9%
Formal	30.0%	28.4%	25.9%
Informal	37.7%	36.0%	32.8%

Note: $\bar{y}$ is average earnings.

Table 3.6 Saving Rates of Young Agents, Extension of Social Security

	Benchmark	Pension extension
Saving rate		
Aggregate	33.9%	33.3%
Formal	28.4%	27.9%
Informal	36.0%	35.4%

3.B Figures

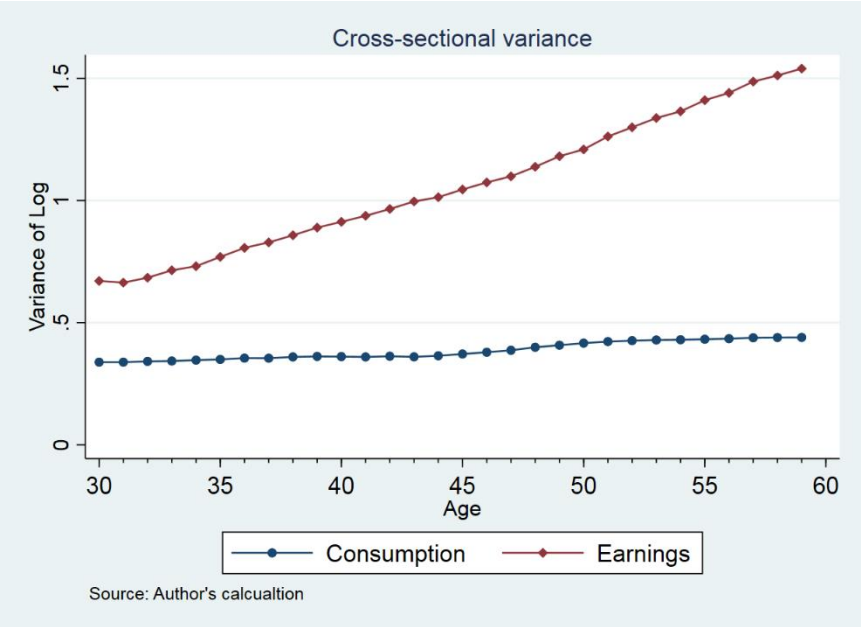


Figure 3.1 Variance of the Logarithm of the Earnings and Consumption.
Source: HSES 2002-2017, NSO

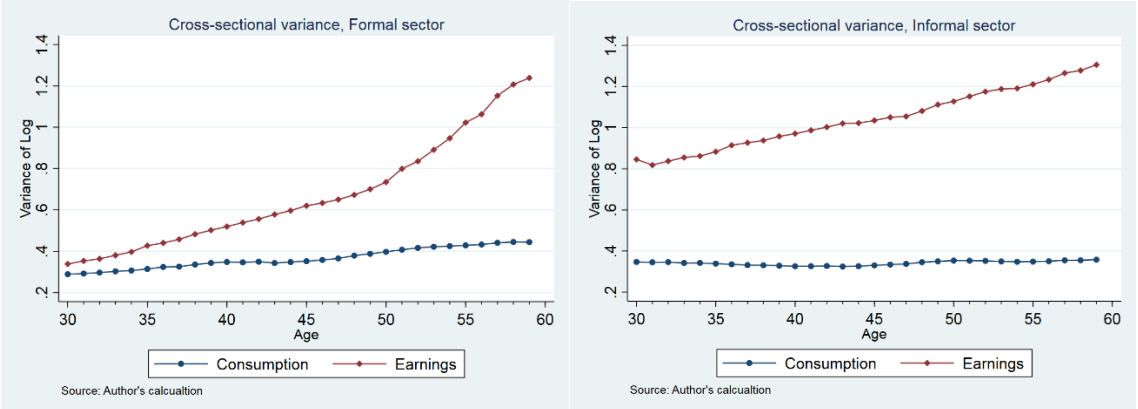


Figure 3.2 Variance of the Logarithm of the Earnings and Consumption by Sector.
Source: HSES 2002-2017, NSO

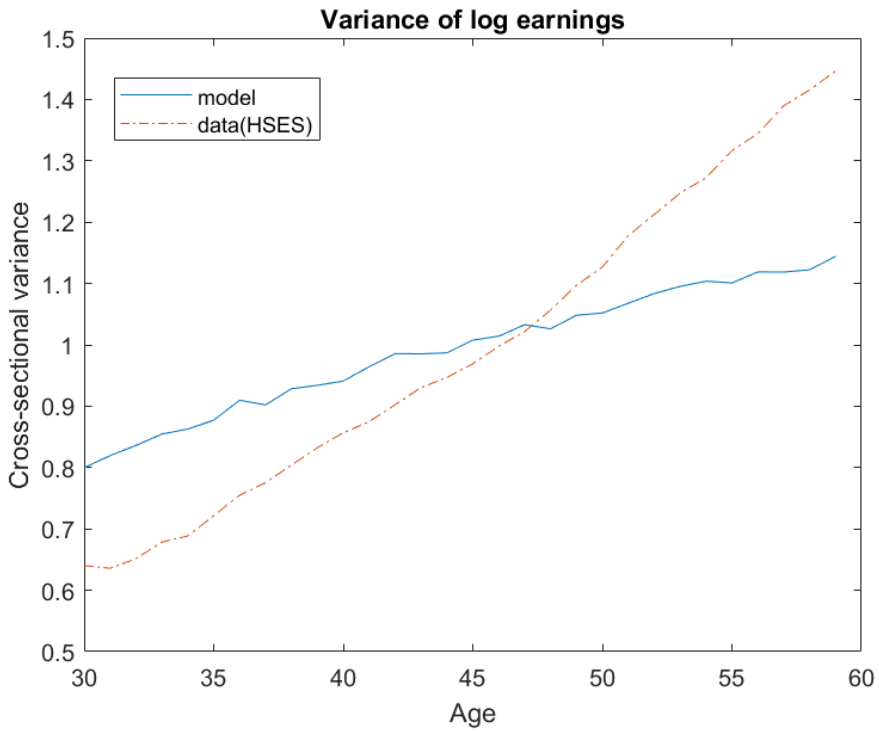


Figure 3.3 Age-profile of Earnings Inequality, Basic model.

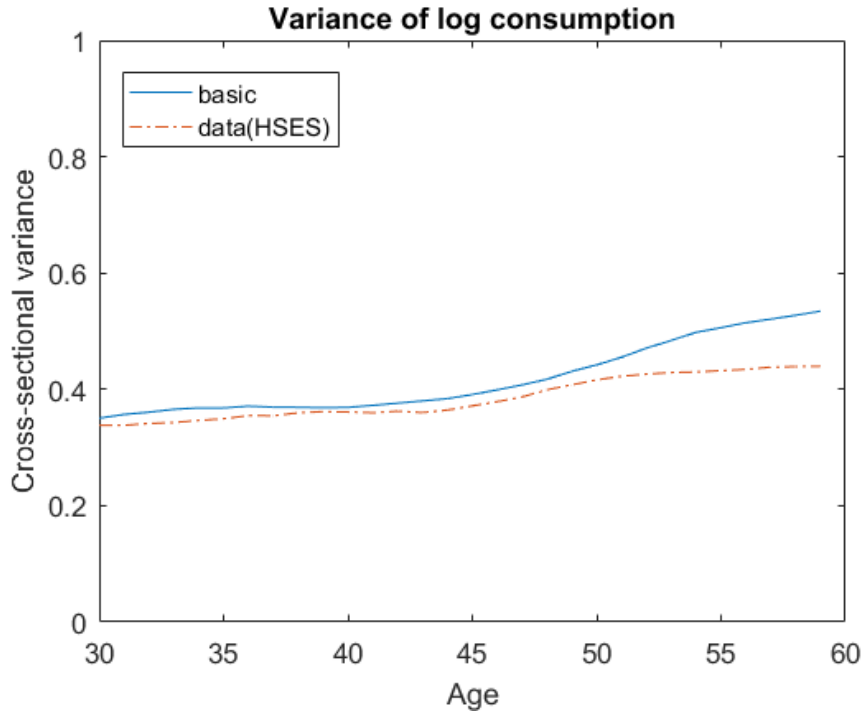


Figure 3.4 Age-profile of Consumption Inequality, Basic model.

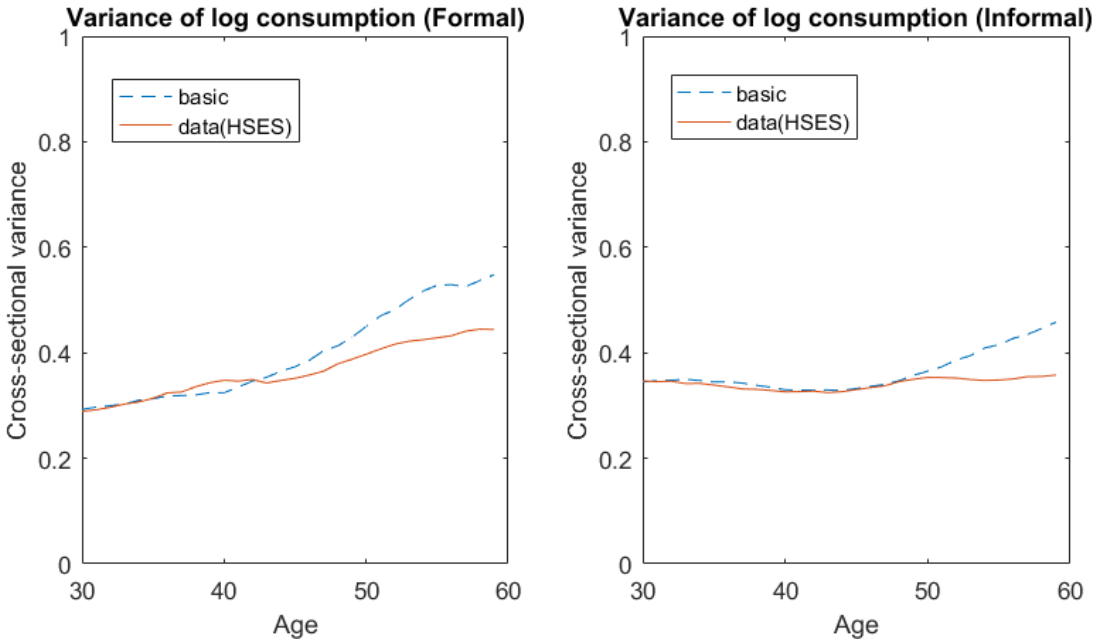


Figure 3.5 Age-profile of Consumption Inequality by Sector, Basic Model.

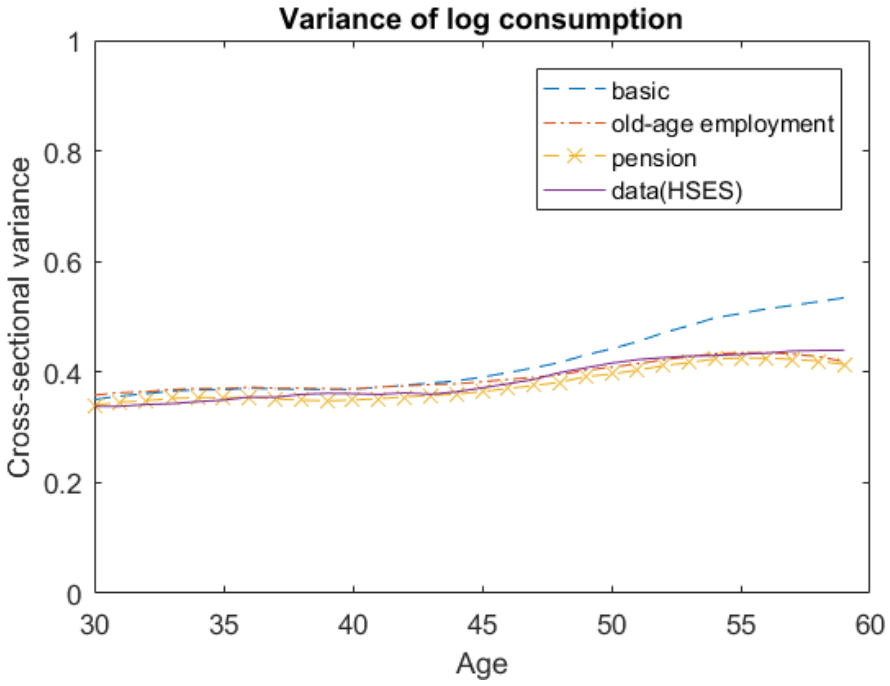


Figure 3.6 Age-profile of Consumption Inequality (calibrated).

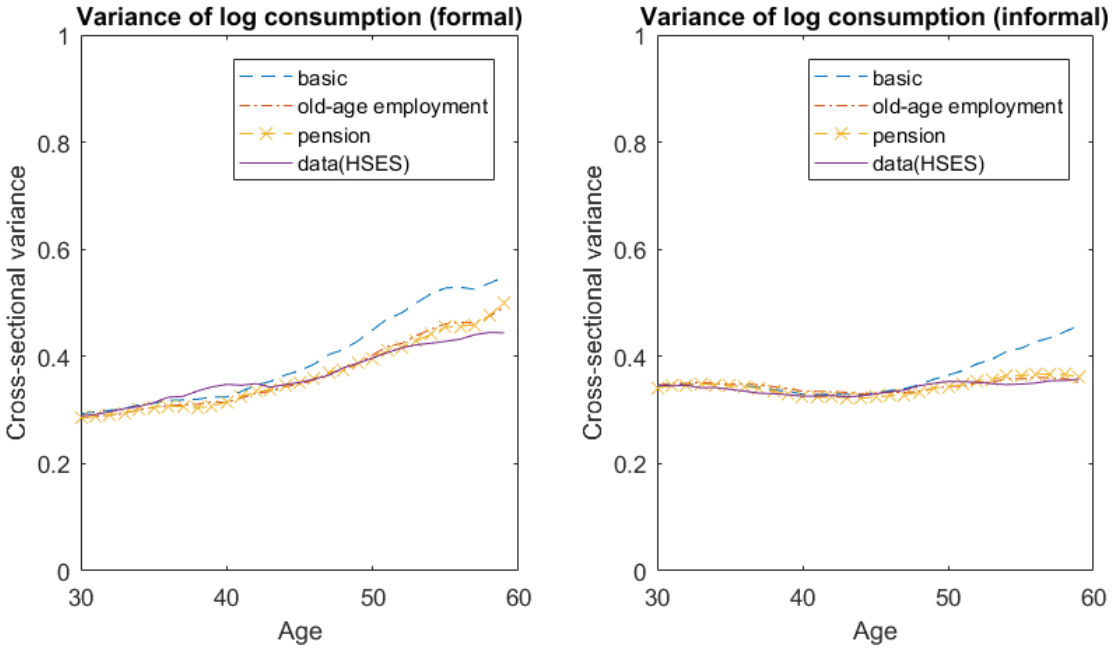


Figure 3.7 Age-profile of Consumption Inequality by Sector (calibrated).

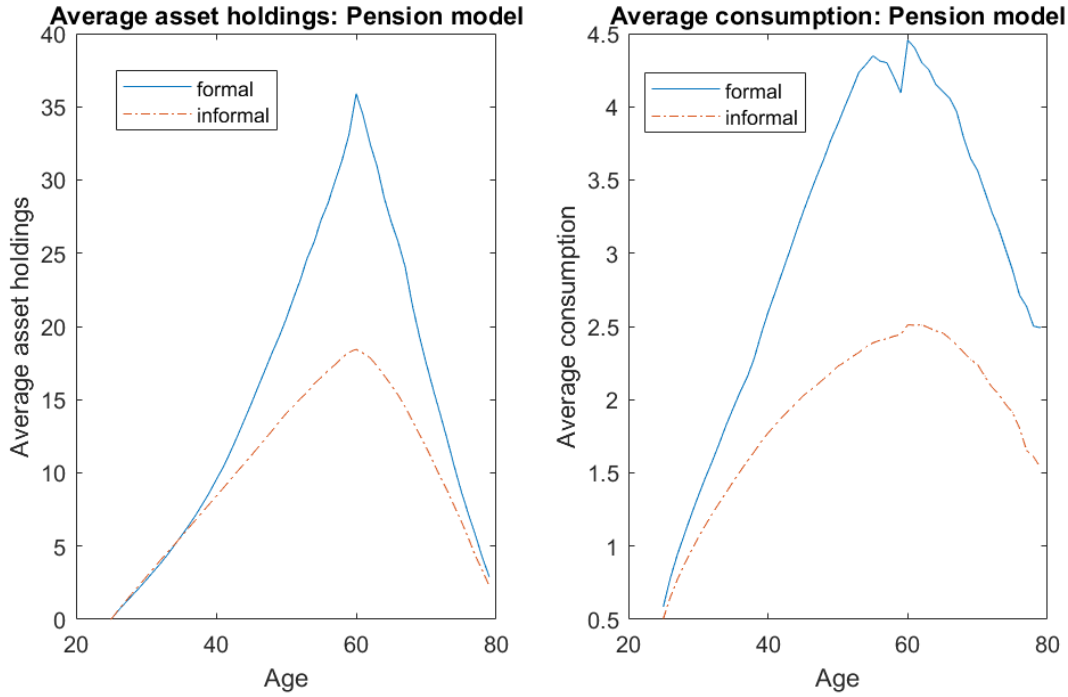


Figure 3.8 Average Asset Holdings and Consumption by Sector, Social Security Model.

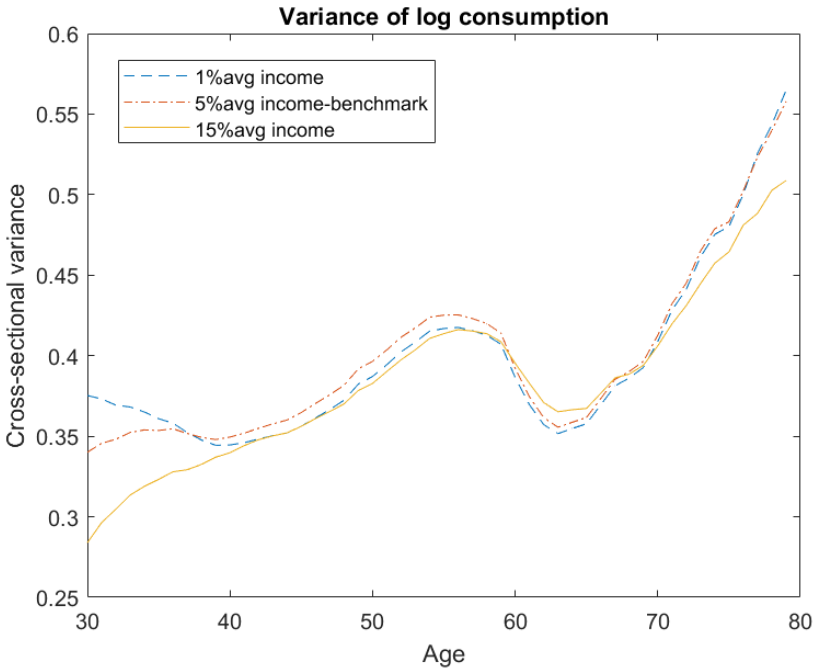


Figure 3.9 Variance of the Logarithm of the Consumption, Extension of Social Welfare.

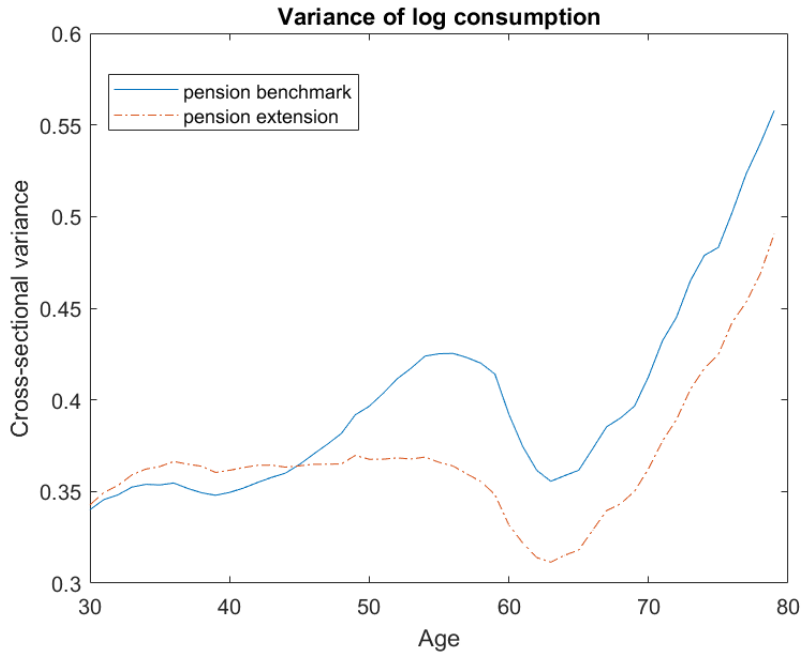


Figure 3.10 Variance of the Logarithm of the Consumption, Extension of Social Security.

Appendix

Social Welfare Programs in Thailand

The social welfare system was recently introduced in Thailand. Households in Thailand can be qualified for two means-tested programs if they meet the government's income and assets test. The first program is the Child Support Grant (CSG) scheme that was first introduced in 2015. The CSG scheme provides a monthly subsidy to babies and toddlers from low-income families. Families with an average monthly income of less than 3,000 Thai baht per person are eligible to receive 600 baht per child per month. The second program is the welfare card program that was implemented in 2017. There are five criteria for the welfare card program. The individual must be (a) a Thai citizen; (b) at least 18 years of age; (c) unemployed or an annual income below 100,000 baht; (d) no financial assets worth more than 100,000 baht; and (e) no owned real estate. The qualified households can receive a monthly allowance of 300 baht per person to shop for necessities at designated shops. In addition, the welfare card provides monthly subsidies for transportation that are 500 baht for buses/BTS/MRT, 500 baht for inter-city public transportation services, and 500 baht for trains.

Following Hubbard (1995), the total means-tested social welfare is calculated based on the representative household, which is a female-headed household with two children and no earnings and assets. This household can receive child subsidy benefits from CSG amounting to 1,200 baht per month and a monthly allowance from the welfare card of 1,800 baht per month. Therefore, the total means-tested social benefits are

approximately 36,000 Thai baht per household year.[24] The per capita subsidy it was 5.5% of the GDP per capita in 2017.[25]

[24] It is equal to (300 + (500 x 3) + (600 x 2)) x 12 = 36,000 Thai baht.
[25] The per capita subsidy is (36,000/3) and GDP per capita in Thailand as of 2017 was 218,200 Thai baht.

CHAPTER 4

Health Insurance and Household Savings[26]

4.1 Introduction

This chapter examines the relationship between health insurance coverage and precautionary saving decisions. Individuals face uncertainties throughout their lives. Medical expenditures is a major uncertainty in one's life. To prevent the impact of future uncertainties and maintain a smooth consumption level, individuals typically withhold a portion of their current income as a precautionary saving and/or purchase health insurance. The standard theory of consumption and precautionary saving suggests that the introduction of insurance reduces the amount of precautionary savings. The expectation that health insurance coverage and savings are partially substitutable has been used to test the existence of a precautionary motive for saving.

However, empirical findings based on private health insurance coverage are not consistent with theoretical expectation and are mostly focused on the US and other developed countries. Starr-McCluer (1996) examine the relationship between health insurance and savings using the US Survey of Consumer Finances data and finds that insured households save more than comparable households without coverage. Similarly, Guariglia and Rossi (2004) use the British Household Panel Survey data to investigate the effect of private medical insurance on household savings. They find a positive correlation between insurance coverage and savings. Both of these studies suggest that

[26] This chapter is a part of a joint project with Minchung Hsu, Pei-Ju Liao and Chang-Ching Lin on health insurance and precautionary savings.

private health insurance does not generally crowd-out savings, and indicate that the precautionary motive for saving is insignificant.

Therefore, this study aims to investigate the effect of health insurance on private savings in Thailand. We focus our analysis towards determining whether households who are not covered by private health insurance save more than insured households as a precaution against the risk of medical expenditures. We define private health insurance to include both self-purchased insurance and employment-based insurance. According to our interpretation of formal/informal workers, this implies that those who are insured or have private health insurance are mainly workers in the formal sector and those who are uninsured are workers in the informal sector.

We select Thailand for our empirical analysis for two main reasons. First, Thailand introduced the Universal Coverage Scheme (UCS) since 2002, which is public health insurance with universal provision and almost free of charge.[27] As the UCS insures all Thai citizens against unexpected health expenditure, there should be no crowding-out of saving by private health insurance. Kirdruang and Glewwe (2018) study the impact of UCS on households' consumption and savings in Thailand and find that the UCS had no impact on savings. However, a recent study by Ushijima (2020) evidences a decrease in household precautionary savings after the introduction of UCS in Thailand.

Second, although the UCS provides health care services at a reduced price, individuals prefer to use private health services because private hospitals typically provide a higher quality of medical treatment and services. In addition, while employment-based insurance covers both private and public hospitals, the UCS covers

[27] Initially, the scheme levied a charge of 30 baht (~$0.75) per service contact. (Limwattananon et al., 2015)

only public hospitals. Moreover, to use the UCS service, an individual has to register at a single, designated public hospital and cannot use the UCS elsewhere. During an emergency, an uninsured person might face unexpected out-of-pocket medical expenditures as they might have to go to a hospital that is not covered by the USC. As a result, uninsured individuals might save for precautionary purpose more than those with insurance and private health insurance might crowd-out saving.

To investigate this, we use data from the Thai Household Socio-Economic Survey (HSES) between 2011 and 2017. For our analysis, we follow the methodology used by Guariglia and Rossi (2004) to analyze the relationship between saving and private health insurance. We start with the Tobit estimation for the decision to save. We then use an Instrumental Variables (IV) technique and a Full Model Maximum Likelihood approach to control for endogeneity.

Our empirical findings give interesting results. We find a positive relationship between saving and health insurance in the Tobit model. While contradicting the theory of precautionary saving, this is in line with findings from the literature. In their study, Guariglia and Rossi (2004) suggest that there might be a bias due to endogeneity caused by individual perceptions of risk. Accordingly, we use the IV Tobit and Full Model Maximum Likelihood approaches to control for endogeneity. We find a negative association, albeit statistically insignificant, between saving and health insurance coverage using the IV Tobit specification. This suggests that the positive relationship between insurance coverage and saving in the Tobit model was the result of an endogeneity bias. We further control for endogeneity using a Full Model Maximum Likelihood technique and find a statistically significant negative relationship between

saving and health insurance coverage. Our findings thus provide strong evidence in support of the theory that private health insurance reduces precautionary saving.

The rest of this chapter is organized as follows. Section 2 presents our dataset and descriptive statistics. Section 3 provides the empirical specification of the Tobit model and its results. Section 4 explains how we control for endogeneity using the IV Tobit specification and a Full Model Maximum Likelihood approach. Section 5 concludes.

4.2 Data and Descriptive Statistics

4.2.1 Data

We used data from the Household Socio-Economic Survey (HSES) between 2011 and 2017.[28] Observations in our dataset are restricted to household heads aged between 25 and 54 years old who are in employment in order to avoid the effect of schooling, retirement, and unemployment on savings. Households with self-employed heads are excluded because it is particularly difficult to distinguish family savings from the firms' investment. Households with heads working in the public sector are also excluded. The sample used in the analysis is made up of 22,406 observations. All income and spending variables are expressed in Thai baht as of 2015.

Monthly household savings are calculated as income minus consumption expenditures.[29] The HSES provides detailed data on household income and expenditure as well as information on whether a household has employment-based health insurance or purchase any non-saving insurance. We use the question "at present, do you receive

[28] The surveys used in the study are from 2011, 2013, 2015, and 2017.
[29] We use average monthly consumption expenditures and average monthly current income from the survey to calculate savings.

the following welfare of medical services?" to identify whether a household is covered by employment-based or private health insurance (EPHI) or not. If a household answer "medical card (social security), private health insurance, or welfare by employer", it is classified as covered by EPHI.

4.2.2 Descriptive Statistics

Descriptive statistics of the discussed variables are summarized in Table 4.1. In terms of income and spending variables, the average monthly household income is 26,438 Thai baht. The average household monthly savings is 4,750 Thai baht and the average health expenditure is 189 Thai baht. Around 76.64% of all households have positive savings and 61.58% are covered by EPHI. Regarding household characteristic variables, the average number of adults and the average number of dependent children in a household are 2.14 and 0.51, respectively. Only 9% of households have mortgage debts. The average age of a household head is around 40 years old.

Descriptive statistics on saving behavior and insurance coverage for various social-demographic groups are presented in Table 4.2. Column 1 shows the percentages of savers, of which 76.64% have positive savings. The proportion of savers increases with the education and income levels of household heads. Households whose heads are aged between 45 and 54 years old, are not married or co-habiting, and have no dependent child have a higher percentage of savers.

Column 2 shows the percentages of households covered by EPHI. Overall, 61.68% of the dataset are covered. The proportions of EPHI coverage show a similar pattern to the percentages of savers in column 1, in that the respondents who are not married or co-habiting and do not have dependent children are more likely to have EPHI coverage. The

percentage of those with EPHI coverage also tends to increase with the income level of household heads.

Columns 3–6 show the percentages of savers who were insured and uninsured and the amount saved by them. The EPHI insured households tend to have higher proportions of savers as well as higher amounts of savings than uninsured households. This pattern holds for the whole dataset and the various social-demographic groups reported in Table 4.2. Overall, the percentage of savers with EPHI coverage is 79.94%, compared with 71.34% of those without. The average monthly savings of those with coverage is 8,879 Thai baht, whereas the average monthly savings for those without is 4,983 Thai baht. The percentage of savers and the amount of savings tends to rise with education and income levels for both insured and uninsured households.

These descriptive statistics show a positive correlation between private health insurance and saving. This suggests that Thai individuals who are uninsured by the EPHI save less for unexpected medical expenditures. In the next section, we follow the methodology outlined by Guarglia and Rossi (2004) to test whether this conclusion is robust.

4.3 Empirical Specification and Results

4.3.1 General Specification

The empirical specifications used in this study follow Guariglia and Rossi (2004). First, the Tobit estimation technique is used to analyze the determinants of individual saving decisions and to study how insurance coverage affect savings decisions. We use this technique because saving is calculated from income minus consumption expenditure,

which could be either a positive or negative value. However, there is a possibility that income is under-reported. The negative saving might be observed because households actually dissave, or they are unintended to under-reported their income. As a result, we truncate all negative savings to zero and use the Tobit technique. The regression equation is as follows:

$$S_i^* = \alpha + \gamma P_i + \Gamma X_i + e_i, \qquad S_i = \begin{cases} 0 \ \ if \ S_i^* \leq 0, \\ S_i^* \ if \ S_i^* > 0, \end{cases} \quad i = 1, \dots, N. \tag{4.1}$$

S_i^* is the true propensity to save of the household, which could be either a positive or negative value. However, negative values are truncated to zero and would not be observed due to censoring. Thus, S_i denotes the observed saving of household *i*, allowing for positive or zero savings. P_i is a dummy variable coded as 1 if household *i* is covered by EPHI and 0 otherwise. X_i are other covariates that are assumed to affect saving, listed in the Appendix. Finally, e_i is an idiosyncratic error term.

4.3.2 Tobit Regression Results

The regression results are shown in column 1 and 2 of Table 4.3. The coefficient of P_i is positive and statistically significant, and shows that those with EPHI tend to save more than those without by 2,457.44 Thai baht. This positive correlation suggests that insurance coverage does not crowd-out private savings, which is consistent with the findings from Guariglia and Rossi (2004). Our results show that savings tend to be higher for households with higher permanent incomes and more employed adults. In contrast, savings tend to be lower for those with older household heads, household heads who are married, and for households with more dependent adults or children living within the household. Interestingly, the coefficient of education and mortgage debt diverges from

our expectations. The households that have heads with higher education degrees tend to save less, whereas those with mortgage debts tend to have more savings.

4.4 Controlling for Endogeneity

It is possible that the results from the Tobit regression might be biased due to endogeneity. Similar to saving, the decision to purchase medical insurance may depend on households' perceptions of risk.[30] Households that are more risk-averse tend to save more for precautionary reasons than less risk-averse households. Thus, instead of having low levels of savings, households with EPHI coverage may also tend to have a higher levels of savings because they are more risk-averse.

4.4.1 Instrumental Variables Tobit Regression and Results

An IV Tobit model is used to account for endogeneity from unobservable factors affecting saving and insurance purchasing decisions. Occupational dummies are selected as instrument variables on the insurance coverage variable (P_i).

The first criterion with which to evaluate whether occupational dummies are suitable instrument variables is that they should be strongly correlated with insurance coverage.[31] As 96% of the insured households receive coverage through the firm by which they are employed, it is reasonable to assume that occupation-related variables are good proxies for insurance coverage. Occupations that are mostly available in the formal

[30] In this study, insurance coverage is an employment-based insurance that is paid by an employer. Therefore, the relevant decision can be to join a firm that provides private health insurance.
[31] Medical insurance subsidies are usually occupation-based. Guariglia and Rossi (2004) also use occupational dummies as proxies for insurance coverage, as 71 percent of the dataset received their insurance coverage through the firms by which they were employed.

sector, such as professionals and technicians, are more likely to be covered by the EPHI than service workers.[32] To test for this requirement, a regression of P_i on the occupational dummies and the other exogenous variables is performed using a Probit method. Wald test (Judge et al., 1985) is then used to test for the joint significant of these instruments. The results show a $\chi^2(4)$ statistic equal to 847.31 with a *p*-value of 0.000, indicating that the occupational dummies generally have a high explanatory power for health insurance coverage.

Another requirement for instrument validity is that they should have no predictive power on savings when a regression includes P_i and other exogenous variables. Using a Tobit model to test the joint significance of the occupational dummies, a regression on all exogenous variables X_i, insurance coverage, and the occupational dummies gives an $F(4, 22376)$ equal to 2.74 with *p*-value of 0.027. Although the *p*-value is not high enough to suggest that the instruments do not directly affect the savings, the instruments are still valid because there is no coefficient related to the occupational dummies that is individually statistically significant in the regression.

We then performed a Basmann (1960) test to test for an over-identifying restriction by regressing the residuals from the IV Tobit model on the occupation dummies and all other exogenous variables and testing the joint significance of the instruments. The regression gives an $F(4, 22376)$ equal to 1.58 with *p*-value of 0.1764, suggesting that the instruments are uncorrelated with the residuals and thus have no predictive power for saving and are legitimate.

[32] More than 90 percent of professionals and technicians work in the formal sector, while 39 percent of service workers work in the formal sector (The informal employment survey 2019, NSO).

Following the procedure illustrated in Newey (1987), the estimated results of the saving equation based on the IV-Tobit model are reported in column 3 and 4 of Table 4.3. Our result differs from those of Guariglia and Rossi (2004) because the coefficient on the insurance coverage (-2289.10) becomes negative, although it is not significant. This suggests that the positive correlation between insurance coverage and savings could be affected by the endogeneity bias. The effects of the other exogenous variables on savings are similar to the Tobit model.

4.4.2 Full Model Maximum Likelihood Regression

An alternative, and more efficient, way to address the endogeneity of insurance coverage is to use a Full Model Maximum Likelihood (FMML) approach. This method considers the interdependence between health insurance coverage and saving decisions.

The detailed specification of the FMML function also follows Guariglia and Rossi (2004). Two sets of equations are used, the insurance coverage equation and the saving equation. The equation determining whether households have EPHI is described as follows:

$$P_i^* = \theta\varpi_i + u_i, \qquad P_i = \begin{cases} 0 \;\; if \; P_i^* \leq 0, \\ 1 \;\; if \; P_i^* > 0, \end{cases} \tag{4.2}$$

where ϖ_i includes variables in X_i as well as occupational dummies. Another equation for saving behavior is described by

$$S_i^* = \alpha + \gamma P_i^* + \Gamma X_i + e_i = \lambda\varpi_i + \varepsilon_i, \tag{4.3}$$

where $\varepsilon_i = \gamma u_i + e_i$ and $S_i = S_i^*$ if $S_i^* > 0$; otherwise, $S_i = 0$. The error terms from the insurance coverage and saving equations are assumed to be jointly normally distributed with a zero mean and a variance-covariance matrix:

$$\begin{bmatrix} \sigma_e^2 & \sigma_{eu} \\ \sigma_{ue} & 1 \end{bmatrix},$$

where the variance of u_i is normalized to 1. The correlation between the error terms, u_i and ε_i, reflects the risk aversion term, which is not observed and is contained in the error terms of both equations. Allowing the covariance between u_i and ε_i to be non-zero prevents the coefficient on insurance coverage in eq. (4.3) from including a risk aversion component, thus removing this bias.

Based on the observed values of S_i and P_i, households can be classified into four groups. Denoting the univariate normal density function with φ, the univariate normal cumulative distribution function with Φ_n, the bivariate normal cumulative distribution function with Φ_b, the variance of ε_i with σ_ε^2, and the covariance between u_i and ε_i with $\sigma_{\varepsilon u}$, and letting $\xi = \sigma_{\varepsilon u}/\sigma_\varepsilon^2$ and $\zeta = \sigma_{\varepsilon u}/\sigma_\varepsilon$, the probabilities related to each group can be written as follows:

Group 1: Households who save and are insured, such that:

$$\Pr(S_i^* > 0; P_i^* > 0) = \Pr(S_i^*)\Pr(P_i^* > 0|S_i^*) = \varphi(\varepsilon_i)\Phi_n\left(\frac{\theta\varpi_i + \xi\varepsilon_i}{\sqrt{1-\zeta^2}}\right).$$

Group 2: Households who do not save and are insured, such that:

$$\Pr(S_i^* \leq 0; P_i^* > 0) = \Phi_b\left(\frac{-\lambda\varpi_i}{\sigma_\varepsilon}, \theta\varpi_i, -\zeta\right).$$

Group 3: Households who save and are not insured, such that:

$$\Pr(S_i^* > 0; P_i^* \leq 0) = \Pr(S_i^*)\Pr(P_i^* \leq 0|S_i^*) = \varphi(\varepsilon_i)\Phi_n\left(\frac{-\theta\varpi_i - \xi\varepsilon_i}{\sqrt{1-\zeta^2}}\right).$$

Group 4: Households who do not save and are not insured, such that:

$$\Pr(S_i^* \leq 0; P_i^* \leq 0) = \Phi_b\left(\frac{-\lambda\varpi_i}{\sigma_\varepsilon}, -\theta\varpi_i, \zeta\right).$$

The log-likelihood function used to estimate the parameters then takes the following form:

$$L = \sum_{i \epsilon Group1} \ln[\Pr(S_i^* > 0; P_i^* > 0)] + \sum_{i \epsilon Group2} \ln[\Pr(S_i^* \leq 0; P_i^* > 0)]$$

$$+ \sum_{i \epsilon Group3} \ln[\Pr(S_i^* > 0; P_i^* \leq 0)] + \sum_{i \epsilon Group4} \ln[\Pr(S_i^* \leq 0; P_i^* \leq 0)].$$

The estimates of the saving equation using the FMML method are reported in Table 4.4. Columns 1 and 2 show the results of the equation for insurance coverage. The probability of having coverage is lower for households with male household heads and those more children. The coefficients of the occupational dummies are all negative. This suggests that other occupations have a lower probability of being covered by the EPHI than "legislators, managers, and professionals."[33] It is higher for married household heads and those at all educational levels. Households with mortgage debt also have a higher probability of being covered by EPHI.

The coefficient of EPHI insurance coverage (-20226.05) is negative and statistically significant. The magnitude of this coefficient implies a large impact of EPHI on saving that conditional on other family and head characteristics, saving of households with EPHI is lower than those without EPHI by 20,226 Thai bahts. The results of the explanatory variables on savings are still similar to those in the standard Tobit and IV-Tobit models, except for the genders of household heads and educational dummies. In particular, the coefficients of all educational dummies become significantly positive, suggesting that saving tends to increase with educational attainment. Our result thus shows that there is evidence of a substitution effect between insurance coverage and

[33] "Legislators, managers, and professionals" is the omitted group.

saving. This supports the existence of precautionary saving as a hedge against health expenditure risks.

4.5 Conclusion

This chapter investigates the association between employment-based/private health insurance coverage and private saving. We ask whether households who do not have EPHI coverage and can face unexpected health care expenditures for private health care services tend to save more than those with insurance. Our findings show that, after controlling for endogeneity, the insured households have significantly lower savings than those without EPHI coverage.

We follow the methodology outlined by Guariglia and Rossi (2004). Our result from the Tobit regression shows a strong positive correlation between EPHI and saving, in line with the findings of both Starr-McCluer (1996) and Guariglia and Rossi (2004). We then use the Instrumental Variables (IV) technique to investigate whether our estimates might have suffered from bias due to endogeneity. We find a negative effect of EPHI on saving, but it is not significant. We then take into account the interdependencies between health insurance coverage and saving decisions using a Full Model Maximum Likelihood approach. The resulting coefficient of insurance coverage on savings is negative and statistically significant. Our findings thus provide evidence for the crowding-out effect of health insurance coverage on private saving. This supports the theory of precautionary saving and confirms the significance of precautionary saving motive to reduce medical expenditure uncertainty.

4.A Tables

Table 4.1 Descriptive Statistics

	Number of observations N=22,406	
	Mean	Standard deviation
Income and spending (Thai baht per month)		
Current income	26,438.06	26,916.45
Savings	4,750.13	16,977.86
Health spending (out-of-pocket)	188.62	765.56
Household characteristics		
No. of adults in a household	2.14	1.00
No. of employed adults in a household	1.74	0.76
No. of children in a household	0.51	0.79
Age of household head	39.76	8.11
	Frequency	%
Saving or not		
Negative/zero	5,234	23.36
Positive	17,172	76.64
EPHI coverage		
Uncovered	8,609	38.42
Covered	13,797	61.58
Mortgage debt		
Others	20,388	90.99
Debtors	2,018	9.01
Marital status		
Married/co-habiting	15,137	67.56
Others	7,269	32.44
Gender		
Female	6,222	27.77
Male	16,184	72.23
Education		
High school dropout or lower	11,889	53.06
High school graduate	4,574	20.41
Some college	1,831	8.17
College and above	4,112	18.35

Table 4.2 Savings and Health Insurance Coverage by Demographic Characteristics, Age, Education, and Income

Number of observations N=22,406

	saver	EPHI coverage	saver EPHI insured	savings EPHI insured	saver EPHI uninsured	savings EPHI uninsured
	(1)	(2)	(3)	(4)	(5)	(6)
All	76.64	61.58	79.94	8,878.82	71.34	4,982.65
Demographic variables						
Married/co-habiting	76.43	59.99	80.08	10,196.59	70.95	5,349.42
Not married/co-habiting	77.08	64.88	79.69	6,328.86	72.27	4,128.45
No dependent child	79.39	66.83	81.50	7,749.31	75.13	4,993.58
With dependent children	71.75	52.21	76.40	11,628.06	66.66	4,967.41
Age						
25-34	77.17	52.21	78.68	5,709.33	73.41	4,943.18
35-44	75.41	64.07	78.68	9,117.36	68.19	4,943.18
45-54	77.59	49.95	82.28	12,414.52	72.91	5,300.98
Education						
High school dropout	74.16	46.79	77.42	5,476.09	71.29	4,647.84
High school graduate	76.28	73.11	79.01	6,314.72	68.86	5,952.47
Some college	77.99	87.82	79.10	8,938.89	69.96	6,087.25
College and above	83.61	79.82	85.59	16,480.65	75.78	5,803.39
Income						
First quartile	60.36	36.72	68.52	2,085.48	53.09	1,266.81
Second quartile	76.55	55.05	80.75	4,199.80	67.34	2,204.40
Third quartile	81.70	71.89	84.59	7,110.39	75.86	3,632.83
Forth quartile	88.81	85.31	91.02	25,221.62	83.52	9,749.09

Notes: Column (1) reports the percentage of savers, Column (2) is the percentage of EPHI coverage in the dataset, Column (3) is the percentage of savers among EPHI insured, Column (4) is the non-zero average monthly savings of the EPHI insured (THB), Column (5) is the percentage of savers who do not have EPHI coverage, and Column (6) is non-zero average monthly savings of those without EPHI coverage (THB).

Table 4.3 Tobit Estimates and Instrumental Variables Tobit Estimates for Saving

	Tobit (1)	Standard error (2)	IV Tobit (3)	Standard error (4)
EPHI insurance	2457.44***	(326.90)	-2289.10	(1533.38)
Permanent income	0.40***	(0.03)	0.42***	(0.01)
Health spending	1.28**	(0.55)	1.33***	(0.16)
Demographic variables				
Age	-601.70***	(176.74)	-547.29***	(162.52)
Age-squared	796.91***	(220.16)	695.96***	(204.75)
No. of emp.	1435.42***	(391.79)	1211.98***	(284.88)
No. of dep adults	-394.89	(265.69)	-428.98**	(202.16)
No. of children	-986.24***	(223.94)	-1295.17***	(199.51)
Male	384.83	(303.18)	141.97	(313.55)
Married/co-habiting	-2001.60***	(375.96)	-1929.23***	(327.75)
Education				
High school graduate	-1379.65***	(354.19)	-738.52*	(411.31)
Some college	-2028.63***	(504.20)	-943.01	(626.90)
College degree or above	-1035.31*	(534.91)	-470.82	(545.07)
Other Variables				
Mortgage debtor	4003.70***	(656.96)	4492.65***	(478.86)
2013 dummy	117.93	(379.71)	168.53	(357.77)
2015 dummy	-733.72**	(321.13)	-682.69*	(358.43)
2017 dummy	-843.79**	(403.76)	-773.41**	(362.22)
Constant	-1071.49	(3892.80)	1474.91	(4000.47)
Industrial dummies	Yes		Yes	
Occupational dummies	Yes		Yes	

Notes: *, **, *** denote 10%, 5%, and 1% significance, respectively. “High school dropout” is the omitted educational group. *Source:* The HSES, 2011-2017.

Table 4.4: Full Model Maximum Likelihood Estimates for EPHI Coverage and Saving

	EPHI coverage (1)	Standard error (2)	Saving (3)	Standard error (4)
EPHI insurance			-20226.05***	(359.01)
Permanent income			0.48***	(0.01)
Health spending	0.00***	(0.00)	1.52***	(0.18)
Demographic variables				
Age	0.05***	(0.01)	-310.36*	(177.18)
Age-squared	-0.07***	(0.01)	277.59	(221.87)
No. of emp.	0.00	(0.02)	347.81	(301.05)
No. of dep adults	-0.02	(0.01)	-491.27**	(221.23)
No. of children	-0.16***	(0.01)	-2436.21***	(190.78)
Male	-0.20***	(0.02)	-800.05**	(333.33)
Married/co-habiting	0.08***	(0.02)	-1581.32***	(358.38)
Education				
High school graduate	0.38***	(0.02)	1797.09***	(391.11)
Some college	0.84***	(0.04)	3218.57***	(574.30)
College degree or above	0.48***	(0.03)	1885.49***	(556.61)
Occupation				
Technicians and associate professionals	-0.26***	(0.05)		
Service, sales and clerical support workers	-0.67***	(0.04)		
Craft and related trades workers	-1.01***	(0.05)		
Plant and machine operators, and assemblers	-0.57***	(0.05)		
Other Variables				
Mortgage debtor	0.42***	(0.04)	6295.26***	(499.23)
2013 dummy	0.09***	(0.03)	369.36	(391.38)
2015 dummy	0.07***	(0.03)	-477.94	(392.08)
2017 dummy	0.09***	(0.03)	-497.36	(395.95)
Constant	-0.08	(0.28)	11071.79***	(4285.14)
Industrial dummies	Yes		Yes	

Notes: *, **, *** denote 10%, 5%, and 1% significance, respectively. "High school dropout" is the omitted educational group. "Legislators, managers, and professionals" is the omitted occupational group. *Source:* The HSES, 2011-2017.

Appendix

Explanatory Variables used in Empirical Analysis

The following covariates are considered in our regression:

- *Household permanent income:* The permanent income is obtained by taking the fitted values from least squared regression of the household's monthly current income on household characteristics and household heads' marital status, gender, age, age-squared, educational dummies, occupational dummies, and the interactions of the latter two groups of dummies with age and age-squared.[34]

- *Family composition:* The variables include the number of employed adults, the number of dependent adults, the number of children, and the marital status and gender of the household head to capture the composition of families. A different number of employed adults or children in families can reflect a difference in household income and consumption.

- *Education:* Education level affect saving decision because those with higher education tend to save more due to higher income. The sample is classified into four groups according to the education level of household head: college and post-graduate graduates, post-secondary graduates, high school graduates, and high school dropouts or lower.

- *Age:* The variables include age and age-squared of the household head to capture the curvature of the saving function. At different ages, people are facing financial

[34] Using permanent income can prevent a problem from a distortion of longer-run relationships between income and savings (Starr-McCluer, 1996). This approach follows Guariglia and Rossi (2004).

constraints and uncertainties differently, which can significantly result in different saving decisions.

- *Employment status:* The employment uncertainty varies among people in different industries and occupations. The variables include industrial dummies and occupational dummies.

- *Health spending:* Total health spending is used as a proxy for health status, which includes all health-related spending and medical expenditure.

- *Other variables:* The mortgage debt variable is included because the ability to save of households with mortgage debts might be lower than those without mortgage debt. The survey year dummy variables are also included to capture the global changes from year to year.

Robustness Check

The Ordinary Least Squares (OLS) Estimation of Savings on EPHI

In this section, we check the robustness of our estimation results by applying the ordinary least squares (OLS) estimation method and compare with the results from the Tobit model. We use the Tobit model to avoid the bias from under-reported household income. However, if the income is precisely reported, we can use the OLS method for the estimation. The regression equation for saving is as follows:

$$S_i = \alpha + \gamma P_i + \Gamma X_i + e_i, \qquad i = 1, \dots, N, \tag{4.4}$$

where S_i is the saving, which is calculated from income minus consumption expenditure. Equation 4.4 is also used for the Instrumental Variable (IV) estimation and the Full Model Maximum Likelihood (FMML) approach.

Table A4.1 exhibits the results from the OLS and IV estimation, and Table A4.2 presents the results from the FMML estimation. The coefficient of EPHI coverage on saving is significantly positive in the OLS model. After controlling for endogeneity, we find a negative effect of EPHI on savings. By using the IV technique, the coefficient of insurance coverage become negative but insignificant. The result from the FMML approach shows a strong negative effect of EPHI on saving. The estimation results are consistent with our Tobit model.

Table A4.1 OLS Estimates and Instrumental Variables OLS Estimates for Saving

	OLS (1)	Standard error (2)	IV (3)	Standard error (4)
EPHI insurance	3239.69***	(267.05)	-1592.31	(1477.67)
Permanent income	0.44***	(0.03)	0.46***	(0.03)
Health spending	1.74***	(0.62)	1.79***	(0.62)
Demographic variables				
Age	-395.83**	(155.37)	-341.07**	(152.97)
Age-squared	519.70***	(193.36)	417.87**	(189.36)
No. of emp.	862.93**	(371.49)	635.26	(393.53)
No. of dep adults	-679.73***	(248.83)	-715.24***	(249.59)
No. of children	-1170.00***	(169.28)	-1484.71***	(205.08)
Male	50.03	(281.79)	-197.39	(291.76)
Married/co-habiting	-1139.17***	(334.00)	-1064.23***	(331.80)
Education				
High school graduate	-1375.90***	(308.99)	-723.84*	(369.39)
Some college	-1766.08***	(458.87)	-661.06	(482.53)
College degree or above	-1179.15**	(544.20)	-607.61	(532.40)
Other Variables				
Mortgage debtor				
2013 dummy	5384.84***	(672.64)	5882.93***	(680.11)
2015 dummy	-308.29	(343.20)	-256.29	(347.49)
2017 dummy	-750.35**	(304.74)	-696.94**	(305.36)
Constant	-834.79**	(357.78)	-760.08**	(356.68)
Industrial dummies	Yes		Yes	
Occupational dummies	Yes		Yes	

Notes: *, **, *** denote 10%, 5%, and 1% significance, respectively. “High school dropout” is the omitted educational group. *Source:* The HSES, 2011-2017.

Table A4.2: Full Model Maximum Likelihood Estimates for EPHI Coverage and Saving (OLS)

	EPHI coverage (1)	Standard error (2)	Saving (3)	Standard error (4)
EPHI insurance			-18293.26***	(311.13)
Permanent income			0.43***	(0.01)
Health spending	0.00***	(0.00)	1.44***	(0.16)
Demographic variables				
Age	0.05***	(0.01)	-219.64	(156.92)
Age-squared	-0.07***	(0.01)	165.90	(196.50)
No. of emp.	0.00	(0.02)	-436.92	(266.65)
No. of dep adults	-0.02	(0.01)	-272.20	(194.98)
No. of children	-0.15***	(0.01)	-1721.48***	(166.57)
Male	-0.21***	(0.02)	-752.22**	(295.73)
Married/co-habiting	0.09***	(0.02)	-1549.32***	(317.87)
Education				
High school graduate	0.37***	(0.02)	1422.59***	(345.59)
Some college	0.81***	(0.04)	2620.86***	(509.66)
College degree or above	0.51***	(0.03)	1126.24**	(495.18)
Occupation				
Technicians and associate professionals	-0.31***	(0.05)		
Service, sales and clerical support workers	-0.70***	(0.04)		
Craft and related trades workers	-1.04***	(0.05)		
Plant and machine operators, and assemblers	-0.64***	(0.04)		
Other Variables				
Mortgage debtor	0.47***	(0.04)	5407.14***	(446.40)
2013 dummy	0.08***	(0.03)	24.31	(345.64)
2015 dummy	0.06**	(0.03)	-731.73**	(345.92)
2017 dummy	0.09***	(0.03)	-843.97**	(349.84)
Constant	-0.15	(0.28)	12279.38***	(3786.98)
Industrial dummies	Yes		Yes	

Notes: *, **, *** denote 10%, 5%, and 1% significance, respectively. "High school dropout" is the omitted educational group. "Legislators, managers, and professionals" is the omitted occupational group. *Source:* The HSES, 2011-2017.

CHAPTER 5

Conclusion and Policy Implications

This dissertation investigates three main issues in economic research by focusing on developing countries. First, we investigate the age-earnings profiles and earnings uncertainties between the formal and informal sectors. Second, we examine the impact of earnings uncertainties from different sectors on earnings and consumption inequality over the life cycle. We also conduct policy experiments to investigate the impact of social insurance policy on inequality and individuals' consumption and saving behavior. Finally, we study the relationship between insurance coverage and precautionary savings.

In chapter 2, we find that in a developing country like Thailand, formal workers have higher earnings than informal workers, and their earnings increases more significantly with age. Moreover, sector effect on the earnings growth is more significant than the effect of schooling. We find that even formal workers with lower education have higher earnings growth than informal workers with higher education.

We further examine the patterns of earnings shocks for workers in different sectors. We find that earnings shocks in the informal sector are greater than in the formal sector. However, the shocks are more persistent in the formal sector, indicating that it takes more time for formal workers to recover from shocks than informal workers. We also explore the difference in earnings shocks between age groups in each sector. We find that earnings shocks increase with age in both sectors. Our results also show that

persistence declines with age in the formal sector while slightly increases with age in the informal sector.

In general, our findings conclude that working in the formal sector offers a substantially higher return than the informal sector in developing countries. Moreover, workers in the informal sector tend to suffer more than workers in the formal sector when they encounter earnings risks. An extension of social insurance, such as social security, would help alleviate the effect of earnings shocks.

In chapter 3, we investigate the inequality in earnings and consumption as well as consumption and saving behaviors over the life cycle in developing countries and ask whether the relationship between consumption inequality and age can be explained by a life cycle model with uninsurable earnings shocks from different sectors of employment. The model is calibrated to match Thai data. We find that key factors accounting for consumption inequality are the idiosyncratic earnings shocks in different sectors, working during old age, and social security. The different earnings shocks for different sectors of employment can explain earnings inequality. We also find that old-age employment and social security are important because it serves as insurance arrangements that are available in developing countries. In the absence of these two features, our model would generate 17% more consumption inequality.

We further propose two policy reforms, that is, the extension of social welfare and extension of the social security system. Our finding indicates that a high level of social welfare cannot reduce consumption inequality significantly, except those very young and very old. We also find that the extension of social security to cover the informal sector can reduce consumption inequality significantly among middle-aged workers and the elderly.

Although old-age employment is available in most developing countries, our study suggests that the development of social security or pension benefits at old age is also essential. Moreover, our policy experiment suggests that the extension of social security to cover the informal sector is also preferred for developing countries.

In chapter 4, we investigate the association between employment-based/private health insurance (EPHI) coverage and private saving in Thailand. We find that, after controlling for endogeneity, households with EPHI coverage have significantly lower savings than those without coverage. Our finding provides evidence for the crowding-out effect of health insurance coverage on saving. This supports the theory of precautionary saving and confirms the significance of precautionary saving motive to reduce medical expenditure uncertainty.

According to our definition of informality, the insured thus refers to workers in the formal sector, and the uninsured are informal workers. Our results imply that informal workers who are not covered by EPHI tend to save more than formal workers as precautionary savings against unexpected health expenditures.

References

Bacchetta, M., Ernst, E., & Bustamante, J. (2009). *Globalization and informal jobs in developing countries.* wto.int.

Basmann, R. (1960). On finite sample distributions of generalized classical identifiability test statistics. *Journal of the American Statistical Association, 55*, 650-659.

Blundell, R., Graber, M., & Mogstad, M. (2015). Labor income dynamics and the insurance from taxes, transfer, and the family. *Journal of Public Economics, 127*, 58-73.

Blundell, R., Pistaferri, L., & Preston, I. (2008). Consumption inequality and partial insurance. *American Economic Review, 98*(5), 1887-1921.

Chen, M. (2005). Rethinking the informal economy: linkages with the formal economy and the formal regulatory environment. *WIDER Research Paper, No.2005/10, ISBN 9291906891*. Helsinki: The United Nations University World Institute for Development Economics Research (UNU-WIDER).

Dasgupta, S., Bhula-or, R., & Fakthong, T. (2015). *Earnings differentials between formal and informal employment in Thailand.* Bangkok: ILO Regional Office for Asia and the Pacific.

De Nardi, M., Fella, G., & Paz-Pardo, G. (2019). Nonlinear household earnings dynamics, self-insurance, and welfare. *Journal of the European Economic Association, 0*(0), 1-39.

Deaton, A., & Paxson, C. (1994). Intertemporal choice and inequality. *Journal of Political Economy, 102*(3), 437-467.

Gong, X., & Van Soest, A. (2002). Wage differentials and mobility in the urban labour market: a panel data analysis for Mexico. *Labour Economics, 9*(4), 513-529.

Guariglia, A., & Rossi, M. (2004). Private medical insurance and saving: Evidence from the British Household Panel Survey. *Journal of Health Economics, 23*, 761-783.

Hubbard, R. G., Skinner, J., & Zeldes, S. P. (1995). Precautionary saving and social insurance. *Journal of Political Economy, 103*, 360-399.

International Labour Organization (ILO). (2003). *Report of the Seventeenth International Conference of Labour Statisticians.* Geneva: ILO.

International Labour Organization (ILO). (2014). *World social protection report.* Geneva: International Labour Office.

International Labour Organization (ILO). (2018). *Women and men in the informal economy: A statistical picture (third edition).* Geneva: International Labour Office.

Judge, G., Griffiths, R., Hill, R., Lütkepohl, H., & Lee, T.-C. (1985). *The Theory and Practice of Econometrics* (second ed.). New York: Wiley & Sons.

Kambourov, G., & Manovskii, I. (2009). Accounting for the changing life-cycle profile of earnings. *Working Paper*.

Karahan, F., & Ozkan, S. (2013). On the persistentce of income shocks over the life cycle: Evidence, theory, and implications. *Review of Economic Dynamics, 16*, 452-476.

Kirdruang, P., & Glewwe, P. (2018). The impact of universal health coverage on households' consumption and savings in Thailand. *Journal of the Asia Pacific Economy, 23*(1), 78-98.

Lagakos, D., Moll, B., Porzio, T., Qian, N., & Schoellman, T. (2018). Life cycle wage growth across countries. *Journal of Political Economy, 126*(2), 797-849.

Limwattananon, S., Neelsen, S., O'Donnell, O., Prakongsai, P., Tangcharoensathien, V., van Doorslaer, E., & Vongmongkol, V. (2015). Universal coverage with supply-side reform: The impact on medical expenditure risk and utilization in Thailand. *Journal of Public Economics, 121*, 79-94.

Meghir, C., & Pistaferri, L. (2004). Income variance dynamics and heterogeneity. *Econometrica, 72*(1), 1-32.

Mincer, J. (1974). *Schooling, Experience and Earnings.* New York: National Bureau of Economic Research.

National Statistical Office. (2017). Report on the 2017 survey of the older persons in Thailand. Bangkok.

National Statistical Office. (2019). The Informal Employment Survey 2019. Bangkok.

Newey, W. (1987). Efficient estimation of limited dependent variable models with endogenous explanatory variables. *Journal of Econometrics, 36*, 231-250.

Rejda, G. E. (2012). *Social insurance and economic security.* New York: M.E.Sharpe.

Starr-McCluer, M. (1996). Health insurance and precautionary savings. *American Economic Review, 86*, 258-295.

Storesletten, K., Telmer, C., & Yaron, A. (2004). Consumption and risk sharing over the life cycle. *Journal of Monetary Economics, 51*, 609-633.

Tansel, A., & Kan, E. (2012). The formal/informal employment earning gap: Evidence from Turkey. *MPRA Paper No.38498*.

Tauchen, G. (1986). Finite state Markov-chain approximation to univariate and vector autoregressions. *Economics Letter, 20*, 177-181.

Ushijima, K. (2020). The impact of Universal Health Coverage Scheme on precautionary savings: Evidence from Thailand. *forthcoming*. doi:10.1086/703236

World Health Organization (WHO). (2011). Sustainable health financing structures and universal coverage. *Resolution No. WH64.9 adopted by 64th World Health Assembly*. Geneva.