AY 2016

Analysis of Returns to Schooling: Empirical Evidence from Thailand

UPALAT KORWATANASAKUL

Major in International Studies 4011S303-3

GRADUATE SCHOOL OF ASIA-PACIFIC STUDIES,
WASEDA UNIVERSITY

PROF. NOBUHIKO FUWA

C.E. PROF. KAZUO KURODA D.E. ASSOC PROF. KAORU NABESHIMA

PROF. MASAHIKO GEMMA

Acknowledgement

This dissertation is made possible by all generous supports and comments from various professors. First and foremost, I would like to thank my academic advisors Professor Nobuhiko Fuwa and Professor Kazuo Kuroda. It has been an honour to be a PhD student under their supervision. I appreciate all their contributions, including time, effort, and thoughtful advice and suggestion on academic and non-academic matters. Furthermore, I owe my deep gratitude to Associate Professor Kaoru Nabeshima and Professor Masahiko Gemma for all valuable comments and suggestions to improve my research.

The members of Fuwa seminar and Kuroda seminar have contributed immensely to my personal and professional development. The seminars have been a source of friendship as well as knowledge. I am especially grateful to Sebastian, Harue, Yen-Tsung, Daniel, Utsumi-san, and Diana for all joys, supports, warmth, and loves.

I gratefully acknowledge all organisations that have been supporting my research and career development. Those organisations include the GSAPS office, MEXT, National Statistics Office (Thailand), especially K.Sukanya (คุณสุกัญญา), K.Kanyawee (คุณกัญญาวีณ์), and Director of Statistical Forecasting Bureau, and the Waseda Writing Center. MEXT Scholarship Program serves as the main funding source that made my PhD study possible.

My time at Waseda University is made enjoyable in large part due to many friends which become a part of my life. I am indebted to Ben, Nuch, Jan, P'Tip, P'Tong, Silp, Shino, and Baek-san who support me in every aspect of my life. Especially Jacinta Bernadette Rico (JB) and Kenji Kaneshiro, I would like to thank them for being my best friends, and also being a wonderful sister and brother of all time.

Lastly, I would like to thank my family for all their love and encouragement. For my parents who raise me with love and support me in all my pursuits. For my brother who takes care of me and makes my life easier. For Kaneshiro family, Kawanishi family, and Nishiyama family which have been supporting me as if I were one of their family members. Thank you so much. And lastly for Shin-chan, thank you very much to walk into my life. With your love and support, everyday has been and will be wonderful.

Upalat Korwatanasakul

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2SLS	Two stage least squares
ATE	Average treatment effect
CCT	Conditional cash transfer
CPI	Consumer price index
DZ	Dizygotic twins
ESDP	Eastern Seaboard Development Programme
FAOSTAT	Food and Agriculture Organization of the United Nations, Statistics Division
FRDD	Fuzzy regression discontinuity design
IV	Instrumental variable
LATE	Local average treatment effect
LFS	Labour Force Survey
MZ	Monozygotic twins
NSO	National Statistical Office
OLS	Ordinary least squares
RCT	Randomised controlled trial
RDD	Regression discontinuity design
UNDP	United Nations Development Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization
WT	Within twin pair
US	United States
UK	United Kingdom
GDP	Gross domestic product
GER	Gross enrolment ratio

CHAPTER 1. INTRODUCTION

The rates of return to schooling is one of the most important topics that economists have been investigating, especially by utilising the Mincerian model. However, there has been a long debate that the ordinary least squares (OLS) estimates from the Mincer equation is possibly biased due to an endogeneity problem. Hundreds of studies with different methods of estimation attempt to deal with the endogeneity bias but they fail to establish a causal effect of schoolings on earnings in the absence of randomised experiment. Even though a randomised controlled trial (RCT) is the ideal estimation method, it is not feasible to conduct the RCT in most of the studies of returns to education. The second-best candidate is the estimation with quasi-experiment, e.g. regression discontinuity design (RDD), differences-in-differences, to name a few. However, the studies utilising those methods of estimation are rare, especially in developing countries where the issue of data scarcity is prevalent. The 1978 compulsory education law change in Thailand offers an unusual event which can be used as a quasi-experiment in estimating the returns to education.

Section 1. RESEARCH QUESTIONS

This study exploits the opportunity of quasi-experiment in Thailand that occurs from the change in compulsory education law to answer these following research questions:

1. Is there a causal relationship between education and earnings? If yes, how large is the effect of education on earnings?
2. What is the most dominant source of bias in the OLS estimation?
3. What is an interplay between the rates of return to schooling and economic development? What is a role of different stages of economic development in explaining a magnitude and a direction of a bias of the estimated returns? And what is a role of the estimated returns in explaining the development process?
4. How does heterogeneity in individuals' demographic characteristics affect the returns to schooling and the magnitude and the direction of the endogeneity bias?

Section 2. RESEARCH OBJECTIVES

First, this study aims to establish the causal relationship between education and earnings by utilising the classic Mincer model. The second objective is to interpret the estimated rates of return to schooling and their bias, both in terms of the magnitude and direction, in the context of Thai social and economic development. Third, this study tries to examine the effect of heterogeneity in demographic characteristics on returns to schooling.

Section 3. RESEARCH CONTRIBUTIONS

This study makes three main contributions in terms of methodology and also substantive aspect in the context of Thailand and, by implication, developing countries in general.

First, the previous literature reveals that there is a quite different pattern of the relative magnitudes of the estimates from OLS and instrumental variables (IV) estimation using compulsory schooling as IV between developed and developing countries. Investigating this difference can contribute to a better understanding of (a) how and when the conventional "ability bias" matters in estimating returns to schooling and (b) the impact of compulsory schooling in different settings.

The second contribution is in terms of substantive aspect in the context of Thailand. As Thailand experienced rapid economic development and structural transformation during 1960-1990, obtaining the rates of return to education in this period helps us better understand the process of Thai economic development as well as the interplay between the rates of return to schooling and the economic development process during 1980 to 1990. The overall social and economic conditions of the development in Thailand are consistent with the general characteristics of other developing countries. Hence, estimating the rates of return to schooling in Thailand, by implication, also provides better understandings on the role of human capital in the process of development in other developing countries. Due to the fact that developing countries possess radically different degrees of market completeness and different quality of institution from those of developed countries, this warrants value for investigation of the returns to schooling in the context of developing countries. This further investigation possibly gives a different economic pattern and implications of the returns to schooling.

Finally, the third contribution is on the construction of the database and discussion of the descriptive analysis for the discrepancies among different demographic characteristics, including gender, cohort, area of residence, region of residence, and economic sector. In addition to the overall estimates of returns to schooling, another important issue is an issue of heterogeneity in educational returns across individuals. Heterogeneity in individuals' demographic characteristics tends to distort the returns to education; for example, the female rates of return to education is likely to be higher than those of male. Hence, it is worth examining heterogeneous returns to schooling from different demographic characteristics.

Section 4. RESEARCH METHODOLOGY

To estimate the rates of return to schooling, this study exploits an opportunity of quasi-experiment in Thailand which occurs from a change in compulsory education law in 1978. The main estimation method is the IV estimation using the pooled cross-sectional

Labour Force Survey (LFS) data from 1986 to 2012.

Fundamentally, the estimation model is based on the Mincer wage equation. The presentation of the models is organised in terms of the analytic order of IV estimation.
First stage least squares regression:

$$\boldsymbol{S_i = \pi_0 + \pi_1 F_i + \pi_2 A_i + \pi_3 A_i^2 + \pi_4 C_i + \pi_5 R_i + \varepsilon_i} \tag{1-1}$$

Reduced form:

$$\boldsymbol{\log y_i = \alpha_0 + \alpha_1 F_i + \alpha_2 A_i + \alpha_3 A_i^2 + \alpha_4 C_i + \alpha_5 R_i + \theta_i} \tag{1-2}$$

OLS regression:

$$\boldsymbol{\log y_i = \beta_0 + \beta_1 S_i + \beta_2 A_i + \beta_3 A_i^2 + \beta_4 C_i + \beta_5 R_i + e_i} \tag{1-3}$$

Second stage least squares regression:

$$\boldsymbol{\log y_i = \gamma_0 + \gamma_1 \widehat{S_i} + \gamma_2 A_i + \gamma_3 A_i^2 + \gamma_4 C_i + \gamma_5 R_i + \vartheta_i} \tag{1-4}$$

Where: Log y_i is log of monthly wages of individual *i*. S_i refers to years of education an individual *i* attended, while $\widehat{S_i}$ refers to the fitted value estimated from the first stage least squares regression. F*i* represents a dummy variable indicating whether an individual complying with the 1978 compulsory education law. Control variables include age of an individual *i* as a proxy of working experience (A*i*), birth cohort dummies (C*i*), and regional dummies (R*i*). ε_i, θ_i, e_i, and ϑ_i represent disturbance terms.

Section 5. RESEARCH FINDINGS

First, the IV estimation indicates that the coefficients of years of schooling are statistically significant and robust across different specifications. This confirms a causal relationship between education attainment and earnings. One additional year of schooling leads to approximately 8 per cent increase in monthly wages.

Second, the result shows that the OLS estimates are greater than those of IV around 3 per cent. This indicates that the net effect of different sources of endogeneity bias leads to the overestimated rates of return to schooling in the OLS regression. In the context of Thailand, the ability bias outweighs other sources of bias, including the discount rate bias and the measurement error bias.

Third, the dominance of ability bias is mainly explained by the inequality of income and educational opportunity during the early period of social and economic development. There are two sources of ability bias, including the ability bias from a selection of more-able child within poor households and the ability bias due to higher financial endowment in rich families. This finding may possibly be generalised to the case of other developing countries, which share similar social and economic context with Thailand.

Fourth, with the disaggregated analyses by different demographic characteristics, it is generally observed higher returns to schooling for women, older cohort, urban area, the North and Northeast regions, and service sector. These results can be explained in the context of Thai development process. In addition, higher bias gap is prevalent in subsamples of women, older cohort, rural area, the North and Northeast regions, and agricultural sector. Again, income and educational opportunity inequality during the early period of social and economic development help explain the higher bias gap in those socially disadvantaged groups. With less educational opportunity and high poverty, socially disadvantaged households tend to choose only most-able children to send to school since they cannot bear costs of education for all children in a household. This is consistent with the general ability bias hypothesis that argues more-able individuals tend to have higher schooling.

Section 6. ORGANISATION OF THE DISSERTATION

This dissertation consists of eight chapters. Chapter 2 provides readers with a comprehensive overview of Thai economic and educational development during 1960-1990. Chapter 3 revisits the Mincer model to show its validity and provides a comprehensive literature review in terms of potential biases from various methods of estimation in the study of returns to education. Chapter 4 presents a general theoretical framework of the analysis of returns to education. Chapter 5 discusses the data, methodology, and identification strategy employed in this study. The empirical results with the discussions from IV estimation and disaggregated analysis of returns to schooling are provided in Chapter 6 and Chapter 7, respectively. Finally, the last Chapter concludes this dissertation, suggests policy implications, discusses research limitations, and provides a future research direction.

CHAPTER 2. THAI ECONOMIC AND EDUCATIONAL DEVELOPMENT

This chapter intends to provide readers with a comprehensive overview of Thai development in terms of economics and education during 1960-1990. The overall background helps understanding the interpretation and discussion of empirical results in the later chapters. The organisation of this chapter is as followings: Section 1 describes the overall economic and educational development in Thailand and shows that the situation of the development in Thailand is consistent with the general characteristics of developing countries suggested by Behrman and Srinivasan (1995). Section 2 provides a background of the demand side and supply side of Thai education. Section 3 portrays a detailed situation of migration in Thailand. Finally, Section 4 concludes this chapter.

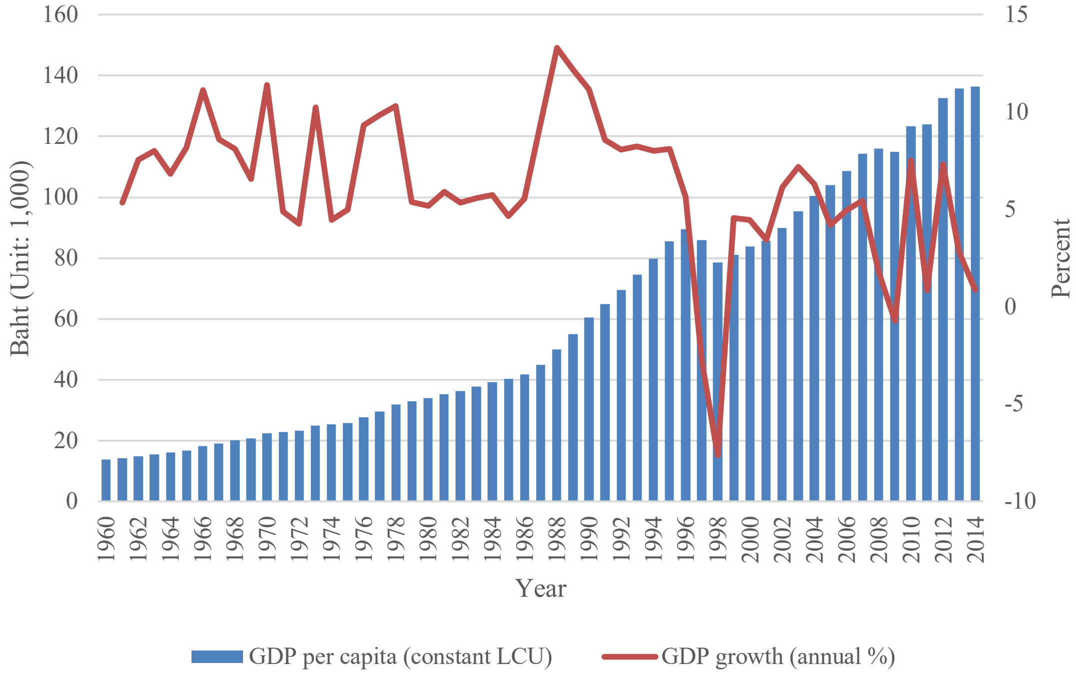
Section 1. THE PROCESS OF DEVELOPMENT IN THAILAND

This section provides a background of the process of development in Thailand, both in terms of economics and education, and illustrates that the situation of the development in Thailand is consistent with the characteristics of developing countries proposed by Behrman and Srinivasan (1995). They suggest that developing countries possess different degrees of market completeness and different quality of institutions from developed countries. While market with a highly competitive level and advanced institutions are observed in developed countries, malfunctioned market and institutions are prevalent in developing countries. The followings are the proposed characteristics of developing economies in comparison with developed ones. First, a majority of population still depends heavily on agricultural sector and rural labour activities. Second, in agricultural sector, it is common for family members to help working without formal pay; therefore, non-wage labour accounts for a large proportion of the total labour force in developing economies. Third, there has been a rapid growth of labour force in developing countries. Fourth, high labour participation rates among 15-64 year olds can be observed due to low human capital investments among young cohorts. Fifth, there are low school enrolment rates and the education gap is in favour of males. Sixth, lower non-labour production inputs per worker are observed in developing countries.

2.1.1. Overall Economic Development and Structural Transformation

Thailand has been one of the fastest growing countries in the world. According to **Figure 2-1**, from 1960 - 2014, the average annual growth rate of real GDP is 6.16 per cent. GDP per capita has been growing substantially during the past decades. Thailand reached its peak in 1988 with 13 per cent of real GDP growth caused by second structural transformation, export-oriented policy, and hit its bottom in 1998 with a negative growth at -7.6 per cent due to the 1997 Asian financial crisis originated from Thailand, so-called Tom Yum Kung crisis. A standard of living has been improving significantly as per capita GDP has been increasing substantially during the past fifty years. However, Thailand was caught in a middle-income trap for a long period of time. With a prolonged political unarrest, the future of Thai economy is still in gloom. From 2006 to present, severe fluctuations in Thai economy can be observed.

Figure 2-1 Real GDP Growth and Per Capita GDP in Thailand, 1960-2014



Source: Author's compilation based on data from World Development Indicators, 2016.

Table 2-1 Period of Economic Growth

Thailand Rates of Growth of GDP and GDP per capita, 1951-2003		
Period	Real GDP growth	Real GDP growth per capita
1951-1986 Pre-boom	6.5	3.9
1987-1996 Boom	9.2	8
1997-1998 Crisis	-6.1	-7.1
1999-2003 Post-Crisis	4.0	3.3
Whole period 1951 to 2003	6.2	4.2

Source: Siriprachai, 2009.

According to Warr (2005), Thai economy can be divided into four main periods, namely pre-boom (1951-1986), boom (1987-1996), crisis (1997-1998), and post-crisis (1999 onwards). **Table 2-1** shows different periods of economic growth in Thailand from 1951-2003. During 1951-1986, this period is characterised as a pre-boom period which the real GDP growth was approximately 6.5 per cent. In this period, the Government put efforts in building basic physical infrastructure, e.g. road, electricity, to name a few, to help facilitating trading and the growth of the economy. The real GDP growth rate jumped to 9.2 per cent during the boom period, 1987-1996. This is considered as an economic miracle growth in Thailand. However, Thailand experienced a financial crisis in 1997 and faced a negative GDP growth at 6.1 per cent, which is the lowest in the Thai economic history. In the post-crisis period, Thailand managed to bounce back and achieved a real GDP growth rate of 4 per cent. The overall average real GDP growth and real GDP growth per capita are around 6.2 per cent and 4.2 per cent, respectively, during 1951-2003.

Figure 2-2 demonstrates Thai income growth for the last five decades in a comparative East Asian perspective. Data on GDP per capita is presented for eight East Asian countries, including Thailand, the Philippines, Malaysia, China, India (South Asia)[1], Korea, Japan, and Singapore. Comparing with other countries, Thai economy is doing fairly well even after the financial crisis. **Figure 2-3** compares real GDP of East Asian countries by indexing level of real GDP to 100. This figure emphasises the fact that Thailand is one of the fastest growing economies during 1986-1996. Thailand shows the highest level of real GDP among other East Asian countries.

The high GDP growth is claimed to be the result of the structural transformation during the pre-boom period, 1951-1986. Thailand went through the structural transformation from a primitive agriculture-based economy to newly industrialised economy. According to **Figure 2-4**, the share in GDP of the agricultural sector sharply declined, from nearly 40 per cent in 1960 to 25.9 per cent in 1970. Owing to industrial development, a crossing point between the share of agriculture and manufacture could be observed. In 1971, it was the first time that the share of manufacture outweighed that of agriculture. The share of agricultural sector gradually dropped to 23.2, 12.8, and 8.5 in 1980, 1990, and 2000, respectively. On the other hand, the share of manufacturing sector raised from 20 per cent in 1960 to 25 per cent in 1970. It kept rising to 28.7, and 37.2 in 1980 and 1990, respectively. The share remained constant after 1990.

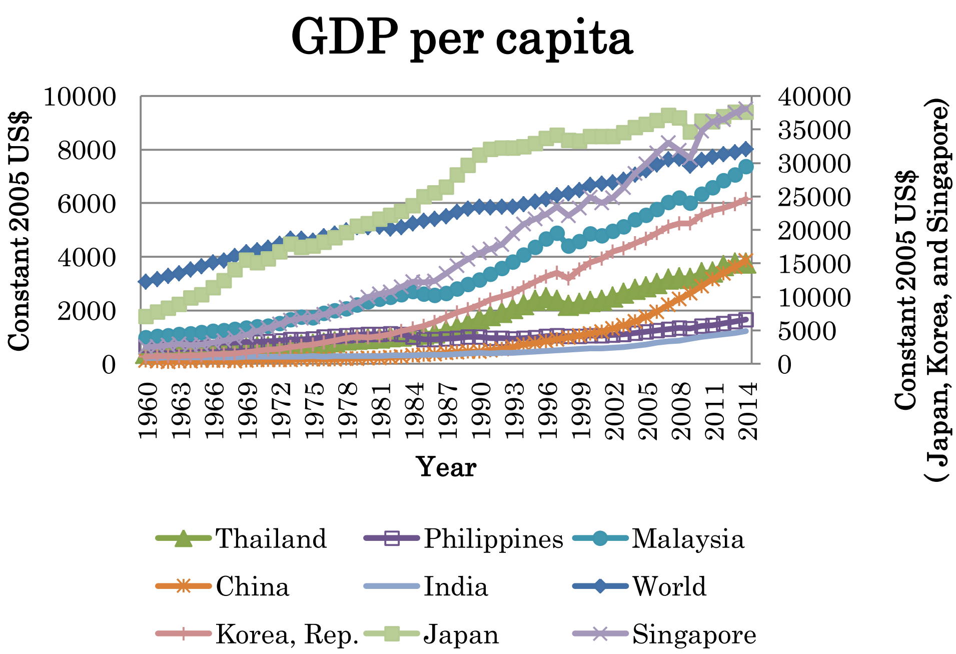
Table 2-2 summarises the sectoral composition of Thailand's shares of GDP and annual GDP growth performance into different decades of development. The agricultural sector played an important role in Thai economy and act as an engine of economic development in the 1960s and 1970s (Suphannachart and Thirawat). As previously discussed, the manufacturing sector surpassed the agricultural sector in 1971 and the role of manufacturing sector became even more important in 1980s and afterwards. In terms of GDP growth, the role of agricultural sector has declined substantially over time.

Table 2-3 indicates the sectoral composition of Southeast Asian countries' shares of GDP in 1981, 1990, and 2003. Although different countries experience different levels of economic

[1] Even though India is located in South Asia, it is included in the analysis as it is considered as a fast growing economy. It is interesting to compare the pattern of its growth to other East Asian countries.

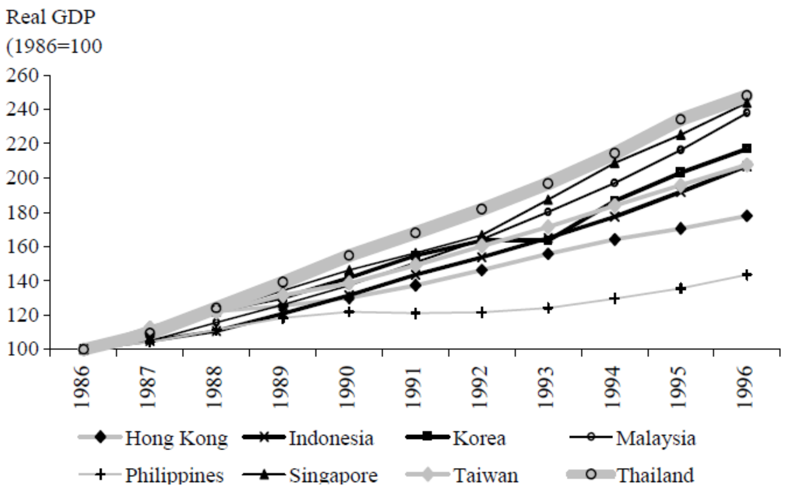
development, every country seems to go through a similar structural transformation. Except for Myanmar and Singapore, the share of GDP by manufacturing sector in every country increases over time, while a drop in the share of GDP by agricultural sector can be observed. Hence, Southeast Asian countries go through structural transformation from agriculture-based economies to industrialised economies in a similar period of time with different paces of development. The middle income and developing countries such as Thailand, Malaysia, Indonesia, and Vietnam have relatively greater structural change, compared to the developed and the least developed countries. For example, in Thailand, the agricultural share of GDP dropped from 21.4 per cent in 1981 to 12.5 per cent in 1990 and declined further to 9.7 per cent in 2003. As Singapore underwent the structural transformation long before other Southeast Asian countries, Singapore is already considered as a developed country moving forward to service-based economy. **Table 2-3** indicates that Singapore's GDP is mainly driven by the service sector, which is roughly 60 per cent of its GDP. On the other hand, Myanmar is considered as a least developed economy heavily based on primitive agricultural sector. Hence, even in a recent year, the share of agricultural sector is still high which constitutes half of its GDP. Overall, **Table 2-3** shows that the decline of agricultural share and the rise of industrial share can be observed simultaneously in most of the Southeast Asian countries. Similar to most of industrialisation process in other countries, this implies that the industrial sector is developed at the expense of the agricultural sector.

Figure 2-2 Income Growth in East Asian Countries during 1960-2014



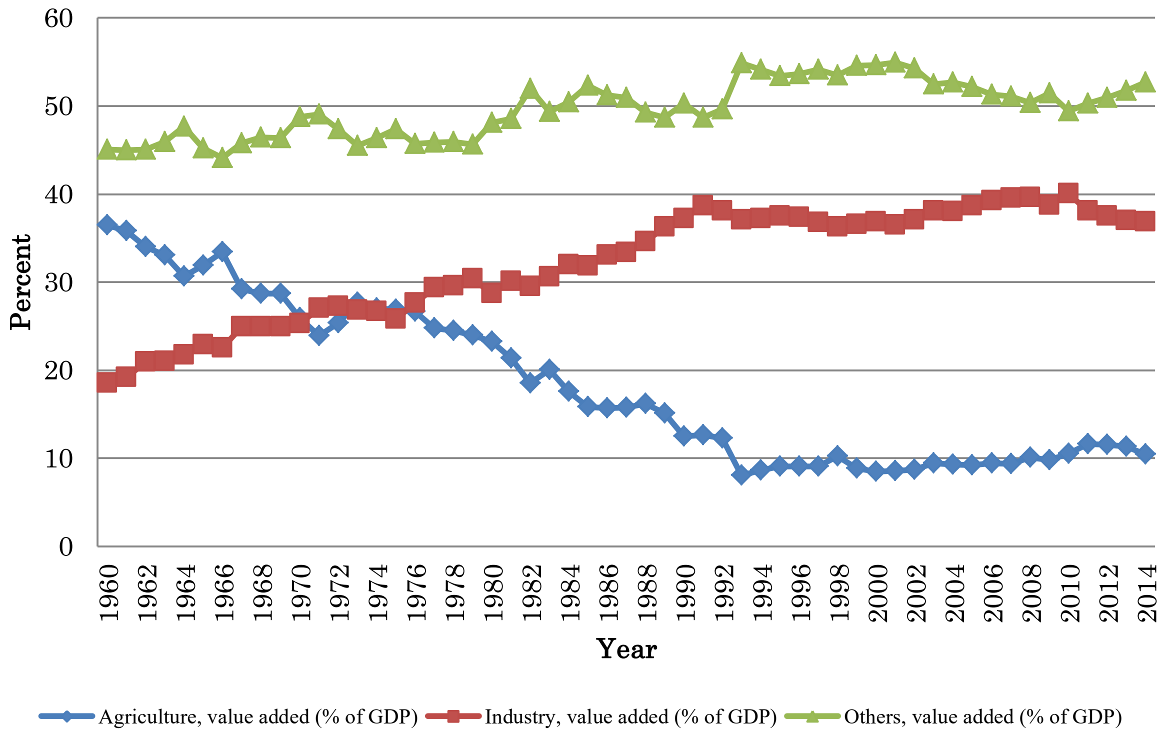
Source: Author's compilation based on data from World Development Indicators, 2016.

Figure 2-3 Real GDP in East Asia, 1986-1996



Source: Warr (2011) based on data from Asian Development Bank.

Figure 2-4 Structural Transformation: Net Output as % of GDP (%), 1960 - 2014



Source: Author's compilation based on data from World Development Indicators, 2016.

Table 2-2 Shares of GDP and Annual GDP Growth Rates by Economic Sector

	1960s	1970s	1980s	1990s	2000s
Shares of GDP (%)					
Agriculture	29.04	24.52	18.40	11.75	9.52
Manufacturing	16.62	20.81	23.98	30.89	38.30
Services	11.49	11.99	13.39	10.81	11.50
Annual GDP Growth Rates (%)					
Agriculture	6.67	4.49	4.27	1.72	1.68
Manufacturing	10.78	9.49	8.77	8.14	5.35
Services	7.96	7.87	7.04	4.40	4.39
All sectors	8.36	6.92	7.24	5.28	4.06

Source: Suphannachart and Thirawat based on data from National Economic and Social Development Board.

Table 2-3 Sectoral Shares in GDP (percentage)*

	Agriculture			Industry			Manufacturing**			Services		
	1981	1990	2003	1981	1990	2003	1981	1990	2003	1981	1990	2003
Brunei	na	na	na	na	na	na	na	na	na	na	na	na
Cambodia	na	55.6	37.2	na	11.2	26.8	na	5.2	19.3	na	33.2	36
Indonesia	23.9	19.4	16.6	41.2	39.1	43.6	12.1	20.7	24.6	34.9	41.5	39.8
Lao PDR	81.2	61.2	48.6	9.9	14.5	25.9	na	9.9	19.2	8.9	24.3	25.5
Malaysia	23.0	15.0	9.1	35.6	41.5	47.0	19.7	23.8	30.2	41.4	43.5	43.9
Myanmar	47.4	57.3	57.1	12.4	10.5	10.5	9.3	7.8	7.8	40.2	32.2	32.4
Philippines	24.9	21.9	14.4	39.2	34.5	32.4	25.5	24.8	22.9	35.9	43.6	53.2
Singapore	1.2	0.3***	0.1***	37.9	32.7	33.0	28.5	25.5	26.3	60.9	67.0	66.9
Thailand	21.4	12.5	9.7	30.1	37.2	44.0	22.6	27.2	35.2	48.5	50.3	46.3
Vietnam	55.0	38.7	21.8	25.0	22.7	40.0	na	12.3	20.8	20.0	38.6	38.2

Note: * The sectoral classification has been revised to be in accordance with the International Standard Industrial Classification of all Economic Activities (ISIC): Agriculture also contains the Forestry and Fishing sub-sectors; Industry includes the Mining and quarrying, Manufacturing, Electricity, gas, and water supply and Construction sub-sectors. All other sub-sectors are classified under Services. Also, apart from figures of Lao PDR in 1981 that are calculated from data at a constant 1986 factor cost, all figures are calculated from data at current market prices.

** Manufacturing is a component of Industry.

*** Comprises the Mining sub-sector, which is normally a component of Industry.; na: not available.

Source: Numnak (2006) based on data from Asian Development Bank (1999) for 1981 and Asian Development Bank (2004) for 1990 and 2003.

Table 2-4 Growth of GDP and its Sectoral Components in Thailand, 1951-2006

	Pre-boom 1968-86	**Boom 1987-1996**	**Crisis 1997-1999**	**Recovery 2000-2006**	**Whole Period 1968-2006**
Total GDP	6.7	9.5	-2.5	5.0	6.4
Agriculture	4.5	2.6	0.1	2.7	3.3
Industry	8.5	12.8	-1.7	6.2	8.4
Services	6.8	9	-3.6	4.3	6.1

Source: Warr (2011) based on data from Bank of Thailand (1951-1986) and National Economic and Social Development Board (1987-2006).

Table 2-5 Total Factor Productivity Growth by Sectors, 1980 to 2002

Average Growth Rates (Per Cent Per Annum)	**Aggregate**	**Agriculture**	**Manufacture**	**Services**
Output	6.01	2.64	8.09	5.53
Raw Labour	2.19	1.50	5.25	3.47
Human Capital	2.49	9.43	11.35	6.90
Physical Capital	9.05	8.50	13.84	18.47
Agricultural Land	1.12	1.12	0	0

Source: Author's compilation adjusted from Warr (2011) based on data from National Economic and Social Development Board.

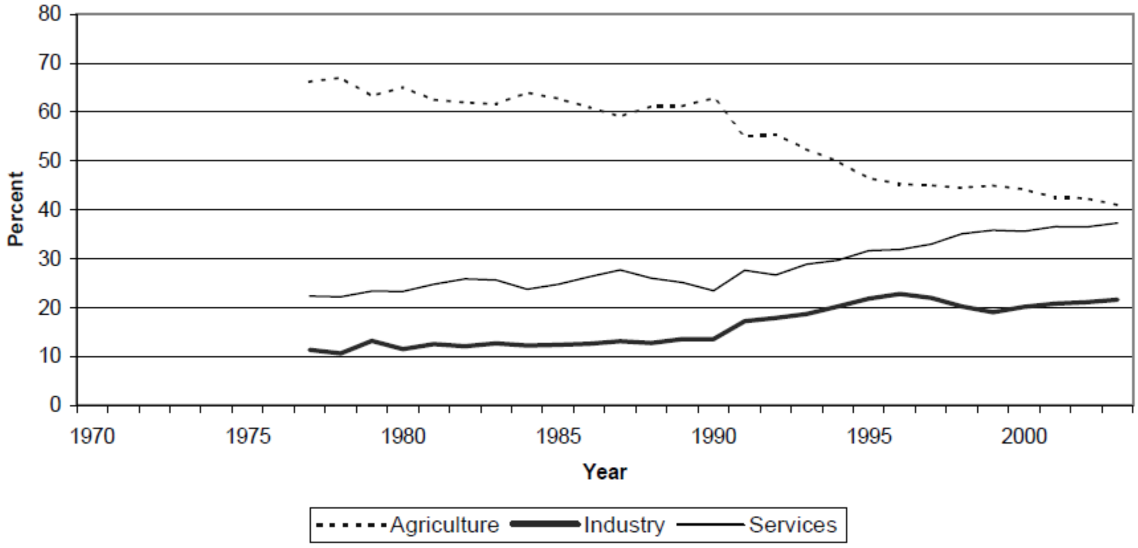
Table 2-4 provides similar information as that of **Table 2-2**. It depicts the sectoral composition of Thailand's annual GDP growth performance into different period of economic development. However, in **Table 2-4**, it is possible to observe resilience of each economic sector after the 1997 financial crisis. Agriculture seems to have the highest resilience as it manages to bounce back to the same growth level achieved before the crisis. However, the manufacturing and the service sectors take a longer time to get back to the growth level achieved before the crisis. Hence, the agricultural sector is the first sector to help stabilising the economic growth after the crisis. Even though the agricultural sector contributes to the GDP less than the manufacturing sector does and indicates the lowest annual GDP growth rate among all sectors, it seems that the agricultural sector has been playing an important role in Thai economic development even after the structural change.

The agricultural sector contributes not only to the economic stability but also to the development of other economic sectors. As previously argued, the manufacturing sector is developed at the expense of the agricultural sector. Resources from the agricultural sector such as labour force, land, and capital have been drawn to other sectors, especially the manufacturing sector, while the agricultural sector manages to maintain its output (Warr, 2011). Since resources from the agricultural sector are allocated to other sectors, there has been insufficient Government investment in research and development regarding domestic agricultural technology. Thus, farmers still use the same old technology in cultivation and slow growth can be observed in agricultural sector. Thai agricultural sector is still backward and traditional. During the period of social and economic development, the agricultural growth is solely driven by expanding cultivated areas (Siriprachai, 2009) rather than utilising agricultural technology advancement. Land productivity remains very low and stable, while labour productivity increases substantially (James, et al., 1987; Timmer, 1991; Watanabe, 1992). An increase in labour productivity leads to a contraction of agriculture as a share of total output. According to **Table 2-5,** it can be observed that the output growth from the agricultural sector (2.64 per cent) is lower than those from the manufacturing (8.09 per cent) and the service sector (5.53 per cent).

Thai agricultural sector plays another crucial role in economic development since it serves as an important source of income for a majority of Thai population. Behrman and Srinivasan (1995) suggest that a majority of population still depends heavily on agricultural sector and rural labour activities in developing countries. This also holds true in the case of Thailand. **Figure 2-5** indicates sectoral employment during 1977-2003. Even though the employment share in the agricultural sector has been declining over time, it still remains substantially larger than the other economic sectors, especially during the period of social and economic development, 1970-1997. This emphasises the fact that agriculture has a far larger role in the economy in terms of the employment. The declining share of agricultural labour is due to the effect of the structural transformation, which labour from the agricultural sector is allocated to the manufacturing sector, especially over the last ten years. The employment share in the agricultural sector fell from roughly 70 per cent in 1977 to 40 per cent in 2003. As a consequence, the shares of the employment in the manufacturing and the service sectors in 2003 have approximately doubled since 1977. Although both sectors' share of employment has increased gradually, they are still less dominant than the agricultural sector. These 30 percentage points of the declining share in the agricultural sector go to the manufacturing sector 20 points and the service sector 10 points. **Table 2-6** provides more detailed information on sectoral employment for modern economic period. It emphasises the importance of the agricultural sector in supporting Thai labour even in the modern economy. Even though its employment share dropped to 42 per cent in 2002, the agricultural sector is still the most dominant sector.

In addition, Thailand also goes through another structural transformation in terms of export and import pattern. Thailand changes from agriculture produce exporter, e.g. rice, to manufactured goods exporter, especially garments and parts and components. Siriprachai (2009) divides Thai industrial development into four phases based on the characteristics of import and export activities. The first period is the period of import substitution (1961-1971). From this period, the Government started to allocate a larger portion of the Government budget towards manufacturing sector, e.g. food processing. The expansion of manufacturing sector was also due to the allocation of other resources from the agricultural sector; therefore, it is often argued that the expansion of the manufacturing sector is at the expense of the agricultural sector. To promote import substitution, the Government protected domestic industries targeting the domestic market. However, the domestic market soon became saturated. Hence, in the second phase the Government shifted its emphasis to export-oriented industries, e.g. textiles. This phase refers to the period of export promotion (1972-1976). Due to the world recession, this industrial development strategy was not successful. The next phase is called the Big Push (1977-1982). The main Government strategy was to focus on development of necessary infrastructure and enhancement of domestic industry. The Eastern Seaboard Development Programme (ESDP) was initiated to build an industrial complex in the East area of Thailand. Moreover, the discovery of natural gas in the Gulf of Thailand helped contributing to the success of the large scale industrial development plan. Since 1983, the Government moved forward to the next phase of industrial development, manufacturing export led growth. In this phase, the Government implemented the development strategy in favour of foreign investors to attract foreign direct investment. Industries such as parts and components, automobiles, and electrical appliances demonstrated a high growth. This contributed mainly to the fast growth of Thai economy at that time. This phase of industrial development corresponded to the boom period of the overall economy. Thailand successfully underwent the second structural transformation from agriculture produce exporter to manufactured goods exporter.

Figure 2-5 Sector Employment Share, 1977-2003



Source: Bosworth (2005) based on data from National Accounts of Thailand, National Economic and Social Development Board.

Table 2-6 Labour Force Status: Thailand (Unit: thousands)

	1986	1987	1988	1989	1990	1995	1997	1998	2002
Labour Force	27,525	28,732	29,614	30,283	30,809	32,175	32,780	32,596	34,292
Employed Persons									
Total	25,220	26,174	27,727	28,061	28,812	30,815	31,714	30,775	33,133
Agriculture	16,070	15,659	17,379	17,020	17,129	14,389	14,315	14,056	13,792
Manufacturing	2,300	2,739	2,611	3,104	3,322	4,609	4,644	4,556	5,257
Construction	678	817	809	947	2,649	2,248	2,502	1,661	1,701
Commerce	2,707	3,086	3,031	3,063	2,935	3,909	4,207	4,257	4,989
Employed Persons (%)									
Agriculture	63.72	59.83	62.67	60.65	59.45	46.69	45.14	45.67	41.63
Manufacturing	9.12	10.46	9.42	11.06	11.53	14.96	14.64	14.80	15.87
Construction	2.69	3.12	2.92	3.37	9.19	7.30	7.89	5.40	5.13
Commerce	10.73	11.79	10.93	10.92	10.19	12.69	13.26	13.83	15.06
Unemployed Persons	1,445	1,795	1,277	1,178	1,061	550	495	1,423	766
Unemployment Rate (%)	5.25	6.25	4.31	3.89	3.44	1.72	1.53	4.37	2.24

Source: Krongkaew, Chamnivickorn, and Nitithanprapas (2006) based on data from Labour Force Survey 1986-2002, National Statistical Office and Thailand Development Indicators, 1990-1999, National Economic and Social Development Board.

2.1.2. Non-Wage Labour and Informal Sector[2]

Table 2-7 Thai Labour Market: Formal and Informal Employment (Per Cent)

	1999	2000	2001	2002
Employment rate	93.7	94.2	94.8	96.4
Share of Employment				
Formal Sector	26.7	28.1	27.5	27.9
Informal Sector	73.3	71.9	72.5	72.1

Source: Krongkaew, Chamnivickorn, and Nitithanprapas (2006) based on data from Labour Force Survey.

Table 2-7 depicts the fact that the informal sector plays an important role in Thai labour market as it dominates more than 70 per cent of the total employment. The share of informal sector is expected to be even higher during the early stage of social and economic development. The informal sector includes both own account workers and informal employers

[2] According to ILO (1993), informal sector is defined as, "The informal sector is broadly characterised as consisting of units engaged in the production of goods or services with the primary objective of generating employment and incomes to the persons concerned. These units typically operate at a low level of organisation, with little or no division between labour and capital as factors of production and on a small scale. Labour relations - where they exist - are based mostly on casual employment, kinship or personal and social relations rather than contractual arrangements with formal guarantees".

such as unpaid family workers, which are usually characterised as non-wage labour. **Table 2-8** indicates employment by work status during 1985-2002. In 1985, the share of unpaid family workers amounts to 42.9 per cent but decreases to 25.6 per cent in 2002. Although the decline of the share of unpaid family workers can be observed, it still covers a large proportion of employment even in the modern economy. During the same period, the share of own account workers remains stable at roughly 30 per cent. Hence, **Table 2-8** points out the quantitative importance of non-wage labour since the share of both unpaid family worker and own account worker involves a majority of Thai employment.

Table 2-8 Employment by Work Status, 1985-2002 (Per Cent)

	1985	1990	1995	1996	1997	1998	1999	2000	2001	2002
Employer	1.0	1.2	2.9	2.5	2.3	2.5	2.9	3.4	2.9	3.2
Government Employee	6.2	6.0	7.5	7.1	7.3	8.4	8.6	8.2	8.5	7.8
Private Employee	19.2	22.5	28.2	30.6	30.3	28.1	29.7	31.4	31.9	32.2
Own Account Worker	30.7	29.8	30.2	30.9	29.8	31.2	31.7	30.1	32.0	31.2
Unpaid Family Worker	42.9	40.5	31.2	28.9	30.3	29.8	27.1	26.9	24.7	25.6

Note: 1) Data are extracted from tables titled 'Employed Persons by Work Status, Industry and Sex' or 'Number of Employed Persons by Work Status, Industry and Sex (Quarter 3)' from the source.
2) From 2001, Own Account Workers also included those in Member of Producers' Cooperatives.
Source: Author's compilation adjusted from Numnak (2006) based on data from National Statistical Office (1987-2002).

It is generally observed that non-wage labour has lower income. **Table 2-9** indicates that almost half of the agricultural labour lived below the poverty line in 1988. Even though headcount ratio of agricultural sector has been declining over time due to the industrial growth, the poverty is still dominant in the agricultural sector. On the other hand, headcount ratios from the other economic sectors are relatively low and demonstrate a declining trend over time. In 2002, less than 10 per cent of labour in the other sectors lived below the poverty line. Although the agricultural sector has the highest resilience among all economic sectors, agricultural labour seems to be substantively vulnerable to the economic shock. After the 1997 financial crisis, headcount ratio in the agricultural sector increased from 19.2 per cent in 1996 to 26.2 per cent in 2002. In contrast, headcount ratios of the other sectors went up around 1-2 per cent and dropped back to the same level in 2000, except for the construction sector which suffers severely from the crisis. As the majority of agricultural labour are either classified as own account worker or unpaid family worker, it can be concluded from **Table 2-9** that, on average, non-wage labour is poorer than wage labour and more vulnerable to economic shock. **Table 2-10** reinforces the previous discussion. It shows that headcount ratios of unpaid family worker and own account worker are the highest among all working status. Nevertheless, in more recent years, the poverty of own account worker becomes less dominant.

In conclusion, Thailand faces the same situation as other developing countries where informal sector and non-wage labour, especially unpaid family workers in agriculture at lower incomes, is much more important (Behrman and Srinivasan, 1995), at least quantitatively as the share of both unpaid family worker and own account worker accounts for a majority of Thai employment. Moreover, non-wage labour tends to have lower income and tend to be more vulnerable to exogenous shocks.

Table 2-9 Headcount Ratios by Types of Enterprises

Types of Enterprises	**1988**	**1990**	**1992**	**1994**	**1996**	**1998**	**2000**	**2002**
Agricultural	45.7	38.6	35.1	25.9	19.2	22.5	26.2	17.6
Manufacturing	16.5	17.6	17.0	7.9	6.8	8.2	6.8	6.3
Construction	15.4	10.8	1.4	4.5	1.4	0.4	6.8	0.6
Trades	11.1	8.2	6.3	3.6	1.6	3.1	3.3	2.2
Services	14.9	11.6	7.3	5.5	3.5	4.3	2.9	2.4

Source: Krongkaew, Chamnivickorn, and Nitithanprapas (2006) based on data from Socio-Economic Surveys

Table 2-10 Headcount Ratios by Employment Status

Working Status	**1994**	**1996**	**1998**	**2000**	**2002**
Employer	15.8	12.2	15.4	18.3	9.3
Own account worker	16.4	10.8	10.4	11.4	9.8
Private employee	8.1	4.6	6.0	5.4	4.8
Government employee	0.9	0.6	0.5	0.9	0.7
Unpaid family worker	25.3	17.5	20.7	25.0	16.3
Unemployed	20.7	8.4	9.5	na	6.7
Economically inactive	13.5	8.6	10.7	10.5	8.0
No occupation	11.6	9.1	14.4	8.4	7.1
Child under 15 and no income from working	21.9	15.2	16.8	19.1	13.2
Total	16.3	11.4	13.0	14.2	9.8

Note: na: not available.
Source: Krongkaew, Chamnivickorn, and Nitithanprapas (2006) based on data from Socio-Economic Surveys

2.1.3. Labour Force and Labour Participation Rate

The other characteristics of developing countries are rapid growth of labour force and high labour force participation. This is due to the fact that young cohort starts working at early ages. Thailand seems to follow these two patterns as well.

Table 2-11 shows that, compared to developed countries such as the US and Japan, the growth of labour force is greater in Thailand. During the fast-growing periods of the US

(1950-60) and Japan (1960-70), labour force growed around 1 per cent in both countries. On the other hand, Thai labour force growth reached 3.4 per cent during the pre-boom and 2.85 per cent during the boom period. In terms of employment, there has been a high growth during the boom period, which is around 4.7 per cent. **Table 2-12** summarises labour force growth by industry and shows that labour force growth is highest in the manufacturing sector during the pre-boom and boom periods. In contrast, the growth of labour force in the agricultural sector is fairly slower. After 1990, negative growth can be observed in the agricultural sector. This is due to labour force reallocation from the agricultural sector to the manufacturing sector during the rapid economic expansion.

Table 2-12 supports the fact that there is a high labour force participation rate in Thailand. The overall labour participation rate and the rate of 15-24 years old move along each other and remain constant for a decade. However, the rate for 15-24 years old starts to decline thereafter. This is possibly a result of educational opportunity expansion. As the decline of participation rate of labour aged 15-24 years begins, the overall labour participation rate also decreases simultaneously. Although the declining portion of the rate for 15-24 years old does not exactly match with that of the overall rate, it is possible to conclude that the high participation rate is partly due to the high participation of the young cohorts. During 1980-1990, high participation rate among labour aged 15-24 years can be observed in Thailand, which means that there is low enrolment in secondary school and higher education. This statement is supported by **Figure 2-7** which illustrates that students tend to enrol in school only until the final grade of primary school (grade 6) and that significantly low numbers of students enrol in secondary school.

Table 2-11 Growth of Thailand's Economy in Different Areas, 1970-2003

Period	GDP	Investment	Capital Stock	Employment	Labour Force		
					Thailand	the US	Japan
1950-60	na	na	na	na	na	1.10	1.3*
1960-70	na	na	na	na	1.92	1.60	1.4**
1970-80	6.7	6.4	5.2	2.1	3.41	2.00	na
1980-90	7.8	10.9	8.2	4.7	2.85	1.20	na
1990-2000	4.5	-2.4	7.9	2.8***	na	1.00	na
1970-1996	7.5	8.9	7.8	2.8	na	na	na
1996-2003	1.5	-8.3	2.0	1.3	na	na	na

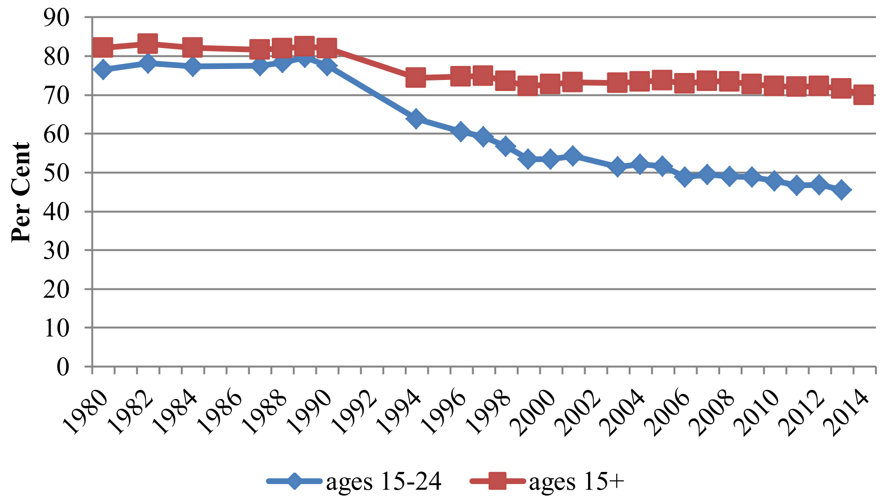
Note: Employment data begin in 1971 and are based on the annual average of a varying number of survey rounds. The capital stock is available only through 2002.
* 1959-62; ** 1962-1967;***1990-1997; na: not available.
Source: Bosworth (2005) based on data from National Accounts of Thailand, National Economic and Social Development Board. Labour force data are from Sarntisart (2000) for Thailand, Toosi (2002) for the US, and Tachi and Yoichi (1969) for Japan.

Table 2-12 Growth of Labour Force by Industry

	1970	1980	1990	1991	1992	1993	1994	1995
Total	1.92	3.41	2.85	-5.26	5.39	-0.37	-1.68	2.16
Agriculture	1.54	2.45	1.61	-16.94	5.62	-5.98	-6.69	-5.21
Manufacturing	3.78	6.72	9.12	16.77	7.51	6.26	0.28	9.96
Other	3.47	6.41	4.48	14.94	4.12	7.06	5.49	9.50

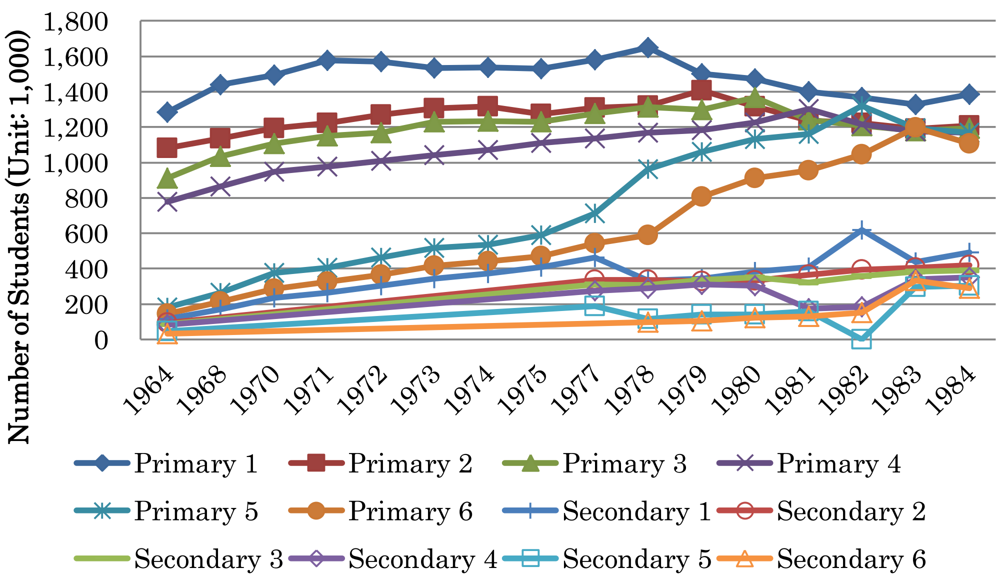
Source: Author's compilation adjusted from Sarntisart (2000) based on data from Labour Force Surveys of multiple years.

Figure 2-6 Labour Force Participation Rate (%)



Source: Author's compilation based on data from World Development Indicators, 2016.

Figure 2-7 Number of Students Enrolment by Education Level (Grade), 1964-1984



Source: Author's compilation based on UNESCO and Ministry of Education (1968-1986).

2.1.4. Human Capital Investment and Gender Inequality

In Thailand, human capital investments are fairly low. According to **Table 2-13**, a majority of labour force, 79 per cent, completes only primary education during the boom period. Roughly 6 per cent still has no education. Even though there is an increase in the share of labour force attaining secondary education and higher education over time, the level is still unsatisfactory even in a more modern economic era. In 2002, 23.6 per cent of labour force possesses secondary education degree, while 11.7 per cent obtain higher education. This leaves the majority of labour force with at most a primary education.

Table 2-13 Percentage of Labour Force by Level of Education Attainment since 1986

	1986	1990	1995	2000	2002
No Education	5.8	5.4	4.1	3.4	3.3
Primary and less	79.0	78.1	74.7	64.6	61.3
Secondary	9.9	11.0	14.3	21.5	23.6
Higher education	4.8	5.4	6.8	10.5	11.7
Other	0.5	0.1	0.1	0.03	0.1

Note: 1) Data are computed from figures in tables titled 'Population 11 Years and Over by Level of Education and Sex', 'Population 13 Years and Over by Level of Education and Sex', or 'Population 15 Years and Over by Level of Education and Sex' of the source. 2) Person in labour force, 11 years of age and over in 1986, 13 years of age between 1990 and 2000, and 15 years of age in 2002.
Source: Numnak (2006) based on data from National Statistical Office (1989; 1993; 1996; 2000b; 2002).

Table 2-14 shows adult literacy rate and net enrolment ratio by region and selected Asian countries in 1997. By international standards, the net secondary enrolment rate in Thailand is in a severe situation. The rate is 47.6 per cent which is the second to the bottom in the region and is lower than the regional average and the world average by roughly 10 per cent and 17.8 per cent, respectively. In contrast, the secondary enrolment rates in developed countries such as Japan and Korea are 99.9 per cent and that of Singapore is 75.6 per cent which is one of the highest rates in the region. It seems that Thai economic growth miracle is driven by a small portion of highly educated human resource. However, high literacy rate owing to universal primary education helps prepare a majority of labour force to function properly in the industrialised process. With low enrolment rate in secondary education and higher education, it seems that Thai social and economic development may not be sustainable. As commonly observed in other Asian and developing countries, there also has been a gender gap in educational opportunity and labour market in Thailand. It is always the case that human capital investments are in favour of males.

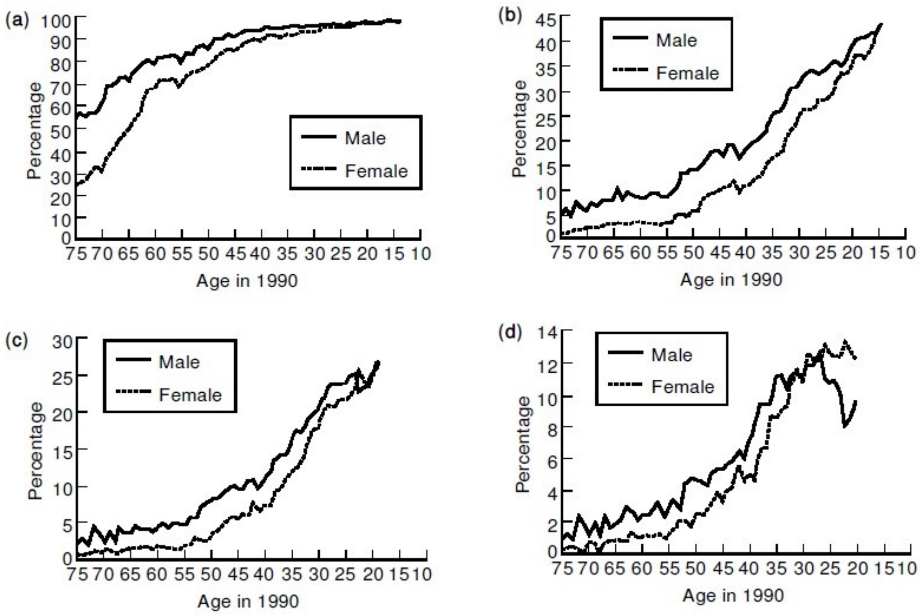
Figure 2-8 indicates that percentage of males who complete a certain educational level is higher than that of females in every level of education, including four years of primary education, lower secondary education, upper secondary education, and higher education. Hence, there is discrimination against women in education at all levels. Gandhi-Kingdon (2002) explains this situation as "unexplained parental discrimination". Given the same level of ability for women and men, women are likely to get less support from their families. In the early stage of social and economic development, market distortions such as gender discrimination prevent women to attend schools. The gender gap in education persists over time and is widest for older cohorts. However, the situation in Thailand has been improving and the gap seems to be closed at all level of education since 1990.

Table 2-14 Adult Literacy Rate and Net Enrolment Ratio by Region and Selected Asian Countries, 1997

	Adult Literacy Rate (%)	Net Primary Enrolment Ratio (%)	Net Secondary Enrolment Ratio (%)
World	na	87.6	65.4
Southeast Asia and the Pacific	na	97.8	58.3
Japan	na	99.9	99.9
South Korea	97.2	99.9	99.9
Singapore	91.4	91.4	75.6
China	82.9	99.9	70.0
India	53.5	77.2	59.7
The Philippines	94.6	99.9	77.8
Malaysia	85.7	99.9	64.0
Thailand	94.7	88.0	47.6
Indonesia	85.0	99.2	56.1
Lao PDR	58.6	73.0	63.4
Cambodia	na	99.9	38.8
Vietnam	91.9	99.9	55.1
Myanmar	83.6	99.3	54.2

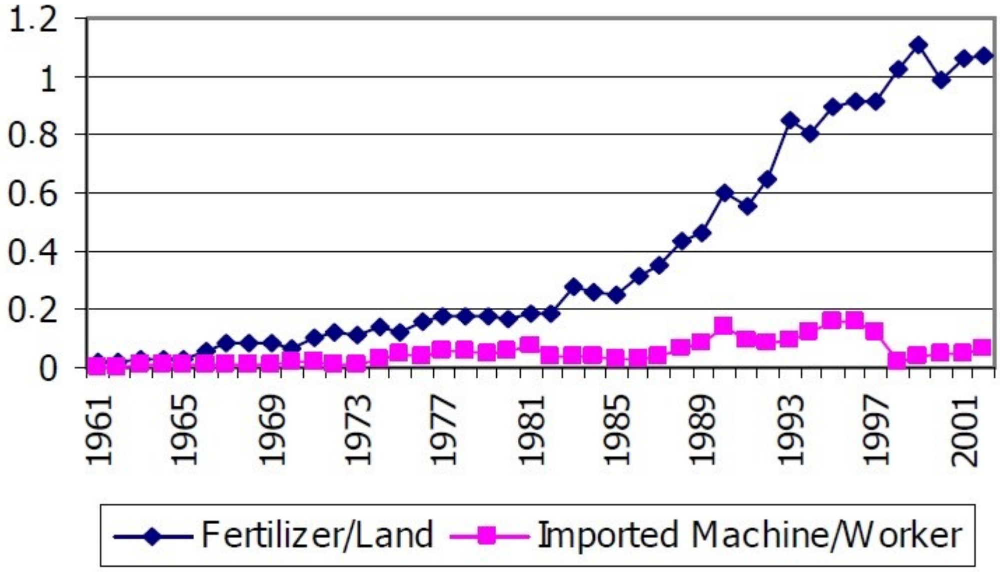
Source: Numnak (2006) based on data from UNDP (1999).

Figure 2-8 The Closing of the Gender Gap in Schooling



Note: Trends in percentage by sex, of persons in Thailand, 1990 (a) completing at least 4 years of schooling, (b) completing at least some lower secondary schooling, (c) with at least some upper secondary schooling and (d) with at least some tertiary level education.
Source: Knodel (1997).

Figure 2-9 Factor Intensity in Agriculture



Source: Nidhiprabha (2005) based on data from FAOSTAT.

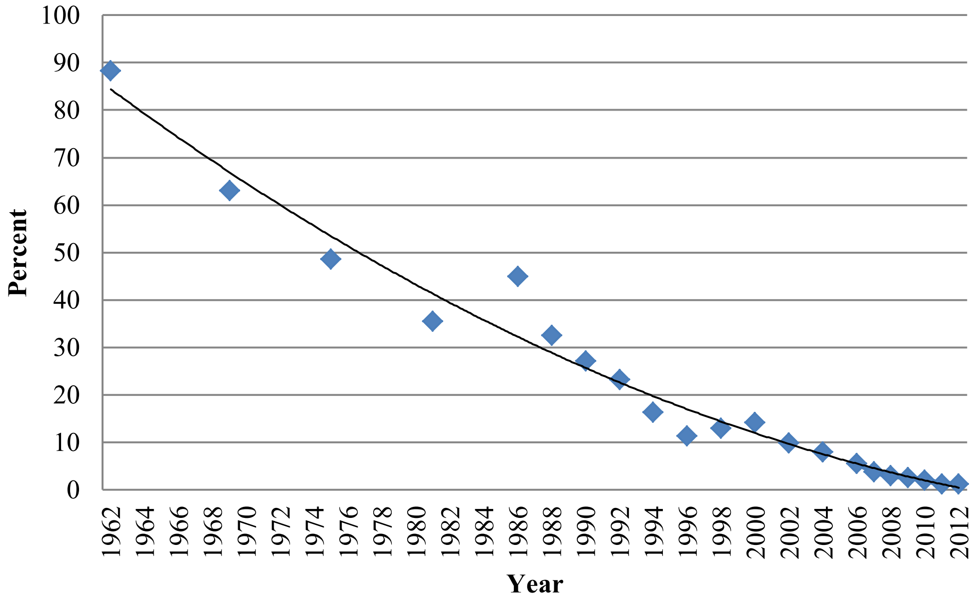
2.1.5. Non-labour Production Inputs per Worker

Thailand has a significantly low non-labour production inputs per worker, especially in the agricultural sector. **Figure 2-9** shows that the ratio between imported machine per worker is almost close to zero for past 40 years (1961-2001). First, this implies that Thai agricultural sector is still backward and traditional. As discussed above, there is slow growth in the agricultural sector as its growth is solely driven by expanding cultivated areas (Siriprachai, 2009) rather than utilising agricultural technology advancement during the period of social and economic development. Second, the ratio also implies that there is abundant of labour in Thailand, especially in the agricultural sector. These two factors contribute to a low non-labour production inputs per worker in Thailand and, in general, in developing countries (Behrman and Srinivasan, 1995).

2.1.6. Poverty Reduction

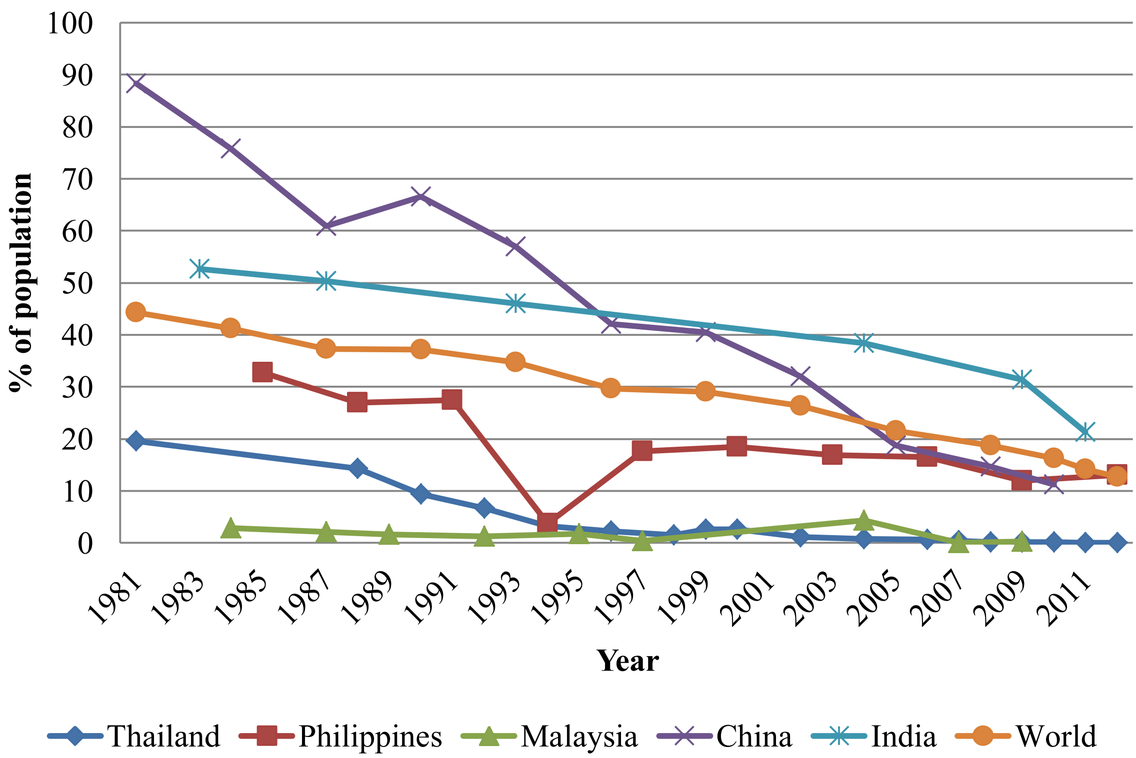
Figure 2-10 shows that Thailand has made enormous progress in reducing poverty. Headcount ratio has declined steadily from 88.3 per cent in 1962 to 1.2 per cent in 2012. However, during 1970-1980, the headcount ratio was roughly 40-50 per cent, which is considered to be relatively high. This implies that 40-50 per cent of the population still lived below the poverty line even in the midst of high economic growth. **Figure 2-11** provides an overview of the poverty incidence of Thailand and other Asian countries for past 30 years. In general, there is a declining trend of poverty in all countries. Compared with other Asian countries, Thailand and Malaysia manage to deal with the poverty incidence very effectively since the headcount ratios are almost close to zero in recent years.

Figure 2-10 Poverty Incidence, 1962-2012 (Headcount Measure, % of total Population)



Source: Author`s compilation based on 1962-2002 data from Siriprachai (2009) and 2004-2012 data from World Development Indicators (2016).

Figure 2-11 Poverty Headcount Ratio at $1.90 a day (2011 PPP)



Source: World Development Indicators (2016).

2.2.1. Enrolment Situation

There has been an increasing trend of enrolment in every education level. **Figure 2-12** illustrates number of student enrolment by education level during 1951-1986. While the enrolment rate of lower primary is the highest among all education levels, the number of student enrolment in other education levels remains substantially low. This implies that most of students drop out after completing the lower primary education which is the compulsory education level (before 1978). According to **Figure 2-7**, between 1977 and 1978, a sharp increase in the number of student enrolment of grade 5 and grade 6 could be observed. This is due to the change of compulsory education law in 1978, which is utilised as an instrument for the estimation.

Table 2-15 indicates the gross enrolment ratio (GER) during 1960-1983. Thailand achieved steady but uneven progress in school enrolment in different education levels. The primary education showed the most impressive demand for education as the enrolment rate reaches 83 per cent in 1960 and almost 100 per cent in 1983. On the contrary, demand for secondary and tertiary education stagnated during 1960-1970. The enrolment rate for tertiary education remained at 2 per cent for a decade. However, from 1980, the increasing demand for secondary and tertiary education could be observed. There was a 12 percentage point increase for both levels of education. This shows a favourable trend in Thai education during the early stage of the development. However, there is plenty of room for improvement in the enrolment situation, in both secondary and tertiary education.

According to **Table 2-15**, gender disparities are less pronounced in the primary education. The share of female students is almost half of the total number of enrolment in primary education. However, the share of female students is still lower than that of male students in every level of education. Moreover, the gender inequality becomes wider and wider in the higher levels of education. Large disparities remain, especially in secondary education (vocational) and higher education. The observed improvement of the gender equality is fairly slow during 1956-1968. The gender disparities in primary education remain constant for a decade. In contrast, the situation has been improving gradually in the secondary and tertiary education which the share of female is approximately 40 per cent of total number of enrolled students. Despite the improvement of gender disparities, it is always the case that human capital investments are in favour of males.

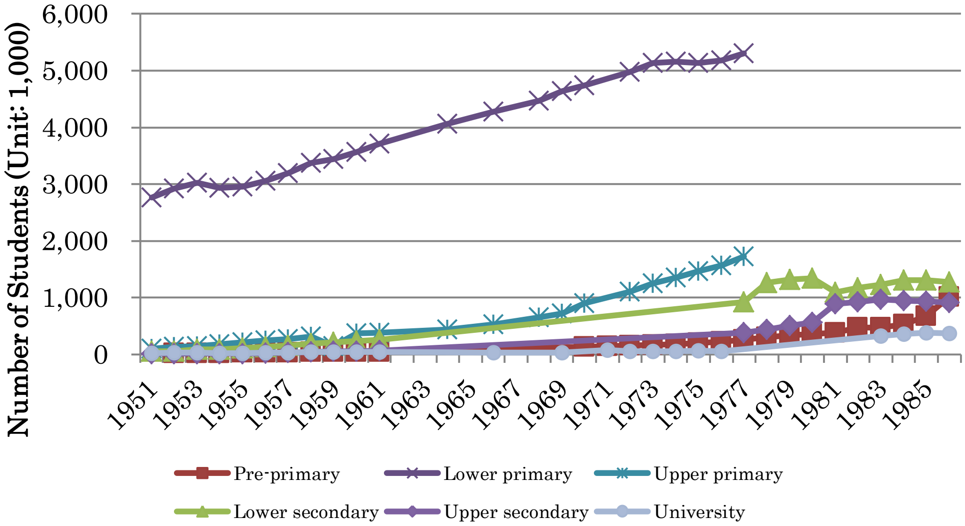
Table 2-8 indicates percentage of male and female enrolment by age and education level from which the situation of enrolment rate after 1968 can be inferred. In terms of primary education, the enrolment rates of cohorts aged 20 and 30 in 1990 represented the enrolment rate of 4th grade students in 1980 and 1970, respectively (**Figure 2-14**[3] panel a). The same inference can also be applied to other education levels. The enrolment rates of cohorts aged 15, 25, and 35 in 1990 represents the enrolment rate of the last grade of lower secondary education (approximately 15 years old) in 1990, 1980, and 1970, respectively (**Figure 2-14** panel b). Moreover, the enrolment rates of cohorts aged 20, 30, and 40 in 1990 represents the enrolment rate of the last grade of upper secondary education (approximately 20 years old) in 1990, 1980, and 1970, respectively (**Figure 2-14** panel c). Lastly, the enrolment rates of cohorts aged 25, 35, and 45 in 1990 represents the enrolment rate of the last year of the university (roughly 25 years old) in 1990, 1980, and 1970, respectively (**Figure 2-14** panel d). The ages used in the deduction are approximated; therefore, Figure 2-14only provides a rough idea of the male and female enrolment rates in different education levels in 1970, 1980 and 1990. **Figure 2-14** indicates that the gender disparities exist in every education level even during the economic development period and that there is a gradual growing demand for each education level during 1970-1990. The enrolment rates in secondary and tertiary education are still unfavourable.

In terms of repetition rate (Figure 2-15), the repetition rates were lower than 10 per cent for all education levels, except for grade 1. The repetition rates were even lower than 5 per cent for the higher grades. It seems that during 1970s it was not common for students to repeat the same grade. One of the reasons was that the Government aimed to quantitatively increase the number of labour force that graduates at least primary education. Moreover, to keep students repeating the same was costly and created a burden on the Government budget.

In conclusion, there has been a rather gradual growing demand in Thai education market for past 40 years (1950-1990). The highest demand for education can be observed in the primary education, while the enrolment rate is substantively low in the other education levels. The lower demand for education in women is another concerned issue. Even though, gender disparities are less pronounced in the primary education, male enrolment rates in other education level are always greater than those of women. The situation becomes better after 1990. Thus, the deep-rooted problem Thailand faces is a persistent low demand for secondary and higher education.

[3] Figure 14 is reproduced based on Figure 8.

Figure 2-12 Number of Student Enrolment by Education Level, 1951-1986



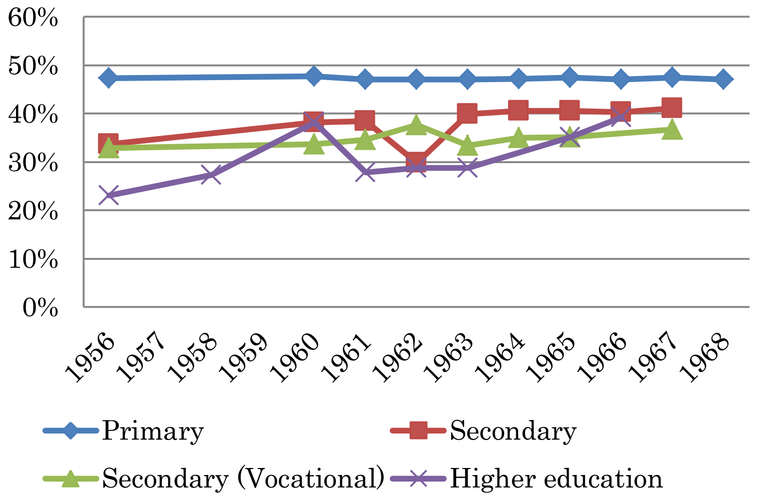
Source: Author's compilation based on UNESCO and Ministry of Education (1965-1988).

Table 2-15 Enrolment Trends, 1960-1983

	(Units)	1960	1970	1980	1983
Primary Level					
Number of Students	(million)	3.936	5.635	7.392	7.272
Number of Teachers	(million)	0.109	0.163	0.299	0.356
Gross Enrolment Ratio	(%)	83	83	96	97
Secondary Level					
Number of Students	(million)	0.311	0.695	1.92	2.192
Number of Teachers	(million)	0.016	0.045	na	na
Gross Enrolment Ratio	(%)	12	17	29	30
Tertiary Level					
Number of Students	(million)	0.046	0.055	0.361	1.12
Number of Teachers	(million)	0.003	0.008	0.03	0.029
Gross Enrolment Ratio	(%)	2	2	14	23

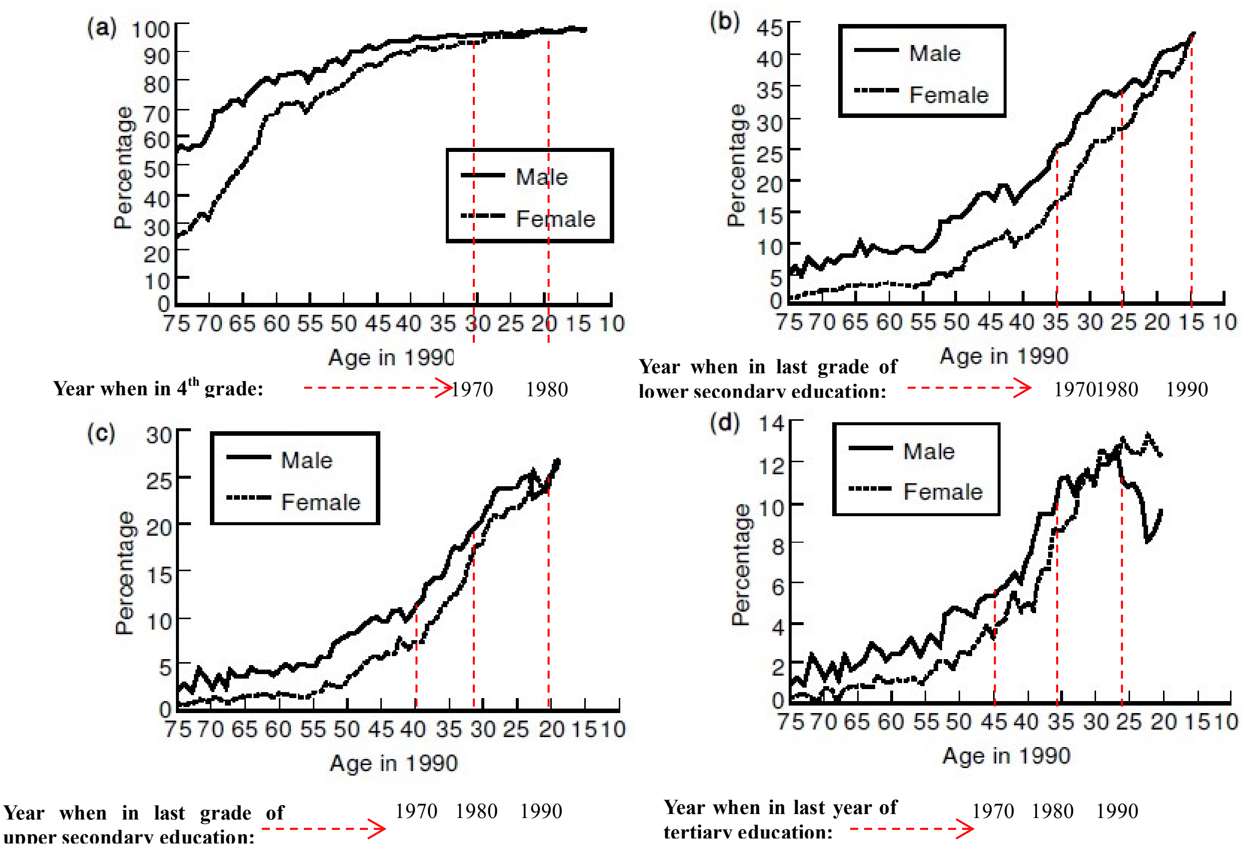
Note: na not available.
Source: World Bank, 1989 based on UNESCO Statistical Yearbooks.

Figure 2-13 Percentage of Female Students by Education Level, 1956-1968



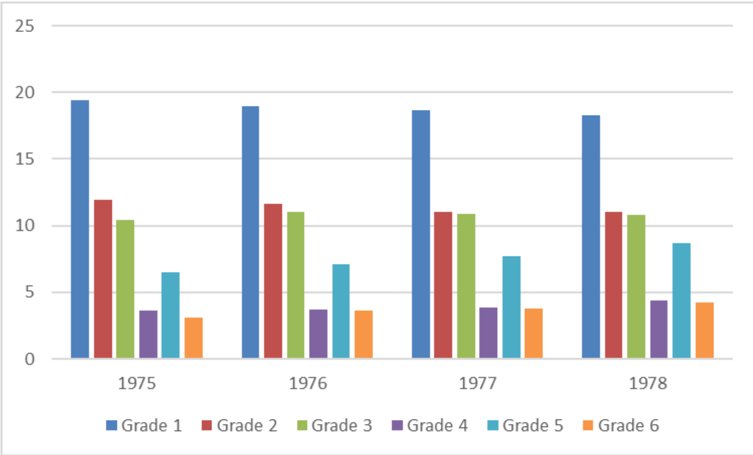
Source: Author's compilation based on UNESCO (1955-1970).

Figure 2-14 Percentage of Male and Female Enrolment by Age and Education Level



Note: Trends in percentage by sex, of persons in Thailand, 1990 (a) completing at least 4 years of schooling, (b) completing at least some lower secondary schooling, (c) with at least some upper secondary schooling and (d) with at least some tertiary level education.
Source: Author's compilation based on Knodel (1997).

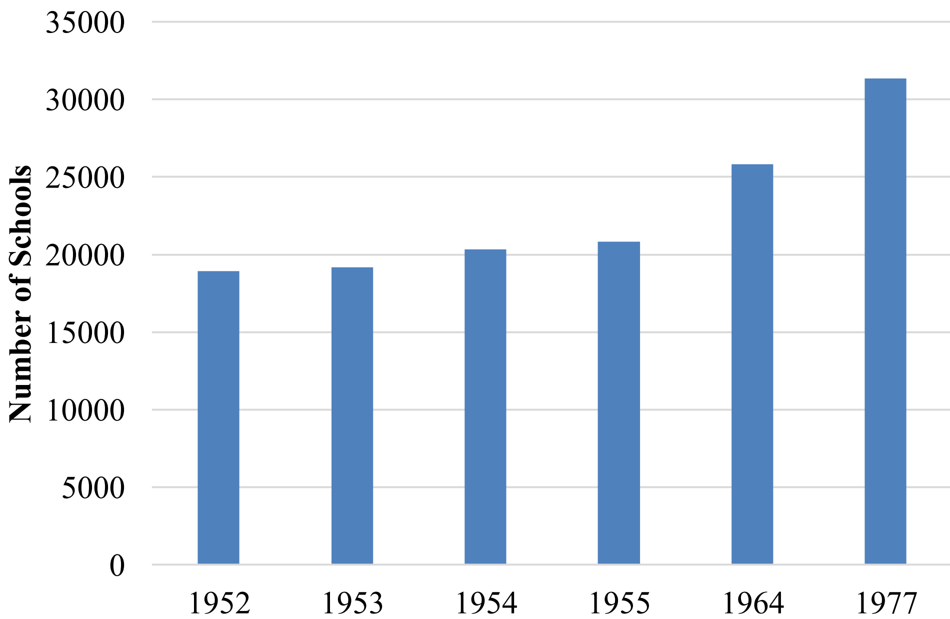
Figure 2-15 Repetition Rate in Primary Education by Year



Source: UIS (2016)

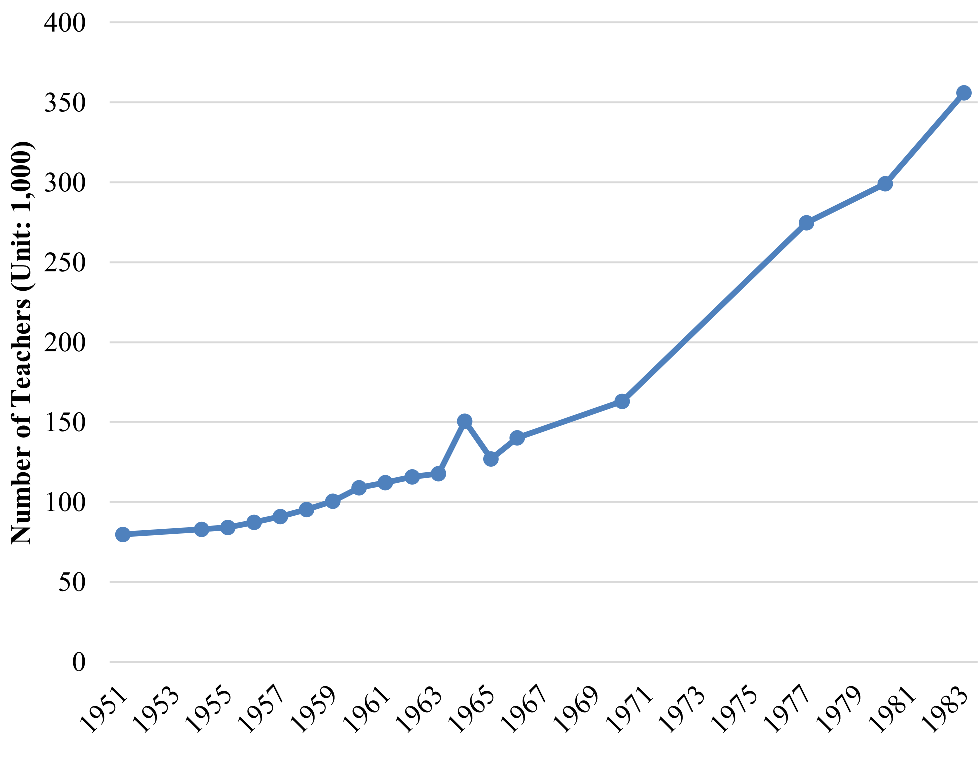
2.2.2. Supply Side of Thai Education

Figure 2-16 Numbers of Primary Schools



Source: Author's compilation based on UNESCO (1953, 1968) and Ministry of Education (1979).

Figure 2-17 Numbers of Teachers in Primary Education



Source: Author's compilation based on UNESCO (1955-1970).

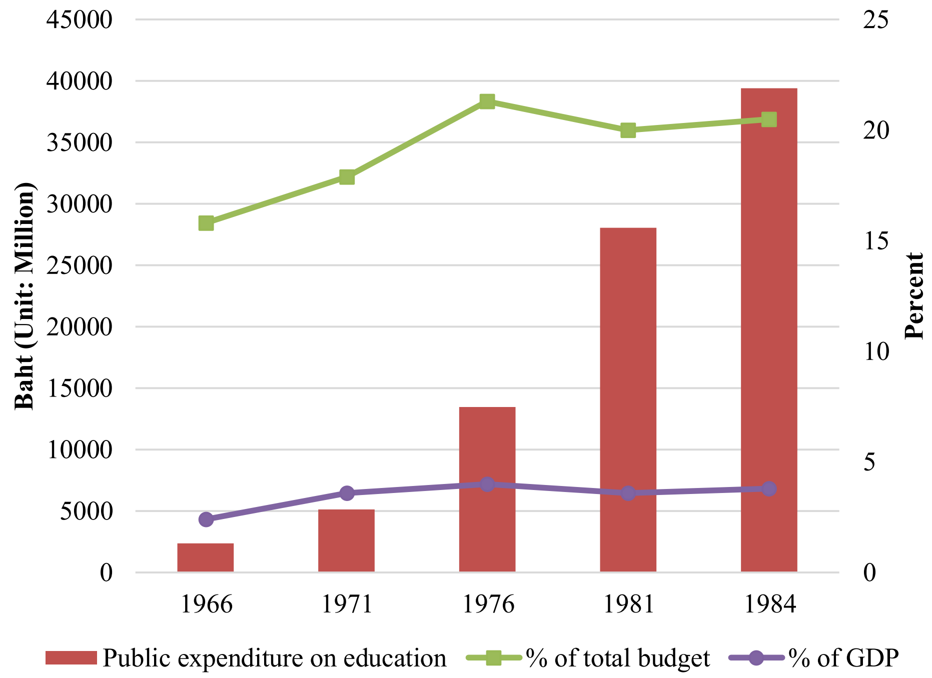
Important indicators of educational supply provided by the Government include numbers of schools and teachers.

Figure 2-16 illustrates the numbers of primary schools during 1952-1977, whereas **Figure 2-17** shows numbers of teaching staffs in primary education during 1951-1983. The trends of these two indicators correspond to each other since more schools mean more teaching staffs needed to fill in teaching positions in those schools. The numbers of schools increased roughly 12.5 per cent and 20 per cent during 1955-1965 and 1965-1975, respectively. Similarly, the numbers of teachers rose approximately 25 per cent between 1955 and 1965. However, during 1965-1975, a sharp increase around 60 per cent in the numbers of teachers could be observed. The numbers of teachers grew roughly 2 times faster than those of schools. The higher growth rate of primary schools and teachers could be observed during the early stage of structural transformation, 1965-1975. The Government raised both supplies of primary schools and teaching staffs to respond to the increasing demand for primary education during the same period. The boost in both demand and supply of education was a part of a national development plan to be ready for the next stage of social and economic development.

A financial resource from the Government played a crucial role behind those increases in physical supplies of education such as school buildings and teaching staffs since public education is financed almost entirely by the central Government. During 1966-1971, there was a small increase in public expenditure on education (**Figure 2-18**). However, after 1971, the sharp increases of the public expenditure on education and its percentage of total budget could be observed. Thailand's education budget amounted to 5,000 million Baht in 1971, while it rose substantially to 13,500 million Baht in 1976. Similarly, the educational expenditure as a share of total Government budget increased about 4 per cent over the same period (17.9 per cent to 21.3 per cent). In addition, public education expenditure as a proportion of GDP increased from 2.4 per cent in 1966 to 4 per cent in 1976. This was due to the fact that the Government had to prepare teaching staffs and all educational facilities to deal with an influx of students after enforcing the new compulsory education law in 1978. This explanation can be supported by the above discussion on the increase in numbers of primary schools and teaching staffs during the same period. In 1981 and thereafter, a high level of public expenditure on education still could be observed. Teachers' wages and maintenance constituted a large portion of the education expenditure, while the Government still had to expand the free compulsory education. In 1981, the education budget was approximately 28,000 million Baht and within three years it jumped to 39,406 million Baht in 1984. According to World Bank (1989), it is reported that 57 per cent of the 1984 budget was spent on primary education, while 16 per cent on secondary and 12 per cent on universities. This corresponds to the different degrees of demand for different education levels discussed in the previous section. The education expenditure as a share of total government budget and

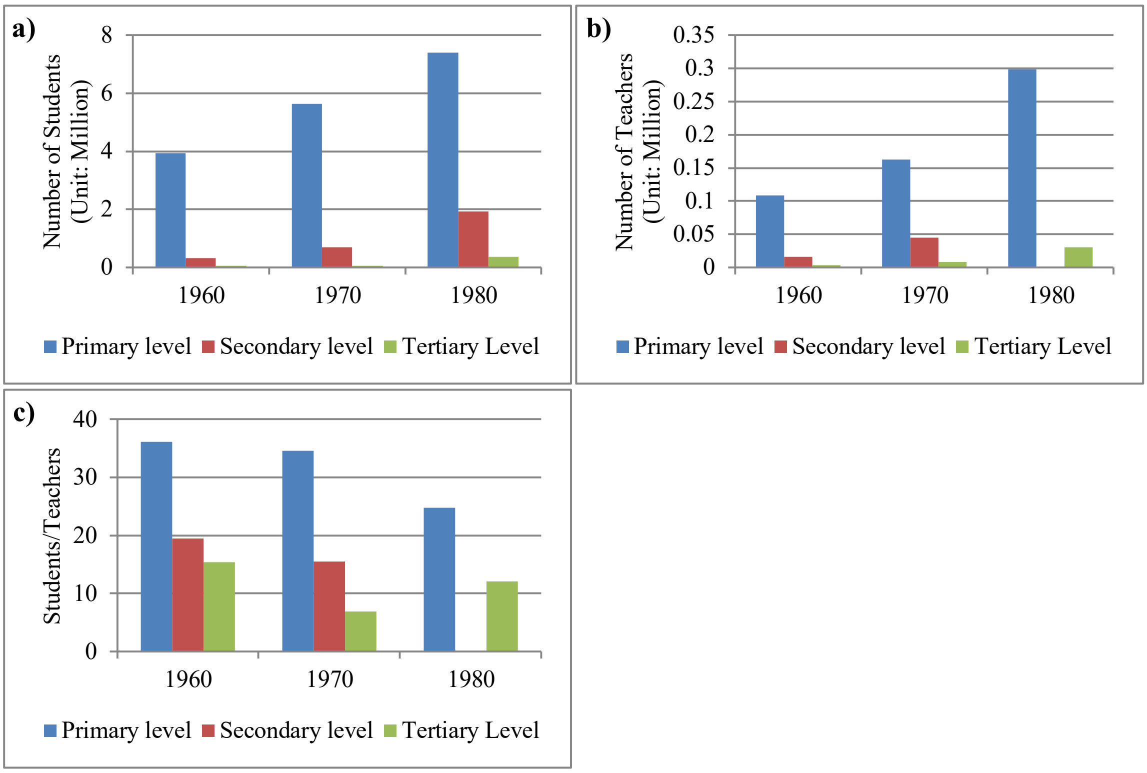
GDP remained constant at roughly 20 per cent during 1976-1984.

Figure 2-18 Public Expenditure on Education, 1966-1984



Source: Author's compilation based on World Bank (1989)

Figure 2-19 Enrolment Trends, Number of Teachers, and Student-Teacher Ratio, 1960-1980



Source: Author's compilation based on World Bank (1989)

Figure 2-19 summarises the situation of the demand side and supply side of Thai education during the period of structural transformation. There has been a rapid expansion of the education system in the primary education for past 20 years (1960-1980), both in terms of student enrolment and number of teachers. In contrast, the demand and supply for secondary and tertiary education grew at a much slower pace. The Government raised both supplies of primary schools and teaching staffs to respond to the increasing demand for primary education while tries to increase the quality of the education. Hence, an improvement of overall education quality could also be observed as the student-teacher ratios kept declining over time across different levels of education. The boost in both demand and supply of education was a part of a national development plan to prepare Thai labour force to be ready for the next stage of social and economic development.

Section 3. MIGRATION

In addition to characteristics of developing countries discussed above, migration is also another important perspective in developing countries. In developed countries, motivation and factors of migration seem to be more sophisticated than those of developing countries, in which the migration is mainly driven by basic needs such as better job opportunities and better quality of education. Nevertheless, migration in developing countries plays a crucial role in economic development as it reallocates labour between the agricultural and the other sectors and promotes urban growth. Thus, the issue of migration warrants further investigation.

The phenomenon of "human capital drain from rural to urban areas" seems to occur in Thailand. Migration streams of migrants can be broken down into six categories, including urban-urban, rural-urban, unknown-urban, rural-rural, urban-rural, and unknown-rural. According to **Table 2-16**, rural migrants (rural areas as migration destination) accounted for 77, 65.2, and 64 per cent of the total migration streams in the period of 1965-70, 1975-80, and 1985-90, respectively. Even though the decline in the number of rural migrants could be seen, the majority of the migration streams concentrated in rural migrants. In contrast, the rural-urban migration stream kept increasing during the period of 1965 – 1990. The rural-urban migration flow made up for 10.5-18.4 per cent. This indicates a growing trend of human capital drain from rural to urban areas.

In terms of education distribution of migrants, **Table 2-17** indicates that rural-urban migrants consist of more educated individuals than those of rural non-migrant. The majority of rural-urban migrants attain at most primary education, which amounts to 59.6 per cent of the rural-urban migration flow, while 76 per cent of the rural non-migrant population complete only primary education. The percentage of persons completes secondary and university of rural-urban migrants (27.4 per cent and 10.2 per cent, respectively) are higher than those of rural non-migrants (11.7 and 2.5 per cent, respectively). The flow of rural-urban migrants is dominated by the group of more educated individuals. In other words, more educated individuals have higher mobility. This pattern can also be observed from each of migration flows, even within rural-rural migration stream.

Table 2-16 Migration Streams of Migrants

Migration Stream	1965-70		1975-80		1985-90	
	Number	Per cent	Number	Per cent	Number	Per cent
Total migrants	**3,331,100**	**100**	**2,947,700**	**100**	**4,026,100**	**100**
Urban migrants	**763,400**	**23.0**	**1,024,900**	**34.8**	**1,448,700**	**36.0**
Urban-urban	297,000	8.9	506,000	17.2	545,100	13.5
Rural-urban	348,000	10.5	420,600	14.3	738,400	18.4
Unknown-urban	118,400	3.6	98,300	3.3	165,200	4.1
Rural migrants	**2,567,700**	**77.0**	**1,922,800**	**65.2**	**2,677,400**	**64.0**
Rural-rural	2,086,700	62.6	1,532,900	52	1,645,100	40.9
Urban-rural	180,400	5.4	278,300	9.4	508,900	12.6
Unknown-rural	300,600	9.0	111,600	3.8	423,400	10.5

Source: Pejaranonda et al, 1995 and Vutthisomboon (1998).

Table 2-17 Percentage Distribution of Educational Attainment of Populations Aged Six Years and Above

	Migrant		Non-migrant	Migrant		Non-migrant
Education	**Urban-Rural**	**Rural-Rural**	**Rural**	**Urban-Urban**	**Rural-Urban**	**Urban**
Males						
No Education	3.5	5.6	9.6	3.2	2.5	6.3
Primary	48.4	71.2	76.0	39.4	59.6	44.4
Secondary	34.6	18.1	11.7	34.0	27.4	32.6
University	13.2	4.5	2.5	23.0	10.2	16.3
Other	0.2	0.6	0.2	0.4	0.3	0.3
Total	100	100	100	100	100	100
Females						
No Education	7.0	8.9	14.2	4.6	4.1	9.4
Primary	54.7	73.9	75.8	42.2	63.6	44.4
Secondary	24.2	12.0	7.8	30.5	21.4	25.4
University	14.1	5.2	2.2	22.7	10.8	16.0
Other	na	na	na	na	0.1	0.1
Total	100	100	100	100	100	100

Note: na: not available.
Source: Author's compilation adapted from Vutthisomboon (1998) based on data from the National Statistical Office, Thailand, 1993.

In terms of regional mobility, both lifetime and temporary migrants tend to move to Bangkok and the Central region (**Table 2-18**). During the development period, 1970-1990, net migration losses were highest in the Northeast and the North regions. On the other hand, the Central area, and Bangkok and metropolitan areas had the highest net migration gains among all regions.

Table 2-18 Regional Net Gains and Losses from Five-year Migration 1955-1990

Current residence	Total	Region of Previous Residence				
		Bangkok	Central	North	Northeast	South
1955-60						
Bangkok	67,045		41,208	5,047	17,855	2,935
Central	-46,643	-41,208		-14,710	15,102	-5,827
North	30,134	-5,047	14,710		21,106	-635
Northeast	-59,809	-17,855	-15,102	-21,106		-5,746
South	9,273	-2,935	5,827	635	5,746	
1965-70						
Bangkok	168,863		83,358	21,909	43,221	20,375
Central	-80,370	-83,358		-10,804	17,290	-3,498
North	6,340	-21,909	10,804		17,790	-345
Northeast	-85,006	-43,221	-17,290	-17,790		-6,705
South	-9,827	-20,375	3,498	345	6,705	
1975-80						
Bangkok	170,400		29,042	22,233	99,602	19,523
Central	49,454	-29,042		14,981	63,748	-233
North	-17,713	-22,233	-14,981		23,120	-3,619
Northeast	-194,815	-99,602	-63,748	-23,120		-8,345
South	-7,326	-19,523	233	3,619	8,345	
1985-90						
Bangkok	365,900		500	65,800	261,100	38,500
Central	293,400	-500		42,500	246,900	4,500
North	-89,300	-65,800	-42,500		26,800	-7,800
Northeast	-553,700	-261,100	-246,900	-26,800		-18,900
South	-16,300	-38,500	-4,500	7,800	18,900	

Source: Vutthisomboon (1998) based on 1955-60, 1965-70 and 1975-80 data from Goldstein and Goldstein (1986); 1985-90 data from the National Statistical Office (1993)

Section 4. THE DEVELOPMENT OF THAI ECONOMY AND EDUCATION IN THE CONTEXT OF THE RATES OF RETURN TO SCHOOLING ANALYSIS

It is possible that the stage of social and economic development may explain a direction of bias in the OLS estimation of returns to schooling. In this study, the dominance of ability bias is mainly explained by the inequality of income and educational opportunity during the early period of social and economic development. At that time, Thailand was characterised as a developing and agriculture-based economy with relatively high poverty rate and income gap. A majority of population still depended heavily on agricultural sector and rural labour activities, which earn relatively lower than other sectors. It was common for family members to help working without formal pay in those sectors. Moreover, there was low school enrolment rate observed during the same period. Overall, it was common to observe the inequality, both in terms of income and access to basic public services such as basic education, between different socio-economic groups. Individuals from agricultural families are socially discriminated, poorer, and having greater immediate need to work. Their opportunity cost of schooling is also higher. It is likely that individuals from agricultural families may quit school at younger ages. Due to high costs of schooling and demand for household labour, agricultural households cannot afford to send all children to school. Thus, they may select only one or few children with the highest ability among all children to receive proper education, while the rest of the children are expected to help household agricultural activities. As a result, children with higher ability tend to have more schooling in agricultural households. This is consistent with the general ability bias hypothesis which argues that more-able individuals are more likely to have higher years of schooling. On the other hand, individuals from a higher socio-economic class are likely to have higher investments in education qualitatively (better school quality) and quantitatively (more years of schooling) (Behrman, 1999). This reflects the ability bias usually discussed in both studies of developed and developing countries. In sum, there are two sources of ability bias in case of Thailand which includes the ability bias from a selection of more-able child within poor households and the ability bias due to higher financial endowment of rich families.

The argument of the ability bias due to a selection of more-able child within poor households is more relevant and specific to the case of Thailand and, perhaps, other developing countries. In other words, this type of ability bias is less relevant in developed countries where poverty and inequality of educational opportunity, especially for the basic education, are less pronounced. Due to high economic development and standard of living, individuals in developed countries have more economic power to afford education for their children regardless of their level of ability. In addition, a universal access to free education is commonly observed, especially for basic education. Therefore, the equal education opportunity is more prevalent in developed countries. Even though there is a tendency that individuals with higher ability are more likely to attain more years of schooling, the ability bias seems to be less dominant in developed countries, at least for the basic level of education. This is possibly the main reason that the upward ability bias is observed in Thailand and other developing countries, but not in developed countries. Instead, OLS estimation from developed countries suffers from downward bias (IV estimates > OLS estimates). Hence, other sources of bias seem to dominate the effect of the upward ability bias in case of developed countries.

Explaining the upward ability bias in terms of inequality of income and educational opportunity during the early period of Thai social and economic development may possibly be generalised to other developing countries, which share similar social and economic context with Thailand. The main result shows that the magnitude of OLS estimates in Thailand is comparable with those of other developing countries. However, the IV estimates are fairly different. As different instruments from different studies define different parameter (Heckman et al., 2006), a similarity of magnitude of IV estimates should not be expected to observe from different estimations. In other words, it is usually the case that IV varies across different studies. Nonetheless, similar instruments and identification strategies correct the same endogeneity bias, given that those studies share similar sample characteristics. Thus, the difference between estimates from OLS and IV estimations should indicate a similar direction of bias whether OLS estimates are higher than those of IV.

Section 5. CONCLUSION

This chapter provides readers with a comprehensive overview of Thai economic and educational development during 1960-1990. Thai economy can be divided into four main periods, namely pre-boom (1951-1986), boom (1987-1996), crisis (1997-1998), and post-crisis (1999 onwards). During 1951-1986, the Government put efforts in building basic physical infrastructure, e.g. road, electricity, to name a few, to help facilitating trading and the growth of the economy. The real GDP growth rate jumped to 9.2 per cent during the boom period, 1987-1996. This is considered as an economic miracle growth in Thailand. However, Thailand experienced a financial crisis in 1997 and faced a negative GDP growth at 6.1 per cent, which is the lowest in the Thai economic history. In the post-crisis period, Thailand managed to bounce back and achieved the real GDP growth rate of 4 per cent. The overall average real GDP growth and real GDP growth per capita were around 6.2 per cent and 4.2 per cent, respectively, during 1951-2003.

The high GDP growth is claimed to be the result of structural transformation during the pre-boom period, 1951-1986. Thailand went through a structural transformation from a primitive agriculture-based economy to newly industrialised economy, which the agricultural sector played a crucial role in reallocating resources to the other economic sectors. Similar to most of industrialisation process in other countries, the industrial sector was developed at the expense of the agricultural sector. In addition, Thailand also went through another structural transformation in terms of export and import pattern. Thailand changed from agriculture produce exporter, e.g. rice, to manufactured goods exporter, especially garments and parts and components.

Overall, the situation of the development in Thailand is consistent with the characteristics of developing countries proposed by Behrman and Srinivasan (1995). They suggest that developing countries possess different degrees of market completeness and institutions from developed countries. The followings are the proposed characteristics of developing economies in comparison with developed economies. First, a majority of population still depends heavily on agricultural sector and rural labour activities. Second, in agricultural sector, it is common for family members to help working without formal pay; therefore, non-wage labour accounts for a large proportion of the total labour force in developing economies. Third, there has been a rapid growth of labour force in developing countries. Fourth, high labour participation rates among 15-64 year olds can be observed due to low human capital investments among young cohorts. Fifth, there are low school enrolment rates and the education gap is in favour of males. Sixth, lower non-labour production inputs per worker are observed in developing countries. Lastly, a majority of Thai population still lives below the poverty line even in the midst of high economic growth during 1970-1980.

In terms of educational development, there had been a rapid expansion of the education system in the primary education for past 20 years (1960-1980), both in terms of student enrolment and number of teachers. In contrast, the demand and supply for secondary and tertiary education grew at a much slower pace. The Government raised both supplies of primary schools and teaching staffs to respond to the increasing demand for primary education while tried to increase the quality of the education. Hence, an improvement of overall education quality could also be observed as the student-teacher ratios kept declining over time across different levels of education. The boost in both demand and supply of education was a part of a national development plan to prepare Thai labour force to be ready for the next stage of social and economic development.

In terms of migration, the rural-urban and interregional migration help facilitating the development process of Thai economy by reallocating human resources between regions and economic sectors. The phenomenon of "human capital drain from rural to urban areas" seems to occur in Thailand. Even though the majority of the migration streams concentrated in rural migrants, there was a growing trend of human capital drain from rural to urban areas during the period of 1965 – 1990. The flow of rural-urban migrants is dominated by the group of more educated individuals. In other words, more educated individuals have higher mobility. This pattern can also be observed from every type of migration stream. Regarding regional mobility, net migration losses were highest in the Northeast and the North regions during

1970-1990, while the Central area, and Bangkok and metropolitan areas had the highest net migration gains among all regions.

In conclusion, Thailand experiences rapid economic development and structural transformation during 1960-1990. Obtaining the rates of return to education in this period helps us better understand the process of Thai economic development as well as the interplay between the rates of return to schooling and economic development in Thailand during 1980s to 1990s. In addition, the overall social and economic conditions of the development in Thailand are consistent with the general characteristics of other developing countries. Hence, by implication, estimating the rates of return to schooling in Thailand also provides better understandings on the role of human capital in the process of development in other developing countries. As developing countries possess different degrees of market completeness and institutions which are radically different from those of developed countries, this warrants value for investigation of the rate of returns to schooling in the context of developing countries. This further investigation possibly gives a different economic pattern and implications of estimated results.

CHAPTER 3. LITERATURE REVIEW

The rates of return to schooling is one of the most important topics that economists have been investigating, especially by utilising the Mincerian model. However, there has been long debate that the OLS estimates from the Mincer equation is possibly biased due to the endogeneity problem. Schultz (1988) suggests that the most concerned issue in the study of returns to education is the issue of possible biases in estimation owing to the endogeneity problem. Later, the attention should be given to the issue of sample selection. Strauss and Thomas (1995) also emphasise the importance of these two issues and posit that they are more relevant in the studies from developing countries.

First, this chapter revisits the Mincer model to show that the Mincer function is alive and well in fitting actual age-wages profiles. Second, the chapter reviews previous literature regarding methodological issues in estimating the rates of return to schooling. The third section surveys previous studies of the rates of return to schooling with a particular focus on studies utilising compulsory education law change as an instrumental variable. Fourth, previous studies on the sample selection are briefly discussed. Finally, the chapter is closed by emphasising the contributions of this research.

Section 1. REVISITING MINCER MODEL

Despite the fact that the Mincer model has been intensively used to find the rates of return to schooling, there is a doubt whether the Mincer model can capture new elements and changes in the current social and economic situation. As the world economy has been growing rapidly, social and economic conditions also have changed tremendously during the past decades. However, the Mincer equation has never changed and is still as simple as it is.

The following is the simple Mincerian earnings equation (Mincer 1958, 1974):

$$\textbf{(3-1)} \qquad \log y_i = \beta_0 + \beta_1 S_i + \beta_2 X_i + \beta_3 X_i^2 + e_i$$

where the log of individual earnings (y_i) is a function of years of education an individual i attended (S_i) and number of years an individual i has worked after completing his/her education (X_i). e_i represents a disturbance term. Economists are interested in estimating the β_1 since it represents an educational premium for wages and/or an internal rate of return to schooling[4].

Card (1999) shows the evidence of the shape of the Mincer function and its specification, that supports the fact that Mincer equation is still alive and well in fitting the actual age-wages

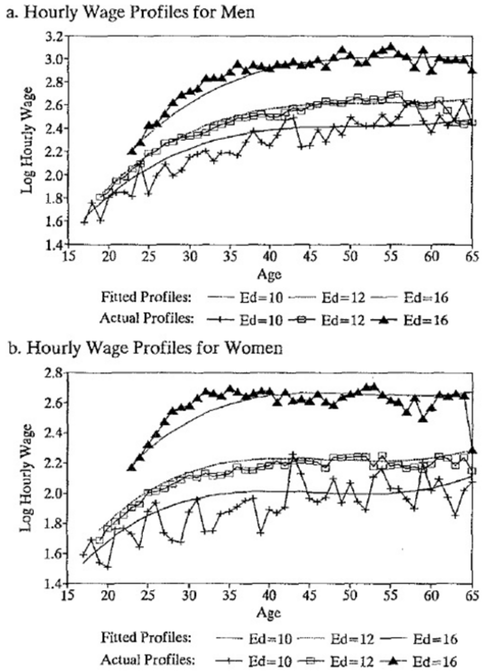
[4] See: Chapter 4 for a more detailed discussion on how to interpret the coefficient of school attainment.

profiles. By utilising the United States (US) pooled samples from 1994-1996 March Current Population Surveys, Figure 3-1 illustrates the age profiles of hourly wages for men (a) and women (b) by education level (10, 12 and 16 years of education). Each data point on the lines represents the average log hourly wages by age. The lines with the marks indicate the actual age-wages profiles, while the smooth lines show the fitted profiles estimated by the Mincer equation. Card (1999) modifies the model specification by adding a cubic term of potential experience (X_i) and a dummy variable indicating whether individuals are black. According to Figure 3-1, the age-wages profiles for US men and women are well-approximated since the fitted lines move along closely with the actual age-wages profiles. However, Mincer's equation understates some portions of the actual profiles of different education groups in the US data.

In the case of Thailand, a similar conclusion can be drawn from the age-wages profiles that also indicate that the estimates from Mincer model reasonably well approximate the actual age-wages profiles. Figure 3-2 shows the age profiles of hourly wages by education levels for Thai workers using pooled samples from the 1986-2012 Labour Force Surveys. The data represents mean log monthly wages by age for individuals with four, six, nine, twelve, sixteen, and twenty-one years of education. Four years of education refer to the minimum years of education required by the 1962 Compulsory Education Act, whereas six years of education represent the minimum compulsory education level enforced by the new 1978 compulsory education law. Other education years refer to the final year of each academic level, including lower secondary level, upper secondary level, undergraduate level (normally four years after graduating from upper secondary school), and graduate level. Plotted lines along with the actual means are the fitted values obtained from Mincer model that includes only a quadratic term of age. Comparisons of the fitted and actual data suggest that the age-earnings profiles for Thai workers are fairly smooth, and reasonably well-approximated. In contrast to the age-wages profiles from the US, the problem in fitting the precise curvature is less pronounced in the case of Thailand.

In conclusion, the age-wages profiles from the US and Thailand confirm that the Mincer model is still alive and well in fitting the actual age-wages profiles. Even though the model specification requires functional adjustment and has some troubles fitting the precise curvature in the case of the US, Mincer model with the OLS estimation can fit the actual data fairly well, especially in the case of Thailand. Instead of using the cubic terms of age, the model used in fitting the Thai age-wage profiles maintains the quadratic terms. Regardless of its simplicity and accuracy in estimation, there has been a long debate that the OLS estimation of Mincer model may yield biased estimates because of the problem of endogeneity and the sample selection. Both issues are discussed thoroughly in the following sections.

Figure 3-1 Age Profiles of Hourly Wages for Men (a) and Women (b), the US



Note: The samples include 102,718 men and 95,360 women aged 16-66 with positive potential experience and average hourly earnings between \$2.00 and \$150.00 in 1995 dollars. 53 per cent of the sample have 10, 12, or 16 years of schooling and are used in graphs. The regression models are fit by gender to all education groups and include a linear education term, a cubic in experience, and a dummy variable for individuals of black race.
Source: Card (1999).

Figure 3-2 Age Profiles of Hourly Wages, Thailand

Note: The regression models include a linear education term and a quadratic in age.
Source: Author's compilation based on Labour Force Survey data, 1986-2012.

Section 2. METHODOLOGICAL ISSUES: POTENTIAL BIASES FROM VARIOUS METHODS OF ESTIMATION

As indicated in the Section 1, Mincerian function fits the data very well and OLS regression has been used exhaustively to estimate rates of return to education. Nevertheless, there has been a long debate that OLS estimates have a possibility to be biased due to the endogeneity problem. This section provides a comprehensive review on potential endogeneity biases from various methods of estimation and their related previous literature.

3.2.1. Standard OLS Estimation and Endogeneity Bias

In general, the standard OLS estimation is possibly biased due to the problem of endogeneity. Endogeneity biases occur when explanatory variables are correlated with the error term in models of estimation. There are three main sources of endogeneity, including simultaneity, measurement errors, and omitted variables. In the context of returns to schooling, the omitted variable is the most debatable source of endogeneity bias (Willis, 1986, Schultz, 1988, Card, 1999, Deere and Vesovic, 2006). The issue of omitted variable bias has been discussed in the literature beginning with Beck's (1964) work. The main concern is that there may be a positive correlation between innate ability and schooling as it is likely that more able individuals tend to invest in more schooling. Without controlling for the ability, estimated returns to schooling include both the effects of education and ability. Hence, this leads to overestimations of the "true" rates of return to schooling. This upward bias commonly refers to "ability bias".

However, the source of upward bias is not only limited to individuals' ability, but also includes family background and quality of schooling (Schultz, 1988, Behrman, 1990, Strauss and Thomas, 1995). As both family background and school quality can be linked to the individuals' ability, the arguments for both factors are similar to that of the ability bias. It is argued that individuals from wealthier families are more likely to invest in more (and better quality) schooling (Behrman and Birdsall, 1983, Strauss and Thomas, 1995). As wages are a function of individuals' ability, it is arguably that parents with high ability earn more, on average, and inherit financial endowments to their children. Hence, it is possible that family background is related to both years of schooling and earnings. Furthermore, it is posited that there is a positive association between quality of school and years of schooling as individuals with higher ability tend to be educated in schools with a higher quality (Behrman and Birdsall, 1983, Strauss and Thomas, 1995). Hence, if these variables are omitted from the estimation equation, the estimates of the rates of return to schooling are possibly upwardly biased. As Behrman (1999) put it, "individuals with higher investments in schooling are likely to be individuals with more ability and more motivation who come from family and community backgrounds that provide more reinforcement for such investments and who have lower marginal private costs for such investments and lower discount rates for the returns from those investments and who are likely to have access to higher quality schools". Family background and school quality factors are all linked back to the ability variable, and probably are easier to be controlled or included in the wage equation. Therefore, more emphasis is on the issue of unobserved ability which is mainly discussed in the literature of the rates of return to education. No agreement has been reached in the recent literature on both the degree and the direction of the ability bias of the estimated rates of return to schooling.

First, the debate regarding the degree of the ability bias can be divided into two schools of thought. On one hand, it is usually argued that there is a substantial large ability bias, especially in the case of developing countries (Behrman, 1999). On the other hand, the ability bias is rather small (Griliches, 1977, Card, 1999) and overstated (Becker, 1964) in the literature of returns to schooling. Many studies have been attempting to correct for the omitted variable bias by utilising different methods of estimation. For example, with fixed effects estimation, samples of twins and siblings have been utilised to deal with the problem of unobserved ability. In general, it is found that OLS estimates are marginally larger than those of the estimation utilising sample of identical twins. Thus, Card (1999) used the findings from the sample of twins to make an argument that as the results from the OLS estimation and studies of identical twins are quantitatively similar, there is a small upward ability bias in the OLS estimation. Moreover, the estimated rates of return from IV estimation based on institutional or policy changes, e.g. changes in compulsory education law, is usually higher than those of OLS estimations by approximately 20–40 per cent (Card, 1999). The findings show that the OLS estimates are downwardly biased, given that the IV estimation deals with the endogeneity properly. This contradicts to the ability bias hypothesis, which states that there is an upward bias in OLS estimates. Nevertheless, Behrman (1999) raises interesting observations from recent studies of returns to schooling mainly from the US to defend his stand. First, there are many studies that try to control for unobserved ability, family background, and school quality in the estimations to obtain the "true" rates of return.[5] The results suggest that the OLS estimates suffer from the ability bias and the impact of education attainment is overestimated 40-100 per cent (Behrman, 1990, 1999). Second, even though the evidence of a strong ability bias is weaker in the studies from the US (Card, 1999), there is still a debate on the rates of return to schooling and also increasing numbers of studies supporting the fact that there is a correlation between school quality and returns to schooling

[5] See: "Revisionist" literatures suggested by Behrman (1999): school quality (Behrman and Birdsall, 1983), unobserved shared family background of adult siblings and of members of the same household (Behrman and Wolfe, 1984; Behrman and Deolalikar, 1993), usually unobserved abilities through new tests (Boissiere et al., 1985; Knight and Sabot, 1990; Glewwe, 1996), selectivity (Schultz, 1988), dropout and repetition rates (Behrman and Deolalikar, 1991), measurement error, school quality and behavioral choices regarding school attendance (Alderman et al., 1996b).

from the US literature. Hence, Behrman (1999) perceives that the estimation from developing countries would further overstate the "true" rates of return to education. Finally, regarding the recent estimates from the IV estimation, Lang (1993) and Card (1999, 2001) posit that the higher estimate from IV is because a policy change may affect only certain subgroups, especially relatively disadvantaged groups, rather than the whole population. Therefore, it is not possible to conclude the effect of ability bias in this case since the estimates do not represent the "real" rates of return from the representative sample.

The direction of ability bias can go either upward or downward. It depends on the direction of correlation between ability and schooling. An upward ability bias means that there is a positive correlation between ability and schooling. This bias generally refers to the "ability bias". The (upward) ability bias argues that more-able individuals are more likely to attain more years of schooling. In contrast, in the case that a correlation between ability and schooling is negative, OLS estimates are downwardly biased. This downward ability bias refers to a "discount rate bias" (Lang, 1993, Card, 1994). It is argued that individuals make a decision to go to school by comparing returns to schooling with foregone returns from labour market (an opportunity cost of attending school). Thus, individuals with higher ability may choose not to go to school, if they can earn high returns from the labour market (high opportunity cost of attending school). In addition, the discount rate bias can also be explained by a preference towards schooling and also financial constraint in attending school. For example, individuals with a financial constraint find the opportunity cost of attending school expensive, and they may opt to earn their living instead of going to school. The argument of discount rate bias is often used to explain the estimated results in recent literature that find that IV estimates are greater than those of OLS estimations. However, as discussed above, a lack of the element of randomization in a quasi-experiment may possibly explain the higher IV estimates.

The issues of measurement errors, especially systematic one, are less emphasised in the literature of returns to schooling, while the problem of simultaneity is not relevant in this study. Regarding measurement errors, it generally assumes that any observation consists of the true value and some random or systematic error value. The estimated returns to schooling with random measurement error tends to become smaller in absolute term (Griliches, 1977, Blackburn and Neumark, 1995). The bias drives the estimates towards zero, which commonly refers to an attenuation bias. However, there is a consensus in the previous literature that the measurement error bias is of modest magnitudes (Schultz, 1988; Card, 1999; Krueger and Lindahl, 2001; Orazem and King, 2008) as years of schooling and other wage determinants are measured relatively well compared with other micro-level variables, e.g. landholding, consumption expenditure, to name a few. In contrast, the systematic error can lead to a more serious problem since it is difficult to be detected and is unable to be analysed statistically. Nevertheless, the issue of systematic measurement error has not been raised in the literature of returns to schooling. Therefore, the issue seems not to be specific to the study in this field. Regarding the problem of simultaneity, it is possible that explanatory variables may correlate with the error in a model in which explanatory variables are determined simultaneously with the dependent variable. Thus, the model gives biased and inconsistent OLS estimates and the bias is generally upward. However, the problem of simultaneity is not relevant in the study of returns to education as individuals' current wages receiving from working in the labour market cannot determine individuals' past levels of schooling. It may be possible that individuals may enter the labour market after finishing compulsory education and then come back to study later. Nevertheless, the practice is not common at least during the period of analysis (1970–1990) in which graduates from primary education are still significantly small.

3.2.2. Proposed Remedies to Deal with the Endogeneity Bias

3.2.2.1 OLS with an Explicit Proxy for Ability

In the previous section, it is clear that OLS estimates possibly suffer from the ability bias. There may be a positive correlation between ability and schooling since more able individuals tend to invest in more schooling. Without controlling for the unobserved ability, estimated returns to schooling include both the effects of education and the ability. Hence, this leads to overestimations of the "true" rates of return to schooling.

There have been several studies[6] utilise different test scores as a proxy for ability, e.g. test scores on cognitive skills, test scores on literacy, to name a few. Table 3-1 summarises the previous studies regarding OLS estimation with a proxy of unobserved ability. The overall results indicate that the OLS estimates of returns to schooling tend to be smaller after introducing ability proxy variables. The magnitude of the returns to schooling is quite low and moderate in developed countries (1.6–6 per cent). Since it is difficult to find a data set with the information on test score, the samples of most studies tend to be very small (under 300 observations). Moreover, after introducing the ability variable, most studies focus on the returns to ability rather than those of years of schooling. Hence, there is no systematic way of analysing the effect of years of schooling on earnings. Although the returns to schooling are observed to be smaller after controlling for the unobserved ability, it is still inconclusive whether the estimated results are unbiased due to aforementioned arguments.

There are several studies criticising against the validity of test score as a proxy of innate ability. Becker's (1964) argues that ability is not one dimensional characteristic since ability may include a wide range of domains, e.g. physical strength, intelligence, creativity, to name a few. Thus, there is no guarantee that those tests can perfectly measure the true ability or they even can capture the true ability at all (Lillard, 1977). Using test scores as a proxy for ability may still result in biased estimates due to the problem of endogeneity. The fact that test scores

[6] See Hanushek and Woessmann (2008) for a comprehensive review on impacts of cognitive skills on individual incomes of developed and developing countries.

are imperfect proxy of the true ability means that some parts of ability still remain in the error term. In other words, there is a problem of measurement error when the imperfect proxy of ability is used in the regression. Previous studies, e.g. Hanushek and Woessmann (2008), Korwatanasakul (2013), argue that ability that influences earnings is mainly from the effect of cognitive skills; therefore, they limit the definition of ability to cognitive skills. Cognitive skills consist of three components, including mathematics, science, and language. In contrast, Heckman, Stixrud, and Urzua (2006) argue that both cognitive and noncognitive skills affect wages. Noncognitive skills, e.g. personality traits, persistence, communication skills, motivation, and teamwork skills, to name a few, are found to substantially influence schooling and wages (Bowles and Gintis, 1976; Edwards, 1976; Jencks, 1979; Wolfe and Johnson 1995; Duckworth and Seligman 2005; Heckman, Stixrud, and Urzua, 2006).[7] Given that a test score is a perfect proxy of ability or cognitive skills, the regression model accounting for only the cognitive skills still possibly gives biased estimated results of returns to schooling. As the disturbance term contains the non-cognitive skills highly correlated with schooling and wages variables, the regression model suffers from the omitted variable bias.

In sum, dealing with the problem of unobserved ability bias, the previous studies try to explicitly add the ability component in the Mincer model by using the test score variable as a proxy for ability. However, test score seems to be an imperfect proxy. This results in a problem of measurement error and, in turn, an endogeneity bias in the regression estimation.

[7] See the survey article by Bowles, Gintis, and Osborne (2001)

Table 3-1 OLS Estimation with a Proxy of Unobserved Ability

Study	Country	Proxy		OLS: Returns to Schooling	OLS with a Proxy	
					Returns to Schooling	Returns to Ability
Bishop (1989)	US	General Intellectual Achievement on Productivity (GIA)		na	0.0584 (9.16)	0.204 (6.85)
O'Neill (1990)	US	The Armed Forces Qualifications Test (AFQT)	Black White	na	0.0465 (4.65) 0.0511 (7.39)	0.0039 (5.31) 0.0019 (4.14)
Blackburn and Neumark (1993, 1995)	US	The Armed Services Vocational Aptitude Battery (ASVAB)		0.058 (0.005)	0.042 (0.006)	n.a.
Murnane, Willett, and Levy (1995)	US	IRT-scaled mathematics	NLS72 Male Female HS&B Male Female	 0.022 (3.63) 0.054 (11.07) 0.044 (5.72) 0.065 (10.63)	 0.013 (1.85) 0.037 (6.74) 0.021 (2.22) 0.037 (5.32)	 0.004 (3.13) 0.009 (7.18) 0,011 (5.06) 0.017 (8.57)
Mulligan (1999)	US	Normalised AFQT		0.046 (.007)	0.029 (0.007)	0.109 (0.011)
Finnie and Meng (2002)	Canada	Literacy Skills in Daily Use (LSUDA)	Male Female	0.0181 (6.12) 0.0388 (9.57)	0.0163 (4.86) 0.0389 (9.15)	0.0014/0.0003 (3.33/2.23) 0.00010/0.0006 (0.79/1.19)[a]
Green and W. Craig Riddell (2003)	Canada	International Adult Literacy Survey (IALS)		0.27 (0.042)	0.21 (0.040)	0.0029 (0.00047)
Boissiere and Knight (1985)	Kenya and Tanzania	Reasoning ability (R) and cognitive achievement (H)	Kenya Tanzania	0.476 (6.70) 0.280 (4.30)	0.192 (2.47) 0.112 (1.42)	R: 0.006 (1.32) H: 0.020 (6.18) R: 0.001 (0.15) H: 0.013 (3.22)
Alderman et al. (1996)	Pakistan	Cognitive skills (C) and Ability (A)		0.046 (4.53)	0.023 (1.32)	C: 0.007 (1.41) A: 0.006 (0.67)

Note: a – (Literacy/Numeracy). na means not available. **Source:** Author's compilation.

3.2.2.2 Fixed Effects Regression Using Twins or Siblings Sample[8]

With fixed effects estimation, samples of twins and siblings have been utilised to deal with the problem of unobserved ability. It argues that these samples, especially a sample of identical (monozygotic) twins, possibly offer a high degree of internal validity, an estimation design that successfully uncovers a causal effect of a variable of interest that is schooling in this case (Bingley, Christensen, and Walker, 2007). The justification of using identical samples is that they are inherited the same genetic traits implying same innate ability and, most likely, the same degree of other endowments, e.g. investment in schooling, peer influences, geographic and sociological influences, to name a few. Hence, after taking a first difference within identical twin pairs, the unobserved ability and other endowments are removed from the estimation model. In other words, within-twin differences in education are uncorrelated with unobservable within-twin differences as the unobserved ability and other endowments are removed from the error term. In terms of external validity (an issue whether the estimated causal relationship can be generalised to other populations or settings), Bingley et al. (2007) argue that there is an external validity in the monozygotic twins estimation as Christensen et al. (2006) find that the distribution of educational outcomes among non-twins and twins samples are similar in nature, in spite of their substantial difference in the distributions of weights at birth.

Table 3-2 shows the estimated results from the studies of monozygotic (MZ) and dizygotic (DZ) twins (identical and fraternal twins). In general, the OLS estimates of general sample are greater than those of within twin pair (WT) sample as estimation from WT sample arguably removes the ability bias, both in studies of MZ and DZ twins. Moreover, IV estimates are likely to be greater than those of OLS with WT sample, but smaller than those of OLS with general sample. This is explained by the hypothesis that OLS estimates from first difference suffer from the classical measurement error bias which downwardly biases estimated results towards zero. However, some studies show contradicting results.

Nevertheless, the literature of fixed effects regression using twins samples is subject to two main criticisms. The first and the most important criticism is that there may be a problem in the internal validity of the twins studies. Neumark (1999) and Bound and Solon (1999) argue that the within differences in schooling may be correlated with the error term since presumption that the omitted ability is entirely made up of a genetic effect and a family effect does not necessarily hold true. If some parts of the omitted ability still remain in the error term after first differencing, the within-twin estimator may still suffer from an omitted variable bias and, in fact, may even be more biased than OLS estimates (Neumark, 1999 and Bound and Solon, 1999). Hence, the problem of endogeneity may still exist in the fixed effects estimation with the samples of twins. The second criticism is in terms of external

[8] See Harmon, Oosterbeek, and Walker (2003) and Bingley, Christensen, and Walker, 2007 for detailed and comprehensive literature review.

validity. As the estimation is based on a small and non-random twins sample, it may be difficult to generalise the estimated causal relationship to the non-twins population. Blachflower and Elias (1993) show strong evidence that, in comparison to non-twins of a similar age and environment, a sample of twins can be quite different. Even though estimation using twins samples indicates the differences between OLS and IV estimates and is expected to deal with the endogeneity bias, due to aforementioned arguments, estimated results still have problems both in terms of internal and external validity. Thus, it is difficult to draw any concrete inferences and conclusions from the studies of twins.

Table 3-2 Studies of Monozygotic (MZ) and Dizygotic (DZ) Twins

Study	Country	Sample	Number of Twin Pairs	OLS	OLS within twin pair	IV within twin pair
MZ twins						
Ashenfelter and Krueger (1994)	US	Twinburg, 1991	147	0.084 (0.014)	0.092 (0.024)	0.129 (0.030)
Berhman et al. (1994)	US	NAS-NRC, 1973	141	0.094[a] (0.011)	0.035 (0.004)	0.101 (0.012)
Miller et al. (1995)	Australia	Australian Twins Register, 1985	602	0.064 (0.002)	0.025 (0.005)	0.048 (0.010)
Ashenfelter and Rouse (1997)	US	Twinburg, 1991-1993	333	0.110 (0.009)	0.070 (0.019)	0.088 (0.025)
Berhman and Rosenzweig (1997)	US	Minnesota Twins Register, 1993	720	0.113[a] (0.005)	0.104 (0.017)	n.a.
Miller, Mulvey, and Martin (1997)	Australia	Australian Twins Register, 1985	Male: 282	0.071[d] (0.003)	0.023 (0.008)	0.033 (0.014)
			Female: 320	0.057[d] (0.002)	0.028 (0.006)	0.058 (0.011)
Rouse (1998)	US	Twinburg, 1991-1993, 1995	453	0.105 (0.008)	0.075 (0.017)	0.110 (0.023)
Isacsson (1999)	Sweden	Swedish Twin Registry, 1990	2,492	0.046 (0.001)	0.022 (0.002)	0.024[b] (0.008)
Isacsson (2004)	Sweden	Swedish Twin Registry, 1990	2,609	0.066[c] (0.009)	0.028[c] (0.009)	0.052[c] (0.036)
Bonjour et al. (2004)	UK	St Thomas' Hospital Twins Register	Female: 187	0.077 (0.001)	0.039 (0.023)	0.077 (0.033)
DZ twins						
Berhman et al. (1994)	US	NAS-NRC and Minnesota, 1973	n.a.	0.073 (0.003)	0.057 (0.005)	n.a.
Miller et al. (1995)	Australia	Australian Twins Register, 1985	568	0.066 (0.002)	0.045 (0.005)	0.074 (0.008)
Miller, Mulvey, and Martin (1997)	Australia	Australian Twins Register, 1985	Male: 164	0.071[d] (0.003)	0.029 (0.011)	0.051 (0.019)
			Female: 158	0.057[d] (0.002)	0.049 (0.007)	0.071 (0.011)
Isacsson (1999)	Sweden	Swedish Twin Registry, 1990	3,368	0.047 (0.001)	0.039 (0.002)	0.053[b] (0.006)
Isacsson (2004)	Sweden	Swedish Twin Registry, 1990	3,601	0.066[c] (0.008)	0.047[c] (0.009)	0.056[c] (0.003)

Note: a – GLS estimate. b – not instrumented but evaluated at a reliability ratio of 0.88. c – evaluated at upper secondary level of schooling. d – pooled DZ and MZ.
Source: Author's modification from Bingley, Christensen, and Walker (2009) based on Bound and salon (1999) and Card (1999).

3.2.2.3 Instrumental Variable (IV) Estimation Using Natural or Quasi-Experiment

Traditional Instrumental Variable

A traditional IV estimation is introduced into the research of returns to schooling to provide estimates of the causal effect of schooling on wages that are arguably free from ability bias. The IV estimation utilises the fact that the instrument variable is correlated with the endogenous explanatory variable (years of schooling), but has no correlation with the error term in the main regression model. In other words, the instrument variable must satisfy two conditions, including strong first stage and exclusion restriction, to be able to draw a causal relationship between years of schooling and wages. Several studies attempt to use various kinds of instruments; however, estimates from the traditional IV estimation possibly suffer from the endogeneity bias. Since the identification strategy cannot be tested by any statistical methods, it tends to be easily refuted by any arguments convincing that there is a correlation between the instrument variable and the error term. Moreover, there is the problem of weak instrument which occurs when the instruments are only weakly correlated with the endogenous variables in the first-stage equation. Bound, Jaeger, and Baker (1993, 1995) put it, "cure can be worse than the disease" since a weak instrument results in biased and inconsistent estimates. The biased estimates tend to be in the same direction as OLS (Bound, Jaeger, and Baker, 1993, 1995; Chao and Swanson, 2005). Furthermore, with weak instruments, statistical hypothesis testing has incorrect size, and confidence intervals are wrong.

IV Estimation with Regression Discontinuity Design (RDD)

An IV estimation with RDD gains its popularity due to its advantages over a traditional IV estimation, especially in terms of its internal validity. As discussed in the previous section, the traditional IV estimation still possibly produces biased estimates since it is difficult to justify the internal validity, both in terms of relevance and exogeneity conditions. Regarding the assumption of relevance, the IV with a quasi-experiment ensures that there is a causal relationship between the event (law change) and the endogenous variable (years of schooling); and between the event (law change) and the outcome (wage) through the effect of the endogenous variable (years of schooling). In terms of exogeneity condition, IV with a quasi-experiment is exogenous to an individual as RDD exploits a particular event, compulsory education law change in this study, being random to individuals. Hence IV with RDD is not correlated with individuals' unobserved variables, e.g., ability and preferences, and has less possibility to have the endogeneity bias. In contrast, the traditional IV estimation is often based on arbitrary identification strategy and there is no statistical way to check the exogeneity. Thus, in terms of internal validity, the IV utilising a quasi-experiment based on interventions in the school system is more convincing than a traditional IV estimation.

However, the IV utilising a quasi-experiment is not without a caveat. Even though the IV estimation with RDD satisfies both conditions, relevance and exogeneity, for the internal validity, the causal relationship may not be generalised to other populations and settings. The issue of generalisation refers to the external validity. In fact, the IV estimation with RDD lacks the element of randomization since it mainly relies on the arbitrary cut-off. Hence, only within a narrow bandwidth around the cut-off, treatment assignment approximates randomization or treatment effect. This method is equivalent to random assignment in a neighbourhood, which explains the reason that the regression discontinuity design is closer to randomised experiment than other quasi-experimental methods. Imbens and Angrist (1994) argue that the RDD estimates the average treatment effect from those who comply with the treatment. This estimated average treatment effect is commonly referred to Local Average Treatment Effect (LATE). Even though it is difficult to estimate Average Treatment Effect (ATE) as a full compliance from the whole population may not happen, the LATE converges to the ATE as the number of compliers becomes an increasingly large proportion of the sample (Oreopoulos, 2006).

Even though a randomised controlled trial (RCT) is the ideal estimation method, it is not possible to conduct RCT in most of the studies of returns to education. To the best of my knowledge, there is only one study from Schultz (2004) in estimating the rates of return to schooling by utilising RCT. Schultz estimates the internal rates of return to schooling from Progresa programme. The Progresa programme was implemented in 495 localities in rural communities with an initial sample size of 40,959 children. The main propose of the programme is to evaluate the effects of conditional cash transfer (CCT) given to rural poor families, both in terms of the school attainment and the returns to schooling. Schultz reports that the rates of return to schooling from the Progresa programme is approximately eight per cent per year. Nevertheless, this estimate is based on a strict set of assumptions regarding the behaviour of the sample. For example, Schultz assumes that "the rural youth migrate to the urban area at age 18, and then work until their retirement at age 65." Unfortunately, the magnitude and the direction of bias from OLS estimation are still inconclusive since there is no information on OLS estimates provided.

From previous arguments, other estimation methods may still suffer from endogeneity bias and RDD is closer to randomized experiment than other quasi-experimental methods e.g. differences-in-differences (Raju and Trias, 2010)[9]. Therefore, the second best candidate among various methods of estimation of returns to schooling is the IV estimation with RDD. The estimated results of RDD is discussed in Section 3, "Why Thailand?: Issues from the Previous Literatures of Natural Experiment with Compulsory Education IV."

[9] This material is meant to be a supporting material for the book called, "Impact Evaluation in Practice" (Gertler, et al., 2010).

Previous Studies of Returns to Schooling Using IV Estimation (Excluding Studies with Compulsory Education IV)

In general, it is found that the IV estimates are greater than those of OLS. **Table 3-3** outlines a list of the selected previous studies of returns to schooling using IV estimation, excluding studies with compulsory education IV. Several instrument variables have been introduced into the literature, ranging from quarter of birth (Angrist and Krueger, 1991), draft lottery number (Angrist and Krueger, 1992), school building project (Duflo, 1999), to name a few. Some of these instruments are purely traditional IVs which are difficult to justify their internal validity, namely college proximity (Card, 1995), family composition (Butcher and Case, 1994; Dearden, 1998), parental income and education (Dearden, 1998; Uusitalo, 1999), location of residence (Uusitalo, 1999), while the rest are based on quasi-experiments utilising a random nature of a particular event to estimate the rates of return to education.

Why Are IV Estimates Greater than Those of OLS?

Heckman, et al. (2006) illustrate that theoretically IV estimates should be smaller than those OLS under the estimation from traditional Mincer model. For simplicity, the age component in the Mincer model is omitted from the analysis.

(3-2) $$\ln \boldsymbol{Y} = \boldsymbol{\alpha} + \boldsymbol{\beta S} + \boldsymbol{U}$$

Under the OLS estimation with the endogeneity problem, the OLS estimates suffer from an upward ability bias. Therefore, the estimated rate of returns $\hat{\beta}_{OLS}$ is greater than the "real" rate of returns β.

(3-3) $$\textbf{plim} \quad \widehat{\boldsymbol{\beta}}_{\boldsymbol{OLS}} = \boldsymbol{\beta} + \frac{\boldsymbol{COV(S,U_0)}}{\boldsymbol{V(S)}} > \boldsymbol{\beta}$$

Given that IV estimation corrects for the endogeneity problem, the estimated rate of returns $\hat{\beta}_{IV}$ should be equal to the "real" rate of returns β.

(3-4) $$\textbf{plim} \quad \widehat{\boldsymbol{\beta}}_{\boldsymbol{IV}} = \boldsymbol{\beta} + \frac{\boldsymbol{COV(Z,U_0)}}{\boldsymbol{COV(Z,S)}} = \boldsymbol{\beta}$$

Hence, the IV estimates are theoretically less than the OLS estimates

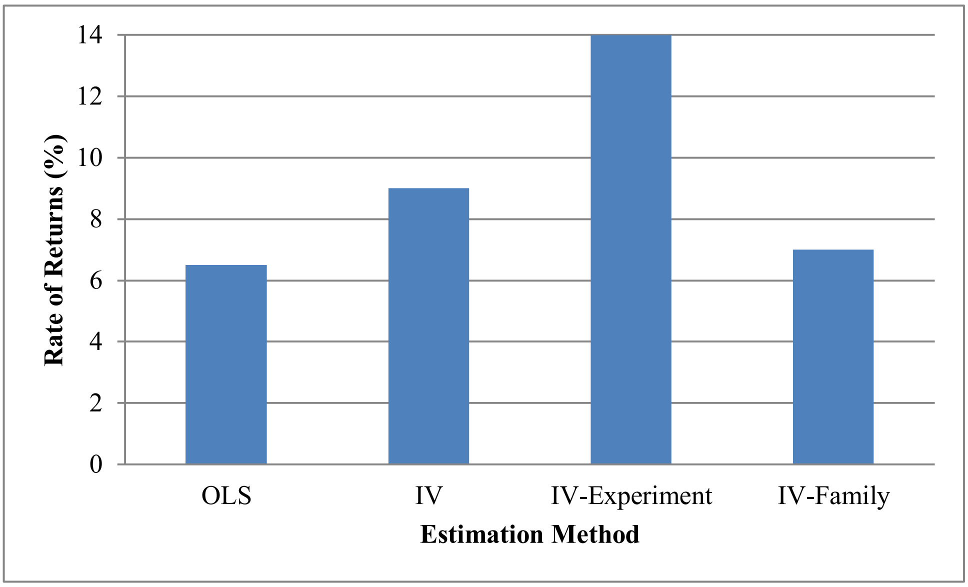
(3-5) $$\widehat{\boldsymbol{\beta}}_{\boldsymbol{IV}} < \widehat{\boldsymbol{\beta}}_{\boldsymbol{OLS}}$$

IV estimates are often higher than those of OLS. Harmon et al. (2003) compare estimated returns to schooling from 942 studies using same data set but different methods of estimation, including OLS estimation, traditional IV estimation, IV estimation based on interventions in the school system, and IV estimation with family background. **Figure 3-3** shows that the average OLS estimate is approximately 6.5 per cent which is the lowest among all estimation. In contrast, the average IV estimate based on education reforms is around 14 per cent indicating the highest rate of return to schooling. The average estimates from the traditional IV and the IV using family background as instruments are also greater than that of OLS estimation. Card (1999) points out that the estimates based on education reform are likely to be at least 20 per cent above the OLS estimates.

Although, as shown above, estimates from IV should theoretically be lower than those of OLS, there are few plausible hypotheses that potentially explain the reasons that IV estimates are greater than those of OLS. First, Bound and Jaeger (1996) suggest that the IV estimation may produce further upward biased results due to the unobserved heterogeneous characteristics between the control and the treatment groups, e.g. family backgrounds, and schooling quality. However, Card (1999) argues that even with the family background controls, the IV estimation still yields a higher estimated result. **Figure 3-3** also supports Card's (1999) argument. Second, the effect of the ability bias in the OLS estimates is substantially smaller than that of the measurement error bias; therefore, the downward bias can be observed from the OLS estimates (Griliches, 1977; Angrist and Krueger, 1991). Nonetheless, the bias gap[10] between OLS and IV is significantly large. According to Card (1999), only ten per cent of the bias gap is explained by the measurement error, and it is unreasonable to conclude that a large positive gap only comes from the effect of the measurement error. A third explanation by Ashenfelter and Harmon (1998) suggests that higher estimates of IV are possibly due to a pure publication bias caused by a tendency to report a large point estimate of the returns to schooling. Finally, Card (1999, 2001) and Lang (1993) posit that the IV estimation yields LATE rather than ATE as the education reform, especially compulsory schooling, tend to only change the behaviour of individuals who would otherwise have low schooling. These individuals are more likely to come from socially and economically disadvantaged groups and have higher returns to schooling than those of the whole population (LATE is greater than ATE). Thus, the IV estimation based on interventions in the school system is expected to give higher rates of return to schooling than those of the OLS estimated from the whole population. In fact, if heterogeneous characteristics, e.g. unobserved ability and cost terms are not observed among individuals in the treatment group, the same level of school attainment and returns to education would be expected among the treatment group members through the effect of the intervention.

[10] Bias gap is defined as a difference between OLS and IV estimates in absolute term.

Figure 3-3 Meta-Analysis of Models with Endogenous Schooling



Note: Total number of estimates is 942 (OLS: 863 studies, IV: 79 studies, IV-Experiment: 17 studies, and IV-Family: n.a.

Source: Harmon et al. (2001) and Harmon et al. (2003).

Table 3-3 Previous Studies of Returns to Schooling Using IV Estimation (Excluding Studies with Compulsory Education IV)

Author	Country	Instrument	Sample	OLS	IV
Angrist and Krueger (1991)	US	Year*Quarter of Birth; State*Quarter of Birth	1970/1980 Census: Men born 1920-29, 1930-39, 1940-49	0.063 (0.000)	0.081 (0.033)
Angrist and Krueger (1992)	US	Draft Lottery Number*Year of Birth	1979-85 Current Population Survey (CPS): Men born 1944-53 (potential Vietnam War draftees)	0.059 (0.001)	0.066 (0.015)
Card (1995)	US	Nearby college in county of residence in 1966	National Longitudinal Survey (NLS): Men aged 14-24 in 1966 sampled as employed in 1976	0.073 (0.004)	0.132 (0.049)
Butcher and Case (1994)	US	Presence of siblings (sisters)	Panel Study of Income Dynamics (PSID) 1985: White women aged 24+	0.091 (0.007)	0.185 (0.113)
Uusitalo (1999)	Finland	Parental income and education,	Finnish Defence Forces Basic	0.089 (0.006)	0.129 (0.018)

		location of residence	Ability Test Data matched to Finnish income tax registers		
Duflo (1999)	Indonesia	Indonesian school building project	Indonesian—males	0.077 (0.001)	0.091 (0.023)
Denny and Harmon (2000)	Ireland	Irish school reforms—abolition of fees for secondary schooling	The Economic and Social Research Institute (ESRI) 1987—males	0.080 (0.006)	0.136 (0.025)
Dearden (1998)	UK	Family composition, parental education	National Child Development Study (NCDS): Men	0.048 (0.004)	0.055 (0.005)

Source: Author's adjustment from Harmon, Oosterbeek, and Walker (2003).

Section 3. WHY THAILAND?: ISSUES FROM THE PREVIOUS LITERATURE OF QUASI-EXPERIMENT WITH COMPULSORY EDUCATION IV

There have been studies from both developed and developing countries estimating the rates of return to schooling by utilising quasi-experiment with compulsory education IV. This section raises debates, issues, and critiques of the previous studies and justifies that Thai data is appropriate in estimating returns to education.

3.3.1. Rates of Return to Schooling Using Compulsory Education as IV from Developed Countries

Table 3-4 summarises the previous literature regarding the rates of return to schooling using compulsory education as IV from developed countries. Harmon and Walker (1995) are the pioneers of the research in this area. They exploit the change in the minimum school leaving age in the UK in 1947 and 1973 as instrumental variables and find that the estimated returns to schooling are approximately 15 per cent, while those of OLS estimation are around 6 per cent. The result shows that the IV estimates are greater than those of OLS. In favour of Harmon and Walker (1995), Oreopoulos (2006) argues that IVs of previous studies affect fewer than 10 per cent of the population exposed to the instrument, whereas the minimum school leaving age in the study of Harmon and Walker (1995) covers almost half of the population. In other words, the estimated local average treatment effect from Harmon and Walker's study (1995) is close to the real average treatment effect. Oreopoulos (2006) improves Harmon and Walker's estimation model (1995) by adding birth cohort control to allow for systematic inter-cohort changes in school attainment. His findings for the UK are consistent with those of Harmon and Walker (1995) in which the IV estimates are higher than those of OLS. Nevertheless, the rates of return to schooling are approximately 10 per cent, which is lower than that of previous study. In terms of the results from the US and Canada where the change of minimum school leaving age seems not to affect most of the population, the IV estimates are substantively greater than those of OLS. This is arguably due to the fact that the LATE in the US and Canada represents the estimated returns to schooling of a socially disadvantaged group which usually has higher returns to schooling. Even though Oreopoulos's (2006) study is more convincing than others in terms of model specification and the estimated results, his estimates may still suffer from the endogeneity bias. In the context of the study based on compulsory education reform, Card (1999) argues that the samples used in developed countries have a high tendency to be affected by the World War II, that makes the treatment and control groups are fundamentally different and biases the results of returns to schooling. Ichino and Winter-Ebmer (2004) find that the average year of schooling of children born during the war time, 1930-1935, is significantly lower than children born in the cohorts that do not experience the war.

From late 2000s, many studies argue that compulsory education has a positive but smaller effect than previous studies, or zero effect on earnings. The second wave of the studies finds statistically significant result with much smaller returns of about 3-6 per cent, compared with those of previous studies, 10-14 per cent. These studies likely use the school-based interventions even if they do not affect most of the population in the analysis, while they try to argue that the compulsory education does not have an impact on returns to schooling. Hence it is common to observe that the returns to schooling from these studies are not statistically significant or even have a negative sign.

In the most recent study of Stephen Jr. and Yang (2014), an interesting methodological issue regarding the model specification has been raised. They find that adding region and school quality control leads the effect of compulsory education on earnings towards zero or even wrong-signed. They argue that without the interaction between region control and year of birth, changes across regions over time of school quality cannot be captured in the estimation. The effects of differences in school quality (and perhaps family background) may cause estimates to be biased since the source of upward bias is not limited to only an individual's ability, but also includes family background and quality of schooling (Schultz, 1988, Behrman, 1990, Strauss and Thomas, 1995).

In conclusion, there are two main debates in the literature from developed countries. The first debate is that the LATE is far from the ATE as the interventions do not affect the majority of the population. Moreover, the IV estimates tend to be higher than those of OLS since the estimated results indicate the estimated returns to schooling of a socially disadvantaged group that usually has higher returns to schooling. The second debate raised by Stephen Jr. and Yang (2014) seems reasonable and is worth further examination as it is consistent with the theoretical discussion regarding the ability bias. Nevertheless, Card (1999) and Ichino and Winter-Ebmer (2004) argue that, due to the effect of the World War, the samples from developed countries may suffer from the endogeneity bias which possibly results in a fundamental difference between treatment and control groups. Thus, it is also worth examining the previous studies from developing countries where the effect of the war is less pronounced. Moreover, as discussed in Chapter 2, developing countries have their own characteristics which are radically different from developed countries. Therefore, this warrants value for their investigation.

Table 3-4 Literature Review of Rates of Return to Schooling Using Compulsory Education as IV from Developed Countries

Author	Country	Identification		Schooling Coefficients	
				OLS	IV/DD
Harmon and Walker (1995) [a]	The United Kingdom	The raising of the school-leaving age from 14 to 15 in 1947, and the raising of the school-leaving age from 15 to 16 in 1973.		0.061 (0.001)	0.153 (0.015)
Oreopoulos (2006, 2008)	The United Kingdom, The United States, and Canada	The raising of the school-leaving age from 14 to 15 in 1947 (UK) and in 1957 (Northern Ireland). A compulsory attendance laws (the United States and Canada).	The United States Canada The United Kingdom Britain	0.078*** (0.0005) 0.099*** (0.0007) 0.085*** (0.002) 0.083*** (0.003)	0.142*** (0.0119) 0.096*** (0.0254) 0.108*** (0.0328) 0.101** (0.0421)
Oosterbeek and Webbink (2007)	The Netherlands	A change in the minimum school leaving age in 1975.	1953-1963	na	-0.004 (0.04)
Pischeke and Wachter (2008)	Germany	An introduction of the ninth grade since 1949.	Qualification and Career Survey Micro Census	0.066 (0.002) 0.074 (0.001)	0.058 (0.038) 0.016 (0.015)
Devereux and Hart (2010)	The United Kingdom	Raising of the school-leaving age from 14 to 15 in 1947 (the UK).	No age control Quartic age control Age dummies control	na na na	0.021 (0.024) 0.019 (0.023) 0.025 (0.027)
Devereux and Fan (2011)	The United Kingdom	The education expansion during 1989 and 1994.	Men Hourly wages Men Weekly earnings Women Hourly wages Women Weekly earnings	0.078*** (0.003) 0.069*** (0.003) 0.096*** (0.003) 0.122 *** (0.002)	0.062*** (0.016) 0.066*** (0.019) 0.053*** (0.014) 0.066*** (0.013)
Clay et al. (2012)	The United States	A compulsory attendance laws		0.080 *** (0.003)	0.114 (0.098)
Grenet (2013)	France and the United Kingdom	The raising of the school-leaving age from 14 to 16 in 1967 (France), and a dummy variable for the raising of the school-leaving age from 15 to 16 in 1972 (the UK).	France Men France Women UK Men UK Women	0.073*** (0.001) 0.087*** (0.001) 0.095*** (0.001) 0.119*** (0.001)	-0.004 (0.029) -0.007 (0.013) 0.069*** (0.029) 0.067*** (0.011)
Stephen Jr. and Yang (2014) [a,b]	The United States	A compulsory attendance laws	All white ages 25-54	0.068 (0.0003)	-0.003 [-0.058 , 0.016]

Note: a no report on statistical significance; b the confidence intervals based on Moreira's CLR test, which are reported in brackets below the 2SLS estimates, allow for the correlation of the error terms within each state of birth/year of birth cell (Stephen

Jr. and Yang, 2014); na – not available.
Source: Author’s compilation.

3.3.2. Rates of Return to Schooling Using Compulsory Education as IV from Developing Countries

Table 3-5 is the summary of the previous literature from developing countries. There are few studies using compulsory education as IV in estimating returns to schooling since the issue of data scarcity is common in most of developing countries. Although the studies from China utilise the same data set, their findings contradict each other. This problem also applies to the case of Turkey. In case of China, the IV estimates from the study of Fang et al. (2012) is around 2 per cent and statistically significant at one per cent, while that of La Vincent and College (2014) is about 6 per cent but not statistically significant. However, both studies share similar magnitude of OLS estimates and show that the IV estimates are lower than those of OLS. Aydemir and Murat (2013) report that the rates of return to schooling in Turkey are approximately 16-30 per cent for the IV estimation. With gender disaggregation, the returns to schooling for men are around 0.5 per cent and not statistically significant, whereas those of women are approximately four per cent and statistically significant at five per cent (Aydemir and Murat, 2015). The estimated results from the latter study are far from those of the original study in 2013. However, both studies still illustrate a similar magnitude of OLS estimates but different bias direction. Interestingly, the studies from China and Turkey indicate the similar trend that the OLS estimates are higher than those of IV. In other words, the upward ability bias can be expected in the OLS estimation from both countries. In Indonesia, the rates of return to schooling estimated from IV estimation are statistically significant and around 13 per cent (Parinduri, 2014). This is consistent with the first wave of the studies from developed countries. However, the method of estimation used in this study is a fuzzy regression discontinuity design (FRDD)[11]. With FRDD, the identification strategy depends on a cut-off, around which there is a discontinuity in the probability of assignment from 0 to 1. Nonetheless, this assumption is not likely to hold; therefore, FRDD possibly produces biased estimates of returns to schooling. Moreover, the result from Parinduri (2014) is not directly comparable to other studies using a sharp regression discontinuity design in the estimation.

In sum, the estimated results of developing countries are different from those of developed countries. The results from developing countries show that the OLS estimates are higher than those of IV estimation, though some of the results are not statistically significant. In other words, the ability bias is dominant in the OLS estimation of developing countries. On the other hand, the estimates from developed countries e.g. US and UK, tend to produce OLS estimates that are lower than those of IV. As Chapter 2 illustrates that the economic and social conditions of developing countries are fundamentally different from those of developed countries, this may explain the different patterns of endogeneity bias from countries with different stages of development. This finding is worth further investigation.

[11] For more detail see Appendix III.

Table 3-5 Literature Review of the Rates of Return to Schooling Using Compulsory Education as IV from Developing Countries

Author	Country	Identification		Schooling Coefficients	
				OLS	IV/DD
Parinduri (2014)	Indonesia	An arbitrary rule that assigned students to a longer school year in 1978-1979.	Log hourly wages Log monthly wages	na	0.13*** (0.04) 0.17*** (0.04)
Fang et al. (2012)	China	A compulsory education law 1986.		0.09 *** (0.004)	0.02*** (0.06)
La and College (2014)	China			0.0842*** (0.00369)	0.0585 (0.0684)
Aydemir and Murat (2013)	Turkey	The extension of compulsory schooling from 5 to 8 years in 1997.	Linear time trend for year of birth Quadratic time trend for year of birth	0.0304*** (0.00395) 0.0304*** (0.00397)	0.301*** (0.0814) 0.158*** (0.0300)
Aydemir and Murat (2015)	Turkey		Men Women	0.050*** (0.002) 0.063*** (0.003)	0.005 (0.014) 0.038** (0.015)

Note: na – not available.
Source: Author's compilation.

Section 4. RATES OF RETURN TO SCHOOLING FROM THAILAND

3.4.1. Overall Estimated Results

The studies of returns to schooling from Thailand are very similar. They share common characteristics, in terms of data, research methodology, and estimated results. Regarding data and research methodology, most of the studies utilise the Labour Force Survey data collected by the National Statistics Office as it is the most comprehensive data containing large representative samples and all variables required by the Mincer equation. However, using the Labour Force Survey data is not without caveats. Both advantages and disadvantages of the Labour Force Survey data are discussed in Chapter 5 Data and Methodology. Either a single year or multiple years of the Labour Force Survey data are used in the literature. The main statistical estimation methods are cross-sectional or pooled cross-sectional OLS regressions without fixing for the endogeneity bias, even in the most recent studies, e.g. Wannakrairoj (2013) and Tangtipongkul (2015). With a similar data set of the same or different years, a comparable magnitude of OLS estimates can be observed among previous studies. The OLS estimates are approximately 10-11 per cent. As mentioned earlier, these estimated returns to schooling highly possibly suffer from the endogeneity bias, especially the ability bias. **Table 3-6** briefly outlines the literature review of the rates of return to schooling from Thailand.

3.4.2. The Returns to Education in Thailand: A Pseudo-Panel Approach

This section thoroughly reviews the study by Warunsiri and McNown (2010). The study offers the most unique and interesting estimation method and estimated results among literature regarding returns to schooling in Thailand. In attempt to solve the endogeneity bias, Warunsiri and McNown (2010) use two types of estimation, including IV estimation with college proximity as an instrument, and estimation pseudo-panel approach, to estimate returns to education in Thailand. They use the Labour Force Survey data from 1986 to 2005.

Following Card (1995) and Uusitalo (1999), they adopt college proximity as an IV by arguing that the existence of college in each area would reduce the costs of schooling and, in turn, increase the number of years of schooling. Furthermore, following Deaton (1985), Warunsiri and McNown (2010) construct synthetic cohorts (pseudo panels) by using a pooled cross sectional LFS data. They treat those cohorts as repeated observations over time and utilise the fixed effects estimation with the (pseudo) panel data to eliminate unobserved ability from the error term. The following sector briefly explains the method of estimation using fixed effects with a pseudo-panel approach.

Mincerian equation:

(3-6) $$\log y_i = \beta_0 + \beta_1 S_i + \beta_2 X_i + \beta_3 X_i^2 + e_i$$

where the log of individual earnings (y_i) is a function of years of education an individual attended (S_i) and number of years an individual has worked after completing his/her education (X_i). e_i represents a disturbance term. Equation (3-7) is the time and individual specific representation of (3-6) where *i* indexed individuals and *t* indexes time periods:

(3-7) $$\log y_{it} = \beta_0 + \beta_1 S_{it} + \beta_2 X_{it} + \beta_3 X_{it}^2 + \alpha_{it} + e_{it}$$

where the log of earnings (y_{it}) of individual *i* at time *t* is a function of years of education an individual *i* attended at time *t* (S_{it}) and number of years an individual *i* has worked after completing his/her education (X_{it}) at time *t*. e_{it} represents a disturbance term of individual *i* at time *t*. The term α_{it} captures unobserved individual heterogeneity that includes ability and motivation. If α_{it} is observable, it can be included directly into the equation. However, in the absence of such information, unobservable variables represented by α_{it} will cause a regression estimation to produce biased results due to the unobserved variable bias.

Deaton (1985) Pseudo Panel:

(3-8) $$\log \overline{y_{ct}} = \beta_0 + \beta_1 \overline{S_{ct}} + \beta_2 \overline{X_{ct}} + \beta_3 \overline{X_{ct}^2} + \overline{\alpha_{ct}} + \overline{e_{ct}}$$

where a subscript *c* refers to birth cohort. Birth cohort is a group of individuals *i* born in the same year. Data on log wages, years of schooling, working experience, and fixed effects of each cohort is the average value from those cohort's members. Under a large sample per each cohort, Verbeek and Nijman (1992, 1993, and 2007) argue that $\bar{\alpha}_{ct}$ can be treated as the unobserved cohort fixed effect α_c.

(3-9) $$\log \overline{y_{ct}} = \beta_0 + \beta_1 \overline{S_{ct}} + \beta_2 \overline{X_{ct}} + \beta_3 \overline{X_{ct}^2} + \alpha_{ct} + \overline{e_{ct}}$$

Hence, after taking a first difference, the unobserved cohort fixed effect α_c is removed from the estimation model. In other words, the regression model is possibly free from the endogeneity bias.

With individual sample, the OLS estimate is around 11 per cent, while the IV estimate is approximately 14 per cent. In contrast, the rates of return to schooling are roughly 15 per cent for IV estimation and fixed effects estimation with pseudo-panel data. The estimates show a downward bias of the returns to education in least squares regressions with individual data. Warunsiri and McNown (2010) argue that owing to the similar results from the IV and pseudo-panel estimations, the problem of endogeneity bias is possibly eliminated from the estimated rates of return to schooling.

Nonetheless, several problems regarding internal validity of both IV and pseudo-panel estimations can be raised. In terms of the IV estimation, since the identification strategy cannot be tested by any statistical methods, it tends to be easily refuted by any arguments convincing that there is a correlation between the instrument variable and the error term. Carneiro and Heckman (2002) and Meghir and Rivkin (2011) argue that the college proximity is potentially correlated with individual ability; therefore, the IV does not satisfy the condition of exclusion restriction. Individuals and schools are not randomly allocated as individuals with higher endowment, both in terms of innate ability and family wealth, tend to live in a more developed city in which schools tend be located. Regarding the pseudo-panel estimation, the assumption that the unobserved ability is time-invariant may not hold true (Ding and Lehrer, 2014) because the heterogeneous ability evolves over the lifecycle. Regardless of time-variant or time-invariant nature of the ability, the panel regression with fixed effect may suffer from further upward bias as running panel regression with fixed effect excludes time-invariant independent variables from the regression, such as region controls and schooling quality. Stephen Jr. and Yang (2014) find that adding region and school quality control leads the effect of compulsory education on earnings towards zero or even wrong-signed. They argue that without the interaction between region control and year of birth, changes across regions over time of school quality cannot be captured in the estimation. The effects of differences in school quality may cause estimates to be biased since the source of upward bias is not limited to only an individual's ability, but also includes family background and quality of schooling (Schultz, 1988, Behrman, 1990, Strauss and Thomas, 1995).

Table 3-6 Literature Review of Rates of Return to Schooling from Thailand

Author	Sample	Method of Estimation	Dependent Variable		Schooling Coefficients
Mark Bluag (1974)	- Random sample of households in metropolitan area (urban) of Thailand (2,000 men). - A quota sample of highly educated people (3,000 men and women) to prevent unknown biases. - 5,000 individuals (1,485 college graduates).	Regressions with stepwise technique. Independent variables: age, education (years), sex, parents' education (sum total in years), language spoken at home, father's occupation, types of schooling, higher education filed, employment, and occupation. (*): standard error.	log earnings		0.3419 (-0.0517) not significant
Futoshi Yamauchi (2002)	- 1994-1996 and 1998-2000 Labour Force Survey. - Only sample from metropolitan Bangkok (less than 40 years of age) to 1) avoid heterogeneities in labour-demand conditions across local labour markets; 2) seasonal fluctuations in production are minor in Bangkok; 3) the share of migrants is large. - Sample sizes are 3,471 and 4,273.	Pooled cross-section regressions. Independent variables: age, education (years), sex, and time and origin fixed effects. (*): absolute t value No report on statistical significance.	log weekly wage	1994-1996 1998-2000	0.0713 (25.42) 0.0734 (42.47)
Moenjak and Worswick (2003)	- 1989 - 1995 Labour Force Survey. - 15 - 60 years of age. - Sample is restricted to individuals who are sons or daughters of the household heads. This suggests that the sample is a selected group, especially if those who live with parents tend to be those who cannot afford to live apart on their earnings. [a]	Pooled cross-section regressions Independent variables: experience, vocational education, region of residence, are of residence, marital status, migration status, survey year and birth cohort. (*): corrected standard errors for self-selection corrected and normal standard errors for OLS.	log hourly earnings	Vocational education Males (self-selection corrected) Males (OLS) Females (self-selection corrected) Females (OLS)	 0.639** (0.146) 0.238** (0.028) 0.494** (0.180)

					0.207** (0.031)
Hawley (2004)	- 1985, 1995, and 1998 Labour Force Survey. - 24-35 years of age. - Sample sizes are approximately 2,717 to 7,655. [b]	OLS Regressions by years. Independent variables: education (years), experience, area of residence, region, and sector of employment. (*): robust standard error.	log monthly earnings	Men 1985 1995 1998 Women 1985 1995 1998	 0.111 *** (0.005) 0.103 *** (0.008) 0.103 *** (0.004) 0.108 *** (0.008) 0.103 *** (0.004) 0.107 *** (0.004)
Futoshi Yamauchi (2004,2005)	- Surveys conducted in some manufacturing industries in the Bangkok region in July to October 2001. - Sample size is 1,867. - The data do not represent worker population in manufacturing as a whole. The sample of plants is biased toward large scale establishments which imply that the estimates of returns to schooling can be larger than the population estimates if there is a positive correlation between establishment size and wages.	OLS Regression. Independent variables (base model): years of schooling, age, sex, company fixed effects. [c] Second model includes parents' schooling, school quality (GAT and SAT), and elementary-school province fixed effects. Third model is the second model excluding company fixed effects. (*): absolute t value. No report on statistical significance.	log base wage	Base model Second model Third model	0.0680 (10.94) 0.0630 (10.84) 0.0730 (10.79)
Warunsiri and McNown	- 1986-2005 Labour Force Survey. - 199,833 individual observations and 440	OLS, IV and WLS (Pseudo-panel) Independent variables: education (years),	log hourly wages	OLS	0.115 ** (0.000250)

(2010)	cohort-year observations.	age, and cohort dummies variables (for WLS). IV: a dummy variable identifying the provinces in which universities or teacher training colleges are located. (*): standard error. ** indicates significant at or below the 0.05 level.		IV WLS WLS-IV WLS (with cohort dummies)	0.141 ** (0.0103) 0.101 ** (0.00729) 0.148 ** (0.0194) 0.151 ** (0.0100)
Wannakrairoj (2013)	2012 Labour Force Survey with sample size of 19,099 individuals.	Cross-sectional regression. Independent variables: education (years), experience, and area of residence. (*): t-stat.	log monthly wages		0.1087 *** (117.56)
Tangtipongkul (2015)	2007-2010 Labour Force Survey with sample size of 200,000 (approximately).	Cross-sectional regression. Independent variables: education (years), and experience. (*): robust standard error.	log monthly earnings	2007 2008 2009 2010	0.1376 *** (0.0005) 0.1337 *** (0.0005) 0.1322 *** (0.0005) 0.1263 *** (0.0006)

Note: a As argued by Bedi and Gaston (1997), given the prevalence of joint families in developing countries, the individuals may simply live with their parents out of convenience in location or because they are participating in family enterprises. It is premature to assume that all the individuals living with their parents are still subsidized by their parents' incomes.
b This study uses data collected during the agricultural season and thereby is more likely to survey people who were at their permanent place of residence. Thailand is primarily an agricultural economy. Many of the farmers migrate to the cities in the dry season to find work (Chalongphob & Yongyuth, 1996). See critiques on advantages and disadvantages of LFS data pp. 275.
c Province fixed effects are included here because origin-specific heterogeneity (mostly unobservable in our data) that affect labour productivity could be important.
Source: Author's compilation.

Section 5. RETURNS TO SCHOOLING AND SELECTION BIAS

Schultz (1988) argues that a bias may occur when sample selection is related to the schooling-wage relationship. Sample selection criteria leading to a bias includes the choice of occupation, labour force participation, or migration, to name a few. In this section, the sample selection bias in returns to schooling is introduced according to individuals' demographic characteristics that are usually related to sample selection criteria such as choice of occupation, labour force participation, and migration, especially in the case of developing economies. Those individuals' characteristics involve gender and area of residence.

3.5.1. Gender

In estimation of returns to education, it is inconclusive whether the "real" female rates of return should be higher or lower than the estimated one. Both Psacharopoulos (1985) and Schultz (1988) argue that the estimated female returns to schooling are either underestimated or overestimated due to the fact that the estimates are from unrepresentative sample of females. Female labour participation rate is low as females tend to concentrate in self-employed occupations or a non-wage labour market. Hence, the sample of estimation from LFS does not cover all females and deems to be unrepresentative. However, Schultz (1988) posits that women may choose to work fewer hours than do men regardless of female labour force participation rate. Therefore, the problem of sample selection may be underestimated.

3.5.2. Area of Residence

Estimates of returns to schooling may be also subject to a sample selection bias due to a rural-urban migration. In general, better educated individuals tend to migrate from rural areas to urban areas due to higher urban real wages (Greenwood (1975); Schwartz (1976); Schultz (1982b, 1988); Orazem and King (2007)). This phenomenon is described as, "human capital drain from rural to urban areas" (Amare, et al., 2012) and can be observed in other developing countries, namely Taiwan (Speare, 1974), Venezuela (Levy and Wadycki, 1974), Papua New Guinea (Ross, 1984). The evidence from the previous studies of Thailand also indicates the same pattern (Prachuabmoh et al., 1979, Tirasawat, 1985). Thus, the estimated urban rates of return do not purely represent the "real" urban rates of return as they also reflect rural immigrants' rates of return.

Without migration, the "real" rates of return to schooling in urban areas should be higher than the estimated rates. Conversely, the "real" rates of return to schooling in rural areas should be lower than the estimated rates. The underlying reason is that, without rural immigrants, urban workers can enjoy higher rents from disequilibrium between demand for and supply of educated workers since there are more job opportunities that educated workers can utilise their skills honed from their higher school years. However, an influx of rural migrants drives up the supply of educated workers in urban areas. Thus, returns to schooling decrease as the supply increases. Psacharopoulos (1987) posits that returns to schooling among different areas will converge through the process of labour migration. In contrast, the "real" rates of return in rural areas should be lower since there is less demand for educated workers in rural areas and there is no absorption of educated workers to urban areas to relieve the excess supply of educated workers in rural areas. Given that there is no migration, rural educated workers enjoy much smaller rents than those from urban areas. Nevertheless better educated individuals are more likely to migrate from rural areas to urban areas owing to higher urban real wages. Therefore, the rural-urban migration can be observed.

There are implicit assumptions underlying above discussion. First, Schultz (1988) and Agesa (2001) argue that net migration flow is characterised by the flow from rural to urban areas owing to higher real wages in urban areas and unbalance between demand for and supply of educated workers in rural areas. With the modern educational expansion, the supply of educated workers in rural areas increases rapidly, while the development of rural labour market cannot keep pace with the increase in educated labour supply. This results in an influx of migration flow from rural to urban areas. The trend can be observed in Latin American countries and Kenya (Schultz, 1988; Agesa, 2001). Second, regarding education of migrants, more educated individuals tend to migrate to urban areas as they can exploit their skills and earn higher wages in the urban labour market. Furthermore, those migrants are more likely to migrate after graduation (Schultz, 1988). The previous study from Colombia finds that the estimated returns to rural schooling indicate extremely downward bias as there is a large number of rural migrants in Colombia (Orazem and King, 2007). On the other hand, Schultz (1998) and Duraisamy (2002) argue that the sample selection bias is moderate in Côte d'Ivoire, Ghana, and India.

On one hand, it is possible that the rates of return to schooling in rural areas may be downwardly biased due to the problem of sample selection. The conventional estimations of returns to schooling usually ignore the selected nature of the rural samples. Rural labour participation rate is substantially low and rural individuals tend to concentrate in self-employed occupations or a non-wage labour market. Hence, the sample of estimation does not cover all rural individuals and deems to be unrepresentative.

However, the effect of bias due to outmigration may reduce the effect of selection bias due to an unrepresentative sample. As argued above, individuals with higher ability tend to migrate to urban areas to find a better schooling or jobs. Moreover, there is evidence showing that the selection bias due to unrepresentative sample may be modest in developing countries, e.g. Côte d'Ivoire (Schultz, 1988), Ghana (Schultz, 1988), and India (Duraisamy, 2002). After correcting for selection bias, Schultz (1988) and Duraisamy (2002) find that estimated returns to schooling are fairly close to uncorrected estimates.

Section 6. RESEARCH CONTRIBUTION

This study makes three main contributions in terms of methodology and also substantive aspect in the context of Thailand and, by implication, developing countries in general.

First, the previous literature reveals that there is a quite different pattern of the relative magnitudes of the estimates from OLS and instrumental variables (IV) estimation using compulsory schooling as IV between developed and developing countries. Investigating this difference can contribute to a better understanding of (a) how and when the conventional "ability bias" matters in estimating returns to schooling and (b) the impact of compulsory schooling in different settings.

The second contribution is in terms of substantive aspect in the context of Thailand. As Thailand experienced rapid economic development and structural transformation during 1960-1990, obtaining the rates of return to education in this period helps us better understand the process of Thai economic development as well as the interplay between the rates of return to schooling and the economic development process during 1980 to 1990. The overall social and economic conditions of the development in Thailand are consistent with the general characteristics of other developing countries. Hence, estimating the rates of return to schooling in Thailand, by implication, also provides better understandings on the role of human capital in the process of development in other developing countries. Due to the fact that developing countries possess radically different degrees of market completeness and different quality of institution from those of developed countries, this warrants value for investigation of the returns to schooling in the context of developing countries. This further investigation possibly gives a different economic pattern and implications of the returns to schooling.

Finally, the third contribution is on the construction of the database and discussion of the descriptive analysis for the discrepancies among different demographic characteristics, including gender, cohort, area of residence, region of residence, and economic sector. In addition to the overall estimates of returns to schooling, another important issue is an issue of heterogeneity in educational returns across individuals. Heterogeneity in individuals' demographic characteristics tends to distort the returns to education; for example, the female rates of return to education is likely to be higher than those of male. Hence, it is worth examining heterogeneous returns to schooling from different demographic characteristics.

Section 7. CONCLUSION

The rates of return to schooling is one of the most important topics that economists have been investigating, especially by utilising the Mincerian equation. Despite the fact that the Mincer model is still alive and well in fitting the actual age-wage data, there has been a long debate that the OLS estimate from Mincer equation is possibly biased due to the endogeneity problem. Hundreds of studies with different methods of estimation attempt to deal with the endogeneity bias but they fail to establish a causal effect of schoolings on earnings in the absence of randomised experiment. Even though a randomised controlled trial (RCT) is the ideal estimation method, it is not feasible to conduct the RCT in most of the studies of returns to education. The second-best candidate, which is close to the RCT, is the estimation with quasi-experiment, e.g. regression discontinuity design (RDD), differences-in-differences, to name a few. However, the studies utilising those methods of estimation are rare, especially in developing countries where the issue of data scarcity is prevalent.

This study mainly focuses on the IV estimation using compulsory schooling as IV. There are two main debates in the literature from developed countries. The first debate is that the LATE is far from the ATE as the interventions do not affect a majority of the population. Thus, the IV estimates tend to be higher than those of OLS since the estimated results indicate the estimated returns to schooling of a socially disadvantaged group that usually has higher returns to schooling. The second debate raised by Stephen Jr. and Yang (2014) is the concern of potential bias from the omitted variables, including school quality and family background. They argue that without the interaction between region control and year of birth, changes across regions over time of school quality cannot be captured in the estimation. The effects of differences in school quality may cause estimates to be biased since the source of upward bias is not limited to innate ability, but also includes family background and quality of schooling (Schultz, 1988, Behrman, 1990, Strauss and Thomas, 1995). Even though there is no specific debate to the literature from developing countries, the overall results show that the estimates between developed and developing countries are different. The results from developing countries show that the OLS estimates are higher than those of IV estimation, while the estimates from developed countries tend to produce OLS estimates that are lower than those of IV. It implies that the ability bias is dominant in the OLS estimation of developing countries.

This study make three main contributions to the existing debates in the literature in terms of methodology and also substantive aspect in the context of Thailand and, by implication, developing countries in general. First, the previous literature reveals that there is a different pattern of the relative magnitudes between OLS and IV estimates using compulsory schooling as IV between developed and developing countries. Investigating this contrast can contribute to a better understanding of (a) how and when the conventional “ability bias” matters in estimating returns to schooling and of (b) the impact of compulsory schooling in different settings. Second, as Thailand experiences rapid economic development and structural transformation during 1960-1990, obtaining the rate of returns to education in this period helps us better understand the process of Thai economic development as well as the interplay between the rate of returns to schooling and the economic development process during 1980 to 1990. This also helps understand the role of human capital in the process of development in other developing countries. Finally, the third contribution is on the construction of the database and discussion of the descriptive analysis for the discrepancies among different demographic characteristics, including gender, cohort, area of residence, region of residence, and economic sector. In addition to the overall estimates of returns to schooling, another important issue is an issue of heterogeneity in educational returns across individuals. Heterogeneity in individuals' demographic characteristics tends to distort the returns to education; for example, the female rates of return to education is likely to be higher than those of male. Hence, it is worth examining heterogeneous returns to schooling from different demographic characteristics.

CHAPTER 4. THEORETICAL FRAMEWORK

Section 1. INTRODUCTION

This chapter briefly presents a general theoretical framework of the analysis of returns to education. It mainly focuses on the Mincer's (1974) human capital earnings function. Willis et al. (1986) quote Griliches' (1977) questions regarding the Mincerian earnings equation in their article as followings:

> "Why it (Mincer model) should work so well …Why should there be a relation like this in the first place? In other words: (a) what interpretation can be given to such an equation? (b) What interpretation can be given to the estimated [schooling] coefficient? (c) Can one expect it to be "stable" across different samples and different time periods?"

This chapter explores Griliches' questions by beginning with a derivation of the Mincerian earnings function from the human capital theory. Secondly, it discusses the implicit assumptions behind the Mincer model and their implications on estimated coefficient of years of schooling.

Due to its simplicity yet strong explaining power, the Mincer's human capital earnings function is utilised widely as a starting point to examine the effects of investment in education, and on-the-job training, especially in the field of economics of education, labour economics, and development economics.

The common Mincerian earnings equation (Mincer 1958, 1974) is:

(4-1)[12]

$$\log y = a + bS + cX + dX^2 + e$$

where y represents incomes/earnings, b can be interpreted as an educational premium for wages and/or an internal rate of return to schooling under a set of certain assumptions that are discussed later in this chapter, S represents years of education an individual attended, X represents number of years an individual has worked after completing his/her education, and e represents the disturbance term.

[12] Although it is commonly written as log y, it is a natural log of y. Thus, it is more accurate to write as ln y instead of log y. However, this study follows the convention and use log y and ln y interchangeably throughout all chapters.

4.2.1. Simplified Derivation

The Mincer model is originally derived from the human capital theory, pioneered by Becker (1964, 1975), Becker and Chiswick (1966), and Mincer (1958, 1962, 1974). In the human capital theory (Becker, 1964) individuals make a decision to trade-off between current and future consumption. Matt (2013) spells out the implicit assumptions underlying Becker (1964) model of human capital as followings: i) individuals choose a certain level of education to maximise the expected present value of the stream of future returns; ii) individuals do not work while being in school; and individuals do not acquire any education while being in the labour market, meaning that time of schooling is independent of time of working; iii) there are only opportunity costs and no direct costs of education; and iv) the effect of experience on earnings is multiplicative. Thus, the model is represented by[13]:

(4-2) $$\sum_{t=1}^{T-s} \frac{w_s - w_{s-1}}{(1+r_s)^t} = w_{s-1} + c_s$$

where w stands for incomes, T represents an amount of time until retirement, c_s represents the costs of education, and r_s is the internal rate of return. If T is approaching the large number, the equation (4-2) can be approximated as:

(4-3) $$\frac{w_s - w_{s-1}}{r_s} = w_{s-1} + c_s$$

The equation (4-3) can be rearranged further if the costs of education, c_s, is sufficiently small.

(4-4) $$r_s \approx \frac{w_s - w_{s-1}}{w_{s-1}} \approx \log w_s - \log w_{s-1}$$

Harmon, Oosterbeek, and Walker (2003) posit that from equation (4-4) the empirical approximation of the human capital theoretical framework can be derived as the Mincer model, equation (4-1).

4.2.2. Derivation from Mincer's Frameworks

Heckman et al. (2006) elaborate the derivation of Mincerian earnings equation in more detail by utilising two different Mincer's frameworks, including the compensating difference model (Mincer, 1958), and the accounting-identity model (Mincer, 1974)

[13] More details of the derivation can be found in the survey article, "The Returns to Education: Microeconomics", written by Harmon, C., Oosterbeek, H., and Walker, I. (2003)

4.2.2.1 The Compensating Difference Model

In the compensating difference model, as certain jobs require different amount of schooling, individuals are compensated for opportunity costs for acquiring different levels of education, forgone earnings while they are in school. This fact explains the reason that everyone earns different amount of earnings over their lifetime. For example, in Thailand, medical students are required to study six years for the undergraduate level, while students from economics department complete their course within four years, meaning that medical students' forgone earnings are higher than those of students from economics school. Hence, in general, medical doctors would earn more since they get compensated for being in the school longer.

The main assumption under the compensating difference model is that individuals are ex ante identical, that implies i) individuals have identical abilities and opportunities; ii) credit markets are perfect and there is no credit constraint; and iii) the environment is perfectly risk free. The compensating differential is determined by equating the expected present value of the stream of future returns, assumed that there are no direct costs while individuals are attending school.[14]

(4-5) $$\mathbf{V}(\boldsymbol{S}) = \boldsymbol{Y}(\boldsymbol{D}) \int_{\boldsymbol{S}}^{\boldsymbol{T}} \boldsymbol{e}^{-\boldsymbol{rt}} \boldsymbol{dt} = \frac{\boldsymbol{Y}(\boldsymbol{S})}{\boldsymbol{r}} (\boldsymbol{e}^{-\boldsymbol{rS}} - \boldsymbol{e}^{-\boldsymbol{rT}})$$

where Y(S) represents the annual earnings of an individual with S years of education, r stands for an interest rate, and T is an amount of time until retirement. The equation (4-5) can be rearranged further by equating earnings streams across schooling and taking logs:

(4-6) $$\ln \boldsymbol{Y}(\boldsymbol{S}) = \ln \boldsymbol{Y}(\mathbf{0}) + \boldsymbol{rS} + \ln((\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{rT}})/(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{r}(\boldsymbol{T}-\boldsymbol{S})}))$$

Similar to the derivation illustrated by Harmon et al. (2003), if T is approaching the large number, the equation (4-6) can be approximated as:

(4-1) $$\ln \boldsymbol{y} = \boldsymbol{a} + \boldsymbol{bS} + \boldsymbol{cX} + \boldsymbol{dX}^{\mathbf{2}} + \boldsymbol{e}$$

Again, this ends up with the common Mincer model, equation (4-1). In this model, it is possible to illustrate that the (internal) rate of return to schooling and the interest rate are equal, which is, in fact, similar to the equation (4-4):

(4-7) $$\ln \boldsymbol{Y}(\boldsymbol{S}) - \ln \boldsymbol{Y}(\mathbf{0}) = \boldsymbol{rS}$$

$$\boldsymbol{b_S} = \boldsymbol{r}$$

[14] More details of the derivation can be found in the handbook chapter, "Chapter 7 Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond", written by Heckman et al. (2006).

4.2.2.2 The Accounting-Identity Model

In the accounting-identity model, it assumes that each individual has his/her own unique identity (ex ante heterogeneity), which implies that the rates of return to education vary in the population to reflect differences in returns. Card (1999) also emphasises that, "the return to education is not a single parameter in the population, but rather a random variable that may vary with other characteristics of individuals, such as family background, ability, or level of schooling". This model is closer to the common Mincerian earnings function as it incorporates the effects of investment in education, and on-the-job training into the model. The incorporation of these two components is motivated by the dynamic human capital investment model of Ben-Porath (1967).

Directly quoted from Heckman et al. (2006), "Let P_t be potential earnings at age t, and express costs of investments in training C_t as a fraction k_t of potential earnings, $C_t = k_t\, P_t$. Let b_t be the average return to training investments made at age t. Potential earnings at t are":

(4-8) $$\boldsymbol{P_t \equiv P_{t-1}(1 + k_{t-1}b_{t-1}) \equiv \prod_{j=0}^{t-1}(1 + b_j k_j)P_0}$$

From the equation (4-8), Heckman et al. (2006) show that the final product of the accounting-identity model is the common Mincer's human capital earnings function, equation (4-1). Through approximation, the equation (4-8) is rearranged into:

Potential earnings: $$\boldsymbol{\ln P_{X+S} \approx \ln P_0 + b_S S + \left(b_0 k + \frac{b_0 k}{2T}\right)X - \frac{b_0 k}{2T}X^2} \quad (4\text{-}9)$$

Replacing the potential earnings into the observed earnings equation:

Observed earnings: $$\boldsymbol{\ln Y(S, X) \approx \ln P_{X+S} - k\left(1 - \frac{X}{T}\right)}$$

(4-10) $$\boldsymbol{\ln Y(S, X) \approx [\ln P_0 - k] + b_S S + \left(b_0 k + \frac{b_0 k}{2T} + \frac{k}{T}\right)X - \frac{b_0 k}{2T}X^2}$$

Mincerian equation: $$\boldsymbol{\ln y = a + bS + cX + dX^2 + e}$$

Comparing the observed earnings equation and Mincerian equation, the intercept, a, represents the first component of the equation (4-10), $[\ln P_0 - k]$, b or b_S is the rate of return to schooling, $\left(b_0 k + \frac{b_0 k}{2T} + \frac{k}{T}\right)$ and $\frac{b_0 k}{2T}$ represent the return to on-the-job training, which are the c and d coefficients in the Mincer model.

Heckman et al. (2006) argue that b_S is an educational premium for wages but not the internal rate of return to schooling. Hence, it cannot use to justify whether we should invest in education or how much we should invest in education. In most applications of the Mincer model, it is assumed that an educational premium for wages and an internal rate of return to schooling are the same across persons and across school years. This point is discussed later in the chapter. **Table 4-1** summarises the differences between the compensating differences model and the accounting-identity model.

Table 4-1 The Differences between the Compensating Differences Model and the Accounting-Identity Model

The compensating differences model (1958)	The accounting-identity model (1974)
Main assumption: Individuals are *ex ante* identical	Main assumption: Individuals are *ex ante* heterogeneous
ρ_s is the same for all individuals	ρ_s varies in the population to reflect heterogeneity in returns
$\rho_s = r = IRR$	$\rho_s \neq IRR$

Source: Author's compilation based on the analysis of Heckman et al. (2006).

Section 3. Assumptions and Implications on Estimated Coefficient of Years of Schooling

Based on a certain set of assumptions, the Mincer model, **equation (4-1)**, looks clean and simple. However, its underlying set of assumptions is often less explicit. A set of assumptions generally implies limitations of a particular model and interpretations of its results such as estimated coefficients of interest. To better understand the Mincer model, it is worth examining its underlying assumptions. Card (1999) describes the set of Mincerian earnings function assumptions as followings: i) the numbers of years of schooling is the only correct measurement of educational attainment; ii) the returns to schooling is constant with respects to years of education, meaning that whether an individual moves up from grade 4 to grade 5 or from grade 5 to grade 6, each additional year of schooling affects earnings indifferently and are perfectly substitutable for each other (Meghir and Rivkin, 2011); iii) time of formal schooling is independent of time of working and their effects are separable; iv) there are only opportunity costs and no direct costs of education; and v) the rate of returns to schooling is constant across individuals (homogeneity assumption). Assumption iii) and iv) are directly derived from the human capital theory (Becker, 1964).

Under the assumption i) and ii), the coefficient b in **equation (4-1)** can be interpreted as an educational premium for wages but not as an *internal* rate of return to schooling. It reflects the effect of schooling in the labour market. The coefficient b is often addressed by the term, "the rate of return to schooling/education", and this may confuse with the term, "the internal rate of return", which is calculated to make a judgment on the investment project. However, after satisfying the additional assumptions iii) and iv), the coefficient b is referred to the internal rate of return (Willis, 1986). The mathematical illustration is out of scope of this study; however, Heckman et al. (2006) provide a thorough analysis on this matter in the Handbook of the Economics of Education.

Regarding the coefficient of years of schooling as a wage premium from education, the Mincer model is still well and alive in estimating the relationship between wage and schooling conditioned by other factors. It is relatively easier to satisfy assumptions i) years of schooling as educational attainment and ii) constant returns. If the research interest focuses on the *average* returns to *years of schooling*, rather than returns from other measurement of education (e.g. cognitive skills), or exact returns to each year of education, assumption i) and ii) are already satisfied. Thus, under these two assumptions, the estimated coefficient of years of schooling is considered as growth of earnings with respect to schooling or a price of schooling from a market wage equation (Heckman, et al., 2006).

However, the main criticisms of Mincer model's assumptions are often on the assumptions iv) cost of education and v) homogeneity of returns across individuals as some studies are interested in estimating the *internal* rate of returns to education. Estimating the internal rates of return for any investment projects have to take into account all costs and incomes involved in the investment project. Nevertheless, the Mincer model explicitly excludes the cost component and assumes that there are only opportunity costs and no direct costs of education. Hence, the estimated coefficient of education attainment cannot be interpreted as the internal rate of return, which is required to consider a complete accounting of the costs and benefits (Heckman, et al. 2006; Meghir and Rivkin 2011). Instead, the estimated coefficient of education attainment should be interpreted as growth of market earnings in terms of schooling. Moreover, with assumption v), the model ignores the heterogeneity in discount rates and uncertainty, which may affect decision making towards schooling. Violation of this assumption possibly results in biased estimation of returns to schooling.

Section 4. CONCLUSION

Even though Mincer model seems unrealistic and some assumptions can be challenged, the estimation based on Mincer model still shows a positive effect of schooling on earnings (Harmon, et al., 2003) and well fit the actual age-wages profiles (Card, 1999). Chapter 3 Section 1 shows that Mincer model with the OLS estimation can fit the actual recent data fairly well in both developed and developing countries. On the other hand, Heckman, et al. (2006) argues that Mincer model can only be used as a good starting point to develop a more complicated analytical framework. They criticise Mincer model on three main bases. First, there has been no decisive estimated rate of returns to schooling. Second, the estimated coefficient of years of schooling is ambiguous and difficult to make a meaningful economic interpretation. Finally, even a theoretically credible method of estimation such as IV estimation with RDD may possibly produce biased estimates as its instruments are generally found to be weak in many cases. A more comprehensive theoretical and empirical discussion can be found in the work of Heckman, Lochner and Todd (2003).

Lastly, this study shall explicitly limit the definition of estimated coefficient of education attainment to a wage premium from education, but still follows the convention by using the term “the rates of return to schooling” to refer to growth of market earnings in terms of schooling.

CHAPTER 5. DATA AND METHODOLOGY

This chapter provides comprehensive discussion regarding the data, methodology, and identification strategy employed in this study. In general, this study exploits an opportunity of quasi-experiment in Thailand which occurs from a change in compulsory education law to estimate the rates of return to schooling. The main estimation method is the IV estimation and the only source of data used in the estimation is the pooled cross-sectional Labour Force Survey (LFS) data from 1986 to 2012.

Section 1. DATA

The data set used in this analysis is drawn from the Thailand's Labour Force Survey (LFS) conducted by the National Statistical Office (NSO) for the years 1986 to 2012. The Thai Labour Force Surveys are collected quarterly[15] on about 80,000 random households for a total of about 200,000 observations per quarter, representing 0.1–0.5 per cent of the total Thai population. The LFS is the only national data set that comprehensively includes information both on demographic characteristics and labour related characteristics.[16]

The sample of estimation in this study is obtained by pooling the 27 consecutive annual LFSs. Only the data from the third quarter of the LFS is used in this study to control for the seasonal migration of agricultural labour. In general, agricultural workers move back and forth between urban manufacturing sector and rural agricultural sector. Nevertheless, they tend to migrate back to rural agricultural sector during the rainy season (Sussangkarn & Chalamwong, 1996), which is the third quarter of the year. Moreover, this study limits the sample to 1,307,988 wage workers aged 15-60 in the year of interview. The age restriction of 15-60 years is imposed because 15-years-old is the minimum legal age that individuals can start working and 60-years-old is usually the retirement age in Thailand. In addition to the age restriction, a birth cohort restriction is also employed in this study. The analysis is limited to individuals born between 1955 and 1985 since these cohorts are the observations around the cut-off for the instrumental variable. Individuals born in 1966 is the first cohort affected by the 1978 compulsory education law. This point is further discussed in the next section. The set of variables in this study covers age, birth cohort, years of schooling, region of residence, area of residence, industrial sector, and estimated monthly wages. A summary of descriptive statistics of these variables is given in **Table 5-1**.

[15] (i) January–March (dry or nonagricultural season), (ii) April–June (the period in which a large group of new workers enter the labor force after graduation), (iii) July–September (rainy and agricultural season), and (iv) October–December.

[16] See Appendix II for sample design.

Table 5-1 Descriptive Statistics

Variable	Description	Mean	Standard Deviation	Min	Max
Age	Age at the time of the survey	32.24	9.65	15	60
Year	Year of Labour Force Survey (1986 = 86, 1987 = 87, …, 2011 = 111, 2012 = 112)	100.62	6.98	86	112
Cohort	Birth cohort (Born in year 1955 = 55, …, born in year 1985 = 85)	68.38	8.28	55	85
Schooling	Number of years of schooling (No school = 0; Primary school grade 1 - 6 = {1, ..., 6}; Lower secondary school grade 7 - 9 = {7, …, 9}; Upper secondary school grade 10 -12 = {10, …, 12}; University 1st – 4th year = {13, …, 16}; Master degree 1st – 2nd year = {17, …, 18}; and Doctoral degree 1st-3rd year = {19, …, 21}	8.37	4.91	0	21
Bangkok	Living in Bangkok Metropolitan Area including Bangkok and the five adjacent provinces of Nakhon Pathom, Pathum Thani, Nonthaburi, Samut Prakan, and Samut Sakhon. (Bangkok = 1, otherwise = 0)	0.15	0.36	0	1
North	Living in North region (North = 1, otherwise = 0)	0.21	0.41	0	1
Northeast	Living in Northeast region (Northeast = 1, otherwise = 0)	0.20	0.40	0	1
South	Living in South region (South = 1, otherwise = 0)	0.21	0.41	0	1
Centre	Living in Central region (Centre = 1, otherwise = 0)	0.23	0.42	0	1
Agriculture	Working in agricultural sector including agriculture, hunting, forestry, fishing, mining, and quarrying. (Agriculture = 1, otherwise = 0)	0.25	0.43	0	1
Manufacture	Working in manufacturing sector including electricity, gas and water supply, construction, and general manufacturing, e.g. food products and beverages, tobacco products, textiles, to name a few. (Manufacture = 1, otherwise = 0)	0.20	0.40	0	1
Service	Working in service sector including wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods; hotels and restaurants; transport, storage and communication; financial intermediation; real estate, renting and business activities; public administration and defence; compulsory social security; education; health and social work; other community, social and personal service activities; and private households with employed persons. (Service = 1, otherwise = 0)	0.30	0.46	0	1
Urban	Living in urban area (Urban = 1; Rural = 0)	0.58	0.49	0	1
Male	Male (Male = 1; Female = 0)	0.48	0.50	0	1
Log wage	Log monthly wage adjusted for inflation using regional headline CPI.	8.80	0.78	3.98	13.88
Compulsory education	Experienced 1978 compulsory education law (Experienced = 1; Not experienced = 0)	0.52	0.50	0	1

Source: Author's compilation based on LFS, 1986-2012.

The dependent variable in the estimation of returns to schooling is "log monthly wage". Monthly wages are calculated from different types of wages reported by each observation. The calculation is done by the National Statistical Office. As this study pools multiple years of data together, the data in nominal value such as monthly wage requires adjustment for inflation. This nominal wage is deflated by the regional headline Consumer Price Index (CPI), which 2011 is used as the reference base year.[17] Finally, monthly wage adjusted for inflation is transformed into a log form.

[17] The data of regional headline CPI index is retrieved from Bureau of Trade and Economic Indices: http://www.price.moc.go.th/price/cpi/index_new_all.asp.

Furthermore, there are two school-related variables including years of schooling and compulsory education. A “years of schooling” variable is one of the main independent variables in the Mincer model. In the Labour Force Survey, the measure of school attainment is not the actual number of years spent at school but the highest degree attained by each individual. Hence, the school attainment variable is recoded into years of schooling ranging from zero (no education) to 21 years for those with PhD degree. On the other hand, compulsory education is a dummy variable indicating whether individuals experienced the 1978 compulsory education law (Experienced = 1; Not experienced = 0). This variable is utilised as the instrument variable in this analysis.

In this study, there are three time-related variables. First, “age” refers to the age at the time of the survey. “age” and its quadratic term are also the main independent variables in the Mincer model. Second, “year” refers to the year conducting the Labour Force Survey. This study pools multiple years of cross-sectional LFS data together and covers the data from 1986 to 2012. Lastly, “cohort” represents year in which individuals were born, birth cohort. Due to the cohort restriction, the oldest cohort is 1955 birth cohort, while the youngest one is the birth cohort born in 1985.

Other variables are included in the estimation model as control variables. First, the regional control variables, namely Bangkok, North, Northeast, South, and Centre are a set of dummy variables indicating the region in which each individual resides. For example, Bangkok is the dummy variable indicating whether individuals living in Bangkok Metropolitan Area, including Bangkok and the five adjacent provinces of Nakhon Pathom, Pathum Thani, Nonthaburi, Samut Prakan, and Samut Sakhon. Second, industrial control variables are also added into the estimation model. It is a set of dummy variables identifying the economic sector in which individuals are working. Agricultural sector includes agriculture, hunting, forestry, fishing, mining, and quarrying, while manufacturing sector covers electricity, gas and water supply, construction, and general manufacturing, e.g. food products and beverages, tobacco products, textiles, to name a few. Service sector involves wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods; hotels and restaurants; transport, storage and communication; financial intermediation; real estate, renting and business activities; public administration and defence; compulsory social security; education; health and social work; other community, social and personal service activities; and private households with employed persons. Furthermore, the variables of area of residence and gender also serve as the control variables in the estimation model. An “urban” variable refers to a dummy variable indicating whether individuals live in the urban area (Urban = 1; Rural = 0). Finally, “male” is a dummy variable indicating whether individuals are male (Male = 1; Female = 0).

Section 2. METHODOLOGY

5.2.1. What Are Impact Evaluation, Natural Experiment and Quasi-Experiment?

Impact evaluation is an evaluation seeking to find a causal effect of a particular programme e.g. policy on a specific outcome by controlling or excluding other factors that may also have an effect on that specific outcome, for example, the impact of merit awards on future academic outcomes (Thistlethwaite and Campbell, 1960)[18]. However, to estimate the causal effect of the programme, other factors must be controlled for, which is almost impossible, infeasible, and unethical (in some cases) to do so. For example, to understand the (causal) effect of education on wages, ideally we can estimate the effect by randomly selecting sample and dividing them into two groups. The first group receives a proper education, while another group is controlled not to have access to education. However, it is unethical to conduct such a programme because education is a fundamental right of each individual. To evaluate this type of programme, researchers tend to rely on a natural experiment or a quasi-experiment. A natural experiment is an empirical study in which individuals are divided into a treatment group and a control group by nature or by other factors outside the control of the researchers. Thus, it is arguable that a natural experiment resembles a randomized experiment[19]. It is, however, very rare to have such a phenomenon. Thus, researchers resort a new method which is similar to the randomized experiment, except that quasi-experiment lacks the element of randomization. The quasi-experiments are introduced into the research field as it is often the case that randomization is not possible or feasible. The quasi-experiments such as regression discontinuity design can help researcher exploit implementation features of the programme to measure its impact by solving problems of unobserved factors, and generating unbiased estimates of the causal effects of a treatment.

5.2.2. What Is Regression Discontinuity?[20]

According to Gertler, P.J., et al (2011), the regression discontinuity design is defined as, "an impact evaluation method that can be used for programs that have a continuous eligibility index with a clearly defined cut-off score to determine who is eligible and who is not". This method relies on knowledge of the selection process under two main conditions: first, the selection criteria, a continuous "score" or index[21], needs to be known and quantifiable; second,

[18] Thistlethwaite and Campbell (1960) are the very first researchers who introduced the method of regression discontinuity design into economic academia. The regression discontinuity design is also discussed in the main text.

[19] For more details see: "Impact Evaluation in Practice" (Gertler, P.J., et al, 2011)
This book provides a good introduction to the topic of impact evaluation and its practice in development.

[20] For basic references regarding regression discontinuity design see: Lee and Lemieux, 2009, 2010, and Imbens and Lemieux, 2008.

[21] There are few different terms to address the continuous indicator, namely an assignment variable, a forcing variable or a running variable.

assignment to "treatment" depends discontinuously on the "score" at a threshold or cut-off. For example, the scholarship programme is implemented with an expectation that students earning scholarship will do better than those who did not receive the scholarship. Given that the allocation of the scholarship based on an observed GPA, let say 3.00, students with GPA from and above 3.00 are eligible for the scholarship, while students with GPA below 3.00 are illegible. It is obvious that students with GPA 4.00 will do better academically than those with GPA 1.50. However, students with GPA just below the cut-off, GPA 3.00, (who did not receive the scholarship) are good comparisons to those just above the cut-off (who did receive the scholarship) since students with GPA 2.99 and students with GPA 3.00 should not be different. It is simply that the group just above and the group just below the cut-off point are very similar in nature, both observable and unobservable characteristics. The only difference between these two groups, treatment group (receiving scholarship) and comparison group (not receiving scholarship), therefore, should be due to the arbitrary cut-off. Hence, within narrow bandwidth around cut-off, treatment assignment approximates randomization or treatment effect. This method is equivalent to random assignment in a neighbourhood, which explains the reason that the regression discontinuity design is closer to randomised experiment than other quasi-experimental methods e.g. differences-in-differences (Raju and Trias, 2010)[22].

There are two types of regression discontinuity design, including sharp discontinuity and fuzzy discontinuity[23]. On the one hand, the main characteristic of sharp discontinuity is that discontinuity precisely determines treatment status. In other words, everyone in both treatment and control groups completely follows the policy rule (perfect compliance). On the other hand, in fuzzy discontinuity design, treatment assignment is probabilistic. It exploits discontinuities in the probability or expected value of treatment conditional on a covariate. Probability of treatment changes discontinuously at the cut-off, but not from zero to one, since not everyone strictly follows the policy rule, which is similar to a randomized experiment with imperfect treatment compliance. Some participants would always find a way to get treated, regardless of treatment assignment; other participants would never participate the treatment programme even if assigned. The cut-off becomes an instrumental variable for treatment status. This point will be elaborated in the discontinuity model identification section. In contrast, in the case of the sharp discontinuity, the probability of treatment changes from zero to one when the sample crosses the cut-off. Consequently, the gap in the relationship between outcome and the assignment variable cannot be interpreted as an average treatment effect. Nevertheless, Hahn et al. (2001) mathematically illustrates that the estimated treatment effect can be interpreted as a local average treatment effect when the assignment variable is

[22] This material is meant to be a supporting material for the book called, "Impact Evaluation in Practice" (Gertler, et al., 2010).
[23] See: Appendix III.

treated as an instrumental variable for treatment status.[24]

5.2.3. Discontinuity Model Identification: (Sharp) Regression Discontinuity

The identification for sharp regression discontinuity is as following:
Treatment status or (actual) treatment indicator, D_i, is a deterministic and discontinuous function of a covariate x_i. Once x_i is known, D_i is known.

(5-1)
$$\boldsymbol{D_i = 1 \leftrightarrow x_i \le x_0.}^{25}$$
$$\boldsymbol{D_i = 0 \leftrightarrow x_i > x_0.}$$

Where:

D_i = 1 if observation *i* receives treatment and
0 if observation *i* does not receive treatment.

x_0 = the cut-off point.

Potential outcomes, $\mathrm{E}[Y_{0i}|X_i]$, can be described by a linear, constant effects model:

(5-2)
$$\boldsymbol{\mathrm{E}[Y_{0i}|X_i] = \alpha + \beta X_i}$$
$$\boldsymbol{Y_{1i} = Y_{0i} + \rho}$$

This results in the regression:

(5-3)
$$\boldsymbol{Y_i = \alpha + \beta X_i + \rho D_i + \eta_i}$$

The key identifying assumption is that D_i is not only correlated with X_i but it is a deterministic function of X_i:

$$\mathrm{E}[Y_{0i}|X_i] \text{ and } \mathrm{E}[Y_{1i}|X_i] \text{ are continuous in } X_i \text{ at } X_0$$

This means that all other unobserved determinants of Y are continuously related to the running variable X, meaning that the only difference between these two groups, treatment group, Y_{1i}, and comparison group, Y_{0i}, is due to the arbitrary cutoff. Hence, a valid counterfactual is formed.

In case of nonlinear relationship, potential outcomes can be described as:

(5-4)
$$\boldsymbol{Y_i = \delta(x_i) + \rho D_i + \eta_i}.$$

---

[24] For more details on mathematical illustration, see: Hahn et al. (2001), Angrist and Pischke (2009), Lee and Lemieux (2009, 2010), and Moscoe, Bor, and Bärnighausen (2015)

[25] The assignment rule can be the other way around depending on the definition of the dummy variable, D_i. For example, $D_i = 0 \leftrightarrow score_i \le x$ and $D_i = 0 \leftrightarrow score_i > x$

Where:

$\delta(x_i)$ = function that is continuous around the cut-off point.

With IV estimation, the first stage is as following:

(5-5) $$\boldsymbol{x_i = \lambda_0 + \lambda_1 D_i + \nu_i}$$

The second stage is:

(5-6) $$\boldsymbol{y_i = \beta_0 + \beta_1 D_i + \delta(x_i) + \varepsilon_i}$$

It is common to approximate $\delta(x_i)$ by using a pth order polynomial:

(5-7) $$\boldsymbol{y_i = \beta_0 + \beta_1 D_i + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_p x_i^p + \varepsilon_i}$$

5.2.4. Advantages and Disadvantages of Regression Discontinuity

5.2.4.1 Advantages

The IV estimation with RDD has a strong internal validity with relatively weaker assumptions. Other methods of estimation such as a traditional IV estimation still possibly produce biased estimates since it is difficult to justify the internal validity, both in terms of relevance and exogeneity conditions. Regarding the assumption of relevance, the IV with RDD ensures that there is a causal relationship between the event (law change) and the endogenous variable (years of schooling); and between the event (law change) and the outcome (wage) through the effect of the endogenous variable (years of schooling). In terms of exogeneity condition, the regression discontinuity yields an unbiased estimate of treatment effect at the discontinuity. The IV with a quasi-experiment is exogenous to an individual as the RDD exploits a particular event, e.g. compulsory education law change, which is random to an individual. Hence the IV with RDD is not correlated with individual unobserved variables, e.g., ability and preferences. As previously discussed, the regression discontinuity exploits the fact that the groups just above and just below the cut-off (selection criteria to treatment) are very similar in nature, both observable and unobservable characteristics. The only difference between these two groups, treatment group and comparison group, should be due to the arbitrary rule or cut-off. The closer to the cut-off is, the more similar the treatment and control groups are. As a result, the comparison between those two groups is as good as if the treatment and comparison groups are chosen by randomized assignment to treatment. Therefore, it has less possibility to have an endogeneity bias and yields the strongest internal validity among the quasi-experimental methods.

5.2.4.2 Disadvantages

However, the IV utilising a quasi-experiment is not without a caveat. Even though the IV estimation with RDD satisfies both conditions of relevance and exogeneity for the internal validity, the causal relationship may not be generalised to other populations and settings. The issue of generalisation refers to the external validity. In fact, the IV estimation with RDD lacks the element of randomization since it mainly relies on the arbitrary cut-off. Hence, only within narrow bandwidth around the cut-off, treatment assignment approximates randomization or treatment effect. This method is equivalent to random assignment in a neighbourhood, which explains the reason that the regression discontinuity design is closer to randomised experiment than other quasi-experimental methods. Imbens and Angrist (1994) argue that RDD estimates the average treatment effect from those who comply with the treatment. This estimated average treatment effect is commonly referred to Local Average Treatment Effect (LATE). Even though it is difficult to estimate Average Treatment Effect (ATE) as a full compliance from the whole population may not happen, the LATE converges to the ATE as the number of compliers becomes an increasingly large proportion of the sample (Oreopoulos, 2006). Nevertheless, this limitation of RDD depends on the research question or evaluation question of interest. If the main question is to verify the effect of the treatment on the whole population, RDD may not be a perfect method of estimation but still produces satisfactory results. On the other hand, if the main question is to verify the treatment effect of a particular group of observations, RDD possibly best serves the purpose.

In addition, RDD implicitly assumes that there are no other rules or policy changes at the same cut-off. As those changes complicate identification strategy, it is not obvious whether the treatment effect comes from the rule identified by RDD or from other changes that happen to occur at the same time. Nonetheless, this issue is rather simple to identify by examining other possible rule or policy changes that may affect the dependent variable of interest and check whether they happen at the same or near the cut-off.

Section 3. A BRIEF HISTORY OF PRIMARY EDUCATION AND COMPULSORY EDUCATION ACT

According to **Table 5-2**, there were several changes to the compulsory education level in Thailand between 1921 and 2002. The first Primary Education Act was issued in 1921 and enforced the first compulsory education in Thailand. Students were required to complete six years of compulsory education. However, this Primary Education Act was applied to only certain areas due to the scarcity of educational resources, namely, schools, teachers, and educational budget. Thereafter, the 1935 Primary Education Act was enacted. After realising the difficulties in achieving a universal primary education, the Government decided to lower the primary education and also the compulsory education level to only four years as the

Government wanted more than half of the school-age population to complete compulsory education. During 1935-1978, there were few trivial changes in the Primary Education Acts, e.g. administration tasks expected to perform by education-related officers. In 1960, the Government launched National Education Reform plan to expand the education system as it realised that students who graduated from four years of primary education are still too young to work and should be educated more to understand cultural, social, and economic context of Thailand. Thus, in 1962, a primary education covered seven years, including four years of lower primary education and three years of upper primary education (**Figure 5-1**). The approximate age to enter the first year of primary education was seven years old. Age range of first year students was between six years old to eight years old. According to the 1962 Primary Education Act, every child of age eight years old had to be in school until age 15 years, unless they graduated from the 4th grade of primary education. Nonetheless, the compulsory education still covered only four years of primary education. It took around 40 years for the Government to change the compulsory education from four years to six years in 1978.

Table 5-2 Thai Education Act, 1921-2002

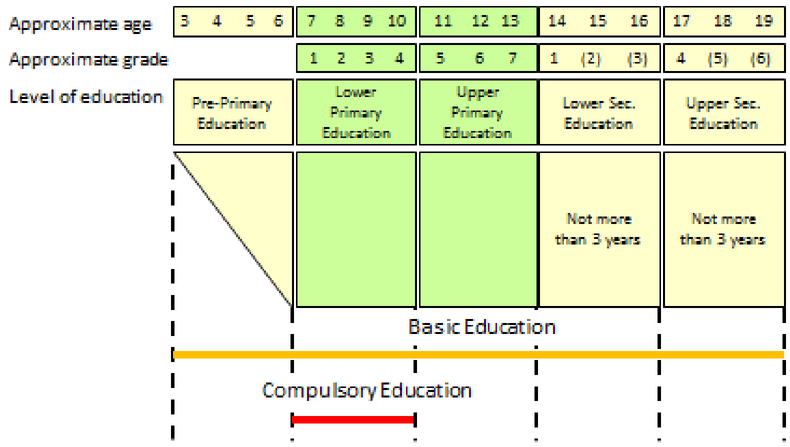
Act	Year	Compulsory School Entering Age	Minimum School Leaving Age	Years of Compulsory Education	Level of Compulsory Education
Primary Education Act	[1] 1921	7 (8-10)	14 [2]	6	Primary education
Primary Education Act	1935	8 (9-10)	15	4	Primary education [3]
Primary Education Act (2)	[4] 1940				
Primary Education Act (3)	1962	8	15	4	Lower primary education [5]
Primary Education Act (4)	[4] 1966				
Primary Education Act (5)	1978	8	15	6	[6] Primary education
Primary Education Act	1980	8	15	6	[7] Primary education
Compulsory Education Act	2002	7	16	9	Lower secondary education [8]

Note: 1 This Act is applied to only certain areas.
2 It is flexible and can be extended.
3 A primary education covers only 4 years in this period as the Government wants more than half of the school-age population to complete compulsory education.
4 There were changes only in the administration tasks.
5 A primary education covers 7 years, including 4 years of lower primary education and 3 years of higher primary education.
6 There is an exception in some areas where the highest level of primary education was only grade 4.
7 There is an exception in some areas where the highest level of primary education was only grade 5.
8 6 years of primary education and 3 years of lower secondary education.

Source: Author's compilation based on Primary Education Act (1921; 1935; 1940; 1962; 1966; 1978; 1980) and

Compulsory Education Act (2002).

Figure 5-1 National Education Reform 1960*



* 2 years before Primary Education Act 1962 was in effect

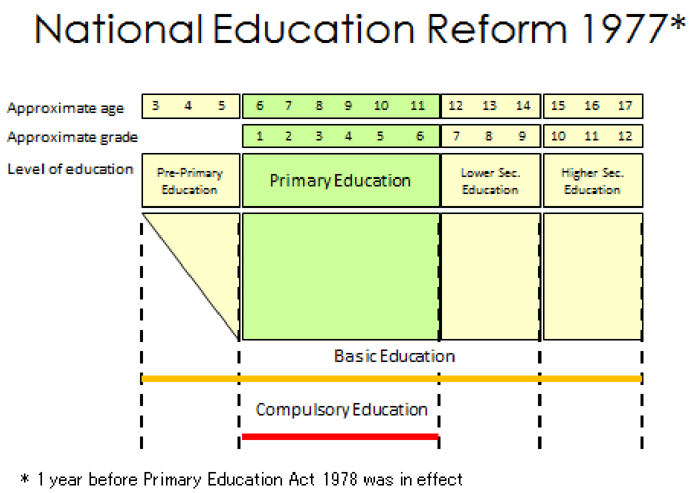
Source: Author's compilation based on information from Ministry of Education.

The 1978 Primary Education Act seemed to have a substantial impact on students' school attainment as it was the first education act with which every individual was forced to comply. As the Government foresaw a growing demand for educated workers in the labour market, it set out a National Education Reform plan to improve the national education system and also revised the level of compulsory education. The main motivation of the plan was to prepare future labour force for the on-going structural transformation from agriculture-based economy to industrialised economy. At that time, the whole educational structure was changed from 4-3-3-2 (**Figure 5-1**) to 6-3-3 (**Figure 5-2**). Previously, Thai education system consisted of four years of lower primary education, three years of upper primary education, and five years of secondary education. In contrast, the 1978 Primary Education Act reduced the total year of primary education to six years of primary education without division between lower and upper level and the secondary education remained the same. In addition, the Government expanded the compulsory education from four years to six years of primary education.

The first birth cohorts affected by the 1978 compulsory education change are cohorts born between 1966 and 1968. To identify the affected cohorts, the timeline between individuals' ages and school year is calculated (**Table 5-3**). As the law was issued and took effect immediately on the fifth of April 1978 which is a month before the 1978 academic year starts. Therefore, every student enrolled in grade 4 in 1977 were forced to move up to grade 5 in 1978. The possible age range of the fourth year students is between nine and elven years old that represent the 1966-1968 cohorts.

During the process of compulsory education expansion process, the Government provided a five-years-adjustment period (1978-1982) to those districts that were not ready for the compulsory education reform. Due to the fact that nation-wide education reform required a number of human resources, financial resources, and time, a universal compulsory education could not be achieved within one year. The readiness of each district did not only limit to the financial issue but also involve technical issues such as classroom building and teacher reallocation. Before 1978, schools were divided into four years of lower primary school and three years of upper primary school. Only large-scale schools such as province-funded schools had comprehensive seven years of primary education so they could easily adjust to the 1978 compulsory education law. However, a majority of schools had either four years of lower primary education or three years of upper primary education. These schools found it difficult to comply with the new compulsory education law. They needed more times to build more classrooms or combined with another school. Moreover, even though the school facilities were ready, human resources such as teachers might not be enough. However, the Government declared that by 1982 every student and every school must comply to the 1978 Primary Education and Compulsory Education Act. This five-years-adjustment period plays an important role in the identification strategy of this study.

Figure 5-2 National Education Reform 1977*



Source: Author's compilation based on information from Ministry of Education.

Table 5-3 The Identification of the First Cohorts Affected by the 1978 compulsory education change

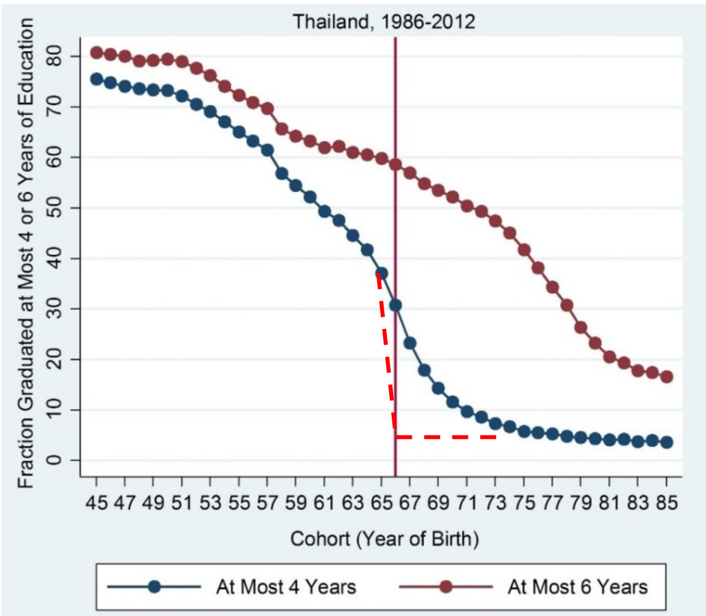
Cohort	Year													
	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978
School Grade										1	2	3	4	5
1969					0	1	2	3	4	5	6	7	8	9
1968				0	1	2	3	4	5	6	7	8	9	10
1967			0	1	2	3	4	5	6	7	8	9	10	11
1966		0	1	2	3	4	5	6	7	8	9	10	11	12
1965	0	1	2	3	4	5	6	7	8	9	10	11	12	13

Source: Author's compilation.

Section 4. IDENTIFICATION STRATEGY

This study exploits an opportunity of quasi-experiment in Thailand that occurs from a change in compulsory education law in 1978 to estimate the rates of return to schooling. In fact, there are a few compulsory education law changes in Thailand (**Table 5-2**). However, the change in 1978 is exploited due to its exceptional effect on education attainment of Thai population. The law change affects almost half of the population; therefore, the estimated LATE is possibly closer to the ATE than those of previous literature. As suggested by Oreopoulos (2006), "the magnitude of this impact provides a rare opportunity to estimate LATE of primary education that come close to population ATEs." Moreover, it is possible to link the finding of returns to schooling to the development process in Thailand. During the compulsory education law change period (1978), Thailand experienced rapid economic development and structural transformation. Estimating the rates of return to education in this period helps understanding Thai economic development as well as the effect of human capital in the process of development in general.

Figure 5-3 Fraction Graduating at Most 4 and 6 Years of Education, 1986-2012, Thailand



Note: The lower line shows the proportion of adults aged 15 to 60 from 1986 to 2012 Labour Force Surveys who report the highest attained level of education is at most 4 years. The upper line shows the proportion of adults aged 15 to 60 who report the highest attained level of education is at most 6 years. The vertical solid line indicates the first cohort affected by the 1978 compulsory education law. The dash line indicates that, in case there was no 5-year-adjustment period, it is expected to see a sharp drop of the fraction graduated at most 4 years of education from 40 per cent to less than 10 per cent. Hence, the vertical dash line represents the discontinuity.
Source: Author's compilation based on Labour Force Survey data 1986-2012.

The change to the compulsory education law in Thailand from four years to six years in 1978 has a powerful and immediate effect that redirected almost half the population of the fourth grade of primary education to stay in school for two more years (until grade 6 which is the final grade of primary education). **Figure 5-3** indicates the effect of the change of compulsory education law on the fraction of individuals graduating at most four years of education. There has been a gradual declining trend of individuals who decide to quit school at grade 4. It is not obvious to see a sharp drop of individuals graduating at most four years of education due to the 5-year-adjustment period of compulsory education expansion project. The lower line shows the proportion of adults aged 15 to 60 who report the highest attained level of education is at most four years, while the upper line shows the proportion of adults aged 15 to 60 who report the highest attained level of education is at most 6 years. The dash line indicates that, in case there was no 5-year-adjustment period, it is expected to see a sharp drop of the fraction graduated at most four years of education from 40 per cent to less than 10 per cent. Hence, the vertical dash line represents the discontinuity and a sharp drop after cohort 1965 can be seen.

In this study, the cut-off or the selection criteria is defined by the introduction of the 1978 compulsory education law and the 5-year-adjustment period of compulsory education expansion project. According to the 1978 compulsory education law, the compulsory education changes from four years to six years. The first cohorts affected by the compulsory education change are the cohorts born between 1966 and 1968. Thus, the cohorts born in 1966 and thereafter are in a treatment group, which is forced to be in school two more years until the end of the 6th grade of primary education. On the other hand, the cohorts born before 1966 are in a control group, that complies with the previous compulsory education law (four years of compulsory education). However, the Government provided a five-years-adjustment period (1978-1982) to those districts that were not ready for the compulsory education reform. Therefore, the treatment group during 1978-1982 includes only individuals from the districts that comply with the 1978 compulsory education law. As discussed above, without the 5-year-adjustment period of compulsory education expansion project, it is expected to see the discontinuity beginning from cohort 1966.

Section 5. ECONOMETRIC MODEL

This section spells out the econometric models employed in the main estimation. Fundamentally, the model is based on the Mincer wage equation, that is represented by the OLS regression model (5-10). The presentation of the models is organised in terms of the analytic order of IV estimation. First, the first stage least square regression is conducted to test the internal validity of the instrument in terms of relevance. The estimated result from first stage least square estimation gives a rough idea whether the instrument is partially and

sufficiently strongly correlated with independent variable of interest, ceteris paribus.
First stage least square regression:

(5-8)
$$\boldsymbol{S_i = \pi_0 + \pi_1 F_i + \pi_2 A_i + \pi_3 A_i^2 + \pi_4 C_i + \pi_5 R_i + \varepsilon_i}$$

where years of education an individual i attended (S_i) is a function of a dummy variable indicating whether an individual experiencing the 1978 compulsory education law (F_i), age of an individual as a proxy of working experience (A_i), birth cohort dummies (C_i), and regional dummies (R_i). ε_i represents a disturbance term.

Furthermore, the reduced form is estimated to see the effect of the instrument on the dependent variable of interest. The reduced form equation is expressed an endogenous variable in terms of exogenous variables.
Reduced form:

(5-9)
$$\boldsymbol{\log y_i = \alpha_0 + \alpha_1 F_i + \alpha_2 A_i + \alpha_3 A_i^2 + \alpha_4 C_i + \alpha_5 R_i + \theta_i}$$

where the log of monthly wages (y_i) of individual i is a function of a dummy variable indicating whether an individual experiencing the 1978 compulsory education law (F_i), age of an individual as a proxy of working experience (A_i), birth cohort dummies (C_i), and regional dummies (R_i). θ_i represents a disturbance term.

OLS regression model is the standard estimation exploited excessively in the literature of returns to education. As it is expected that the OLS estimates suffers from the omitted variable bias, the estimated results from OLS regression is used mainly to compare and contrast with those of IV estimation that possibly provides more reliable estimates.

OLS regression:

(5-10)
$$\boldsymbol{\log y_i = \beta_0 + \beta_1 S_i + \beta_2 A_i + \beta_3 A_i^2 + \beta_4 C_i + \beta_5 R_i + e_i}$$

where the log of monthly wages (y_i) of individual i is a function of years of education an individual attended (S_i), age of an individual as a proxy of working experience (A_i), birth cohort dummies (C_i), and regional dummies (R_i). e_i represents a disturbance term.

Theoretically, after estimating the first stage least square, the two stage least squares (2SLS) is performed to obtain the estimated coefficient of the instrument. However, in this study, 2SLS is not explicitly estimated.

Second stage least squares regression:

(5-11)
$$\boldsymbol{\log y_i = \gamma_0 + \gamma_1 \widehat{S_\iota} + \gamma_2 A_i + \gamma_3 A_i^2 + \gamma_4 C_i + \gamma_5 R_i + \vartheta_i}$$

where the log of monthly wages (y_i) of individual i is a function of the fitted value estimated from the first stage least square regression. ($\widehat{S_\iota}$), age of an individual as a proxy of working experience (A_i), birth cohort dummies (C_i), and regional dummies (R_i). ϑ_i represents a disturbance term.

CHAPTER 6. INSTRUMENTAL VARIABLE (IV) ESTIMATION USING THAI DATA

This chapter presents the empirical results from IV estimation by using Thai Labour Force Survey data. The results include the estimates from the first stage least square regression, the reduced form, the OLS regression (years of schooling on log wage), and the IV regression. The results are discussed by comparing and contrasting the estimated returns to schooling among different model specifications, among different estimation methods (mainly between OLS and IV estimations), and with the results from previous literatures. At the end of the chapter, a discussion of estimated results and endogeneity bias in the OLS regression is provided.

Section 1. EMPIRICAL RESULTS

6.1.1. First Stage

Table 6-1 shows the estimated results of compulsory education law on education attainment. Each regression includes controls for a birth cohort quadratic polynomial or birth cohort dummies, regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (six years of compulsory education). Column (3) to (5) also include age controls: a quartic polynomial and fixed effects where indicated. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort.

According to **Table 6-1**, the coefficients of compulsory education are statistically significant and robust across different specifications. The first-stage effect for all specifications is considerably powerful and it is possible to reject the hypothesis that the coefficients are zero. The compulsory education law change leads to six additional years of schooling which is roughly corresponding to the first year of the upper secondary school (grade 10). Even though the new compulsory education law increases the minimum year of schooling from four years to six years, it seems that the impact of the law is beyond its expected initial impact. Lochner and Moretti (2004), and Oreopoulos (2003) also find that the compulsory education law has an additional effect on education attainment beyond its minimum requirements for North America.

Table 6-2 indicates the estimated results of compulsory education law on education attainment from other studies. The coefficients of compulsory education are also statistically significant in other countries. However, the compulsory education variable (IV) is defined differently among the previous studies. The instruments can refer to minimum school-leaving age or specified minimum compulsory level of education. Moreover, the dependent variable from different studies can also refer to either the number of years of schooling or school-leaving age.

The first-stage regressions indicate that the effects of compulsory education on the years of schooling are quite small in most studies, for example, in the United States and Canada, raising the minimum school-leaving age by one year increases the total number of years of completed schooling by only about 0.11 to 0.13 years. This may imply that compulsory education does not affect most of the sample in those countries. Moreover, in UK and France, the effects of compulsory education on the age left full-time education are approximately 0.5 and 0.2 year, respectively. This does not exactly correspond to the compulsory education law that forces students to stay in schools one year longer.

Table 6-1 Estimated Effect of Compulsory Education Law on Education Attainment, Thailand, Ages 15-60, 1986-2012

	(1)	(2)	(3)	(4)	(5)
			(First Stage)		
	Dependent Variable: Number of Years of Schooling				
Compulsory Education	4.358***	6.385***	6.395***	6.413***	6.057***
	(0.713)	(0.087)	(0.105)	(0.105)	(0.084)
Fixed Effects:					
Regional Controls	No	Yes	Yes	Yes	Yes
Birth Cohort	Quartic	Yes	Yes	Yes	Yes
Additional Controls:	None	None	Age Quadratic	Age Quadratic	Age Quadratic
				Gender	Gender
					Urban
Observations	1,307,988	1,307,016	1,307,016	1,307,016	1,307,016
R-squared	0.091	0.123	0.137	0.138	0.192

Note: The dependent variables are number of years of schooling. Each regression includes controls for a birth cohort quartic polynomial or birth cohort dummies, regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (six years of compulsory education). Column (3) to (5) also include age controls: a quadratic polynomial and fixed effects where indicated. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. Robust standard errors in parentheses. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively.

Source: Author's compilation based on LFS 1986-2012.

Table 6-2 Estimated Effect of Compulsory Education Law on Education Attainment from Previous Studies

	First Stage		
	Dependent Variable: Number of Years of Schooling	**Instruments**	**Observations**
United States [a]	0.110 *** (0.0070)	Minimum school-leaving age	2,814,203
United States [b]	0.097 ※ (0.036)	Compulsory level of education (7 years)	3,680,223
Canada [a]	0.130 *** (0.0154)	Minimum school-leaving age	854,243
Spain [c]	0.820 (0.433)	Minimum school-leaving age	1,355
China [d]	0.790 *** (0.11)	Compulsory level of education (9 years)	11,271
China [e]	1.150 *** (0.241)	Compulsory level of education (9 years)	10,956
Turkey [f]	Women: 0.615 *** (0.082)	Compulsory level of education (8 years)	41,562
	Men: 0.955 *** (0.049)		91,756
	Dependent Variable: School-leaving age	**Instruments**	**Observations**
France [g]	Women: 0.272 *** (0.035)	Minimum school-leaving age	47,670
	Men: 0.264 *** (0.025)		50,927
United Kingdom[a]	0.489 *** (0.049)	Minimum school-leaving age	82,908
Britain [a]	0.436 *** (0.064)	Minimum school-leaving age	73,954

Note: (Robust) standard errors in parentheses. ***, **, and * indicate $p<0.01$, $p<0.05$, and $p<0.1$, respectively. ※ No reported statistically significant level.
Source: Author`s compilation based on the information from a Oreopoulos (2008); b Stephen and Yang (2014); c Pon and Gonzalo (2002); d Fang et al. (2012); e La Vincent and College (2014); f Aydemir and Murat (2015); g Grenet (2013).

The result from **Table 6-1** shows much larger effect of compulsory education law in Thailand than those of other countries. The plausible reason is the timing, which the compulsory education law was issued. For example, Thai 1978 compulsory education law was issued in the period before the fast-growing economic development, while there was no universal education at that time. The majority of school-age children were still out of school and the compulsory education law forces them to stay in schools. Thus, a large increase in completed years of schooling and also a sharp drop of sample with education at most four years can be observed. In contrast, in other countries, especially developed countries, e.g. the United Stated, Canada, to name a few, the compulsory education law was issued when those countries were in the stage of development that the majority of school-age children are already enrolled in schools. For instance, Psacharopoulos (2006) argues that the compulsory education law has an impact on only would-be dropouts in case of the United States. Hence, the compulsory education law affects only a certain group of school-age children.

6.1.2. Reduced Form

Table 6-3 shows the reduced form estimates of the effects of the compulsory education law change on monthly wages. The reduced form estimation use similar specifications as those of the first-stage regression, except for that the dependent variable is log monthly wage, instead of education attainment. The coefficients of compulsory education are statistically significant and robust across different specifications. The change in compulsory education level has a very large effect on the monthly wages. It yields approximately 36 per cent to 50 per cent increase in the monthly wages. The fit predicts that average wages increase for the cohorts that come after the change of the law. On the other hand, reduced form effects of compulsory education law on wages and earnings from previous studies are significantly smaller than those of Thailand (**Table 6-4**). Only the study of Harmon and Walker (1995) shows a similar effect of the compulsory education on earnings[26]. However, the coefficients of compulsory education are statistically significant in all previous studies.

Table 6-3 Estimated Effect of Compulsory Education Law on Log Monthly Wages, Thailand, Ages 15-60, 1986-2012

	(1)	(2)	(3)	(4)	(5)
			(Reduced Form)		
			Dependent Variable: Log Monthly Wages		
Compulsory Education	0.355***	0.523***	0.547***	0.537***	0.476***
	(0.071)	(0.033)	(0.029)	(0.030)	(0.026)
Fixed Effects:					
Regional Controls	No	Yes	Yes	Yes	Yes
Birth Cohort	Quartic	Yes	Yes	Yes	Yes
Additional Controls:	None	None	Age Quadratic	Age Quadratic Gender	Age Quadratic Gender Urban
Observations	1,307,988	1,307,016	1,307,016	1,307,016	1,307,016
R-squared	0.017	0.088	0.125	0.133	0.199

Note: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort quartic polynomial or birth cohort dummies, regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Column (3) to (5) also include age controls: a quadratic polynomial and fixed effects where indicated. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. Robust standard errors in parentheses. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively.
Source: Author's compilation based on LFS 1986-2012.

[26] The results from Table 3and Table 4 may not be directly comparable. The comparison of those results should be made with caution as the measurement and the definition of dependent variables are different.

The large estimated effect of compulsory education in Thailand can be explained by the fact that there is a wide income gap between individuals with different levels of education attainment. According to the result from the first-stage estimation, the compulsory education induces students to study, on average, additional six years, corresponding to grade 10. Hence, the cohorts affected by the compulsory education law tend to have much higher level of education. For individuals who have less than six years of education (compulsory education level), an average monthly wage is approximately 5,900 baht, while an average monthly wage of those with secondary education is about 10,000 baht (Labour Force Survey, 1986-2012). The wage of secondary education graduates is at least 40 per cent higher than that of individuals with less than primary education. This is consistent with the estimated results of **Table 6-4**, which indicates that the 36-50 per cent increase in the monthly wages is due to the compulsory education.

Table 6-4 Reduced Form Effects of Compulsory Education Law on Wages and Earnings from Previous Studies

	Reduced Form	
	Dependent Variable: Log Weekly Wage	Observations
United States [a]	0.016 ***	
	(0.0015)	2,814,203
	Dependent Variable: Log Annual Wage	Observations
Canada [a]	0.012 ***	854,243
	(0.0037)	
United Kingdom [a]	0.053 ***	82,908
	(0.017)	
Britain [a]	0.047 **	73,954
	(0.018)	
	Dependent Variable: Log Earnings	Observations
United Kingdom [b]	15 years old: 0.541※	34,336
	(0.055)	
	16 years old: 0.110※	
	(0.076)	

Note: (Robust) standard errors in parentheses. ***, **, and * indicate p<0.01, p<0.05, and p<0.1, respectively.
※ No reported statistically significant level.
Source: Author`s compilation based on the information from a Oreopoulos (2008); b Harmon and Walker (1995).

6.1.3. OLS and IV Estimation

The OLS results are shown in **Table 6-5**. The dependent variable is log monthly wages. Each regression includes controls for a birth cohort, regional dummies, age controls, and a dummy for the compulsory education law. The OLS point estimates of the coefficients of the years of schooling are statistically significant and robust across different specifications. The rates of return to schooling from the OLS estimation are approximately 11 per cent and consistent with those of the previous study from Thailand (Warunsiri and McNown, 2010)[27]. It confirms that the data quality of both studies is fairly similar and a comparison can possibly be directly made between these two studies. **Table 6-7** shows OLS, IV-RD, and fixed effect panel estimates of the returns to schooling from previous studies. Even though the OLS estimates of Thailand are higher than those of developed countries, the magnitudes are still comparable. The OLS results among different developed countries are similar in terms of magnitude of the coefficients and statistical significance level. The estimates are approximately 8 to 10 per cent and statistically significant at 1 per cent. The OLS estimates from developing countries are lower than those of developed countries. However, the OLS estimate from China is around 9 per cent, which is consistent to overall estimates from other countries.

Table 6-5 OLS Returns to Schooling Estimates for Log Monthly Wages, 15-60 Years Old, 1986-2012

	(1)	(2)	(3)	(4)	(5)
			(OLS)		
	Dependent Variable: Log Monthly Wages				
Year of Schoolings	0.113***	0.111***	0.112***	0.112***	0.109***
	(0.00366)	(0.00378)	(0.00425)	(0.00412)	(0.00389)
Fixed Effects:					
Regional Controls	No	Yes	Yes	Yes	Yes
Birth Cohort	Quartic	Yes	Yes	Yes	Yes
Additional Controls:	None	None	Age Quadratic	Age Quadratic Gender	Age Quadratic Gender Urban
Observations	1,307,988	1,307,016	1,307,016	1,307,016	1,307,016
R-squared	0.528	0.567	0.602	0.614	0.621

Note: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort quartic polynomial or birth cohort dummies, regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Column (3) to (5) also include age controls: a quadratic polynomial and fixed effects where indicated. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. Robust standard errors in parentheses. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively.
Source: Author's compilation based on LFS 1986-2012.

The IV estimates of returns to schooling are shown in **Table 6-6**. All discussed controls in the OLS estimation are also included in the IV estimation. The coefficients of years of schooling are statistically significant and robust across different specifications. This confirms a causal relationship between education attainment and earnings. One additional year of schooling leads to approximately 8 per cent increase in monthly wages which is lower than those of Warunsiri and McNown (2010) by half (**Table 6-7**). Although the OLS estimates from **Table 6-5** and Warunsiri and McNown's study (2010) are similar, the IV estimates from both studies indicate significantly contradicting results. The IV estimates from this study is

[27] See Table 6-7 OLS, IV-RD, and FE panel estimates of the returns to schooling from previous studies

less than but somewhat consistent with those of Canada and the United Kingdom, which are approximately 10 per cent. The rates of return to schooling from Warunsiri and McNown's study (2010) are around 14 per cent, which is similar to those of the United States. These rates of return are the highest among all studies, while the lowest rates of return are from China, 2 per cent. The overall IV results from developing countries are roughly 2-5 per cent, which is lower than those of developed countries. Some of the estimates from developing countries are even not statistically significant.

Table 6-6 IV Returns to Schooling Estimates for Log Monthly Wages, 15-60 Years Old, 1986-2012

	(1)	(2)	(3)	(4)	(5)
			(IV)		
			Dependent Variable: Log Monthly Wages		
Years of Schooling	0.0813***	0.0819***	0.0856***	0.0837***	0.0787***
	(0.00548)	(0.00424)	(0.00354)	(0.00363)	(0.00352)
Fixed Effects:					
Regional Controls	No	Yes	Yes	Yes	Yes
Birth Cohort	Quartic	Yes	Yes	Yes	Yes
Additional Controls:	None	None	Age Quadratic	Age Quadratic Gender	Age Quadratic Gender Urban
Observations	1,307,988	1,307,016	1,307,016	1,307,016	1,307,016
R-squared	0.486	0.533	0.575	0.581	0.588

Note: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort quartic polynomial or birth cohort dummies, regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Column (3) to (5) also include age controls: a quadratic polynomial and fixed effects where indicated. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. Robust standard errors in parentheses. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively.
Source: Author's compilation based on LFS 1986-2012.

Comparing the OLS estimates with those of IV from **Table 6-5** and **Table 6-6**, the result shows that the OLS estimates are greater than those of IV around 3 per cent. This indicates the upward bias (overestimation) in the OLS regression, which is consistent with the argument of ability bias in returns to schooling. According to **Table 6-7**, the studies of Canada, China, and Turkey also indicate that the OLS estimates are larger than those of IV. The rates of return to schooling in those countries are also overestimated in the OLS estimation. However, the difference between the OLS and IV estimates are very subtle in Canada (0.099 - 0.096 = 0.003). On the contrary, the other previous studies show that the OLS results are lower than those of IV. It implies that the returns to schooling are underestimated in the OLS regression model. In the United States, the IV estimates are double than those of the OLS, while the differences between OLS and IV estimates are relatively moderate in the United Kingdom and Thailand (Warunsiri and McNown, 2010). Warunsiri and McNown (2010) conclude that the estimates from IV and fixed effect panel are similar in magnitude; therefore, the problem of endogeneity bias is fixed in their estimations. Different from other studies, they use the existence of a university or teacher training college within a province as the instrument in the IV estimation. Meghir and Rivkin (2011) argues that such IV adopted by Warunsiri and McNown (2010) may be correlated with individual ability due to the non-random nature of individuals and school allocation. This implies that the estimates of IV are possibly subject to the endogeneity bias.

Table 6-7 OLS, IV-RD, and FE Panel Estimates of the Returns to Schooling from Previous Studies

	OLS	IV-RD	FE Panel
	Dependent Variable: Log Weekly Wage		
United States [a]	0.078 ***	0.142 ***	na
	(0.0005)	(0.0119)	na
	Dependent Variable: Log Annual Wage		
Canada [a]	0.099 ***	0.096 ***	na
	(0.0007)	(0.0254)	na
United Kingdom [a]	0.085 ***	0.108 ***	na
	(0.002)	(0.0328)	na
Britain [a]	0.083 ***	0.101 ***	na
	(0.003)	(0.0421)	na
	Dependent Variable: Log Annual Earning		
China [b]	0.09 ***	0.02 ***	na
	(0.004)	(0.06)	na
China [c]	0.0842 ***	0.0585	na
	(0.00369)	(0.0684)	na
Turkey [d]	0.050***	0.005	na
Male	(0.002)	(0.014)	na
Female	0.063***	0.038**	na
	(0.003)	(0.015)	na
	Dependent Variable: Log Hourly Wage		
Thailand [e]	0.115 **	0.148 **	0.151 **
	(0.000250)	(0.0194)	(0.0100)

Notes: a Region, birth cohort, age quartic, survey year, and gender are controlled in the regressions. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. Robust standard errors in parentheses. ***, **, and * indicate $p<0.01$, $p<0.05$, and $p<0.1$, respectively.
e Age quadratic, and cohort dummies are included in the regressions. The instrument variable is a dummy variable identifying the provinces in which universities or teacher training colleges are located. Standard errors in parentheses. ** indicates that coefficients are significant at or below the 0.05 level. na – not available.
Source: Author's compilation based on the information from a Oreopoulos (2008), b Fang et al. (2012), c La Vincent and College (2014), d Aydemir and Murat (2015), and e Warunsiri and McNown (2010).

This main finding of this study is consistent with the theoretical discussion regarding the size of IV coefficients compared to that of OLS which IV estimates should be smaller than those of OLS estimation (Heckman, et al., 2006)[28]. Furthermore, the finding also supports Behrman's (1999) view towards the ability bias. Behrman (1990, 1999) argues that there is a substantial large ability bias, especially in the case of developing countries. A number of

[28] See: Chapter 2 for more details

studies controlling for unobserved ability, family background, and school quality find that the OLS estimates suffer from the ability bias which the impact of years of schooling is overestimated 40-100 per cent.

Section 2. DISCUSSION OF THE ESTIMATED RESULTS

6.2.1. Methodological Discussions

6.2.1.1 Dominant Source of Bias

In general, the OLS estimation has a possibility to be biased due to the problem of endogeneity. Endogeneity biases occur when an explanatory variable is correlated with the error term. There are three main sources of endogeneity, including simultaneity, measurement errors, and omitted variables. In the context of returns to schooling, the omitted variable is the most frequently debated source of endogeneity bias (Willis, 1986, Schultz, 1988, Card, 1999, Deere and Vesovic, 2006).

The result shows clearly that the effect of ability bias (upward bias) outweighs the effect of discount rate bias (downward bias) since the difference between the OLS estimates and the IV estimates indicates an upward bias in the OLS estimation (OLS estimates > IV estimates). Hence, a positive correlation between ability and schooling is stronger than that of a negative one (discount rate bias).

The measurement error bias is substantially low and possibly offset by the ability bias in this analysis. It is found that the random (or classical) measurement errors tend to bias towards zero (Griliches, 1977; Angrist and Krueger, 1991). Thus, if the measurement error is significantly dominant, the downward bias should have been observed in the estimated results (IV estimates > OLS estimates). The plausible explanation is that as years of schooling and other wage determinants are measured relatively well compared with other micro-level variables, it is possible that the measurement error bias is of modest magnitudes (Schultz, 1988; Card, 1999; Krueger and Lindahl, 2001; Orazem and King, 2008).

Regarding the simultaneity bias, as comprehensively discussed in Chapter 3, the bias is not relevant in the study of returns to education as individuals' current wages received from working in the labour market cannot determine individuals' past levels of education. It is possible that individuals may enter the labour market after finishing compulsory education and then come back to study later. Nevertheless, the practice is not common, especially during the early period of economic development.

In conclusion, the results from OLS estimation indicate that a net effect of different sources of bias leads to the overestimated rates of return to schooling in the context of Thailand (OLS estimates > IV estimates). This implies that the ability bias outweighs other sources of bias, including the discount rate bias and the measurement error bias.

6.2.1.2 Methodological Implications on the Issues of Estimation

The finding rules out certain issues that may occur in the case of estimates with downward bias (IV estimates > OLS estimates). First, the upward ability bias implies that the IV estimation does not suffer from further unobserved heterogeneity biases. Bound and Jaeger (1996) argue that, in the case of downward bias, the IV estimation may produce further upward biased results due to the unobserved heterogeneous characteristics between the control and the treatment groups, e.g. family backgrounds, and schooling quality. The second implication is that the measurement error bias is substantially low in this analysis. As discussed earlier, if the measurement error is significantly dominant, the downward bias should have been observed in the estimated results (IV estimates > OLS estimates). Lastly, the upward ability bias rules out the argument that LATE estimated from the IV method is far from ATE. Card (1999, 2001) and Lang (1993) suggest that LATE estimates from the IV method tend to be higher than OLS estimates due to the fact that the instrument variable, especially compulsory schooling, affects only a particular group of population. It is often argued that the compulsory schooling may affect only socially and economically disadvantaged individuals who tend to have high returns to schooling. Nevertheless, this argument does not apply to the estimated results of this study as the OLS estimates are higher than those of IV. Thus, it may imply that the instrument variable in this study affects the majority of the population and the estimated results (LATE) are probably close to the ATE.

6.2.1.3 Remaining Puzzle

The contradiction between the IV estimates from this study and the estimates from panel with fixed effects (Warunsiri and McNown, 2010) is puzzling. In general, these two studies use similar data set with similar statistical quality as the OLS estimates from both studies indicate similar magnitude of the estimated coefficients and the level of significance. However, after correcting for the endogeneity bias with different estimation methods, the results are different not only in terms of magnitude of returns to schooling, but also the direction of the bias.

As discussed in Chapter 3, the panel regression with fixed effect may suffer from further upward bias since running panel regression with fixed effects excludes time-invariant independent variables from the regression, such as region controls and schooling quality. Stephen Jr. and Yang (2014) find that adding region and school quality control leads the effect of compulsory education on earnings towards zero or even wrong-signed. They argue that without the interaction between region control and year of birth, changes across regions over time of school quality cannot be captured in the estimation. The effects of differences in school quality may cause estimates to be biased since the source of upward bias is not limited to only an individual's ability, but also includes family background and quality of schooling (Schultz, 1988; Behrman, 1990; Strauss and Thomas, 1995).

6.2.2. The Upward Ability Bias in Thai Context

According to the result of estimation from previous section, it shows that OLS estimates are higher than those of IV estimation which implies the upward ability bias in the OLS estimates. The upward ability bias can also be observed in other studies from developing countries such as China and Turkey. On the other hand, studies from developed countries, e.g. US and UK, tend to produce OLS estimates that are lower than those of IV. It is possible that the stage of social and economic development may explain a direction of bias in the OLS estimation of returns to schooling.

In this study, the dominance of ability bias is mainly explained by the inequality of income and educational opportunity during the early period of social and economic development. At that time, Thailand was characterised as a developing and agriculture-based economy with relatively high poverty rate and income gap. A majority of population still depended heavily on agricultural sector and rural labour activities, which earn relatively lower than other sectors. It was common for family members to help working without formal pay in those sectors. Moreover, there was low school enrolment rate observed during the same period. Overall, it was common to observe the inequality, both in terms of income and access to basic public services such as basic education, between different socio-economic groups. Individuals from agricultural families are socially discriminated, poorer, and having greater immediate need to work. Their opportunity cost of schooling is also higher. It is likely that individuals from agricultural families may quit school at younger ages. Due to high costs of schooling and demand for household labour, agricultural households cannot afford to send all children to school. Thus, they may select only one or few children with the highest ability among all children to receive proper education, while the rest of the children are expected to help household agricultural activities. As a result, children with higher ability tend to have more schooling in agricultural households. This is consistent with the general ability bias hypothesis which argues that more-able individuals are more likely to have higher years of schooling. On the other hand, individuals from a higher socio-economic class are likely to have higher investments in education qualitatively (better school quality) and quantitatively (more years of schooling) (Behrman, 1999). This reflects the ability bias usually discussed in both studies of developed and developing countries. In sum, there are two sources of ability bias in case of Thailand which includes the ability bias from a selection of more-able child within poor households and the ability bias due to higher financial endowment of rich families.

The argument of the ability bias due to a selection of more-able child within poor households is more relevant and specific to the case of Thailand and, perhaps, other developing countries. In other words, this type of ability bias is less relevant in developed countries where poverty and inequality of educational opportunity, especially for the basic education, are less pronounced. Due to high economic development and standard of living, individuals in developed countries have more economic power to afford education for their children regardless of their level of ability. In addition, a universal access to free education is commonly observed, especially for basic education. Therefore, the equal education opportunity is more prevalent in developed countries. Even though there is a tendency that individuals with higher ability are more likely to attain more years of schooling, the ability bias seems to be less dominant in developed countries, at least for the basic level of education. This is possibly the main reason that the upward ability bias is observed in Thailand and other developing countries, but not in developed countries. Instead, OLS estimation from developed countries suffers from downward bias (IV estimates > OLS estimates). Hence, other sources of bias seem to dominate the effect of the upward ability bias in case of developed countries.

Explaining the upward ability bias in terms of inequality of income and educational opportunity during the early period of Thai social and economic development may possibly be generalised to other developing countries, which share similar social and economic context with Thailand. The main result shows that the magnitude of OLS estimates in Thailand is comparable with those of other developing countries. However, the IV estimates are fairly different. As different instruments from different studies define different parameter (Heckman et al., 2006), a similarity of magnitude of IV estimates should not be expected to observe from different estimations. In other words, it is usually the case that IV varies across different studies. Nonetheless, similar instruments and identification strategies correct the same endogeneity bias, given that those studies share similar sample characteristics. As a result, the difference between estimates from OLS and IV estimations should indicate a similar direction of bias whether OLS estimates are higher than those of IV.

Section 3. CONCLUSION

This chapter presents the empirical results from IV estimation by using Thai Labour Force Survey data. The results include the estimates from the first stage least square regression, the reduced form, the OLS regression (years of schooling on log wage), and the IV regression. First, the first stage estimation shows that the coefficients of compulsory education are statistically significant and robust across different specifications. The compulsory education law change leads to six additional years of schooling which is roughly corresponding to the first year of the upper secondary school (grade 10). Even though the new compulsory education law increases the minimum year of schooling from four years to six years, it seems that the impact of the law is beyond its expected initial impact. Second, the reduced form finds that the coefficients of compulsory education are statistically significant and robust across different specifications. The change in compulsory education level has a very large effect on the monthly wages. It yields approximately 36 per cent to 50 per cent increase in the monthly wages. The large estimated effect of compulsory education in Thailand can be explained by the fact that there is a wide income gap between individuals with different levels of education attainment. Third, the OLS point estimates of the coefficients of the years of schooling are statistically significant and robust across different specifications. The rates of return to schooling from the OLS estimation is approximately 11 per cent and consistent with that of the previous study from Thailand (Warunsiri and McNown, 2010). Finally, the IV estimation indicates that the coefficients of years of schooling are statistically significant and robust across different specifications. One additional year of schooling leads to approximately 8 per cent increase in monthly wages. Comparing the OLS estimates with those of IV, the result shows that the OLS estimates are greater than those of the IV around 3 per cent. This indicates the upward bias (overestimation) in the OLS regression, which is consistent with the argument of ability bias in returns to schooling.

This main finding of this study is consistent with the theoretical discussion showing that IV estimates should be smaller than those of OLS estimation (Heckman, et al., 2006). Furthermore, the finding also supports Behrman's (1999) arguments regarding the ability bias. Behrman (1990, 1999) argues that there is a substantial large ability bias, especially in the case of developing countries. A number of studies controlling for unobserved ability, family background, and school quality find that the OLS estimates suffer from the ability bias which the impact of years of schooling is overestimated 40-100 per cent.

There are three main concluding remarks from the discussion of the results in terms of methodological perspectives. First, the results from OLS estimation indicate that a net effect of different sources of bias leads to the overestimated rates of return to schooling in the context of Thailand (OLS estimates > IV estimates). This implies that the ability bias outweighs other sources of bias, including the discount rate bias and the measurement error bias. Second, the finding rules out certain issues that may occur in the case of estimates with downward bias (IV estimates > OLS estimates), including further unobserved heterogeneity biases, the classical measurement error bias, and the argument that LATE estimated from IV method is far from ATE. Finally, there still a remaining puzzle as the IV estimates from this study and the estimates from panel with fixed effects (Warunsiri and McNown, 2010) are different not only in terms of magnitude of returns to schooling, but also the direction of the bias, even though these two studies use similar data set with similar statistical quality.

The last section concludes that the dominance of ability bias is mainly explained by the inequality of income and educational opportunity during the early period of social and economic development and there are two sources of ability bias in case of Thailand that includes the ability bias from a selection of more-able child within poor households and the ability bias due to higher financial endowment of rich families. However, the argument of the ability bias due to a selection of more-able child within poor households is more relevant and specific to the case of Thailand and, perhaps, other developing countries. In other words, this type of ability bias is less relevant in developed countries where poverty and inequality of educational opportunity, especially for the basic education, are less pronounced. Even though there is a tendency that individuals with higher ability are more likely to attain more years of schooling, the ability bias seems to be less dominant in developed countries, at least for the basic level of education. This is possibly the main reason that the upward ability bias is observed in case of Thailand and other developing countries, while the OLS estimation from developed countries suffers from downward bias (IV estimates > OLS estimates). Hence, the other sources of bias seem to dominate the effect of the upward ability bias in case of developed countries. Furthermore, explaining the upward ability bias in terms of the inequality of income and educational opportunity during the early period of Thai social and economic development may also possibly be generalised to the case of other developing countries, which share similar social and economic context with Thailand.

CHAPTER 7. FURTHER ANALYSIS: DISAGGREGATED ANALYSIS OF RETURNS TO SCHOOLING

In addition to the overall estimates of returns to schooling, another important issue is an issue of heterogeneity in educational returns across individuals. In Chapter 2, several education trends in Thailand during the period of analysis (1950-1990) are portrayed. Coupled with international social mandates and financial aids, the structural transformations raise the demand for educated workers. Thus, there has been a rapid expansion of the education system in the primary education for past 20 years (1960-1980), both in terms of student enrolment and number of teachers. However, the enrolment trend indicates that there is inequality in educational opportunities among individuals from different socio-economic groups. For example, the female enrolment was significantly lower than those of men during the period of the structural transformation. Ideally, given same levels of education, ceteris paribus, same returns to education should be observed among individuals regardless of their characteristics. However, heterogeneity in individuals' demographic characteristics tends to distort the returns to education; for example, the female rates of return to education is likely to be higher than those of male. Hence, it is worth examining heterogeneous returns to schooling from different demographic characteristics.

This section presents disaggregated analyses of returns to schooling and their results by dividing the sample into five demographic domains, including gender, birth cohort, area of residence (urban and rural areas), region, and occupational industry. The analysis with these broad domains refers to a one-level disaggregated analysis of returns to schooling, in short a one-level disaggregation. Nevertheless, each individual possesses multiple characteristics and those combinations of characteristics give different levels of returns, for example, females living in a rural area may earn less than those from an urban area. Thus, to better understand the issue of heterogeneity in educational returns, interactions between different demographic characteristics are introduced into the disaggregated analysis. The analysis refers to a two-level disaggregated analysis of returns to schooling or a two-level disaggregation. The sample is disaggregated based on different combinations of two demographic characteristics. **Table 7-1** illustrates an analytical matrix indicating both one-level and two-level disaggregated analyses. "1.x" refers to one-level disaggregated analysis, whereas "2.x" refers to two-level disaggregated analysis. There are some overlapping aspects presented in the analytical matrix, e.g., Gender-Cohort and Cohort-Gender. "See" represents an overlapping section and indicates the section that has been previously discussed. In addition, "na" indicates the section that the results are not significantly different from the main results or they are consistent with the main results. The disaggregated analyses serve at least three main purposes. First, the analyses provide a better understanding of the issue of heterogeneity in returns to schooling. Second, as a by-product, the analyses serve as a robustness check among different sub-samples based on different demographic domains. In addition, further analysis on aggregation biases can be examined to explain the biases in the OLS estimates (Warunsiri and McNown, 2010).

The disaggregated analyses are not without caveats. Schultz (1988) argues that stratifications of samples based on heterogeneous demographic characteristics lead to selection bias in estimating returns to education, especially in case of developing countries. Developing countries share similar economic and social phenomena, namely prevalence of self-employment and non-wage labour, imbalanced economic development among different areas and regions, and gender-based segregation in occupational choices. These phenomena affect individuals' decision making towards investments in education and returns to education. Each demographic domain has its own specific issues discussed in the following sections. With comprehensive investigation of the process of economic development, the disaggregated analyses indicate the estimated results and suggest trends and directions of the returns to schooling across different demographic characteristics.

Table 7-1 Analytical Matrix

	Gender	**Cohort**	**Area**	**Region**	**Economic Sector**
Gender	**1.1 Gender**	na	See 2.1.1 Gender-Area	See 2.1.2 Gender-Region	See 2.1.3 Gender-Economic Sector
Cohort	na	**1.2 Cohort**	na	na	na
Area	*2.1.1 Gender-Area*	na	**1.3 Area**	*2.2.1 Region-Area*	na
Region	*2.1.2 Gender-Region*	na	na	**1.4 Region**	na
Economic Sector	*2.1.3 Gender-Economic Sector*	na	na	*2.2.2 Region-Economic Sector*	**1.5 Economic Sector**

Note: "1.x" refers to one-level disaggregated analysis.
"2.x" refers to two-level disaggregated analysis.
"See" represents an overlapping section.
"na" means not available due to the fact that the results are not significantly different from the main results or they are consistent with the main results.

Source: Author's compilation.

Section 1. ONE-LEVEL DISAGGREGATED ANALYSIS OF RETURNS TO SCHOOLING

A one-level disaggregation refers to an estimation with different subsamples stratified by certain demographic characteristics, namely gender, cohort, area of residence, region, and occupational industry. **Table 7-2** shows the estimated results of the one-level disaggregated analyses. The overall results are consistent with the results from the main analysis in which the upward bias in cross-sectional OLS regressions can be observed. It seems that the ability bias is still dominant in the disaggregated results. The bias gap refers to the difference between OLS estimate and the IV estimate. It is approximately 2 per cent across subsamples.

Table 7-2 Disaggregated Analysis of OLS and IV Returns to Schooling

Dependent Variables	OLS	IV	Bias Gap	Sample Size	Comparison: Returns to Schooling	Comparison: Bias Gap
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016		
all workers	(0.00425)	(0.00354)				
Log monthly wages,	0.108***	0.0808***	0.0272	662,753	Female > Male	Female > Male
male	(0.00439)	(0.00374)				
Log monthly wages,	0.116***	0.0856***	0.0304	644,263		
female	(0.00391)	(0.00367)				
Log monthly wages,	0.126***	0.0872***	0.0388	812,522	Old > Young	Old > Young
cohort 1955-1970	(0.00354)	(0.00347)				
Log monthly wages,	0.101***	0.0856***	0.0154	1,016,749		
cohort 1961-1985	(0.00377)	(0.00354)				
Log monthly wages,	0.108***	0.0849***	0.0231	856,617	Urban > Rural	Rural > Urban
urban	(0.00369)	(0.00169)				
Log monthly wages,	0.105***	0.0709***	0.0341	450,399		
rural	(0.0053)	(0.00622)				
Log monthly wages,	0.0956***	0.0751***	0.0205	162,399		
BKK	(0.00204)	(0.00116)				
Log monthly wages,	0.124***	0.0990***	0.025	256,236	North, Northeast > Others	Northeast, North > Others
North	(0.00439)	(0.00436)				
Log monthly wages,	0.140***	0.0977***	0.0423	298,137		
Northeast	(0.0038)	(0.00683)				
Log monthly wages,	0.0928***	0.0733***	0.0195	222,181		
South	(0.00473)	(0.00343)				
Log monthly wages,	0.0973***	0.0766***	0.0207	369,035		
Centre	(0.00459)	(0.00225)				
Log monthly wages,	0.101***	0.0590***	0.042	428,699	Service > Manufacture > Agriculture	Agriculture > Manufacture > Service
Agricultural sector	(0.00479)	(0.0044)				
Log monthly wages,	0.0936***	0.0706***	0.023	238,274		
Manufacturing sector	(0.00468)	(0.00285)				
Log monthly wages,	0.102***	0.0825***	0.0195	637,105		
Service sector	(0.00325)	(0.00078)				

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author's compilation based on LFS 1986-2012.

7.1.1. Gender

7.1.1.1 Rates of Return to Schooling

Table 7-3 Disaggregated Analysis of OLS and IV Returns to Schooling by Gender

Dependent Variables	OLS	IV	Bias Gap	Sample Size
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Log monthly wages,	0.108***	0.0808***	0.0272	662,753
male	(0.00439)	(0.00374)		
Log monthly wages,	0.116***	0.0856***	0.0304	644,263
female	(0.00391)	(0.00367)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author's compilation based on LFS 1986-2012.

According to **Table 7-3**, in terms of gender, the OLS coefficients on years of schooling for men and women are 0.108 and 0.116, respectively, while those of the IV estimation are 0.081 for men and 0.086 for women. With disaggregation by gender, the upward ability bias in the returns to education estimated from the OLS regression remains. This disaggregation shows the rates of return to schooling for women are marginally higher than those of men and they are somehow similar in magnitude. This is consistent with most of previous studies, e.g., Chiswick (1976), Warunsiri and Mcnown (2010), Parinduri (2014), Hawley (2004), Devereux and Fan (2011), and Grenet (2013) (**Table 7-4**), which the rates of return to education of women are higher than or comparable to those of men. It seems that the result shows a transition towards gender equality in access to education and labour market.

Figure 7-1 and **Figure 7-2** illustrate that women in Thailand earn on average substantially less than do men. However, it is common to observe the rates of return to schooling for women exceed those of men. The reason is simply that the concept of rates of return is rather relative (Psacharopoulos, 1985). Education helps women expand their job opportunities, both in terms of access to labour market and occupational choice (Psacharopoulos, 1988, 1992, Dougherty, 2005). As women have much more limited access to both education and labour market, investing in one additional year of schooling for women would allow them to earn much more than those with less education. On the other hand, it is quite common among men to attain more education and to have a better job position in the labour market; therefore, investing just one additional year of schooling would marginally increase their returns to schooling. Moreover, as the occupational choice for women expands, women move away from lower-paid traditionally female jobs to higher-paid ones. This also contributes to the fact that the female rates of return to schooling are greater than those of male.

Table 7-4 Summary of Returns to Education by Gender

Author	Method of Estimation	Women vs Men	Women	Men
1. Thailand				
Chriswick (1976)[a]	OLS	>	0.130	0.091
Amornthum and Chalamwong (2001)[b]	OLS	<	na	na
Hawley (2004)	OLS	≈	0.103***-0.108***	0.103***-0.111***
Warunsiri and Mcnown (2010)	Fixed effects with panel data	>	0.178**	0.126**
This study	IV with RDD	>	0.0856***	0.0808***
2. Developed Countries				
Devereux and Fan (2011), UK	IV with RDD: Hourly wages IV with RDD: Hourly earnings	< ≈	0.053*** 0.066***	0.062*** 0.066***
Grenet (2013), France and UK	IV with RDD (France) IV with RDD (UK)	… ≈	-0.007 0.067***	-0.004 0.069***
3. Developing Countries				
Fang, et al. (2012), China	IV with RDD	<	0.10*	0.51**
La Vincent and College (2014), China	IV with RDD	<	0.0367 (not significant)	0.0935 (not significant)
Parinduri (2014), Indonesia	IV with fuzzy RDD	>	0.26**	0.14*
Aydemir and Murat (2015), Turkey	IV with RDD	>	0.038**	0.005 (not significant)

Note: a Level of statistical significance is not reported.
b No information on the estimates. The direction of the inequality sign is taken from Warunsiri and Mcnown (2010).
***, **, and * indicate p<0.01, p<0.05, and p<0.1, respectively. na means "not available".

Source: Author's compilation based on LFS 1986-2012.

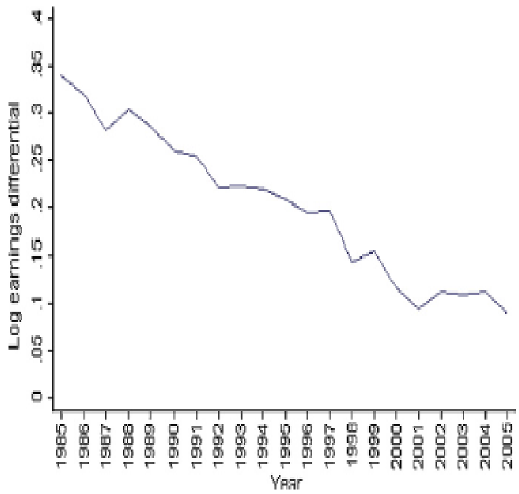
7.1.1.2 Ability Bias (Bias Gap)

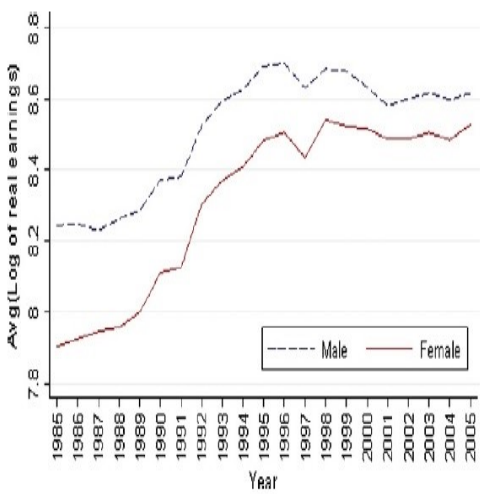
According to **Table 7-3**, the bias gap for women is around 3 per cent, while for men is approximately 2.7 per cent. The magnitude of the male bias gap is consistent with that of Warunsiri and McNown (2010). The bias gaps indicate that there is an upward ability bias in the OLS estimates in both genders. Even though the estimated returns to schooling are higher for women, the difference between female and male bias gaps is fairly small (0.0304 – 0.0272 = 0.0032). Hence, there is no obvious difference between female and male bias gaps. In contrast, the bias gaps for women and for men are 0.05 and 0.02, respectively in the study of Warunsiri and McNown (2010). The bias gap is greater for women than for men (3 per cent). Although the estimated results of the female returns are different, both studies indicate that the bias gap of women is larger than that of men.

Even though the magnitude of the difference between male and female bias gap is small, it seems that the ability bias is more dominant for women. In Thailand, the gender gap in educational attainment at all levels has closed in 1990 (**Figure 2-14**). This implies that during the economic transformation, 1970-1990, there has been a gender gap in all levels of education. エラー! 参照元が見つかりません。**Figure 2-14** shows a percentage of female students of total number of students enrolled in each education level. Even though the situation has been improving over time, gender gap in all levels of education can be observed. In the primary education female students are still less than half of the total number of enrolled students, whereas in the secondary education female represents approximately 40 per cent of the total number of enrolled students. The situation is even worse in vocational secondary schools and higher education. There was discrimination against women in education at all levels; therefore, women seem to face higher opportunity costs in attaining education. Gandhi-Kingdon (2002) explains this situation as "unexplained parental discrimination": given the same level of ability, women would get less support from their families. Compared to men, women are socially disadvantaged. In the early stage of social and economic development, market distortions such as gender discrimination prevent socially disadvantaged groups, e.g. women, to attend schools. Hence, only women with higher ability and better family background are given an opportunity to attend schools. As Behrman (1999) put it, "individuals with higher investments in schooling are likely to be individuals with more ability and more motivation who come from family and community backgrounds that provide more reinforcement for such investments and who have lower marginal private costs for such investments and lower discount rates for the returns from those investments and who are likely to have access to higher quality schools". This is consistent with the ability bias hypothesis which states that more-able individuals are more likely to attain more years of schooling. Hence, the ability bias is more dominant in the female OLS estimates than those of men.

However, the estimated result also implies that there is a trend towards equal educational opportunity between women and men as the differences of the estimated returns to schooling and the bias gaps between women and men is substantially small. As previously discussed, gender gaps in educational attainment at all levels have closed since 1990 (Knodel, 1997). Hence, it is highly possible that the movement towards gender equality starts even before 1990. The situation has also been improving in the secondary education where female is approximately 40 per cent of total number of enrolled students. In other words, women and men face similar opportunity costs in attending the primary and secondary education. Moreover, economic development during 1980s raises a demand for labour, which provides an opportunity for women to enter the labour market. Through modernisation, gender roles have changed overtime (Nakavachara, 2010). Women have more chance to enrol in the education system partly to meet the demand for high-skilled workers. These facts reflect in the results of disaggregated analysis by gender, both in terms of IV estimates and the bias gaps, which are marginally small between women and men.

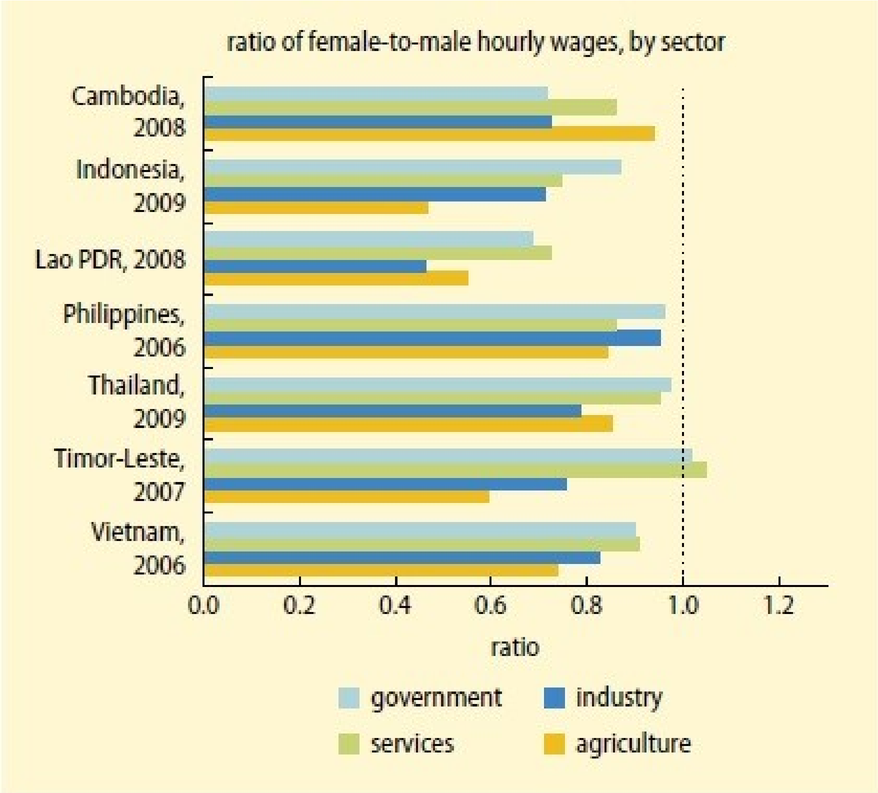
Figure 7-1 Gender Earnings Gap in Thailand (1985-2005)





Source: Nakavachara, 2010

Figure 7-2 Ratio of Female-to-Male Hourly Wages by Sector



Note: World Bank staff estimates using household income and expenditure surveys.
Source: World Bank, 2012.

7.1.1.3 Selection Bias

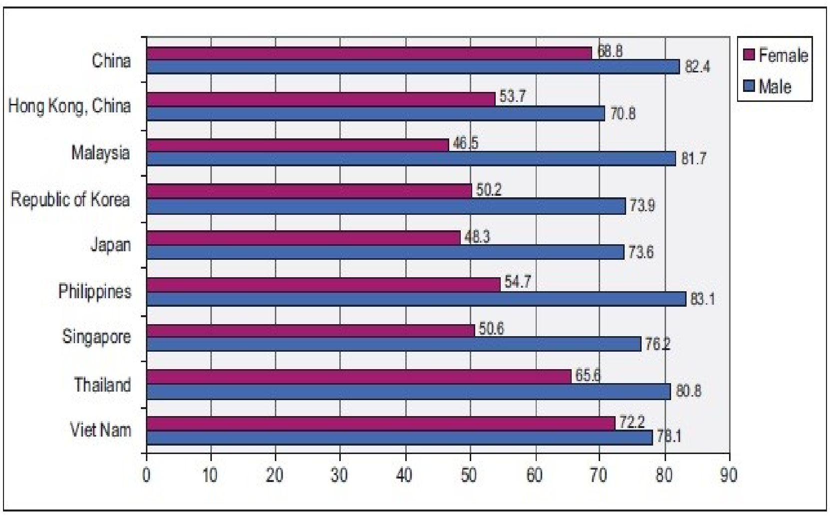
In fact, the "real" female rates of return can be either higher or lower than the estimated ones due to the high probability of sample selection bias. The conventional estimations of returns to schooling usually ignore the selected nature of the female samples. Both Psacharopoulos (1985) and Schultz (1988) argue that the estimated female returns to schooling can be either ooverestiamted or underestimated due to the fact that the estimates are from unrepresentative sample of females. Female labour participation rate is fairly low and females tend to concentrate in self-employed occupations or a non-wage labour market. According to **Table 7-5**, in Thailand, the female percentage in labour force is approximately 46 per cent, which accounts for less than half of the total number of labour force. In addition, a labour force participation rate, on average, indicates that 66 per cent of the total number of female population participates in the labour market, while male participation rate is approximately 81 per cent (**Figure 7-3**). Hence, the sample of estimation does not cover all females and deems to be unrepresentative.

Nevertheless, Schultz (1988) posits that women may choose to work fewer hours than do men regardless of female labour force participation rate. Hence, the problem of sample selection may be even more severe. However, Thai patterns of distribution of working hours for the self-employed by gender are unique among developing countries. Normally, the distribution of working house is relatively flat and likely to be evenly distributed, while in the case of Thailand hours spent working for both genders are heavily centred in the category of 40 hours per week and above (Lee, McCann, and Messenger, 2007). **Table 7-6** indicates that the number of individuals working more than 40 hours per week account for 79 per cent for women and 83 per cent for men. Thai women and men tend to spend time working for long hours; therefore, the argument of Schultz may not apply to the case of Thailand.

7.1.1.4 Conclusion

Within the wage labour market, the estimated returns to schooling for women are higher than those for men. It implies that women face higher opportunity costs in attaining education and participating in the labour market. This is possibly due to the gender discrimination practice during the early period of social and economic development. Moreover, the estimated result shows that women have a higher bias gap than that of men. This implies that the ability bias is more dominant in the female OLS estimates since gender discrimination causes the situation in which only more-able women can obtain higher level of education. Nevertheless, the result also indicates that there is a movement towards equality of access to education and also labour market between women and men as the difference between female and male returns to schooling is fairly small. Through the process of social and economic development, women have more chance to enrol in the education system partly to meet the demand for high-skilled workers and partly to meet the international expectation towards modernisation. This reflects in the results of disaggregated analysis by gender, both in terms of IV estimates and the bias gaps, in which the gaps of each result are marginally small between women and men.

Figure 7-3 Labour Force Participation Rates of Men and Women (15+), 2005



Source: Haspels and Majurin, 2008.

Table 7-5 Share of Female Wage Employment, 1996-2010

Year	Share of Female Wage Employment	
	To Total Female Employment (%)	To Total Wage Employment (%)
1996	34.2	40.8
1997	35.2	42.4
1998	35.3	43.5
1999	37.5	43.9
2000	38.8	44.1
2001	39.9	42.3
2002	39.3	41.4
2003	40.2	42.3
2004	42.9	44.2
2005	42.7	44.7
2006	42.9	44.0
2007	42.9	43.9
2008	42.6	43.8
2009	42.1	43.4
2010	41.5	44.7

Source: International Labour Office, 2013.

Table 7-6 Average Hours Worked per Week for Male and Female Workers (Wage and Salary Sector)

Types of Jobs	Hours	Female		Male	
		1995	2000	1995	2000
Paid Employees	1-9	0.13	0.38	0.04	0.16
	10-19	0.73	1.13	0.55	1.07
	20-29	2.12	2.73	1.46	2.88
	30-34	1.85	1.98	1.57	2.53
	35-39	12.88	14.68	11.48	13.10
	40-49	42.44	47.90	38.53	42.99
	50+	39.84	31.20	46.36	37.27
	Total	100.00	100.00	100.00	100.00
Self-Employed	1-9	0.51	0.67	0.29	0.41
	10-19	2.27	2.94	1.64	2.71
	20-29	6.64	7.11	5.88	5.64
	30-34	2.61	2.80	2.22	2.36
	35-39	10.87	9.59	7.12	6.56
	40-49	23.47	25.63	20.63	21.26
	50+	53.63	51.25	62.21	61.06
	Total	100.00	100.00	100.00	100.00
Total Employment	1-9	0.35	0.54	0.17	0.30
	10-19	1.63	5.16	1.11	1.95
	20-29	4.77	2.14	3.76	4.37
	30-34	2.30	2.44	1.91	2.44
	35-39	11.70	11.86	9.22	9.58
	40-49	31.32	35.55	29.25	31.31
	50+	47.93	42.32	54.58	50.07
	Total	100.00	100.00	100.00	100.00

Source: Lee, McCann, and Messenger (2007).

7.1.2. Cohort

7.1.2.1 Rates of Return to Schooling

Table 7-7 Disaggregated Analysis of OLS and IV Returns to Schooling by Cohort

Dependent Variables	OLS	IV	Bias gap	Observations
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Log monthly wages,	0.126***	0.0872***	0.0388	812,522
cohort 1955-1970	(0.00354)	(0.00347)		
Log monthly wages,	0.101***	0.0856***	0.0154	1,016,749
cohort 1961-1985	(0.00377)	(0.00354)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author's compilation based on LFS 1986-2012.

In terms of cohort, the OLS coefficients on years of schooling for the older cohort (cohorts born in 1955-1970) and the younger cohort (cohorts born in 1961-1985) are 0.126 and 0.101, respectively, while those of the IV estimation are 0.0872 for the older cohort and 0.0856 for the younger one (**Table 7-7**). The OLS estimates from cohort disaggregation still indicate the upward ability bias. (IV estimates are bigger than OLS estimates), which is consistent to the standard estimates (the first row of **Table 7-7**). This disaggregation shows the rates of return to education for cohort 1955-1970 is marginally higher than that of cohort 1961-1985. It seems that the result shows a transition towards a lower equilibrium of demand for and supply of educated labour. It is the equilibrium in which the returns to schooling become smaller as the supply of educated workers shifts to the right (an increase in labour supply) at a faster rate than the shift of demand for educated workers (an increase in labour demand).

Psacharopoulos (1985, 1993) argues that as education keeps expanding over time, diminishing returns should be observed. In other words, ceteris paribus, there is a negative correlation between education expansion and returns to schooling (Psacharopoulos, 1981, 1982). At early stages of development, a rapid economic expansion creates rents or surplus earnings for a small proportion of educated population to enjoy (Psacharopoulos, 2011). Educated workers can enjoy new rents created by the new demand. As a consequence, the expansion of education produces a supply of educated workers to meet a demand for educated workers and, therefore, the rents are eliminated. It is this interaction between the demand and the supply of skilled workers which, in turn, determines the rates of return to education. A rapid increase in the supply responding towards the shift in the demand could lead to a decline of returns to schooling (an elimination of the rents). As the process of economic development deepens, it is also possible that the rates of return remain constant if the demand keeps pace with the increasing supply (Psacharopoulos, 1981, 1988, and Welch, 1971). Tinbergen (1975) explains this situation in terms of a race between education and technology, shifts to the right of the supply of and the demand for educated labour.

Table 7-8 Demand and Supply of Skilled Workers 1960-1980

Skilled Workers	
1960 Stock	80,000
1980 Stock	288,000
Average Annual Demand	
For Replacement	9,300
For Expansion	10,400
Total Skilled Workers Required Annually	19,700

Source: World Bank, 1989

The above hypothesis possibly explains the reason that the rates of return to education for cohort 1955-1970 is marginally greater than that of cohort 1961-1985. On one hand, the older cohort (1955-1970) experienced the early stages of economic development characterised by the sharp increasing demand of educated workers (**Table 7-8**). **Figure 2-4** shows that, during 1960-1990, there was a rapid growth in the industrial sector, that required more educated workers to sustain the growth of this sector and the whole economy. This created rents or surplus earnings for a small group of educated population at that time to enjoy. Hence, the rates of return to schooling are higher for the older cohort. On the other hand, the younger cohort (1961-1985) experienced the prosperous stages of economic development that the excess demand of educated workers was relieved by the expansion of the supply of educated workers due to the Government's education expansion policy. After 1990, the contribution of industrial sector to the overall GDP remained constant (**Figure 2-14**). This implies a constant growing demand for educated workers. The education premium becomes smaller as the supply of educated workers keeps increasing. Thus, the estimated result indicates a transition towards equilibrium of demand for and supply of educated labour in Thailand which the returns to schooling become smaller since the supply of educated workers becomes larger and keeps pace with the increasing demand for educated workers.

7.1.2.2 Ability Bias (Bias Gap)

According to **Table 7-7**, the bias gap of the older cohort is approximately 4 per cent, while that of the younger cohort is around 1.5 per cent. The ability bias is more dominant in the older cohort, cohort 1955-1970. The magnitude of ability bias is possibly higher for the older cohort as they experienced the beginning stage of economic development, which there was still a much more limited access to education. At this period, most of families were under poverty. Hence, they might select only the most-able child among other children to study at school since they could not afford to send all children to school. This was due to high costs of schooling and demand for household labour. Moreover, schools were mainly concentrated in Bangkok and large cities. This imposed higher indirect costs of education, both in terms of time and transportation, to those who live far from the centre. As a result of financial constraint, only children with higher ability tended to be selected to study at schools. This is consistent with the ability bias hypothesis which states that more-able individuals are more likely to have higher years of schooling. On the other hand, the younger generation enjoyed an era of a higher level of economic development with an intensive education expansion. Due to an economic development and a higher standard of living, families gained more economic power to afford education for their children regardless of their level of ability. In addition, a universal access to education promoted free compulsory education and expansion of schools to remote areas; therefore, both direct and indirect costs of schooling have been reduced. The younger generation faced smaller costs of schooling and had higher opportunities to acquire education. Thus, an equality of access to education among the young generation reduces the ability bias in the estimation. As a result, the ability bias is more dominant in the OLS estimates of the older generation than those of younger generation. In addition, the higher ability bias implies that there is an inequality of opportunity in access to education between the older and younger generations.

The data from **Table 7-9** and **Figure 7-4** support the above argument. **Table 7-9** shows poverty incidence in Thailand during 1962-2002. It illustrates that during the social and economic period (1960-1980) Thailand experienced very high poverty incidence both in rural and urban areas. In 1962, the aggregate poverty incidence was as high as 88 per cent. This supports the above argument regarding the poverty and low standard of living in the older generation. Since 1990, the aggregate poverty incidence has become moderate and even lower than 10 per cent after 2000. This implies that the younger generation is benefited from economic development and suffer from poverty substantially lesser than those from older generation. Moreover, **Figure 7-4** indicates that there was much lower number of students enrolled in each education level for the older cohort, especially after four years of compulsory education (1962 Compulsory Education Act). This may be explained by the fact that poverty is more prevalent among the older cohort; therefore, individuals from the older cohort have less opportunity to attend school. As shown in **Table 7-9**, most of families are under poverty for the older cohort. Thus, it is possible that only the most-able child is allowed to study at school, while the rest is expected to help household work. On the other hand, the younger generation enjoys an era of a higher level of economic development and standard of living. This possibly explains a higher number of enrolments at each education level of the younger cohort and supports the fact that the younger generation face smaller costs of schooling and have higher opportunities to acquire education than those of the older generation.

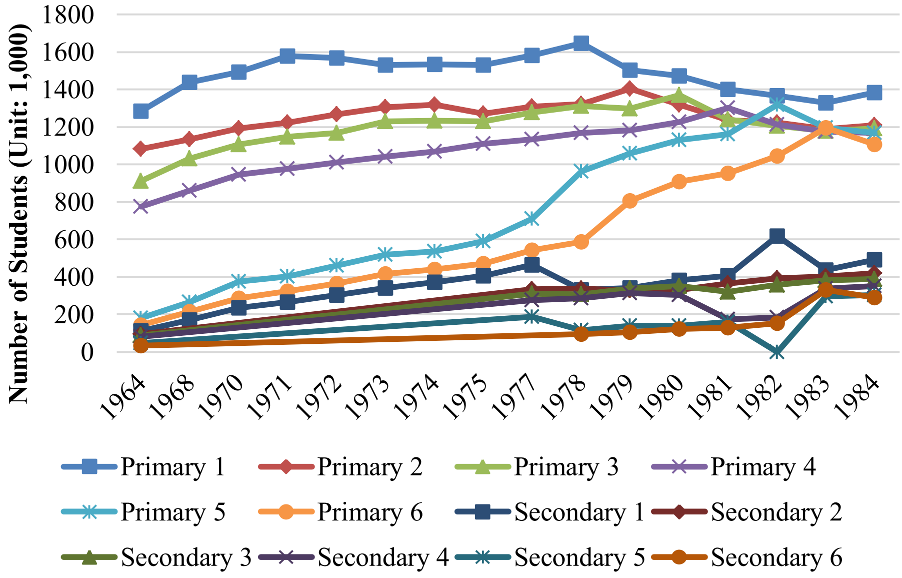
Table 7-9 Thailand: Poverty Incidence, 1962 to 2002 (Headcount Measure, Per Cent of Total Population)

	Aggregate	Rural	Urban
1962	88.3	96.4	78.5
1969	63.1	69.6	53.7
1975	48.6	57.2	25.8
1981	35.5	43.1	15.5
1986	44.9	56.3	12.1
1988	32.6	40.3	12.6
1990	27.2	33.8	1.6
1992	23.2	29.7	6.6
1994	16.3	21.2	4.8
1996	11.4	14.9	3
1998	12.9	17.2	3.4
2000	14.2	21.5	3.1
2002	9.8	12.6	3
Poverty share 2000	100	92.6	7.4
Population share 2000	100	68.4	31.6

Note: Poverty incidence means the number of poor within a reference population group expressed as a proportion of the total population of that group. The headcount measure of aggregate poverty incidence is the percentage of the total population whose incomes fall below a poverty line held constant over time in real terms; rural poverty is the percentage of the rural population whose incomes fall below a poverty line held constant over time in real terms, and so forth. Poverty share means the number of poor within a reference population group expressed as a proportion of the total number of poor within the whole population. Population share means the population of a reference group expressed as a proportion of the total population of that group

Source: Siripachai (2009) based on Development Evaluation Division, National Economic and Social Development Board, Bangkok and Medhi (1993).

Figure 7-4 Number of Students Enrolment by Education Level (Grade), 1964-1984



Source: Author's compilation based on UNESCO and Ministry of Education (1968-1986).

7.1.2.4 Conclusion

The estimated rates of return to schooling for the older cohort exceed those of the younger cohort within the wage labour market. This implies that the older generation faces higher opportunity costs in attaining education. This is possibly due to the interaction between the demand for and the supply of skilled workers in different periods of social and economic development experienced by the older and younger cohorts. At early stages of development, a rapid economic expansion creates rents or surplus earnings for a small proportion of educated population to enjoy. In contrast, the younger cohort enjoys smaller rents as the supply of educated workers keeps pace with the increasing demand in the later stage of development. In addition, the result also implies that there is a transition towards a lower equilibrium of the demand and supply of educated labour. The lower equilibrium indicates the lower returns to schooling since the supply of educated workers increase at a faster rate than does the demand for educated workers.

Furthermore, the bias gap for the older cohort is larger than that of the younger cohort. This implies that the ability bias is more dominant in the OLS estimates of the older generation. At the early stage of development, poverty and low standard of living, e.g. undeveloped school system influence household's decision making in sending their children to schools. It is likely that most of households may have to select only the most-able child to study at school due to the high direct and indirect costs of schooling and demand for household labour. Hence, there is a tendency that schools accumulate more-able children, which arguably causes the higher ability bias in the OLS estimates of the older cohort. Furthermore, the higher ability bias also implies that there is an inequality of opportunity to access education between the older and younger generations.

7.1.3. Areas of Residence (Urban and Rural Areas)

7.1.3.1 Rates of Return to Schooling

Regarding areas of residence, the OLS coefficients on years of schooling for the urban area and the rural area are 0.108 and 0.105, respectively, while those of IV estimation are 0.085 for the urban area and 0.071 for the rural area (**Table 7-10**). The OLS estimates are greater than those of IV; therefore, the upward ability bias still persists in both estimates of the urban and rural areas. There is a consistency between the estimates of disaggregated analysis and those of the standard estimates (the first row of **Table 7-10**). The results from this study and that of Warunsiri and McNown (2010) are at least qualitatively similar in terms of the direction. Both studies report that the estimates of the urban area are higher than those of the rural area (**Table 7-11**). Moreover, the result is consistent with other studies both from Thailand (Chriswick, 1983) and other countries, e.g. Indonesia (Parinduri, 2014), Kenya (Agesa, 2001), China (de Brauw, Rozelle and Zhang, 2005), and Mexico (Schultz, 2004).[29]

Table 7-10 Disaggregated Analysis of OLS and IV Returns to Schooling by Area of Residence

Dependent Variables	OLS	IV	Bias gap	Observations
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Log monthly wages,	0.108***	0.0849***	0.0231	856,617
urban	(0.00369)	(0.00169)		
Log monthly wages,	0.105***	0.0709***	0.0341	450,399
rural	(0.0053)	(0.00622)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author's compilation based on LFS 1986-2012.

As there are more sophisticated jobs that need skilled workers in the urban labour market, urban individuals have more chance to utilise their skills than do those living in rural areas (Warunsiri and McNown, 2010). On the contrary, jobs in rural areas require less skill; therefore, rural skilled workers cannot exploit their skills honed through higher

[29] Studies from Kenya (Agesa, 2001), China (de Brauw, Rozelle and Zhang, 2005), and Mexico (Schultz, 2004) are not included in **Table 7-11 Summary of Return to Education by Areas of Residence**Table 7-11 since the method of estimation is different from those reported in the table. Table 7-11 mainly focuses on the studies using IV with RDD. However, regardless of the estimation methods, all studies from Thailand are included in the table.

education (Jamison and Lau, 1982). In addition, the lesser quality of schools in rural areas also contributes to the lower returns to schooling (Behrman and Birdsall, 1983). Thus, the result implies that there is an inequality of job opportunities and school quality between rural and urban areas, in which better job opportunities and school with higher quality tend to concentrate in urban areas.

7.1.3.2 Ability Bias (Bias Gap)

The bias gap of urban areas is 0.023, whereas that of rural areas is approximately 0.034. In contrast, Warunsiri and McNown (2010) find that the bias gaps of urban and rural areas are 0.074 and 0.029, respectively. The magnitude of the rural bias gap (0.034) is somewhat consistent with that of Warunsiri and McNown (2010), 0.029. However, **Table 7-10** shows that the bias gap of rural areas is greater than that of urban areas, whereas the bias gap is greater for urban areas (7 per cent) in the study of Warunsiri and McNown (2010). The result of this study indicates that the ability bias is more dominant in rural areas. This implies that there is an inequality of opportunity in access to education between the rural and the urban areas, which supports the main IV estimated results showing that the urban IV estimate is greater than that of rural areas.

Table 7-11 Summary of Return to Education by Areas of Residence

Author	Method of Estimation	Urban vs Rural	Urban	Rural
1. Thailand				
Chriswick (1983)	OLS	>	BKK: 0.079***	BKK: 0.025
Warunsiri and Mcnown (2010)	Fixed effects with panel data	>	0.189**	0.142**
This study	IV with RDD	>	0.0849***	0.0709***
2. Others				
Fang, et al. (2012), China	IV with RDD	<	0.14	0.18***
La Vincent and College (2014), China	IV with RDD	<	-0.139 (not significant)	0.116 (not significant)
Parinduri (2014), Indonesia	IV with fuzzy RDD	>	0.22*	0.16**

Source: Author’s compilation.

The higher ability bias in rural samples can be explained by the fact that rural residents have lower opportunity in access to schooling than do urban residents. Rural residents can be regarded as one of socially disadvantaged groups as they are socially discriminated, more credit constrained and having greater immediate need to work. Rural residents have limited access to schools as most schools are concentrated in urban areas during the beginning of social and economic development period. Transportation system and basic infrastructure such as roads are still under development. Some areas even do not have access to such basic infrastructure. This contributes to a higher cost of attending school. Aside from the limited access to school, rural residents tend to be poorer than those living in urban areas. As rural residences' main source of income is from agriculture, the average income is smaller than those of other sectors. Furthermore, agriculture is labour-intensive, and family-oriented; therefore, labour from family members is required. Rural agricultural households exploit their own family members including their school-age children to cultivate the land. Hence, rural children have a greater immediate need to work and find that the opportunity cost of schooling is high. Due to higher costs of education and low income, rural households cannot afford to send their children to school. However, there is a tendency that, with the hope that an educated child will bring fortune to the family, poor households choose only one or few children with the highest ability among others to receive a proper education. As a result, children with higher ability tend to be selected to study at schools. This is consistent with the ability bias hypothesis which argues that more-able individuals are more likely to have higher years of schooling. On the other hand, urban residents have better access to schooling and are less credit constrained. Urban households can afford to send their children to schools regardless of ability level of their children. Thus, the urban OLS estimates suffer from ability bias at a lesser extent than those of the rural sample. In sum, inequality of access to education between rural and urban residents causes the higher ability bias in the OLS estimation and possibly explains the situation which a rural bias gap is greater than that of urban areas.

Table 7-9 supports the argument that rural residents are poorer than urban residents. In 1962, the rural poverty incidence accounts for 96.4 per cent, while the urban poverty incidence is approximately 78.5 per cent. Moreover, the gap between rural and urban poverty incidence keeps expanding over time. In 1990, the rural poverty incidence reduces to 33.8 per cent, whereas that of urban areas is only 1.6 per cent. Hence, it can be observed that rural residents have always been poorer than urban residents. Moreover, **Table 7-12** supports the fact that rural residents are lower educated than those from urban areas. There are approximately 88 per cent of rural residents who have only primary education or less, while just 2 per cent hold a university degree. In contrast, around 50 per cent of urban residents have only primary education or less, whereas 16 per cent hold a university degree. Warr (2011) reports that 94.7 per cent of the total number of poor people in 2002 have only primary education or less, while just 0.31 per cent hold a university degree. Hence, poor people tend to have low education level. This implies a negative correlation between poverty and education. Combining the information from **Table 7-9**, **Table 7-12**, and Warr's report (2011), it can be concluded that, with a high level of poverty, rural residents possibly have lower opportunity to access schooling than do urban residents and, therefore, are less educated than urban residents.

Table 7-12 Percentage Distribution of Educational Attainment of Populations Aged Six Years and Above in Rural and Urban Areas, 1993

Education	Rural	Urban
Males		
No Education	9.6	6.3
Primary	76.0	44.4
Secondary	11.7	32.6
University	2.5	16.3
Other	0.2	0.3
Total	100	100
Females		
No Education	14.2	9.4
Primary	75.8	44.4
Secondary	7.8	25.4
University	2.2	16.0
Other	na	0.1
Total	100	100

Note: na: not available.
Source: Author's compilation adapted from Vutthisomboon (1998) based on data from the National Statistical Office, Thailand, 1993.

7.1.3.3 Selection Bias

The area-disaggregated estimates may be subject to a sample selection bias. A potential selection bias may occur due to a rural-urban migration. In general, better educated individuals tend to migrate from rural areas to urban areas due to higher urban real wages (Greenwood (1975), Schwartz (1976), Schultz (1982b, 1988), Orazem and King (2007)). This phenomenon is described as, "human capital drain from rural to urban areas" (Amare, et al., 2012) and can be observed from other developing countries, namely Taiwan (Speare, 1974), Venezuela (Levy and Wadycki, 1974), Papua New Guinea (Ross, 1984). Evidence from the previous studies of Thailand also indicates the same pattern (Prachuabmoh et al., 1979, Tirasawat, 1985). Thus, the estimated urban rates of return does not purely represent the "real" urban rates of return as it also reflects rural immigrants' rate of returns.

Without migration, the "real" rates of return to schooling in urban areas should be higher than the estimated rates. Conversely, the "real" rates of return to schooling in rural areas should be lower. The underlying reason is that, without rural immigrants, urban workers can enjoy higher rents from disequilibrium between demand for and supply of educated workers since there are more job opportunities that educated workers can utilise their skills honed from their higher school years. However, an influx of rural migrants drives up the supply of educated workers in urban areas. Thus, returns to schooling decrease as the supply increases. Psacharopoulos (1987) posits that the returns to schooling among different areas will converge through the process of labour migration. In contrast, the "real" rates of return in rural areas should be lower since there is less demand for educated workers in rural areas and there is no absorption of educated workers to urban areas to relieve the excess supply of educated workers in rural areas. Given that there is no migration, rural educated workers enjoy much smaller rents than those from urban areas. Owing to higher urban real wages, better educated individuals are more likely to migrate from rural areas to urban areas; therefore, the rural-urban migration can be observed.

There are implicit assumptions underlying above discussion. First, Schultz (1988) and Agesa (2001) argue that net migration flow is characterised by a flow from rural to urban areas owing to higher real wages in urban areas and unbalance between demand for and supply of educated workers in rural areas. With the modern educational expansion, the supply of educated workers in rural areas increases rapidly, while the development of rural labour market cannot keep pace with the increase in educated labour supply. This results in an influx of migration flow from rural to urban areas. The trend can be observed in Latin American countries and Kenya (Schultz (1988), Agesa (2001)). Second, regarding education of migrants, more educated individuals tend to migrate to urban areas as they can exploit their skills and earn higher wages in the urban labour market. Those migrants are more likely to migrate after graduation (Schultz, 1988).

Table 7-13 Migration Streams of Migrants

Migration Stream	1965-70		1975-80		1985-90	
	Number	Per cent	Number	Per cent	Number	Per cent
Total migrants	**3,331,100**	**100**	**2,947,700**	**100**	**4,026,100**	**100**
Urban migrants	**763,400**	**23.0**	**1,024,900**	**34.8**	**1,448,700**	**36.0**
Urban-urban	297,000	8.9	506,000	17.2	545,100	13.5
Rural-urban	348,000	10.5	420,600	14.3	738,400	18.4
Unknown-urban	118,400	3.6	98,300	3.3	165,200	4.1
Rural migrants	**2,567,700**	**77.0**	**1,922,800**	**65.2**	**2,677,400**	**64.0**
Rural-rural	2,086,700	62.6	1,532,900	52	1,645,100	40.9
Urban-rural	180,400	5.4	278,300	9.4	508,900	12.6
Unknown-rural	300,600	9.0	111,600	3.8	423,400	10.5

Source: Pejaranonda et al, 1995 and Vutthisomboon (1998).

In Thailand, the phenomenon of "human capital drain from rural to urban areas" seems to occur as well. Migration streams of migrants can be broken down into six categories, including urban-urban, rural-urban, unknown-urban, rural-rural, urban-rural, and unknown-rural. According to **Table 7-13**, rural migrants (rural areas as migration destination) represents 77, 65.2, and 64 per cent of the total migration streams in the period of 1965-70, 1975-80, and 1985-90, respectively. Even though the decline in the number of rural migrants can be seen, the majority of the migration streams concentrate in rural migrants. In contrast, the rural-urban migration stream keeps increasing during the period of 1965–1990. The rural-urban migration flow makes up for 10.5-18.4 per cent. This indicates a growing trend of human capital drain from rural to urban areas.

In terms of education distribution of migrants, **Table 2-17** supports the second assumption of Schultz (1988). Rural-urban migrants consist of more educated individuals than those of rural non-migrant. The majority of rural-urban migrants attain at most primary education, which amounts to 59.6 per cent of the rural-urban migration flow, while 76 per cent of the rural non-migrant population complete only primary education. The percentage of persons completes secondary and university of rural-urban migrants (27.4 per cent and 10.2 per cent, respectively) are higher than those of rural non-migrants (11.7 and 2.5 per cent, respectively). The flow of rural-urban migrants is dominated by the group of more educated individuals. In other words, more educated individuals have higher mobility. This pattern can also be observed from each of migration flows, even within rural-rural migration stream.

Taking into account the rural-urban migration, the area-disaggregated estimates suffer from a sample selection bias as both assumptions of higher mobility among more educated persons and increasing rural-urban migration flow hold true in case of Thailand. The previous study from Colombia finds that the estimated returns to rural schooling indicate extremely downward bias as there is a large number of rural migrants in Colombia (Orazem and King, 2007). On the other hand, Schultz (1998) and Duraisamy (2002) argue that the sample selection bias is moderate in Côte d'Ivoire, Ghana, and India. It is worth examining the degree of sample selection by comparing corrected estimated returns to schooling with uncorrected ones in the case of Thailand. Nonetheless, with the Labour Force Survey data, it is not possible to test these empirical questions but the issues warrant further investigation.

7.1.3.4 Conclusion

The estimated result shows that returns to schooling in urban areas are greater than those of rural areas. This implies that there is an inequality of job opportunities and school quality between rural and urban areas, in which better job opportunities and school with higher quality tend to concentrate in urban areas. Urban labour market provides more sophisticated jobs that need skilled workers than does rural market. Thus, urban individuals have more chance to utilise their skills than do those living in rural areas (Warunsiri and McNown, 2010). In addition, the lesser quality of schools also contributes to the lower returns to schooling in rural areas as the wage reflects workers' productivity (Behrman and

Birdsall, 1983).

Moreover, the estimated result indicates that the ability bias is more dominant in rural areas as the bias gap for the rural residents is larger than that of the urban residents. This also supports the fact that there is an inequality of opportunity to access education between the rural and urban areas, which is consistent with the main IV estimated results showing that the urban IV estimate is greater than that of rural areas. The higher ability bias in rural samples can be explained by the fact that rural residents have lower opportunity to access schooling than do urban residents. The factors contribute to the lower opportunity include social discrimination and high poverty. Poor households tend to choose only the most-able child to send to school. This is consistent with the ability bias hypothesis which argues that more-able individuals are more likely to have higher years of schooling. On the other hand, urban households can afford to send their children to schools regardless of ability level of their children. Thus, the urban OLS estimates suffer from ability bias at a lesser extent than those of the rural sample.

Table 7-14 Percentage Distribution of Educational Attainment of Populations Aged Six Years and Above

	Migrant		Non-migrant	Migrant		Non-migrant
Education	**Urban-Rural**	**Rural-Rural**	**Rural**	**Urban-Urban**	**Rural-Urban**	**Urban**
Males						
No Education	3.5	5.6	9.6	3.2	2.5	6.3
Primary	48.4	71.2	76.0	39.4	59.6	44.4
Secondary	34.6	18.1	11.7	34.0	27.4	32.6
University	13.2	4.5	2.5	23.0	10.2	16.3
Other	0.2	0.6	0.2	0.4	0.3	0.3
Total	100	100	100	100	100	100
Females						
No Education	7.0	8.9	14.2	4.6	4.1	9.4
Primary	54.7	73.9	75.8	42.2	63.6	44.4
Secondary	24.2	12.0	7.8	30.5	21.4	25.4
University	14.1	5.2	2.2	22.7	10.8	16.0
Other	na	na	na	na	0.1	0.1
Total	100	100	100	100	100	100

Note: na: not available.

Source: Author's compilation adapted from Vutthisomboon (1998) based on data from the National Statistical Office, Thailand, 1993.

7.1.4. Region

7.1.4.1 Rates of Return to Schooling

Table 7-15 Disaggregated Analysis of OLS and IV Returns to Schooling by Region

Dependent Variables	OLS	IV	Bias gap	Observations
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Log monthly wages,	0.0956***	0.0751***	0.0205	162,399
BKK	(0.00204)	(0.00116)		
Log monthly wages,	0.124***	0.0990***	0.025	256,236
North	(0.00439)	(0.00436)		
Log monthly wages,	0.140***	0.0977***	0.0423	298,137
Northeast	(0.0038)	(0.00683)		
Log monthly wages,	0.0928***	0.0733***	0.0195	222,181
South	(0.00473)	(0.00343)		
Log monthly wages,	0.0973***	0.0766***	0.0207	369,035
Centre	(0.00459)	(0.00225)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author's compilation based on LFS 1986-2012.

In terms of region, the disaggregation shows the rates of return to education for the northern and north eastern regions are higher than those of Bangkok and metropolitan area, and the central and southern regions. The OLS coefficients on years of schooling for the northern and north eastern regions are 0.124 and 0.140, respectively, while those of IV estimation are 0.099 for the North and 0.098 for the North East (**Table 7-15**). The OLS and IV estimates of BKK, the South, and the Centre are on average around 9 per cent and 7 per cent, respectively. In all regions, the upward ability bias in the returns to education still persists in the disaggregation analysis. This supports the results of standard estimation where IV estimates are bigger than OLS estimates.

Table 7-16 shows a summary of returns to education by region from previous studies. The result of this study is qualitatively consistent with those of previous literature. First, the previous study from Thailand (Chriswick, 1983) indicates that returns to schooling from the northern and north eastern regions are among the highest, at least for urban areas. Moreover, the order of the returns to schooling from different regions is similar to that of this study. The returns of BKK are higher than those of the Central region, and the returns of the South region are the lowest. Second, the study from Canada (Oreopoulos, 2006) shows that the rate of returns is around 16 per cent after removing observations from the large and developed provinces, namely Quebec and Ontario. Thus, the remaining observations are from less developed regions which are comparable to the North and Northeast regions in Thailand. In contrast, the rates of return are about 10 per cent after removing observations from the Western provinces, which means that the remaining observations are from Quebec and Ontario. These two provinces are considered as developed provinces which are comparable to Bangkok and metropolitan areas and the Central region in Thailand. Hence, the result from Canada indicates that the rates of return to schooling from less developed regions are greater than those of more developed regions. This is consistent with the finding of this study which indicates that the returns from the North and Northeast regions, the least developed and poorest regions, are higher than those from the other regions which are relatively more developed.

Table 7-16 Summary of Returns to Education by Region

Thailand					
OLS estimation	**Bangkok and Metropolitan**	**Northeast**	**North**	**South**	**Central**
Urban	0.079***	0.093***	0.088***	0.059***	0.070***
Rural	0.025	0.040***	0.064***	0.020***	0.050***
Canada					
Provinces Excluded	**Maritimes**	**Quebec**	**Ontario**	**Quebec & Ontario**	**Western provinces**
IV Estimation	0.127 ***	0.166 **	0.117 **	0.163 ***	0.102 ***

Note: ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively.
Source: Thailand: Chriswick (1983), Canada: Oreopoulos (2006).

The main puzzle is that returns to schooling are higher in the agriculture-dominated North and Northeast than in the regions where non-agricultural opportunities are more widely available[30]. According to the cross-country analysis, Psacharopoulos (1981) finds that the returns to schooling in developing countries are greater than that of more developed countries. The North and the Northeast can be considered as developing economies since both are the poorest regions and still mainly depend on agriculture sector as a main source of income. Psacharopoulos (1983, 1985), and Montenegro and Patrinos (2013) explain the situation in terms of a relative scarcity of human-to-physical capital

[30] In fact, the higher average monthly wages are observed in relatively more developed regions, e.g. Bangkok and the Central region, rather than those from less developed regions, e.g., the Northeast and North regions. It should be noted that the concept of rate of returns is rather relative and the rate of returns refer to the rents or education premium created from the interaction of demand for and supply of educated workers.

within the regions. In less developed regions, there is a scarcity of skilled or educated labour; therefore, a small group of educated labour can enjoy the rents or education premium created from the process of economic expansion. On the other hand, the rents are relatively smaller in more developed regions due to a larger supply of educated labour. In addition, the scarcity of educated labour in less developed regions is exacerbated by the out-migration. In long run, the returns to schooling among different regions will converge through the process of labour migration (Psacharopoulos, 1987).

It seems that the finding from regional disaggregation contradicts that of urban-rural disaggregation. The regional rates of return show that less developed regions provide higher returns, while the rates of return from urban-rural disaggregation indicate that there are higher returns in more developed areas (urban areas). The previous section argues that as there are more sophisticated jobs that need skilled workers in the urban labour market (more developed area), urban individuals have more chance to utilise their skills than do those living in rural areas (Warunsiri and McNown, 2010). Moreover, the higher quality of schools in urban areas also contributes to the higher returns to schooling (Behrman and Birdsall, 1983). However, the above lines of logic may not apply to the estimates from the regional disaggregation

It is possible that a migration may explain the contradicting patterns between regional returns and area of residence's returns. The migration stream is not simply limited to a rural-urban migration but also includes regional migration. The interaction between these two dimensions of area and region of residence possibly complicates the estimation. The plausible explanation is that the argument used in urban-rural disaggregation is weaker in the case of regional disaggregation due to the coexistence of urban and rural areas within each region. Even though the Northeast and North regions are the least developed regions in Thailand, their urban areas help absorb some of educated labour by providing opportunities for more educated workers to work in more sophisticated jobs.

The higher returns to schooling in North and Northeast regions can be explained by a regional migration and a movement of demand for and supply of educated labour. In terms of regional mobility, both lifetime and temporary migrants tend to move to Bangkok and the Central region (**Table 7-17**). During 1970-1990, net migration losses are highest in the Northeast and the North regions. On the other hand, the Central area, and Bangkok and metropolitan areas have the highest net migration gains among all regions. Second, **Table 7-13** illustrates that the migration flow is dominated by the rural-rural and rural-urban migration streams. Although the decline in the number of rural migrants can be seen, the majority of the migration streams still concentrate in rural migrants. On the other hand, the rural-urban migration flow makes up around 10.5-18.4 per cent during the same period. It seems that the interregional migration flows from the Northeast and North regions to others and the rural-rural migration stream are quantitatively more dominant and significant than the rural-urban migration flows. This implies that there is an increase in the supply of educated labour in regions absorbing educated migrants, mainly in rural areas. This increasing supply leads to a reduction of rents or education premium used to be enjoyed by a smaller group of educated workers in those migrants-absorbing regions. Thus, a decline of returns to schooling can be seen in those regions, including Bangkok, the Central region and the South region. In contrast, outmigration reduces the supply of educated labour in the Northeast and North regions. Hence, a small group of educated population who keep working in the regions from which migration is originated can enjoy higher rents or surplus earnings created by urban development in those regions. An increase in returns to schooling can be observed. In sum, the rate of returns to schooling in the North and Northeast regions are higher than those of other regions owing to the migration and the interaction of demand and supply within each region.

Table 7-17 Regional Net Gains and Losses from Five-year Migration 1955-1990

Current residence	Total	Region of Previous Residence				
		Bangkok	Central	North	Northeast	South
1955-60						
Bangkok	67,045		41,208	5,047	17,855	2,935
Central	-46,643	-41,208		-14,710	15,102	-5,827
North	30,134	-5,047	14,710		21,106	-635
Northeast	-59,809	-17,855	-15,102	-21,106		-5,746
South	9,273	-2,935	5,827	635	5,746	
1965-70						
Bangkok	168,863		83,358	21,909	43,221	20,375
Central	-80,370	-83,358		-10,804	17,290	-3,498
North	6,340	-21,909	10,804		17,790	-345
Northeast	-85,006	-43,221	-17,290	-17,790		-6,705
South	-9,827	-20,375	3,498	345	6,705	
1975-80						
Bangkok	170,400		29,042	22,233	99,602	19,523
Central	49,454	-29,042		14,981	63,748	-233
North	-17,713	-22,233	-14,981		23,120	-3,619
Northeast	-194,815	-99,602	-63,748	-23,120		-8,345
South	-7,326	-19,523	233	3,619	8,345	
1985-90						
Bangkok	365,900		500	65,800	261,100	38,500
Central	293,400	-500		42,500	246,900	4,500
North	-89,300	-65,800	-42,500		26,800	-7,800
Northeast	-553,700	-261,100	-246,900	-26,800		-18,900
South	-16,300	-38,500	-4,500	7,800	18,900	

Source: Vutthisomboon (1998) based on 1955-60, 1965-70 and 1975-80 data from Goldstein and Goldstein (1986); 1985-90 data from the National Statistical Office (1993)

7.1.4.2 Ability Bias (Bias Gap)

According to **Table 7-15**, the bias gap of the North East and the North are about 4 and 2.5 per cent, respectively, whereas those of other regions are approximately 2 per cent. The ability bias is more dominant in the North East and the North. This implies that the problem of inequality of opportunity in access to education is more prevalent in the northeast and northern regions where the economic and educational systems are less developed.

The Northeast and the North share similar social and economic characteristics. First, both regions are agriculture-dominated and considered to be least developed among all regions. A high poverty rate is also observed in both regions. During the structural transformation, the Government put more emphasis on the industrial sector concentrated in the Central area, the South, and Bangkok and metropolitan areas. Similar to most industrialisation process in other countries, the industrial sector is developed at the expense of the agricultural sector. As these two regions are mainly agriculture-based, less development budget and effort from the Government is allocated to develop basic infrastructure and transportation system. Thus, basic infrastructure and transportation system lag behind those of other regions. In terms of education, the Government implements educational development plan to prepare a supply of skilled workers since it foresees a large increase in a demand during the industrialisation. Even though the Government has a duty to provide a universal compulsory education throughout the country, priorities are given to certain regions due to limited Government budget and a fast-growing demand for skilled labour in those regions. As the industrial sector is mainly developed outside the North and the Northeast, the pilot education expansion project is implemented first in the Central and the South regions. As a result, northern and north eastern households have relatively limited access to schools. Moreover, the problem is aggravated by the undeveloped basic infrastructure which increases the indirect cost of schooling. For all discussed reasons, northern and north eastern households face higher direct and indirect costs of attending school than do residents from other regions.

The OLS estimates from the Northeast and the North possibly suffer from ability bias due to high poverty and high costs of attending school. In less developed regions, a majority of households are under a tight financial constraint and may not be able to afford education for every child in their households. There is a high possibility that only more-able children are likely to be enrolled in schools. This leads to the fact that schools tend to be crowded with more-able students. Hence, an upward ability bias in returns to schooling is relatively greater in less developed economies. This is consistent with the ability bias hypothesis arguing that more-able individuals are more likely to attain more years of schooling and the argument explaining the higher ability in rural areas in the previous section.

7.1.4.3 Selection Bias

One concern regarding the region disaggregated analysis is that the analysis may face a similar problem of selection bias as does the area of residence disaggregated analysis (urban-rural). A potential selection bias may occur from interregional migration. In general, better educated individuals tend to migrate from less developed regions to more developed regions due to higher real wages and better job opportunities. This is similar to the case of rural-urban migration which human capital drain from rural to urban areas can be observed. Hence, the estimated rates of return of more developed regions do not purely represent the "real" rates of return as they also reflect immigrants' rates of return. Without migration, the "real" rates of return to schooling in more developed regions should be higher than the estimated rates since domestic workers can enjoy higher rents from disequilibrium between demand for and supply of educated workers. With immigration of educated workers, the rents or education premium from disequilibrium is smaller due to an increase in the supply of educated workers. In contrast, the "real" rates of return to schooling in less developed regions should be lower than the estimated rate since there is less demand for educated workers in less developed regions. Given that there is no migration, they enjoy much smaller rents than those from more developed regions. However, the sample selection bias due to interregional migration is possibly smaller since there is absorption of educated workers to urban areas within each region. This intraregional migration helps relieve the influx of educated workers into more developed regions.

7.1.4.4 Conclusion

The estimated result indicates that returns to schooling in less developed regions, namely the North and the Northeast, are greater than those of more developed regions, e.g. Bangkok and metropolitan areas. This contradicts the finding from urban-rural disaggregation. The main puzzle is that returns to schooling are higher in the agriculture-dominated North and Northeast than in the regions where non-agricultural opportunities are more widely available. It is possible that a migration may explain the contradicting patterns between the regional rates of return and the area of residence's rates of return. The higher returns to schooling in the North and the Northeast regions can be explained by a regional migration and a movement of demand for and supply of educated labour. The interregional migration flows from the Northeast and the North regions to the other regions lead to an increase in the supply of educated labour and, in turn, a reduction of rents or education premium used to be enjoyed by a smaller group of educated workers in regions that absorb educated migrants. Thus, a decline of returns to schooling can be seen in those regions, including Bangkok, the Central region and the South region. In contrast, outmigration reduces the supply of educated labour in the Northeast and the North regions. Hence, an increase in returns to schooling can be observed in both regions. In addition, the estimated result suggests that there is an inequality of job opportunities and access to schooling across regions.

Furthermore, the estimated result indicates that the ability bias is more dominant in the Northeast and the North regions as the bias gaps for these two regions are larger than those for the other regions. The higher ability bias in estimates from less developed regions can be explained by the fact that residents of those regions have lower opportunity to access to schooling than do residents from more developed regions since northern and north eastern households face higher direct and indirect costs of attending school. Thus, in less developed regions most households are under poverty and may not be able to afford education for every child in their households. There is a possibility that only more-able children are likely to be enrolled in schools, which is consistent with the ability bias hypothesis arguing that more-able individuals are more likely to attain more years of schooling. On the other hand, households from more developed regions can afford to send their children to schools regardless of children's ability. Thus, the OLS estimates from more developed regions suffer from ability bias at a lesser extent than those of less developed regions. In conclusion, an upward ability bias in returns to schooling is relatively greater in less developed economies. This finding also supports the fact that there is an inequality of opportunity to access education across regions.

7.1.5. Economic Sector

7.1.5.1 Rates of Return to Schooling

Table 7-18 Disaggregated Analysis of OLS and IV Returns to Schooling by Industrial Sector

Dependent Variables	OLS	IV	Bias gap	Observations
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Log monthly wages,	0.101***	0.0590***	0.042	428,699
Agricultural sector	(0.00479)	(0.0044)		
Log monthly wages,	0.0936***	0.0706***	0.023	238,274
Manufacturing sector	(0.00468)	(0.00285)		
Log monthly wages,	0.102***	0.0825***	0.0195	637,105
Service sector	(0.00325)	(0.00078)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author's compilation based on LFS 1986-2012.

In terms of economic sector, the OLS coefficients on years of schooling for agricultural workers, manufacturing workers, and service workers are approximately 0.10, 0.09, and 0.10, respectively, while those of IV estimation are about 0.06, 0.07, and 0.08, respectively (**Table 7-18**)[31]. The difference between the estimated returns to schooling for the agricultural sector and the service sector is the largest, which is approximately 2 per cent, while the difference between other estimated returns is around 1 per cent. In each subsample of different economic sectors, there exists the upward ability bias which OLS estimates are higher than those of IV. Moreover, the returns to schooling are lowest in the agricultural sector, but highest in the service sector. This supports the results of standard estimation and implies that there is inequality of job opportunities across different economic sectors since each economic sector provides different degree of returns to schooling. Better job opportunities tend to be available in service and manufacturing sectors in which modern education matters in developing workers' skills related to their work.

[31] Agricultural sector includes agriculture, hunting, forestry, fishing, mining, and quarrying, while manufacturing sector refers to electricity, gas and water supply, construction, and general manufacturing, e.g. food products and beverages, tobacco products, textiles, to name a few. in addition, service sector consists of wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods; hotels and restaurants; transport, storage and communication; financial intermediation; real estate, renting and business activities; public administration and defence; compulsory social security; education; health and social work; other community, social and personal service activities; and private households with employed persons.

The higher returns to schooling in the manufacture and the service sectors can be explained by the fact that most of sophisticated jobs concentrate in these two economic sectors. This is similar to the line of logic explaining the higher estimated returns in urban areas. As jobs in the manufacture and the service sectors are mainly characterised as relatively more sophisticated jobs that need skilled workers, workers in these sectors have more chance to utilise their skills and knowledge gained from schooling than do those working in the agricultural sector. On the other hand, traditional jobs in the agricultural sector require less skill and education. Educated workers cannot exploit their skills honed through higher education. The conventional wisdom is that education is not important in traditional agriculture and education matters in agriculture only when there is new agricultural technology or knowledge to learn from (Schultz, 1964). In other words, human capital is more important for productivity growth in the modern sector rather than the traditional sector. It is usually found that education helps workers in the modern agricultural sector respond faster to the new technology, including new seeds, cultivating techniques, fertilizers, and animal breeds (Welch, 1970; Huffman, 1977; Besley and Case, 1993; Foster and Rosenzweig, 1996, 2004; Abdulai and Huffman, 2005). The result implies that there is an inequality of job opportunities across different economic sectors. Better job opportunities mostly concentrate in the service sector, whereas the agriculture sector provides the least job opportunities in terms of returns to schooling and opportunity to exploit skills learnt from higher level of education.

Table 7-19 Total Factor Productivity Growth by Sectors, 1980 to 2002

Average Growth Rates (Per Cent Per Annum)	Aggregate	Agriculture	Manufacture	Services
Output	6.01	2.64	8.09	5.53
Raw Labour	2.19	1.50	5.25	3.47
Human Capital	2.49	9.43	11.35	6.90
Physical Capital	9.05	8.50	13.84	18.47
Agricultural Land	1.12	1.12	0	0

Source: Author's compilation adjusted from Warr (2011) based on data from National Economic and Social Development Board.

There is evidence to show that Thailand also follows the pattern explained above. In Thailand, the agricultural sector is rather backward and traditional since farmers are slow to adopt new cultivating technology and there is less support from the Government. First, during the period of social and economic transformation, the agricultural growth is solely driven by expanding cultivated areas (Siriprachai, 2009) rather than utilising agricultural technology advancement. Land productivity remains very low and stable, while labour productivity increases substantially (James, et al., 1987; Timmer, 1991; Watanabe, 1992). An increase in labour productivity leads to a contraction of agriculture as a share of total output. It can be observed that the output growth from the agricultural sector (2.64 per cent) is lower than those from the manufacture and the service sectors, 8.09 and 5.53 per cent, respectively (**Table 7-19**). Second, there has been insufficient attention from the Government to invest in research and development regarding domestic agricultural technology. As a result, farmers still use the same traditional technology in cultivation. The significance of the agricultural sector has been declining since 1970s, while the manufacture and the service sectors have been growing (**Figure 2-4**). During the period of industrialisation, the Government budget is mainly distributed to the manufacturing sector; therefore, it is argued that the industrial sector is developed at the expense of the agricultural sector. This directly affects the growth of agricultural sector. Some may argue that even without domestic research and development it is still possible to utilise agricultural technology from abroad. However, Siamwalla (1996) argues that those imported agricultural technology cannot be directly applied to Thai agricultural sector without necessary adjustments owing to different physical and economic environment in Thailand. Without technical assistance from the Government, Thai farmers cannot develop or even adopt new agricultural technology in their cultivation. In sum, it can be concluded that Thai agricultural sector is rather traditional and backward; therefore, workers in the agricultural sector possibly marginally benefit from a modern education. As argued by Schultz (1964), education is not important in traditional agriculture and education matters in agriculture only when there is new agricultural technology or knowledge to learn from.

In addition, the higher returns to schooling in the service sector may also possibly be explained by the interaction between the demand for and supply of high skilled and educated workers during the early period of economic and social development. Service workers are required to possess relatively higher skills and knowledge than those from other sectors, especially those from the agricultural sector. During the structural transformation period, a rapid growth of technology advancement causes excess demand for high skilled labour and excess supply of unskilled labour. These characteristics of early economic development may well explain the situation in which the returns to schooling for manufacturing and service workers are higher than those of the agricultural sector. Excess demand for high skilled labour creates rents or education premium in the manufacturing and the service sectors to a small group of high skilled labour (more educated labour) to enjoy. On the other hand, excess supply of lower skilled labour reduces education premium in the agricultural sector. Hence, the rate of return to schooling is highest in the service sector in which workers can enjoy the education premium due to the scarcity of high skilled labour supply. In contrast, the lowest returns to schooling would be expected to observe in the agricultural sector where education premium is small owing to the excess supply of unskilled workers.

7.1.5.2 Ability Bias (Bias Gap)

Table 7-18 shows that the bias gap of the agricultural sector is 0.042, while those of the manufacture and the service sectors are 0.023 and 0.020, respectively. The result indicates that the ability bias is most dominant in the agricultural sector. This implies that there is an inequality of opportunity in access to education across different economic sectors.

Inequality of access to education across different economic sectors seems to be influenced by other individual characteristics rather than economic sector itself. As ability bias hypothesis argues that more-able individuals tend to attain more schooling, it implies that ability bias is determined when individuals are still at school age (long before entering labour market). Individual characteristics, e.g. gender, area of residence, and cohort, influence households' decision making whether to send their children to school. In contrast, economic sector is a characteristic assigned to individuals after entering a labour market; therefore, it cannot be considered as one of the determinants of ability bias. However, in the context of disaggregated analysis by economic sector, area of residence chiefly explains ability bias within each economic sector. **Table 7-20** shows that workers from different economic sectors seem to have a clear pattern of residence and working area. Agricultural workers tend to live and work in rural areas, while a majority of workers from manufacture and service sectors concentrate in urban areas. On the other hand, a specific pattern of gender or cohort for different economic sectors cannot be observed as the proportion of each gender and each cohort is similar across different sectors.

The same argument used to explain the higher rural ability bias may also explain the situation which the bias gap in the agricultural sector is the highest among all economic sectors. Workers from different sectors follow a certain pattern of location of habitation. In general, a majority of wage farmers live and work in rural areas. As Thai agricultural sector is very labour-intensive and family-oriented, workers in agricultural sector are usually from poor farmer households. They either have their own land to cultivate or serve as employees for other farmer households who own cultivating land. Those who serve as wage employees are from poorer families residing in rural areas, whereas a majority of workers in manufacturing and service sector are mainly from urban areas (**Table 7-20**). Agricultural wage workers who mainly reside in rural areas have limited access to schools because most schools are concentrated in urban areas during the early period of social and economic development. Moreover, transportation system and basic infrastructure such as roads are still under development. This contributes to a higher cost of attending school. Aside from the limited access to school, agricultural workers tend to be poorer than those working in other sectors. **Table 7-21** indicates that agricultural workers earn the least among all sectors. During the period of industrialisation, the significance of agricultural sector has been declining, while the manufacturing and service sectors have been growing. It is argued that the industrial sector is developed at the expense of the agricultural sector. This directly affects income level of agricultural workers. Moreover, a large supply of unskilled labour in agricultural sector also drives the equilibrium wage downward. Due to higher costs of education and lower wages, agricultural households inevitably have to choose only one child or few children to study at schools. It is often the case that parents make their judgment based on children's ability. The most-able children are more likely to be selected to get proper education at schools. This is consistent with the ability bias hypothesis arguing that more-able individuals are more likely to have higher years of schooling. On the other hand, other sectors' workers who tend to live in urban areas have better access to schooling and are less credit constrained. Those workers can afford to send their children to schools regardless of ability level of their children. Hence, it can be observed that the ability bias is more dominant in the OLS estimates from agricultural sector than those from other sectors through the effect of residence area.

Table 7-20 Percentage of Workers in Different Economic Sectors by Areas of Residence

Economic Sector	**Areas of Residence**	
	Rural	**Urban**
Agriculture	54.76	45.24
Manufacture	38.58	61.42
Service	36.93	63.07

Source: Author's compilation based on LFS 1986-2012.

Table 7-21 Average Monthly Wages by Economic Sectors

Economic Sector	**Average Monthly Wages (Baht)**
Agriculture	8,690
Manufacture	8,720
Service	13,000

Source: Author's compilation based on LFS 1986-2012.

7.1.5.4 Conclusion

In terms of economic sector, the returns to schooling in the manufacture and the service sectors are higher than those of the agricultural sector since there are more sophisticated jobs in these two sectors. Hence, better job opportunities are available in the service and manufacture sectors which a modern education matters. On the contrary, traditional jobs in the agricultural sector require less skill and education; therefore, educated workers cannot exploit their skills honed through the modern education. The conventional wisdom is that education is not important in traditional agriculture and education matters in agriculture only when there is new agricultural technology or knowledge to learn from (Schultz, 1964).

In addition, the estimated result indicates that the ability bias is more dominant in less developed economic sector. The ability bias is largest in the agricultural sector and narrowest in the service sector. However, it is not possible to directly explain ability bias by the disaggregated analysis of economic sector. The reason is that ability bias is determined when individuals are still at school age, but economic sector is a characteristic assigned to individuals after entering a labour market. Inequality of access to education across different economic sectors can be explained by other individual characteristics rather than economic sector itself. It can be observed that workers from different economic sectors have a clear pattern of residence and working area. Agricultural workers tend to live and work in rural areas, while a majority of workers from manufacture and service sectors concentrate in urban areas. It is possible to utilise the same argument used to explain the urban-rural ability bias in the discussion of the bias gap of different economic sectors. Due to higher costs of education and lower wages in rural areas, agricultural households tend to choose only the most-able children to send to school. This is consistent with the ability bias hypothesis which argues that more-able individuals are more likely to have higher years of schooling. Other sectors' workers who tend to live in urban areas have better access to schooling and are less credit constrained. Those workers can afford to send their children to schools regardless of ability level of their children. Thus, the OLS estimates from the manufacture and the service sectors suffer from ability bias at a lesser extent than those from the agricultural sector.

7.1.6. Conclusion

With disaggregation by different demographic characteristics, the OLS estimates are greater than those of IV; therefore, the upward ability bias still persists even after estimating the rate of returns to schooling with different subsamples. This is consistent with the result from the standard estimation presented in Chapter 6.

Table 7-22 Summary of One-Level Disaggregated Results

	Rate of Returns	**Bias Gap**
Gender	Female > Male	Female > Male
Cohort	Older cohort > Younger cohort	Older cohort > Younger cohort
Area	Urban > Rural	Rural > Urban
Region	North, Northeast > Others	Northeast, North > Others
Economic Sector	Service > Manufacture > Agriculture	Agriculture> Manufacture > Service

Source: Author's compilation.

Table 7-22 summarises the results from the one-level disaggregated estimations with different subsamples stratified by gender, cohort, area of residence, region, and economic sector. First, in terms of gender, the rate of returns to schooling is marginally higher for women than for men. This implies a movement towards gender equality in access to education and also labour market in Thailand. Second, regarding the cohort disaggregation, the rate of returns is also marginally greater in the older than the younger generations. This is explained by different stages of economic development to which each generation belongs. At early stages of development, a rapid economic expansion creates rents or surplus earnings for a small proportion of educated population to enjoy (Psacharopoulos, 2011). Educated workers can enjoy new rents created by the new demand. As a consequence, the expansion of education produces a supply of educated workers to meet a demand for educated workers. Therefore, the rents are eliminated. As the process of economic development deepens, it is also possible that the rate of returns remains constant if the demand keeps pace with the increasing supply. Third, the urban rate of returns exceeds that of rural areas due to availability of more sophisticated jobs that need skilled workers in the urban labour market and higher quality of schools in urban areas. Urban individuals have more chance to utilise their skills acquired from better quality schools than do those living in rural areas (Behrman and Birdsall, 1983; Warunsiri and McNown, 2010). Fourth, in terms of regional disaggregation, the higher returns to schooling in least developed regions, the Northeast and the North regions, can be explained by a regional migration and a movement of demand for and supply of educated labour. The dominant effect of the interregional migration flows from the Northeast and the North regions to Bangkok, the Central, and South regions leads to an increase in the supply of educated labour and a reduction of returns to schooling in those labour-absorbing regions. In contrast, outmigration reduces the supply of educated labour in the Northeast and the North regions; therefore, returns to schooling increase in those regions. Finally, an excess demand for high skilled labour in the early stage of economic development creates rents or education premium in the manufacturing and the service sector to a small group of more educated labour to enjoy. Hence, the rate of returns to schooling is higher in the service and the manufacturing sectors, while the lowest returns to schooling would be expected to observe in the agricultural sector. Moreover, the higher returns to schooling in manufacturing and service sectors can be explained by the relevancy of modern education to job description in those sectors. Workers have more chance to exploit their skills learnt from modern school in more developed sector, while there is less chance to apply the modern knowledge in a sector such as agriculture in which technology is still traditional and backward.

Table 7-22 also reports the bias gap from the one-level disaggregated analysis. First, even though the magnitude of the difference between male and female bias gap is small, it seems that the ability bias is more dominant for women. In the early stage of social and economic development, market distortions such as gender discrimination prevent women to attend schools. Hence, only women with higher ability and better family background are given an opportunity to attend schools. This is consistent with the ability bias hypothesis which states that more-able individuals are more likely to attain more years of schooling. However, the estimated result also implies that there is a trend towards equal educational opportunity between women and men as the differences of the estimated returns to schooling and the bias gaps between women and men is substantially small. Second, the ability bias is more dominant in the OLS estimates of the older generation. At the early stage of development, poverty and low standard of living influence household's decision making in sending their children to schools. It is likely that most of households may have to select only the most-able children to study at school due to the high direct and indirect costs of schooling and demand for household labour. Third, in terms of area of residence, the estimated result indicates that the ability bias is more dominant in rural areas. Similar results can be observed in disaggregated analysis by region and economic sector. The estimated results show that the ability bias is more dominant in the North East and the North regions, the least developed regions in Thailand, and in the agricultural sector, least developed economic sector both in terms of output growth and technology. The common characteristic is the stage of development of each unit, namely area, region, and economic sector. Less social and economic development implies higher bias gap or higher ability bias in the OLS estimates. A similar line of reasoning explained the ability bias in cohort disaggregated analysis can be applied to disaggregated analyses by area of residence, region, and economic sector. With less educational opportunity and high poverty, households tend to choose only the most-able children to send to school as they cannot bear costs of education for all children in the household. This is consistent with the ability bias hypothesis arguing that more-able individuals tend to have higher education.

In conclusion, with the disaggregated analyses by different demographic characteristics, it is generally observed higher returns to schooling for women, older cohort, urban area, the North and the Northeast regions, and service sector. These results can be explained in the context of Thai development process. In addition, higher bias gap is prevalent in subsamples of women, older cohort, rural area, the North and the Northeast regions, and agricultural sector. Again, income and educational opportunity inequality during the early period of social and economic development help explain the higher bias gap in those socially disadvantaged groups. With less educational opportunity and high poverty, socially disadvantaged households tend to choose only most-able children to send to school since they cannot bear costs of education for all children in the household. This is consistent with the general ability bias hypothesis which argues that more-able individuals tend to have higher schooling.

Section 2. TWO-LEVEL DISAGGREGATED ANALYSIS

A two-level disaggregated analysis refers to an estimation based on different combinations of two demographic characteristics, e.g., gender-cohort, gender-region, cohort-industry, to name a few (**Table 7-23**). As each individual possesses multiple characteristics with different levels of return to schooling, interactions between different demographic characteristics in the two-level disaggregated analysis provide a better understanding towards the issue of heterogeneity in educational returns across individuals. Some of the estimates are relatively smaller or higher than those of the main results, but still comparable. In summary, the overall results are consistent with the results from the main analysis in which the upward bias in cross-sectional OLS regressions can be observed, and the ability bias is still dominant in the disaggregated results[32]. Some estimates from gender and region disaggregation with other characteristics show contradictions with the main results. Hence, results from two-level disaggregated analysis by gender and region with other demographic characteristics are discussed in this section (**Table 7-23**). The full estimated results of two-level disaggregated analyses of other demographic characteristics are provided in the Appendix IV.

Table 7-23 Analytical Matrix of Two-Level Disaggregated Analysis

	Gender	**Region**
Gender	na	See 2.1.2 Gender-Region
Cohort	na	na
Area	*2.1.1 Gender-Area*	*2.2.1 Region-Area*
Region	*2.1.2 Gender-Region*	na
Economic Sector	*2.1.3 Gender-Economic Sector*	*2.2.2 Region- Economic Sector*

Note: "2.x" refers to two-level disaggregated analysis.
"See" represents an overlapping section.
"na" means not available due to the fact that the results are not significantly different from the main results; they are consistent with the main results; or they are already discussed in the one-level disaggregated analysis.
Source: Author's compilation.

[32] Appendix V provides full estimated results of two-level disaggregated analyses.

7.2.1. Gender

7.2.1.1 Gender-Area of Residence

Table 7-24 Two-Level Disaggregated Analysis of OLS and IV Returns to Schooling by Gender and Area of Residence

Dependent Variables	OLS	IV	Bias Gap	Observations
Base Model				
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Gender				
Log monthly wages,	0.108***	0.0808***	0.0272	662,753
Male	(0.00439)	(0.00374)		
Log monthly wages,	0.116***	0.0856***	0.0304	644,263
Female	(0.00391)	(0.00367)		
Area of Residence				
Log monthly wages,	0.108***	0.0849***	0.0231	856,617
Urban	(0.00369)	(0.00169)		
Log monthly wages,	0.105***	0.0709***	0.0341	450,399
Rural	(0.0053)	(0.00622)		
Gender and Area of Residence				
Log monthly wages, urban male	0.102***	0.0776***	0.0244	424,146
	(0.00354)	(0.00192)		
Log monthly wages, rural male	0.101***	0.0687***	0.0323	238,607
	(0.00581)	(0.00657)		
Log monthly wages, urban female	0.113***	0.0871***	0.0259	432,471
	(0.00360)	(0.00210)		
Log monthly wages, rural female	0.106***	0.0668***	0.0392	211,792
	(0.00440)	(0.00613)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author's compilation based on LFS 1986-2012.

Regarding gender and area of residence, the OLS coefficients on years of schooling for each subsample are approximately 0.10-0.11, while those of IV estimation are around 0.067-0.087. The magnitude of OLS and IV coefficients is consistent with those from the one-level analysis. The OLS estimates are greater than those of IV; therefore, the upward ability bias still persists in the estimates from two-level analysis by gender and area of residence. Regardless of gender, the result shows that the urban estimated returns are higher than those of the rural estimation. This pattern follows the estimated result of the one-level analysis. As the urban labour market provides more sophisticated jobs that need skilled workers than does the rural market, urban individuals have more chance to utilise their skills honed from the modern education (Warunsiri and McNown, 2010). However, comparing the results between women and men in the same area of residence, the result shows a contradicting pattern. While the returns from urban women are greater than those of urban men, rural women have smaller returns to schooling than those of rural men ($0.0668 < 0.0687$). Overall, most of the estimates and all bias gaps of disaggregated analysis are consistent with those of the standard estimates (the first row of **Table 7-24**). Considering the area of residence, it seems that a tendency of gender equality in access to education and labour market between women and men is higher in rural areas. The magnitude of estimated returns between women and men are fairly similar.

7.2.1.2 Gender-Region

In terms of gender and region, the overall estimates of returns to schooling and the bias gaps of disaggregated analysis are consistent with those of the standard estimates and the estimates from one-level disaggregated analysis (**Table 7-25**). Regardless of gender, the OLS and IV estimated results show that the returns to schooling for the North and the Northeast regions are the highest among all regions. This pattern can be found in the main estimated result and also the results from one-level aggregated analysis by region. The higher returns to schooling in the North and the Northeast regions can be explained by a regional migration and a movement of demand for and supply of educated labour. Outmigration reduces the supply of educated labour in the Northeast and the North regions; therefore, there is an increase of rents or education premium enjoyed by a smaller group of educated workers in both regions. On the other hand, regardless of region, most of the OLS and IV estimates of the returns to schooling for women exceed those of men. This is consistent with the main results and implies that women face higher opportunity costs in attaining education. This is possibly due to the gender discrimination practice during the early period of social and economic development. Moreover, the estimated result shows that women have a higher bias gap than that of men. This implies that the ability bias is more dominant in the female OLS estimates since gender discrimination causes a situation in which only more-able women can obtain higher level of education.

However, comparing the results between women and men in the North and the Central regions, the estimated results show a contradicting finding against the overall estimates. In both regions, the results show that the estimated returns to schooling are higher for men. In contrast, without considering the region, the results indicate that female returns are greater than those of men. This contradicting pattern may indicate that there is a higher tendency of gender equality in access to education and labour market between women and men in the North and the Central regions as the magnitude of estimated returns between women and men are very similar. It is worth noting that the disaggregated analyses suffer from a selection bias in estimating returns to education, especially in case of developing countries. Thus, the results should be interpreted with care.

Table 7-25 Two-Level Disaggregated Analysis of OLS and IV Returns to Schooling by Gender and Region

Dependent Variables	OLS	IV	Bias gap	Observations
Base model				
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Gender				
Log monthly wages,	0.108***	0.0808***	0.0272	662,753
male	(0.00439)	(0.00374)		
Log monthly wages,	0.116***	0.0856***	0.0304	644,263
female	(0.00391)	(0.00367)		
Region				
Log monthly wages,	0.0956***	0.0751***	0.0205	162,399
BKK	(0.00204)	(0.00116)		
Log monthly wages,	0.124***	0.0990***	0.025	256,236
North	(0.00439)	(0.00436)		
Log monthly wages,	0.140***	0.0977***	0.0423	298,137
Northeast	(0.0038)	(0.00683)		
Log monthly wages,	0.0928***	0.0733***	0.0195	222,181
South	(0.00473)	(0.00343)		
Log monthly wages,	0.0973***	0.0766***	0.0207	369,035
Centre	(0.00459)	(0.00225)		
Gender and Region				
Log monthly wages, male	0.0945***	0.0656***	0.0289	82,071
Bangkok and Metropolitan	(0.00214)	(0.00261)		
Log monthly wages, male	0.121***	0.101***	0.02	126,355
Northern area	(0.00458)	(0.00604)		
Log monthly wages, male	0.131***	0.0871***	0.0439	151,593
North eastern area	(0.00401)	(0.00425)		
Log monthly wages, male	0.0862***	0.0659***	0.0203	116,980
Southern area	(0.00503)	(0.00500)		
Log monthly wages, male	0.0959***	0.0750***	0.0209	186,181
Central area	(0.00462)	(0.00324)		
Log monthly wages, female	0.0968***	0.0796***	0.0172	80,328
Bangkok and Metropolitan	(0.00195)	(0.00118)		
Log monthly wages, female	0.125***	0.0938***	0.0312	129,881
Northern area	(0.00414)	(0.00459)		
Log monthly wages, female	0.148***	0.101***	0.047	146,544
North eastern area	(0.00366)	(0.0101)		
Log monthly wages, female	0.0998***	0.0774***	0.0224	105,201
Southern area	(0.00422)	(0.00273)		
Log monthly wages, female	0.0981***	0.0738***	0.0243	182,854
Central area	(0.00435)	(0.00232)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author's compilation based on LFS 1986-2012.

7.2.1.3 Gender-Economic Sector

Table 7-26 Two-Level Disaggregated Analysis of OLS and IV Returns to Schooling by Gender and Economic Sector

Dependent Variables	OLS	IV	Bias gap	Observations
Base model				
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Gender				
Log monthly wages,	0.108***	0.0808***	0.0272	662,753
Male	(0.00439)	(0.00374)		
Log monthly wages,	0.116***	0.0856***	0.0304	644,263
Female	(0.00391)	(0.00367)		
Economic Sector				
Log monthly wages,	0.101***	0.0590***	0.042	428,699
Agricultural sector	(0.00479)	(0.0044)		
Log monthly wages,	0.0936***	0.0706***	0.023	238,274
Manufacturing sector	(0.00468)	(0.00285)		
Log monthly wages,	0.102***	0.0825***	0.0195	637,105
Service sector	(0.00325)	(0.00078)		
Gender and Economic Sector				
Log monthly wages, male	0.0944***	0.0596***	0.0348	230,334
agricultural and elementary workers	(0.00488)	(0.00381)		
Log monthly wages, male	0.0990***	0.0687***	0.0303	138,554
non-agricultural manual workers	(0.00522)	(0.00284)		
Log monthly wages, male	0.0903***	0.0709***	0.0194	293,022
desk, service, and intellectual workers	(0.00290)	(0.00126)		
Log monthly wages, female	0.104***	0.0476***	0.0564	198,365
agricultural and elementary workers	(0.00437)	(0.00514)		
Log monthly wages, female	0.0831***	0.0675***	0.0156	99,720
non-agricultural manual workers	(0.00335)	(0.00286)		
Log monthly wages, female	0.109***	0.0850***	0.024	344,083
desk, service, and intellectual workers	(0.00329)	(0.00286)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author's compilation based on LFS 1986-2012.

Regarding gender and economic sector, the OLS coefficients on years of schooling for each subsample are approximately 0.08-0.10, while those of the IV estimation are around 0.05-0.09 (**Table 7-26**). The magnitude of the OLS and IV coefficients is consistent with those from the one-level analysis. The OLS estimates are greater than those of the IV; therefore, the upward ability bias still persists in the estimates from the two-level analysis by gender and economic sector. Regardless of gender, the result shows that the returns to schooling for agricultural sector are lowest, while those for service sector are highest. The results support the results of the standard estimation and those of the one-level analysis. It implies that there is inequality of job opportunities across different economic sectors since better job opportunities tend to be available in service and manufacturing sectors in which modern education matters. Regardless of economic sector, only the result from service sector is consistent with the standard estimation in which the female returns to education are higher than those of men. It argues that education helps women expand their job opportunities, both in terms of access to labour market and occupational choice (Psacharopoulos, 1988, 1992, Dougherty, 2005). As women have much more limited access to both education and labour market, investing in one additional year of schooling for women would allow them to earn much more than those with less education. However, the returns to education are higher for men in agricultural and manufacturing sectors. This contradicts the results of the base model and the one-level analysis.

The contradicting results in agricultural and manufacturing sectors are not obvious. The female and male rates of return in manufacturing sector are substantially close to each other since the difference between estimated returns of women and men are fairly marginal (0.0687-0.0675 = 0.0012). This implies that there is a trend towards gender equality in the manufacturing sector in which women and men face similar opportunity costs to receive education and enter labour market. However, there is no obvious reason that explains the higher male returns to schooling in agricultural sector.

7.2.2. Region

7.2.2.1 Region-Area of Residence

Table 7-27 shows the two-level disaggregated analysis of OLS and IV returns to schooling by region and area of residence. The overall estimates of returns to schooling and the bias gaps of disaggregated analysis are consistent with those of the standard estimates and the estimates from one-level disaggregated analysis. First, the OLS estimates are higher than those of the IV estimation. This confirms that the upward ability bias still persists in the estimates from the two-level analysis. Second, regardless of region, the urban rates of return to schooling are greater than those of the rural sample. This is because urban labour market provides more sophisticated jobs that need skilled workers than does rural market. Thus, urban individuals have more chance to utilise their skills than do those living in rural areas (Warunsiri and McNown, 2010). Moreover, regardless of area of residence, the OLS and IV estimated results show that the returns to schooling for the North region are the highest among all regions, while the estimates from other regions preserve the ranking of the rates of returns illustrated in the main estimated result and the results from the one-level aggregated analysis by region, except for the estimates from rural north eastern region. The higher returns to schooling in North and Northeast regions can be explained by a regional migration and a movement of demand for and supply of educated labour. The interregional migration flows from the Northeast and North regions to the other regions leads to an increase in the supply of educated labour and, in turn, a reduction of rents or education premium used to be enjoyed by a smaller group of educated workers in the regions that absorb educated migrants. Thus, a decline of returns to schooling can be seen in those regions, including Bangkok, the Central region and the South region. In contrast, outmigration reduces the supply of educated labour in the Northeast and the North regions. Hence, an increase in returns to schooling can be observed in both regions. In addition, the estimated result also suggests that there is an inequality of job opportunities and access to schooling across regions.

Taking into account of area of residence, the estimated returns to schooling of urban areas are consistent with the main finding which indicates that the returns to schooling for the North and the Northeast regions are the highest among all regions. However, this pattern does not hold true in the case of rural Northeast region where the returns to schooling decreases to the lowest level among all regions. The highest level of returns to schooling of the urban North and Northeast can possibly be explained by the fact that residents from both the North and Northeast may benefit relative more from urbanisation. Compared to other regions, the North and Northeast are relatively less developed. Urbanisation helps residents in those regions expand their job opportunities, both in terms of access to labour market and occupational choice. As residents from these two regions have much more limited access to both education and sophisticated labour market, investing in one additional year of schooling for those residents would allow them to earn much more than those with less education. In contrast, it is quite common among residents from relatively more developed regions, e.g. Bangkok and metropolitan, to attain more education and to have better job opportunities; therefore, investing just one additional year of schooling would marginally increase their returns to schooling. Furthermore, urbanisation helps expanding the occupational choices for those from less developed regions. Thus, they may move away from traditional lower-paid jobs to higher-paid ones. This also contributes to the higher returns to schooling for urban north and northeast residents. On the other hand, jobs in rural areas require less skill; therefore, rural skilled workers cannot exploit their skills honed through higher education (Jamison and Lau, 1982). Hence, residents from the rural areas have lower returns to education. The result shows that the returns to schooling from the rural northeast region are the lowest among other rural regions. As the rural northeast region is the least developed area and jobs in this area requires less skills, residents find it difficult to apply their knowledge gained from education to their jobs. Thus, the rate of returns to education is the lowest among other rural regions. Nonetheless, the reason that the North region still has the highest returns is not obvious. The result suggests that residents in the rural North region benefit from education more than those from the other rural regions.

7.2.2.2 Region-Economic Sector

The two-level disaggregated analysis by region and economic sector (**Table 7-28**) indicates that the overall estimates of returns to schooling and the bias gaps follow those of the standard estimates and estimates from the one-level disaggregated analysis. First, the upward ability bias can still be observed in the estimates from two-level analysis as the rates of return to schooling from OLS estimation exceed those of IV estimation. Second, regardless of region, the rates of return to schooling from the agricultural sector is the lowest among all economic sectors, whereas that from the service sector is the highest. The higher returns to schooling in the manufacture and service sectors can be explained by the fact that better job opportunities mostly concentrate in the service sector, whereas the agriculture sector provides the least job opportunities in terms of returns to schooling and opportunity to exploit skills learnt from higher level of education. Furthermore, regardless of economic sector, the OLS and IV estimated results show that most of the returns to schooling from different regions preserve the ranking of the rates of return estimated from the one-level aggregated analysis, except for the estimates from the Northeast region. Nevertheless, the rates of return to schooling from the North and Northeast regions are still the largest in the service sector. A regional migration and a movement of demand for and supply of educated labour possibly explain the higher returns to schooling in the North and Northeast regions. As the interregional migration flows from the Northeast and North regions to others reduce the supply of educated labour in the home regions, an increase in rents or education premium enjoyed by a smaller group of educated workers can be observed. On the other hand, Bangkok, the Central region and South region absorb immigrants into their labour market. Thus, a decline of returns to schooling can be seen in those regions due to a reduction in education premium induced by the higher supply of educated workers.

Taking into account of economic sector, the estimated returns to schooling of agricultural and manufacturing sectors portray contradicting results towards the main finding which indicates that the returns to schooling for the North and Northeast regions are the highest among all regions. In the two-level analysis, the rates of return of the Northeast regions are the lowest in agricultural and manufacturing sectors. The fact that residents from both the North and the Northeast may benefit relative more from the service sector may explain the situation in which the returns to schooling for the North and the Northeast regions in the service sector are higher than those of other regions. In general, the North and Northeast are less developed and poor regions where agriculture is the main source of income. While the supply of educated workers keep expanding, technology advancement in the agricultural sector cannot keep pace with the growth of educated labour supply. Thai agricultural sector still exploits traditional cultivating technique. However, Schultz (1964) argues that education does not matter in the traditional agriculture in which there is no new agricultural technology or knowledge to learn from. Hence, development of service sector helps residents in those regions expand their job opportunities, both in terms of access to labour market and occupational choice because better job opportunities are available in the service and manufacturing sectors in which modern education matters in developing workers' skills related to those jobs. Investing just one additional year of schooling would substantially increase their returns to schooling. On the contrary, service sector is more prevalent in more developed regions, e.g. Bangkok and metropolitan; therefore, investing one additional year of schooling would marginally increase their returns to schooling.

7.2.2.3 Conclusion

The overall result from the two-level disaggregated analysis supports the results from the base model and the one-level disaggregated analysis. The result indicates that the upward ability bias persists in all analyses. Moreover, with disaggregated analyses by a combination of different demographic characteristics, it generally observes higher returns to schooling for women, older cohort, urban area, the North and Northeast regions, and service sector. In addition, higher ability bias (bias gap) is dominant in subsamples of women, older cohort, rural area, the North and the Northeast regions, and agricultural sector. This implies that there is educational inequality in access to schooling in those demographic characteristics. Individuals with those characteristics have difficulties to attain education due to less educational opportunity and poverty. It is usually the case that decision regarding child's schooling mainly made by his/her parents. Parents made their decision based on current demographic characteristics of their children and their households. Hence, households with poverty and poor educational opportunity tend to select one or few children with higher ability to study at school as they cannot afford to send all children to school. This is consistent with the ability bias hypothesis which argues that more-able individuals tend to have higher schooling. Lastly, it is worth noting that the more detailed the disaggregated analysis is, the higher the risk of sample selection bias. The estimated results should be interpreted with caution.

Table 7-27 Two-Level Disaggregated Analysis of OLS and IV Returns to Schooling by Region and Area of Residence

Dependent Variables	OLS	IV	Bias gap	Observations
Base model				
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Region				
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Log monthly wages,	0.0956***	0.0751***	0.0205	162,399
BKK	(0.00204)	(0.00116)		
Log monthly wages,	0.124***	0.0990***	0.025	256,236
North	(0.00439)	(0.00436)		
Log monthly wages,	0.140***	0.0977***	0.0423	298,137
Northeast	(0.0038)	(0.00683)		
Log monthly wages,	0.0928***	0.0733***	0.0195	222,181
South	(0.00473)	(0.00343)		
Log monthly wages,	0.0973***	0.0766***	0.0207	369,035
Centre	(0.00459)	(0.00225)		
Area of Residence				
Log monthly wages,	0.108***	0.0849***	0.0231	856,617
urban	(0.00369)	(0.00169)		
Log monthly wages,	0.105***	0.0709***	0.0341	450,399
rural	(0.0053)	(0.00622)		
Region and Area				
Log monthly wages, urban	0.0962***	0.0758***	0.0204	143,357
Bangkok and Metropolitan	(0.00198)	(0.00143)		
Log monthly wages, urban	0.121***	0.0997***	0.0213	162,960
Northern area	(0.00394)	(0.00235)		
Log monthly wages, urban	0.129***	0.0965***	0.0325	192,139
North eastern area	(0.00331)	(0.00336)		
Log monthly wages, urban	0.0912***	0.0726***	0.0186	133,456
Southern area	(0.00392)	(0.00245)		
Log monthly wages, urban	0.0957***	0.0773***	0.0184	225,565
Central area	(0.00419)	(0.00103)		
Log monthly wages, rural	0.0876***	0.0680***	0.0196	19,042
Bangkok and Metropolitan	(0.00291)	(0.00237)		
Log monthly wages, rural	0.115***	0.0818***	0.0332	93,276
Northern area	(0.00509)	(0.00766)		
Log monthly wages, rural	0.142***	0.0645***	0.0775	105,998
North eastern area	(0.00563)	(0.0143)		
Log monthly wages, rural	0.0838***	0.0660***	0.0178	88,725
Southern area	(0.00554)	(0.00481)		
Log monthly wages, rural	0.0938***	0.0695***	0.0243	143,470
Central area	(0.00509)	(0.00435)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.

Source: Author's compilation based on LFS 1986-2012

Table 7-28 Two-Level Disaggregated Analysis of OLS and IV Returns to Schooling by Region and Economic Sector

Dependent Variables	OLS	IV	Bias gap	Observations
Base model				
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Region				
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Log monthly wages,	0.0956***	0.0751***	0.0205	162,399
BKK	(0.00204)	(0.00116)		
Log monthly wages,	0.124***	0.0990***	0.025	256,236
North	(0.00439)	(0.00436)		
Log monthly wages,	0.140***	0.0977***	0.0423	298,137
Northeast	(0.0038)	(0.00683)		
Log monthly wages,	0.0928***	0.0733***	0.0195	222,181
South	(0.00473)	(0.00343)		
Log monthly wages,	0.0973***	0.0766***	0.0207	369,035
Centre	(0.00459)	(0.00225)		
Economic Sector				
Log monthly wages,	0.101***	0.0590***	0.042	428,699
Agricultural sector	(0.00479)	(0.0044)		
Log monthly wages,	0.0936***	0.0706***	0.023	238,274
Manufacturing sector	(0.00468)	(0.00285)		
Log monthly wages,	0.102***	0.0825***	0.0195	637,105
Service sector	(0.00325)	(0.00078)		
Region and Economic Sector				
Log monthly wages, Bangkok and Metropolitan	0.0916***	0.0590***	0.0326	18,206
agricultural and elementary workers	(0.00198)	(0.00341)		
Log monthly wages, Northern area	0.108***	0.0693***	0.0387	99,008
agricultural and elementary workers	(0.00486)	(0.00587)		
Log monthly wages, North eastern area	0.136***	0.0407***	0.0953	122,801
agricultural and elementary workers	(0.00528)	(0.00903)		

Log monthly wages, Southern area	0.0729***	0.0573***	0.0156	80,575
agricultural and elementary workers	(0.00485)	(0.00449)		
Log monthly wages, Central area	0.0884***	0.0628***	0.0256	108,148
agricultural and elementary workers	(0.00490)	(0.00412)		
Log monthly wages, Bangkok and Metropolitan	0.101***	0.0731***	0.0279	50,096
non-agricultural manual workers	(0.00372)	(0.00241)		
Log monthly wages, Northern area	0.0960***	0.0760***	0.02	34,490
non-agricultural manual workers	(0.00548)	(0.00330)		
Log monthly wages, North eastern area	0.0971***	0.0598***	0.0373	32,626
non-agricultural manual workers	(0.00499)	(0.00657)		
Log monthly wages, Southern area	0.0837***	0.0644***	0.0193	26,171
non-agricultural manual workers	(0.00520)	(0.00937)		
Log monthly wages, Central area	0.0889***	0.0693***	0.0196	95,064
non-agricultural manual workers	(0.00498)	(0.00417)		
Log monthly wages, Bangkok and Metropolitan	0.0931***	0.0791***	0.014	91,729
desk, service, and intellectual workers	(0.00174)	(0.00241)		
Log monthly wages, Northern area	0.114***	0.0944***	0.0196	122,639
desk, service, and intellectual workers	(0.00344)	(0.00278)		
Log monthly wages, North eastern area	0.113***	0.0929***	0.0201	142,589
desk, service, and intellectual workers	(0.00303)	(0.00255)		
Log monthly wages, Southern area	0.0899***	0.0702***	0.0197	115,376
desk, service, and intellectual workers	(0.00368)	(0.00183)		
Log monthly wages, Central area	0.0949***	0.0774***	0.0175	165,532
desk, service, and intellectual workers	(0.00375)	(0.00188)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author’s compilation based on LFS 1986-2012.

CHAPTER 8. CONCLUDING REMARKS

Section 1. SUMMARY BY CHAPTER

Chapter 2 provides readers with a comprehensive overview of Thai economic and educational development during 1960-1990. Thai economy can be divided into four main periods, namely pre-boom (1951-1986), boom (1987-1996), crisis (1997-1998), and post-crisis (1999 onwards). During 1951-1986, the Government put efforts in building basic physical infrastructure, e.g. road, electricity, to name a few, to help facilitating trading and the growth of the economy. The real GDP growth rate jumped to 9.2 per cent during the boom period, 1987-1996. This is considered as an economic miracle growth in Thailand. However, Thailand experienced a financial crisis in 1997 and faced a negative GDP growth at 6.1 per cent, which is the lowest in the Thai economic history. In the post-crisis period, Thailand managed to bounce back and achieved the real GDP growth rate of 4 per cent. The overall average real GDP growth and real GDP growth per capita were around 6.2 per cent and 4.2 per cent, respectively, during 1951-2003.

The high GDP growth is claimed to be the result of structural transformation during the pre-boom period, 1951-1986. Thailand went through a structural transformation from a primitive agriculture-based economy to newly industrialised economy, which the agricultural sector played a crucial role in reallocating resources to the other economic sectors. Similar to most of industrialisation process in other countries, the industrial sector was developed at the expense of the agricultural sector. In addition, Thailand also went through another structural transformation in terms of export and import pattern. Thailand changed from agriculture produce exporter, e.g. rice, to manufactured goods exporter, especially garments and parts and components.

Overall, the situation of the development in Thailand is consistent with the characteristics of developing countries proposed by Behrman and Srinivasan (1995). They suggest that developing countries possess different degrees of market completeness and institutions from developed countries. The followings are the proposed characteristics of developing economies in comparison with developed economies. First, a majority of population still depends heavily on agricultural sector and rural labour activities. Second, in agricultural sector, it is common for family members to help working without formal pay; therefore, non-wage labour accounts for a large proportion of the total labour force in developing economies. Third, there has been a rapid growth of labour force in developing countries. Fourth, high labour participation rates among 15-64 year olds can be observed due to low human capital investments among young cohorts. Fifth, there are low school enrolment rates and the education gap is in favour of males. Sixth, lower non-labour production inputs per worker are observed in developing countries. Lastly, a majority of Thai population still lives below the poverty line even in the midst of high economic growth during 1970-1980.

In terms of educational development, there had been a rapid expansion of the education system in the primary education for past 20 years (1960-1980), both in terms of student enrolment and number of teachers. In contrast, the demand and supply for secondary and tertiary education grew at a much slower pace. The Government raised both supplies of primary schools and teaching staffs to respond to the increasing demand for primary education while tried to increase the quality of the education. Hence, an improvement of overall education quality could also be observed as the student-teacher ratios kept declining over time across different levels of education. The boost in both demand and supply of education was a part of a national development plan to prepare Thai labour force to be ready for the next stage of social and economic development.

In terms of migration, the rural-urban and interregional migration help facilitating the development process of Thai economy by reallocating human resources between regions and economic sectors. The phenomenon of "human capital drain from rural to urban areas" seems to occur in Thailand. Even though the majority of the migration streams concentrated in rural migrants, there was a growing trend of human capital drain from rural to urban areas during the period of 1965 – 1990. The flow of rural-urban migrants is dominated by the group of more educated individuals. In other words, more educated individuals have higher mobility. This pattern can also be observed from every type of migration stream. Regarding regional mobility, net migration losses were highest in the Northeast and the North regions during 1970-1990, while the Central area, and Bangkok and metropolitan areas had the highest net migration gains among all regions.

In conclusion, Thailand experiences rapid economic development and structural transformation during 1960-1990. Obtaining the rates of return to education in this period helps us better understand the process of Thai economic development as well as the interplay between the rates of return to schooling and economic development in Thailand during 1980s to 1990s. In addition, the overall social and economic conditions of the development in Thailand are consistent with the general characteristics of other developing countries. Hence, by implication, estimating the rates of return to schooling in Thailand also provides better understandings on the role of human capital in the process of development in other developing countries. As developing countries possess different degrees of market completeness and institutions which are radically different from those of developed countries, this warrants value for investigation of the rate of returns to schooling in the context of developing countries. This further investigation possibly gives a different economic pattern and implications of estimated results.

Chapter 3 revisits the Mincer model to show that the Mincer function is alive and well in fitting actual age-wages profiles. Second, the chapter reviews previous literature regarding methodological issues in estimating the rates of return to schooling. The third section surveys previous studies of the rates of return to schooling with a particular focus on studies utilising compulsory education law change as an instrumental variable. Fourth, previous studies on the sample selection are briefly discussed. Finally, the chapter is closed by emphasising the contributions of this research.

The rates of return to schooling is one of the most important topics that economists have been investigating, especially by utilising the Mincerian equation. Despite the fact that the Mincer model is still alive and well in fitting the actual age-wage data, there has been a long debate that the OLS estimate from Mincer equation is possibly biased due to the endogeneity problem. Hundreds of studies with different methods of estimation attempt to deal with the endogeneity bias but they fail to establish a causal effect of schoolings on earnings in the absence of randomised experiment. Even though a randomised controlled trial (RCT) is the ideal estimation method, it is not feasible to conduct the RCT in most of the studies of returns to education. The second-best candidate, which is close to the RCT, is the estimation with quasi-experiment, e.g. regression discontinuity design (RDD), differences-in-differences, to name a few. However, the studies utilising those methods of estimation are rare, especially in developing countries where the issue of data scarcity is prevalent.

This study mainly focuses on the IV estimation using compulsory schooling as IV. There are two main debates in the literature from developed countries. The first debate is that the LATE is far from the ATE as the interventions do not affect a majority of the population. Thus, the IV estimates tend to be higher than those of OLS since the estimated results indicate the estimated returns to schooling of a socially disadvantaged group that usually has higher returns to schooling. The second debate raised by Stephen Jr. and Yang (2014) is the concern of potential bias from the omitted variables, including school quality and family background. They argue that without the interaction between region control and year of birth, changes across regions over time of school quality cannot be captured in the estimation. The effects of differences in school quality may cause estimates to be biased since the source of upward bias is not limited to innate ability, but also includes family background and quality of schooling (Schultz, 1988, Behrman, 1990, Strauss and Thomas, 1995). Even though there is no specific debate to the literature from developing countries, the overall results show that the estimates between developed and developing countries are different. The results from developing countries show that the OLS estimates are higher than those of IV estimation, while the estimates from developed countries tend to produce OLS estimates that are lower than those of IV. It implies that the ability bias is dominant in the OLS estimation of developing countries.

This study make three main contributions to the existing debates in the literature in terms of methodology and also substantive aspect in the context of Thailand and, by implication, developing countries in general. First, the previous literature reveals that there is a different pattern of the relative magnitudes between OLS and IV estimates using compulsory schooling as IV between developed and developing countries. Investigating this contrast can contribute to a better understanding of (a) how and when the conventional "ability bias" matters in estimating returns to schooling and of (b) the impact of compulsory schooling in different settings. Second, as Thailand experiences rapid economic development and structural transformation during 1960-1990, obtaining the rate of returns to education in this period helps us better understand the process of Thai economic development as well as the interplay between the rate of returns to schooling and the economic development process during 1980 to 1990. This also helps understand the role of human capital in the process of development in other developing countries. Finally, the third contribution is on the construction of the database and discussion of the descriptive analysis for the discrepancies among different demographic characteristics, including gender, cohort, area of residence, region of residence, and economic sector. In addition to the overall estimates of returns to schooling, another important issue is an issue of heterogeneity in educational returns across individuals. Heterogeneity in individuals' demographic characteristics tends to distort the returns to education; for example, the female rates of return to education is likely to be higher than those of male. Hence, it is worth examining heterogeneous returns to schooling from different demographic characteristics.

Chapter 4 briefly presents a general theoretical framework of the analysis of returns to education. It mainly focuses on the Mincer's (1974) human capital earnings function. Even though Mincer model seems unrealistic and some assumptions can be challenged, the estimation based on Mincer model still shows a positive effect of schooling on earnings (Harmon, et al., 2003) and well fit the actual age-wages profiles (Card, 1999). Chapter 3 Section 1 shows that Mincer model with the OLS estimation can fit the actual recent data fairly well in both developed and developing countries. On the other hand, Heckman, et al. (2006) argues that Mincer model can only be used as a good starting point to develop a more complicated analytical framework. They criticise Mincer model on three main bases. First, there has been no decisive estimated rate of returns to schooling. Second, the estimated coefficient of years of schooling is ambiguous and difficult to make a meaningful economic interpretation. Finally, even a theoretically credible method of estimation such as IV estimation with RDD may possibly produce biased estimates as its instruments are generally found to be weak in many cases. A more comprehensive theoretical and empirical discussion can be found in the work of Heckman, Lochner and Todd (2003). Lastly, this study shall explicitly limit the definition of estimated coefficient of education attainment to a wage premium from education, but still follows the convention by using the term "the rates of return to schooling" to refer to growth of market earnings in terms of schooling.

Chapter 5 provides comprehensive discussion regarding the data, methodology, and identification strategy employed in this study. In general, this study exploits an opportunity of quasi-experiment in Thailand which occurs from a change in compulsory education law to estimate the rates of return to schooling. The main estimation method is the IV estimation and the only source of data used in the estimation is the pooled cross-sectional Labour Force Survey (LFS) data from 1986 to 2012.

Chapter 6 presents the empirical results from IV estimation by using Thai Labour Force Survey data. The results include the estimates from the first stage least square regression, the reduced form, the OLS regression (years of schooling on log wage), and the IV regression. First, the first stage estimation shows that the coefficients of compulsory education are statistically significant and robust across different specifications. The compulsory education law change leads to six additional years of schooling which is roughly corresponding to the first year of the upper secondary school (grade 10). Even though the new compulsory education law increases the minimum year of schooling from four years to six years, it seems that the impact of the law is beyond its expected initial impact. Second, the reduced form finds that the coefficients of compulsory education are statistically significant and robust across different specifications. The change in compulsory education level has a very large effect on the monthly wages. It yields approximately 36 per cent to 50 per cent increase in the monthly wages. The large estimated effect of compulsory education in Thailand can be explained by the fact that there is a wide income gap between individuals with different levels of education attainment. Third, the OLS point estimates of the coefficients of the years of schooling are statistically significant and robust across different specifications. The rates of return to schooling from the OLS estimation is approximately 11 per cent and consistent with that of the previous study from Thailand (Warunsiri and McNown, 2010). Finally, the IV estimation indicates that the coefficients of years of schooling are statistically significant and robust across different specifications. One additional year of schooling leads to approximately 8 per cent increase in monthly wages. Comparing the OLS estimates with those of IV, the result shows that the OLS estimates are greater than those of the IV around 3 per cent. This indicates the upward bias (overestimation) in the OLS regression, which is consistent with the argument of ability bias in returns to schooling.

This main finding of this study is consistent with the theoretical discussion showing that IV estimates should be smaller than those of OLS estimation (Heckman, et al., 2006). Furthermore, the finding also supports Behrman's (1999) arguments regarding the ability bias. Behrman (1990, 1999) argues that there is a substantial large ability bias, especially in the case of developing countries. A number of studies controlling for unobserved ability, family background, and school quality find that the OLS estimates suffer from the ability bias which the impact of years of schooling is overestimated 40-100 per cent.

There are three main concluding remarks from the discussion of the results in terms of methodological perspectives. First, the results from OLS estimation indicate that a net effect of different sources of bias leads to the overestimated rates of return to schooling in the context of Thailand (OLS estimates > IV estimates). This implies that the ability bias outweighs other sources of bias, including the discount rate bias and the measurement error bias. Second, the finding rules out certain issues that may occur in the case of estimates with downward bias (IV estimates > OLS estimates), including further unobserved heterogeneity biases, the classical measurement error bias, and the argument that LATE estimated from IV method is far from ATE. Finally, there still a remaining puzzle as the IV estimates from this study and the estimates from panel with fixed effects (Warunsiri and McNown, 2010) are different not only in terms of magnitude of returns to schooling, but also the direction of the bias, even though these two studies use similar data set with similar statistical quality.

The last section concludes that the dominance of ability bias is mainly explained by the inequality of income and educational opportunity during the early period of social and economic development and there are two sources of ability bias in case of Thailand that includes the ability bias from a selection of more-able child within poor households and the ability bias due to higher financial endowment of rich families. However, the argument of the ability bias due to a selection of more-able child within poor households is more relevant and specific to the case of Thailand and, perhaps, other developing countries. In other words, this type of ability bias is less relevant in developed countries where poverty and inequality of educational opportunity, especially for the basic education, are less pronounced. Even though there is a tendency that individuals with higher ability are more likely to attain more years of schooling, the ability bias seems to be less dominant in developed countries, at least for the basic level of education. This is possibly the main reason that the upward ability bias is observed in case of Thailand and other developing countries, while the OLS estimation from developed countries suffers from downward bias (IV estimates > OLS estimates). Hence, the other sources of bias seem to dominate the effect of the upward ability bias in case of developed countries. Furthermore, explaining the upward ability bias in terms of the inequality of income and educational opportunity during the early period of Thai social and economic development may also possibly be generalised to the case of other developing countries, which share similar social and economic context with Thailand.

In Chapter 7, with disaggregation by different demographic characteristics, the OLS estimates are greater than those of IV; therefore, the upward ability bias still persists even after estimating the rate of returns to schooling with different subsamples. This is consistent with the result from the standard estimation presented in Chapter 6.

First, in terms of gender, the rate of returns to schooling is marginally higher for women than for men. This implies a movement towards gender equality in access to education and also labour market in Thailand. Second, regarding the cohort disaggregation, the rate of returns is also marginally greater in the older than the younger generations. This is explained by different stages of economic development to which each generation belongs. At early stages of development, a rapid economic expansion creates rents or surplus earnings for a small proportion of educated population to enjoy (Psacharopoulos, 2011). Educated workers can enjoy new rents created by the new demand. As a consequence, the expansion of education produces a supply of educated workers to meet a demand for educated workers. Therefore, the rents are eliminated. As the process of economic development deepens, it is also possible that the rate of returns remains constant if the demand keeps pace with the increasing supply. Third, the urban rate of returns exceeds that of rural areas due to availability of more sophisticated jobs that need skilled workers in the urban labour market and higher quality of schools in urban areas. Urban individuals have more chance to utilise their skills acquired from better quality schools than do those living in rural areas (Behrman and Birdsall, 1983; Warunsiri and McNown, 2010). Fourth, in terms of regional disaggregation, the higher returns to schooling in least developed regions, the Northeast and the North regions, can be explained by a regional migration and a movement of demand for and supply of educated labour. The dominant effect of the interregional migration flows from the Northeast and the North regions to Bangkok, the Central, and South regions leads to an increase in the supply of educated labour and a reduction of returns to schooling in those labour-absorbing regions. In contrast, outmigration reduces the supply of educated labour in the Northeast and the North regions; therefore, returns to schooling increase in those regions. Finally, an excess demand for high skilled labour in the early stage of economic development creates rents or education premium in the manufacturing and the service sector to a small group of more educated labour to enjoy. Hence, the rate of returns to schooling is higher in the service and the manufacturing sectors, while the lowest returns to schooling would be expected to observe in the agricultural sector. Moreover, the higher returns to schooling in manufacturing and service sectors can be explained by the relevancy of modern education to job description in those sectors. Workers have more chance to exploit their skills learnt from modern school in more developed sector, while there is less chance to apply the modern knowledge in a sector such as agriculture in which technology is still traditional and backward.

In terms of the bias gap from the one-level disaggregated analysis, even though the magnitude of the difference between male and female bias gap is small, it seems that the ability bias is more dominant for women. In the early stage of social and economic development, market distortions such as gender discrimination prevent women to attend schools. Hence, only women with higher ability and better family background are given an opportunity to attend schools. This is consistent with the ability bias hypothesis which states that more-able individuals are more likely to attain more years of schooling. However, the estimated result also implies that there is a trend towards equal educational opportunity between women and men as the differences of the estimated returns to schooling and the bias gaps between women and men is substantially small. Second, the ability bias is more dominant in the OLS estimates of the older generation. At the early stage of development, poverty and low standard of living influence household's decision making in sending their children to schools. It is likely that most of households may have to select only the most-able children to study at school due to the high direct and indirect costs of schooling and demand for household labour. Third, in terms of area of residence, the estimated result indicates that the ability bias is more dominant in rural areas. Similar results can be observed in disaggregated analysis by region and economic sector. The estimated results show that the ability bias is more dominant in the North East and the North regions, the least developed regions in Thailand, and in the agricultural sector, least developed economic sector both in terms of output growth and technology. The common characteristic is the stage of development of each unit, namely area, region, and economic sector. Less social and economic development implies higher bias gap or higher ability bias in the OLS estimates. A similar line of reasoning explained the ability bias in cohort disaggregated analysis can be applied to disaggregated analyses by area of residence, region, and economic sector. With less educational opportunity and high poverty, households tend to choose only the most-able children to send to school as they cannot bear costs of education for all children in the household. This is consistent with the ability bias hypothesis arguing that more-able individuals tend to have higher education.

In conclusion, with the disaggregated analyses by different demographic characteristics, it is generally observed higher returns to schooling for women, older cohort, urban area, the North and the Northeast regions, and service sector. These results can be explained in the context of Thai development process. In addition, higher bias gap is prevalent in subsamples of women, older cohort, rural area, the North and the Northeast regions, and agricultural sector. Again, income and educational opportunity inequality during the early period of social and economic development help explain the higher bias gap in those socially disadvantaged groups. With less educational opportunity and high poverty, socially disadvantaged households tend to choose only most-able children to send to school since they cannot bear costs of education for all children in the household. This is consistent with the general ability bias hypothesis which argues that more-able individuals tend to have higher schooling.

Moreover, the overall result from the two-level disaggregated analysis supports the results from the base model and the one-level disaggregated analysis. The result indicates that the upward ability bias persists in all analyses. Moreover, with disaggregated analyses by a combination of different demographic characteristics, it generally observes higher returns to schooling for women, older cohort, urban area, the North and the Northeast regions, and service sector. In addition, higher ability bias (bias gap) is dominant in subsamples of women, older cohort, rural area, the North and the Northeast regions, and agricultural sector. This implies that there is educational inequality in access to schooling in those demographic characteristics. Individuals with those characteristics have difficulties to attain education due to less educational opportunity and poverty. It is usually the case that decision regarding child's schooling mainly made by his/her parents. Parents made their decision based on current demographic characteristics of their children and their households. Hence, households with poverty and poor educational opportunity tend to select one or few children with higher ability to study at school as they cannot afford to send all children to school. This is consistent with the ability bias hypothesis which argues that more-able individuals tend to have higher schooling. Lastly, it is worth noting that the more detailed the disaggregated analysis is, the higher the risk of sample selection bias. The estimated results should be interpreted with caution.

Section 2. POLICY IMPLICATIONS

First, according to the results from the first stage regression, the impact of the compulsory education law can go beyond its expected initial impact due to the timing that the compulsory education law was issued. For example, Thai 1978 compulsory education law was issued in the period before the fast-growing economic development, while there was no universal education at that time. The majority of school-age children were still out of school and the compulsory education law forced them to stay in schools. Thus, a large increase in completed years of schooling and also a sharp drop of sample with education at most four years can be observed. Hence, the Government should consider this unexpected effect of the compulsory education. To meet the demand for further schooling, the Government should provide sufficient educational supplies, including teaching staffs, classrooms, and educational budget for the grades specified by the law and grades beyond.

Second, according to the reduced form regression, the results show that the change in compulsory education level has a very large effect on the monthly wages. It implies that education has a positive effect on earnings and empowers individuals to have a better standard of living. Moreover, the compulsory education at the primary level helps reducing income gap. For individuals who have less than six years of education (compulsory education level), an average monthly wage is approximately 5,900 baht, while an average monthly wage of those with secondary education is about 10,000 baht (Labour Force Survey, 1986-2012). The wage of secondary education graduates is at least 40 per cent higher than that of individuals with less than primary education. However, the Government should also pay attention to the marginalised groups that do not have access to the compulsory education. Otherwise, the compulsory education possibly widens the income gap. The ideal situation would be that the Government provides free compulsory education. In addition, it might be the case that proving free and compulsory education in higher levels of education, e.g. secondary education and tertiary education, may further reduce the income gap between the rich and poor.

Some policy implications can be drawn from the disaggregated analysis with different subsamples stratified by gender, area of residence, and economic sector. Third, in terms of gender, the rate of returns to schooling is marginally higher for women than for men. This implies a movement towards gender equality in access to education and also labour market in Thailand. Fourth, the urban rate of returns exceeds that of rural areas due to availability of more sophisticated jobs that need skilled workers in the urban labour market and higher quality of schools in urban areas. Urban individuals have more chance to utilise their skills acquired from better quality schools than do those living in rural areas (Behrman and Birdsall, 1983; Warunsiri and McNown, 2010). Thus, providing better quality schools and more job opportunities, especially those required more skills, in rural areas possibly increase the returns to schooling in those areas. Fifth, in terms of economic sector, the higher returns to schooling in manufacturing and service sectors can be partly explained by the relevancy of modern education to job description in those sectors. Workers have more chance to exploit their skills learnt from modern school in more developed sector, while there is less chance to apply the modern knowledge in a sector such as agriculture in which technology is still traditional and backward. Thus, integrating agricultural education into the curriculum or providing a concrete linkage between the modern knowledge, e.g. science and technology, and agriculture may help students from agricultural households learn how to apply those modern knowledges into their agricultural production and change their household's farming practice. Moreover, informal education could play an important role in further educating farmers after graduating from the compulsory education. As technologies keep developing and moving forward, farmers also need to keep learning and catching up with those new technologies to stay competitive in the market. Sixth, higher bias gap in subsamples of women, older cohort, rural area, the North and the Northeast regions, and agricultural sector implies that there is income and educational opportunity inequality. During the early period of social and economic development, the Government needs to carefully balance between the economic growth and income distribution.

Finally, the overall social and economic conditions of the development in Thailand are consistent with the general characteristics of other developing countries. Hence, estimating the rates of return to schooling in Thailand, by implication, also provides better understandings on the role of human capital in the process of development in other developing countries. Due to the fact that developing countries possess radically different degrees of market completeness and different quality of institution from those of developed countries, this warrants value for investigation of the returns to schooling in the context of developing countries. This further investigation possibly gives a different economic pattern and implications of the returns to schooling.

Section 3. LIMITATIONS AND FUTURE RESEARCH

8.3.1. Limitations

First, the main analysis may encounter the problem of sample selection due to the data limitation. The LFS provides the information only on wages; therefore, non-wage labours are not included in the analysis. Schultz (1988) argues that a bias may occur when sample selection is related to the schooling-wage relationship. Sample selection criteria leading to a bias includes the choice of occupation, labour force participation, or migration, to name a few. Moreover, gender and area of residence may relate to occupation choices and labour participation. Hence, this may further bias our main results.

Second, the contradiction between the IV estimates from this study and the estimates from panel with fixed effects (Warunsiri and McNown, 2010) is puzzling. In general, these two studies use similar data set with similar statistical quality as the OLS estimates from both studies indicate similar magnitude of the estimated coefficients and the level of significance. However, after correcting for the endogeneity bias with different estimation methods, the results are different not only in terms of magnitude of returns to schooling, but also the direction of the bias. As discussed in Chapter 3, the panel regression with fixed effect may suffer from further upward bias since running panel regression with fixed effects excludes time-invariant independent variables from the regression, such as region controls and schooling quality. Stephen Jr. and Yang (2014) find that adding region and school quality control leads the effect of compulsory education on earnings towards zero or even wrong-signed. They argue that without the interaction between region control and year of birth, changes across regions over time of school quality cannot be captured in the estimation. The effects of differences in school quality may cause estimates to be biased since the source of upward bias is not limited to only an individual's ability, but also includes family background and quality of schooling (Schultz, 1988; Behrman, 1990; Strauss and Thomas, 1995).

Finally, the main puzzle is that returns to schooling are higher in the agriculture-dominated North and Northeast than in the regions where non-agricultural opportunities are more widely available[33]. This contradicts that of urban-rural disaggregation. The regional rates of return show that less developed regions provide higher returns, while the rates of return from urban-rural disaggregation indicate that there are higher returns in more developed areas (urban areas). The previous section argues that as there are more sophisticated jobs that need skilled workers in the urban labour market (more developed area),

[33] In fact, the higher average monthly wages are observed in relatively more developed regions, e.g. Bangkok and the Central region, rather than those from less developed regions, e.g., the Northeast and North regions. It should be noted that the concept of rate of returns is rather relative and the rate of returns refer to the rents or education premium created from the interaction of demand for and supply of educated workers.

urban individuals have more chance to utilise their skills than do those living in rural areas (Warunsiri and McNown, 2010). Moreover, the higher quality of schools in urban areas also contributes to the higher returns to schooling (Behrman and Birdsall, 1983). However, the above lines of logic may not apply to the estimates from the regional disaggregation. It is possible that a migration may explain the contradicting patterns between regional returns and area of residence's returns. The migration stream is not simply limited to a rural-urban migration but also includes regional migration. The interaction between these two dimensions of area and region of residence possibly complicates the estimation. The plausible explanation is that the argument used in urban-rural disaggregation is weaker in the case of regional disaggregation due to the coexistence of urban and rural areas within each region. Even though the Northeast and North regions are the least developed regions in Thailand, their urban areas help absorb some of educated labour by providing opportunities for more educated workers to work in more sophisticated jobs.

8.3.2. Future Research

The first priority of the future research should be dealing with those existing limitations, including the problem of sample selection and remaining puzzles. The problem of sample selection can be solved by the data with better quality. On the other hand, the remaining puzzle might be solved by resorting a different analytic model or framework. Introducing interaction terms of different demographic characteristics in the main model may help solving the remaining puzzles in this study.

Second, due to the fact that developing countries possess radically different degrees of market completeness and different quality of institution from those of developed countries, this warrants value for investigation of the returns to schooling in the context of developing countries. Different economic patterns and implications of the returns to schooling may be observed. However, this is an empirical question and it is worth further investigation. The emergence of large-scale microeconomic datasets from developing countries in last decades provides a chance for researchers to examine this issue in other developing countries or in other social and economic settings so that the findings from this study can be compared and contrasted with.

Lastly, there is another interesting phenomenon indicating that there seems to be a lack of increase in average wages as well as returns to education across cohorts for a long period of time, even in the midst of economic development and structural transformation. Suggested further examination would be to conduct wages decomposition by separating age, cohort, and year effects on the wages. This may provide us a better understanding of the mechanism of the constant or changing average wages over time.

Section 4. CONCLUSION

To estimate the rates of return to schooling, this study exploits an opportunity of quasi-experiment in Thailand which occurs from a change in compulsory education law in 1978. The law change affects almost half of the population; therefore, the estimated LATE is possibly closer to the ATE than those of previous literature. The main estimation method is the IV estimation using the pooled cross-sectional Labour Force Survey data from 1986 to 2012.

This study makes three main contributions in terms of methodology and also substantive aspect in the context of Thailand and, by implication, developing countries in general. First, the previous literature reveals that there is a different pattern of the relative magnitudes between OLS and IV estimates using compulsory schooling as IV between developed and developing countries. Investigating this contrast can contribute to a better understanding of (a) how and when the conventional "ability bias" matters in estimating returns to schooling and of (b) the impact of compulsory schooling in different settings. Second, as Thailand experiences rapid economic development and structural transformation during 1960-1990, obtaining the rate of returns to education in this period helps us better understand the process of Thai economic development as well as the interplay between the rate of returns to schooling and the economic development process during 1980 to 1990. This also helps understand the role of human capital in the process of development in other developing countries. Finally, the third contribution is on the construction of the database and discussion of the descriptive analysis for the discrepancies among different demographic characteristics, including gender, cohort, area of residence, region of residence, and economic sector. In addition to the overall estimates of returns to schooling, another important issue is an issue of heterogeneity in educational returns across individuals. Heterogeneity in individuals' demographic characteristics tends to distort the returns to education; for example, the female rates of return to education is likely to be higher than those of male. Hence, it is worth examining heterogeneous returns to schooling from different demographic characteristics.

This study finds that the IV estimation with RDD indicates that the coefficients of years of schooling are statistically significant and robust across different specifications. This confirms a causal relationship between education attainment and earnings. One additional year of schooling leads to approximately 8 per cent increase in monthly wages. Second, the result shows that the OLS estimates are greater than those of IV around 3 per cent. This indicates that the net effect of different sources of endogeneity bias leads to the overestimated rate of returns to schooling in the OLS regression. In the context of Thailand, the ability bias outweighs other endogeneity biases, including the discount rate bias and the measurement error bias. Third, the dominance of ability bias is mainly explained by the inequality of income and educational opportunity during the early period of social and economic development. There are two sources of ability bias, including the ability bias from a selection of more-able child within poor households and the ability bias due to higher financial and better genetic endowment in rich families. This finding may possibly be generalised to the case of other developing countries, which share similar social and economic context with Thailand. Lastly, with the disaggregated analyses by different demographic characteristics, it is generally observed higher returns to schooling for women, older cohort, urban area, the North and the Northeast regions, and service sector. These results can be explained in the context of Thai development process. In addition, higher bias gap is prevalent in subsamples of women, older cohort, rural area, the North and the Northeast regions, and agricultural sector. Again, income and educational opportunity inequality during the early period of social and economic development help explain the higher bias gap in those socially disadvantaged groups. With less educational opportunity and high poverty, socially disadvantaged households tend to choose only most-able children to send to school since they cannot bear costs of education for all children in the household. This is consistent with the general ability bias hypothesis which argues that more-able individuals tend to have higher schooling.

In conclusion, the interplay between the rates of return to schooling and economic development helps us better understand both the role of different stages of economic development in explaining the magnitude and the direction of bias of the estimated returns; and the role of the estimated returns in explaining the development process. This may also help understanding the role of human capital in the process of development in other developing countries. However, this is an empirical question and it is worth further investigation. The emergence of large-scale microeconomic datasets from developing countries in last decades provides a chance for researchers to examine this issue in other developing countries or in other social and economic settings so that the findings from this study can be compared and contrasted with. Furthermore, there is a puzzle indicating that there seems to be a lack of increase in average wages as well as returns to education across cohorts for a long period of time, even in the midst of economic development and structural transformation. Suggested further examination would be to conduct wages decomposition by separating age, cohort, and year effects on the wages. This may provide us a better understanding of the mechanism of the constant or changing average wages over time.

REFERENCES

[1] Acemoglu, Daron, and Joshua Angrist. 2000. "How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws." National Bureau of Economic Research Macroeconomics Annual 15: 9–59.

[2] Angrist, Joshua D, and Alan B Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" JOUR. The Quarterly Journal of Economics 106 (4). Oxford University Press: 979–1014. doi:10.2307/2937954.

[3] Aydemir, Abdurrahman, and Kirdar Murat. 2015. "Low Wage Returns to Schooling in Developing Country: Evidence from a Major Policy Reform in Turkey." Institute for the Study of Labour Discussion Paper 9274.

[4] Aydemir, Abdurrahman, and Kirdar Murat. 2013. "Estimates of the Economic Return to Schooling in a Developing Country: Evidence from a Major Policy Reform in Turkey." Munich Personal RePEc Archive Working Paper 51938.

[5] Behrman, Jere, and T.N. Srinivasan. 1995. "Part 9: Policy Reform, Stabilization,structural Adjustment and Growth." Handbook of Development Economics 3: 2467–96. doi:10.1016/S1573-4471(95)30109-7.

[6] Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2005. "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." JOUR. The American Economic Review 95 (1). American Economic Association: 437–49. http://www.jstor.org/stable/4132690.

[7] Blaug, Mark. 1974. "An Economic Analysis of Personal Earnings in Thailand." Economic Development and Cultural Change 23 (1): 1–31.

[8] Blaug, Mark. 1976. "The Rate of Return on Investment in Education in Thailand." The Journal of Development Studies 12 (2): 270–83.

[9] Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001. "The Determinants of Earnings: A Behavioral Approach." JOUR. Journal of Economic Literature 39 (4): 1137–76. http://www.aeaweb.org/articles?id=10.1257/jel.39.4.1137.

[10] Brunello, Giorgio, and Raffaele Miniaci. 1999. "The Economic Returns to Schooling for Italian Men. An Evaluation Based on Instrumental Variables." Labour Economics 6 (4): 509–19. doi:10.1016/S0927-5371(98)00013-X.

[11] Callan, Tim, and Colm Harmon. 1999. "The Economic Return to Schooling in Ireland." Labour Economics 6 (4): 543–50. doi:10.1016/S0927-5371(99)00035-4.

[12] Card, David. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." Conometrica 69 (5): 1127– 60.

[13] Card, David. 1999. "The Causal Effect of Education on Earnings, in Orley Ashenfelter and David Card, Eds.,." Handbook of Labor Economics 3A: 1801–63.

[14] Card, David. 1994. "Earnings, Schooling, and Ability Revisited." JOUR. National Bureau of Economic Research Working Paper Series No. 4832. doi:10.3386/w4832.

[15] Carneiro, Pedro, James J Heckman, and Edward J Vytlacil. 2011. "Estimating Marginal Returns to Education." JOUR. American Economic Review 101 (6): 2754–81. http://www.aeaweb.org/articles?id=10.1257/aer.101.6.2754.

[16] Chao, John C, and Norman R Swanson. 2005. "Consistent Estimation with a Large Number of Weak Instruments." JOUR. Econometrica 73 (5). [Wiley, Econometric Society]: 1673–92. http://www.jstor.org/stable/3598886.

[17] Christensen, Kaare, Inge Petersen, Axel Skytthe, Anne Maria Herskind, Matt McGue, and Paul Bingley. 2006. "Comparison of Academic Performance of Twins and Singletons in Adolescence: Follow-up Study." JOUR. BMJ: British Medical Journal 333 (7578). BMJ: 1095–97. http://www.jstor.org/stable/40699994.

[18] Clay, Karen, Jeff Lingwall, and Melvin Stephens Jr. 2012. "Do Schooling Laws Matter? Evidence from the Introduction of Compulsory Attendance Laws in the United States." JOUR. National Bureau of Economic Research Working Paper Series No. 18477. doi:10.3386/w18477.

[19] Deaton, Angus S. 2009. "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development." JOUR. National Bureau of Economic Research Working Paper Series No. 14690. doi:10.3386/w14690.

[20] Deere, Donald R., and Jelena Vesovic. 2006. "Chapter 6 Educational Wage Premiums and the U.S. Income Distribution: A Survey." Handbook of the Economics of Education 1: 255–306. doi:10.1016/S1574-0692(06)01006-3.

[21] Denny, Kevin, and Colm Harmon. 2000. "Educational Policy Reform and the Returns to Schooling from Instrumental Variables." IFS Working Paper 00/07.

[22] Devereux, Paul J., and Wen Fan. 2011. "Earnings Returns to the British Education Expansion." Economics of Education Review 30 (6): 1153–66. doi:10.1016/j.econedurev.2011.03.004.

[23] Dickson, Matt. 2013. "The Causal Effect of Education on Wages Revisited." Oxford Bullatin of Economics and Statistics 75 (4): 477–98.

[24] Dougherty, Christopher. 2005. "Why Are the Returns to Schooling Higher for Women than for Men?" JOUR. The Journal of Human Resources 40 (4). [University of Wisconsin Press, Board of Regents of the University of Wisconsin System]: 969–88. http://www.jstor.org/stable/4129547.

[25] Duckworth, Angela L., and Martin E. P. Seligman. 2005. "Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents." Psychological Science 16 (12): 939–44.

[26] Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." JOUR. The American Economic Review 91 (4). American Economic Association: 795–813.

http://www.jstor.org/stable/2677813.

[27] Edwards, Richard C. 1976. “Individual Traits and Organizational Incentives: What Makes a ‘Good’ Worker?” JOUR. The Journal of Human Resources 11 (1). [University of Wisconsin Press, Board of Regents of the University of Wisconsin System]: 51–68. doi:10.2307/145073.

[28] Fang, Hai, Karen N Eggleston, John A Rizzo, Scott Rozelle, and Richard J Zeckhauser. 2012. “The Returns to Education in China: Evidence from the 1986 Compulsory Education Law.” JOUR. National Bureau of Economic Research Working Paper Series No. 18189. doi:10.3386/w18189.

[29] Figlio, D., K. Karbownik, and K.G. Salvanes. 2016. “Chapter 2 – Education Research and Administrative Data.” In Handbook of the Economics of Education, 5:75–138. doi:10.1016/B978-0-444-63459-7.00002-6.

[30] Gomez-Castellanos, Luisa, and George Psacharopoulos. 1990. “Earnings and Education in Ecuador: Evidence from the 1987 Household Survey.” Economics of Education Review 9 (3). Pergamon: 219–27. doi:10.1016/0272-7757(90)90003-N.

[31] Grenet, J. 2013. “Is Extending Compulsory Schooling Alone Enough to Raise Earnings? Evidence from French and British Compulsory Schooling Laws.” Scandinavian Journal of Economics 115 (1): 176–210.

[32] Hanushek, Eric A, and Ludger Woessmann. 2008. “The Role of Cognitive Skills in Economic Development.” JOUR. Journal of Economic Literature 46 (3): 607–68. http://www.aeaweb.org/articles?id=10.1257/jel.46.3.607.

[33] Harmon, Colm, Hessel Oosterbeek, and Ian Walker. 2003. “The Returns to Education: Microeconomics.” Journal of Economic Surveys 17 (2).

[34] Harmon, Colm, and Ian Walker. 1995. “Estimates of the Economic Return to Schooling for the United Kingdom.” JOUR. The American Economic Review 85 (5). American Economic Association: 1278–86. http://www.jstor.org/stable/2950988.

[35] Hawley, Joshua D. 2004. “Changing Returns to Education in Times of Prosperity and Crisis, Thailand 1985–1998.” Economics of Education Review 23 (3): 273–86. doi:10.1016/j.econedurev.2003.11.010.

[36] Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2006. “Chapter 7 Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond.” Handbook of the Economics of Education 1: 307–458. doi:10.1016/S1574-0692(06)01007-5.

[37] Heckman, James J, and Sergio Urzua. 2009. “Comparing IV With Structural Models: What Simple IV Can and Cannot Identify.” JOUR. National Bureau of Economic Research Working Paper Series No. 14706. doi:10.3386/w14706.

[38] Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social

Behavior." JOUR. Journal of Labor Economics 24 (3). The University of Chicago Press: 411–82. doi:10.1086/504455.

[39] Ichino, Andrea, and Rudolf Winter - Ebmer. 2004. "The Long - Run Educational Cost of World War II." JOUR. Journal of Labor Economics 22 (1). The University of Chicago Press: 57-87. doi:10.1086/380403.

[40] Imbens, Guido W. 2010. "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)." JOUR. Journal of Economic Literature 48 (2): 399–423. http://www.aeaweb.org/articles?id=10.1257/jel.48.2.399.

[41] Imbens, Guido W, and Joshua D Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." JOUR. Econometrica 62 (2). [Wiley, Econometric Society]: 467–75. doi:10.2307/2951620.

[42] Kaewkwan tangtipongkul. 2015. "Rates of Return to Schooling in Thailand." Asian Development Review 32 (2): 38–64.

[43] Kugler, Bernardo, and George Psacharopoulos. 1989. "Earnings and Education in Argentina: An Analysis of the 1985 Buenos Aires Household Survey." Economics of Education Review 8 (4). Pergamon: 353–65. doi:10.1016/0272-7757(89)90021-6.

[44] La, Vincent, and Dartmouth College. 2014. "Does Schooling Pay? Evidence from China." Munich Personal RePEc Archive Working Pape 54578.

[45] Lange, Fabian, and Robert Topel. 2006. "Chapter 8 The Social Value of Education and Human Capital." Handbook of the Economics of Education 1: 459–509. doi:10.1016/S1574-0692(06)01008-7.

[46] Levin, Jesse, and Erik J.S Plug. 1999. "Instrumenting Education and the Returns to Schooling in the Netherlands." Labour Economics 6 (4): 521–34. doi:10.1016/S0927-5371(99)00033-0.

[47] Lillard, Lee A. 1977. "Inequality: Earnings vs. Human Wealth." JOUR. The American Economic Review 67 (2). American Economic Association: 42–53. http://www.jstor.org/stable/1807219.

[48] Meghir, Costas, and Mårten Palme. 2005. "Educational Reform, Ability, and Family Background." JOUR. The American Economic Review 95 (1). American Economic Association: 414–24. http://www.jstor.org/stable/4132688.

[49] Meghir, Costas, and Steven Rivkin. 2011. "Chapter 1 – Econometric Methods for Research in Education." In Handbook of the Economics of Education, 3:1–87. doi:10.1016/B978-0-444-53429-3.00001-6.

[50] Moenjak, Thammarak, and Christopher Worswick. 2003. "Vocational Education in Thailand: A Study of Choice and Returns." Economics of Education Review 22 (1): 99–107. doi:10.1016/S0272-7757(01)00059-0.

[51] Nakavachara, Voraprapa. 2010. "Superior Female Education: Explaining the Gender Earnings Gap Trend in Thailand." Journal of Asian Economics 21 (2): 198–218.

doi:10.1016/j.asieco.2009.09.006.
[52] Oded, Galor. 2011. "Chapter 5 – Inequality, Human Capital Formation, and the Process of Development." In Handbook of the Economics of Education, 4:441–93. doi:10.1016/B978-0-444-53444-6.00005-5.
[53] Oosterbeek, Hessel, and Dinand Webbink. 2007. "Wage Effects of an Extra Year of Basic Vocational Education." Economics of Education Review 26 (4): 408–19. doi:10.1016/j.econedurev.2006.07.013.
[54] Oreopoulos, Philip. 2006. "Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter." JOUR. American Economic Review 96 (1): 152–75. http://www.aeaweb.org/articles?id=10.1257/000282806776157641.
[55] Oreopoulos, Philip. 2007. "Do Dropouts Drop out Too Soon? Wealth, Health and Happiness from Compulsory Schooling." Journal of Public Economics 91 (11): 2213–29. doi:10.1016/j.jpubeco.2007.02.002.
[56] Oreopoulos, Philip. 2006. "The Compelling Effects of Compulsory Schooling: Evidence from Canada." JOUR. The Canadian Journal of Economics / Revue Canadienne d'Economique 39 (1). [Wiley, Canadian Economics Association]: 22–52. http://www.jstor.org/stable/3696130.
[57] Ou, Dongshu. 2013. "Education for All: Quasi-Experimental Estimates of the Impacts of Compulsory Primary Education in Hong Kong." JOUR. Asia Pacific Education Review 14 (3): 267–83. doi:10.1007/s12564-013-9271-z.
[58] Parinduri, Rasyad A. 2014. "Do Children Spend Too Much Time in Schools? Evidence from a Longer School Year in Indonesia." Economics of Education Review 41: 89–104. doi:10.1016/j.econedurev.2014.05.001.
[59] Paul Schultz, T. 2004. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program." Journal of Development Economics 74 (1): 199–250. doi:10.1016/j.jdeveco.2003.12.009.
[60] Peracchi, Franco. 2006. "Chapter 5 Educational Wage Premia and the Distribution of Earnings: An International Perspective." Handbook of the Economics of Education 1: 189–254. doi:10.1016/S1574-0692(06)01005-1.
[61] Pischke, J. S., and T von Wachter. 2008. "Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation." The Review of Economics and Statistics 90 (3): 592–98.
[62] Pons, Empar, and Maria Teresa Gonzalo. 2002. "Returns to Schooling in Spain: How Reliable Are Instrumental Variable Estimates?" Labour 16 (4).
[63] Pritchett, Lant. 2006. "Chapter 11 Does Learning to Add up Add up? The Returns to Schooling in Aggregate Data." Handbook of the Economics of Education 1: 635–95. doi:10.1016/S1574-0692(06)01011-7.
[64] Psacharopoulos, G. 1999. "Human Capital and Earnings: British Evidence and a Critique

(1979) BT - Tackling Inequality." CHAP. In , edited by Richard Layard, 126–54. London: Palgrave Macmillan UK. doi:10.1057/9780230375284_8.

[65] Psacharopoulos, George. 1981. "Returns to Education: An Updated International Comparison." JOUR. Comparative Education 17 (3). Taylor & Francis, Ltd.: 321–41. http://www.jstor.org/stable/3098688.

[66] Psacharopoulos, George. 1972. "Rates of Return to Investment in Education around the World." JOUR. Comparative Education Review 16 (1). [University of Chicago Press, Comparative and International Education Society]: 54–67. http://www.jstor.org/stable/1186342.

[67] Psacharopoulos, George. 1989. "Time Trends of the Returns to Education: Cross-National Evidence." Economics of Education Review 8 (3). Pergamon: 225–31. doi:10.1016/0272-7757(82)90002-4.

[68] Psacharopoulos, George. 1983. "Education and Private versus Public Sector Pay." Labour and Society 8 (2): 123–34.

[69] Psacharopoulos, George. 1985. "Returns to Education: A Further International Update and Implications." JOUR. The Journal of Human Resources 20 (4). [University of Wisconsin Press, Board of Regents of the University of Wisconsin System]: 583–604. doi:10.2307/145686.

[70] Psacharopoulos, George. 1988. "Education and Development: A Review." JOUR. The World Bank Research Observer 3 (1). Oxford University Press: 99–116. http://www.jstor.org/stable/3986525.

[71] Psacharopoulos, George, Ana Maria Arriagada, and Eduardo Velez. 1992. "Earnings and Education among Self-Employed Males in Colombia." JOUR. Bulletin of Latin American Research 11 (1). [Wiley, Society for Latin American Studies (SLAS)]: 69–89. doi:10.2307/3338600.

[72] Psacharopoulos, George, and Francis Steier. 1988. "Education and the Labor Market in Venezuela, 1975–1984." Economics of Education Review 7 (3). Pergamon: 321–32. doi:10.1016/0272-7757(88)90004-0.

[73] Rouse, Cecilia Elena. 2000. "[How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws]: Comment." JOUR. NBER Macroeconomics Annual 15. University of Chicago Press: 68–72. doi:10.2307/3585385.

[74] Rubinstein, Yona, and Yoram Weiss. 2006. "Chapter 1 Post Schooling Wage Growth: Investment, Search and Learning." Handbook of the Economics of Education 1: 1–67. doi:10.1016/S1574-0692(06)01001-4.

[75] Stephens Jr., Melvin, and Dou-Yan Yang. 2014. "Compulsory Education and the Benefits of Schooling." JOUR. American Economic Review 104 (6): 1777–92. http://www.aeaweb.org/articles?id=10.1257/aer.104.6.1777.

[76] Vieira, José A.C. 1999. "Returns to Education in Portugal." Labour Economics 6 (4): 535–

41. doi:10.1016/S0927-5371(99)00034-2.

[77] wannakrairoj, wit. 2013. “The Effect of Education and Experience on Wages: The Case Study of Thailand in 2012.” Southeast Asian Journal of Economics 1 (1): 27–48.

[78] Warunsiri, Sasiwimon, and Robert McNown. 2010. “The Returns to Education in Thailand: A Pseudo-Panel Approach.” World Development 38 (11): 1616–25. doi:10.1016/j.worlddev.2010.03.002.

[79] Wendi M. Boxx. 2014. “Addressing Educational Opportunities for Rural Kanjars.” Münchner Beiträge Zur Bildungsforschung 28.

[80] Wolfe, Raymond N., and Scott D. Johnson. 1995. “Personality as a Predictor of College Performance.” Educational and Psychological Measurement 55 (2): 177–85.

[81] Yamauchi, Futoshi. 2005. “Why Do Schooling Returns Differ So Much? Observations and Puzzles from Thailand and the Philippines.” Economic Development and Cultural Change 53 (4): 959–81.

APPENDIX I A COMPREHENSIVE SUMMARY OF LITERATURE: IV ESTIMATION AND RETURNS TO SCHOOLING

#	Author	Country	Sample	Method of Estimation and Instrument		Schooling Coefficients	
						OLS	IV/DD
1.	Angrist and Krueger (1991)	The United States	1970 and 1980 Census Data, Men.	Instruments are quarter of birth interacted with year of birth. Controls include quadratic in age and indicators for race, marital status, and urban residence	1920-1929 cohort in 1970 1930-1939 cohort in 1980 1940-1949 cohort in 1980	0.070 (0.000) 0.063 (0.000) 0.052 (0.000)	0.101 (0.033) 0.060 (0.030) 0.078 (0.030)
2.	Harmon and Walker (1995)	The United Kingdom	1978 - 1986 data of the British Family Expenditure Survey, Men.	Dependent variable is log earnings. Instruments are a dummy variable for the raising of the school-leaving age from 14 to 15 in 1947, and a dummy variable for the raising of the school-leaving age from 15 to 16 in 1973. Controls are a quadratic in age, area dummies, and year dummies. (*): s.e. No report on statistical significance.		0.061 (0.001)	0.153 (0.015)
3.	Levin and Plug (1999)	The Netherlands	1952 and 1983 data of the Brabant Survey, Men. 1994 data of the OSA panel Survey	Dependent variable is log of net hourly wages. Instruments are quarter of birth interacted with year of birth and change in the minimum school leaving age (MSLA) Controls include IQ, quadratic experience, quadratic age, and marital status. Season of birth is rejected by Sargan's	Brabant Survey OSA panel Survey Season of birth MSLA	0.029 *** (0.004) 0.036 *** (0.002) na na	na na 0.040 *** (0.029) 0.064 **

				identification test. MSLA is not statistically significant in Hausman's *t*-test on exogeneity. (*): s.e.			(0.027)
4.	Callan and Harmon (1999)	Ireland	1987 Survey of Income Distribution, Poverty and Usage of State Services, Men aged 18-64.	Dependent variable is log hourly wages. Instruments are a dummy variable for the raising of the school-leaving age from 14 to 15 in 1972, and a dummy variable for the introduction of free secondary education in the mid-1960s Controls include quadratic age, geographic dummies, marital status, and the occupation-specific unemployment rate. (*): s.e. No report on statistical significance. The instruments are not valid See Bound et al. (1995) Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak	na	na	na
5.	Brunello and Miniaci (1999)	Italy	1993 and 1995 data of the Bank of Italy (BI) survey on the income and wealth of Italian households, Male household heads	Dependent variable is log real hourly wages. Instrument is a dummy indicating whether individuals born since 1951 which is expected to pick up the effects of the 1969 expansion of the higher education enrolment opportunity. Controls include age, areas, town sizes and year. No report on standard error.	na	0.048*** (na)	0.057*** (na)

6.	Vieira (1999)	Portugal	1986 and 1992 data of Quados de Pessoal , Male workers.	Dependent variable is log hourly wages Instruments are changes in the compulsory level of education in 1956 and 1964. Controls include quadratic age and region dummies. (*): *t*-value	Compulsory education 1956 Compulsory education 1964	0.0779* (160)	0.0150 (1.16) 0.0534* (3.90)
7.	Acemoglu and Angrist (2000)	The United States	1950-1990 Census data, White men.	Instruments are a) dummies interacting quarter of birth and year of birth (QOB); b) dummies indicating state- and year-specific child labour laws (SOB-CL); and c) dummies indicating state- and year-specific compulsory attendance laws (SOC-CA). Controls include state-of-residence effects, year effects, year-of-birth effects, and state-of-birth effects. The results are only from specification with dummies indicating state- and year-specific compulsory attendance laws (SOC-CA). No report on statistical significance.	1960-1980 1950-1980 1950-1990 1950 1960 1970 1980 1990	0.073 (0.0003) 0.068 (0.0003) 0.075 (0.0003) 0.055 (0.002) 0.069 (0.001) 0.076 (0.001) 0.075 (0.001) 0.102 (0.001)	0.092 (0.044) 0.099 (0.052) 0.081 (0.023) na na na na na
8.	Denny and Harmon (2000)	Ireland	1987 data of the household survey by the Economic and Social Research Institute , Men.	**Natural experiment** Dependent variable is log gross hourly wages. Instrument is a dummy variable for the introduction of free secondary education in the mid-1960s Controls include quadratic age, trend in secondary participation, urban, marital status, union member, and parental class.		0.0793 (0.1224)	0.1360 (0.0251)

				(*): s.e. No report on statistical significance.			
9.	Pons and Teresa Gonzalo (2002)	Spain	1991 data of the Survey of Structure, Conscience and Biography of Class (ECBC) and 1994 data of the Household Panel of the European Union (PHOGUE), Male workers.	Dependent variable is log of the net hourly wages. Instruments are season of birth and change in the minimum school leaving age (MSLA) in 1970. Controls include quadratic age, tenure, and dummies of geographical areas. (*): robust s.e.	ECBC PHOGUE	0.064 ** (0.004) 0.059 ** (0.002)	-0.015 (0.089) -0.016 (0.041)
10	Meghir and Palme (2005)	Sweden	1961 and 1966 data of the Individual Statistics (IS) project of the Institute for Education at the University of Gothenburg, Men and women.	**Differences-in-differences** Dependent variable is average annual log-pre-tax earnings Instruments are cohort and municipality assigned to the reformed system. Controls include cohort, a set of dummy variables indicating the municipality, time dummies, and observable characteristics of individual students. (*): s.e. No report on statistical significance.	1948 and 1953 cohorts	na	0.0142 (0.89)
11	Oreopoulos (2006)	Canada	1971, 1981, 1986, 1991, 1996, and 2001 data of the Censuses, Men and women.	**Regression Discontinuity** Dependent variables are log annual income and earnings. Instruments are dummy variables for the raising of the school-leaving age. Controls are birth cohort, province, census year, and provincial cohorts. (*): Huber-White s.e.	Log income Log earnings	0.127 *** (0.004) 0.115 *** (0.007)	0.121 *** (0.013) 0.070 *** (0.008)
12	Oreopoulos (2006, 2008)	The United Kingdom,	1983 - 1998 data of the UK General Household	**Regression Discontinuity** Dependent variable is log annual earnings. Log	The United States	0.078 *** (0.0005)	0.142 *** (0.0119)

		The United States, and Canada	Surveys (GHS); 1985 - 1998 data of the Northern Ireland Continuous Household Surveys; 1950-2000 data of US Censuses; and 1971-2001 data of Canada Censuses, Men and women.	weekly earnings for the United States. Instruments are dummy variables for the raising of the school-leaving age from 14 to 15 in 1947 (UK) and in 1957 (Northern Ireland). Controls are birth year, region, survey year, sex, and a quartic in age. The US and Canada results also include a dummy variable for race, and state/provincial controls. (*): s.e.	Canada The United Kingdom Britain	0.099 *** (0.0007) 0.085 *** (0.002) 0.083 *** (0.003)	0.096 *** (0.0254) 0.108 *** (0.0328) 0.101 ** (0.0421)
13	Oreopoulos (2007)	The United Kingdom, The United States, and Canada	1983 - 1998 data of the UK General Household Surveys (GHS); 1985 - 1998 data of the Northern Ireland Continuous Household Surveys; 1950-2000 data of US Censuses; and 1971-2001 data of Canada Censuses, Men and women.	**Regression Discontinuity** Dependent variable is log annual earnings. Log weekly earnings for the United States. Instruments are dummy variables for the raising of the school-leaving age from 14 to 15 in 1947 (UK) and in 1957 (Northern Ireland). Controls are birth year, region, survey year, sex, and a quartic in age. The US and Canada results also include a dummy variable for race, and state/provincial controls. (*): Huber-White s.e.	The United States Canada The United Kingdom	0.133 *** (0.0118) 0.088 *** (0.0027) 0.078 *** (0.0024)	0.142 *** (0.0119) 0.084 *** (0.0267) 0.158 *** (0.0491)
14	Oosterbeek and Webbink (2007)	The Netherlands	1995 data of the Wage Structure Survey (LSO), Men.	**Differences-in-differences.** Dependent variable is log hourly wages. Instrument is a change in the minimum school leaving age (MSLA) in 1975. Controls are quadratic age.	1953-1963	na	-0.004 (0.04)
15	Pischeke and Wachter (2008)	Germany	1979, 1985-86, 1991-92, and 1998-99 data of the	**Differences-in-differences** Dependent variable is log hourly wages. Instrument is an introduction of the ninth	Qualification and Career Survey	0.066 (0.002)	0.058 (0.038)

			Qualification and Career Survey (QaC) and 1989, 1991, 1993, and 1995-2004 data of the Micro Census, Men and Women	grade since 1949. Controls are gender, a quartic in age, year, state of residence, and year of birth. (*): s.e. No report on statistical significance.	Micro Census	0.074 (0.001)	0.016 (0.015)
16	Devereux and Hart (2010)	The United Kingdom	1979 - 1998 data of the General Household Survey (GHS) and 1975 - 2001 data of the New Earnings Survey Panel Dataset (NESPD), Men and Women.	**Regression Discontinuity** Dependent variable is log weekly earnings. Instrument is a raising of the minimum school leaving age in 1947. Controls are age, gender, and year-of-birth. (*): robust s.e. All coefficients are not statistically significant.	No age control Quartic age control Age dummies control	na na na	0.021 (0.024) 0.019 (0.023) 0.025 (0.027)
17	Devereux, and Fan (2011)	The United Kingdom	1997 - 2009 data of the Quarterly Labour Force Survey (QLFS), Men and women.	**Regression Discontinuity (closely related)** Dependent variables are log hourly wages and log weekly earnings. Instrument is a dummy variable for the education expansion during 1989 and 1994. Controls are a quartic in year-of-birth, a quartic in age, white, the quarterly unemployment rate, and cohort size. (*): robust s.e.	Men Hourly wages Weekly earnings Women Hourly wages Weekly earnings	 0.078 *** (0.003) 0.069 *** (0.003) 0.096 *** (0.003) 0.122 *** (0.002)	 0.062 *** (0.016) 0.066 *** (0.019) 0.053 *** (0.014) 0.066 *** (0.013)
18	Clay, Lingwall, and Stephens Jr. (2012)	The United States	1940 data of the Census, Men.	**Following Acemoglu and Angrist (2000),.** Dependent variable is log weekly wages. Instruments are a compulsory attendance laws. Controls are a quadratic state time trend, race, nativity, and native English speaker.		0.080 *** (0.003)	0.114 (0.098)

				(*): s.e.			
19	Fang et al. (2012)	China	1997, 2000, 2004, and 2006 data of the China Health and Nutrition Survey (CHNS), Men and Women. CHNS contains data from 9 of China's provinces. These 9 provinces account for 44% of China's total population.	**Natural experiment** Dependent variable is log annual earnings. Earnings include annual wage, income from businesses, faring, gardening, fishing, livestock, and retirement benefits. Instrument is a compulsory education law 1986. Controls are a quadratic age, gender, ethnic status, marital status, urban residency, province, self-reported health status, and survey wave. (*): s.e.		0.09 *** (0.004)	0.02 *** (0.06)
20	Grenet (2013)	France and the United Kingdom	1990 - 2002 data of the French Enquête Emploi (EE) survey (France) and 1993 - 2006 data of the Quarterly Labour Force Survey (QLFS) (UK), Men and women.	**Regression Discontinuity** Dependent variable is log hourly wages. Instruments are a dummy variable for the raising of the school-leaving age from 14 to 16 in 1967 (France), and a dummy variable for the raising of the school-leaving age from 14 to 15 in 1972 (UK). Controls are a quartic in age, a full set of survey year dummies, and a fourth-order polynomial in school cohort (in 2SLS). (*): s.e.	France Men Women The UK Men Women	0.073 *** (0.001) 0.087 *** (0.001) 0.095 *** (0.001) 0.119 *** (0.001)	-0.004 (0.029) -0.007 (0.013) 0.069 *** (0.029) 0.067 *** (0.011)
21	Dickson (2013)	The United Kingdom	1991-2005 data of the British Household Panel Survey, Men.	**Regression Discontinuity** Dependent variable is log hourly wages. Instrument is a raising of the minimum school leaving age (RoSLA) in 1972. Controls are quadratic age, quadratic		0.046 *** (0.003)	0.102 ** (0.051)

				year-of-birth, region dummies, ethnicity dummies, survey period dummies, dummy for lived with both natural parents from birth to age 16, each parents' occupational class dummies. (*): robust s.e.			
22	Aydemir and Murat (2013)	Turkey	2000 - 2010 data of the Turkish Income and Expenditure Surveys (TIES), Men.	**Regression Discontinuity** Dependent variable is log hourly wages. Instrument is the variation in the years of compulsory schooling across different birth Cohorts. Controls are year of birth cohort trend, age, year, relation to household head, marital status, type of employment, occupation, and sector of employment. (*): robust s.e.	Linear time trend for year of birth Quadratic time trend for year of birth	0.0304 *** (0.00395) 0.0304 *** (0.00397)	0.301 *** (0.0814) 0.158 *** (0.0300)
23	Parinduri (2014)	Indonesia	2007 data of the Indonesia Family Life Survey (IFLS), Men and Women.	**Fuzzy Regression Discontinuity** Dependent variable is log hourly and monthly wages.. Instrument is an arbitrary rule that assigned students to a longer school year in 1978-1979. Controls are gender, and age. (*): robust s.e.	Log hourly wages Log monthly wages	Na na	0.13 *** (0.04) 0.17 *** (0.04)
24	Stephen Jr. and Yang (2014)	The United States	1960-1980 data of the US Censuses of Population, native-born Whites of 1905 – 1954 birth cohorts.	Dependent variable is log weekly wages. Instruments are schooling laws. Controls are quartic age, year of birth, state of birth, region, gender, census year and quality of schooling. The second coefficients result from the inclusion of region and year of birth	White males ages 40-49 White males	0.073 / 0.073 (0.0005) / (0.0005) 0.063 /	0.095 / -0.020 [0.064 , 0.126] / [-0.163 , 0.060] 0.097 /

				interaction term. [*] is the confidence intervals based on Moreira's CLR test	ages 25-54 All white ages 25-54 White ages 25-54 born Non-south White ages 25-54 born South	0.063 (0.0004) / (0.0004) 0.068 / 0.068 (0.0003) / (0.0003) 0.067 (0.0004) 0.069 (0.0004)	-0.014 [0.080 , 0.117] / [-0.066 , 0.021] 0.105 / -0.003 [0.083 , 0.123] / [-0.058 , 0.016] -0.009 [-0.031 , 0.001] 0.019 [-0.097 , 0.085]
25	La and College (2014)	China	1997, 2000, 2004, and 2006 data of the China Health and Nutrition Survey (CHNS), Men and Women. CHNS contains data from 9 of China's provinces. These 9 provinces account for 44% of China's total population.	**Natural experiment** Dependent variable is log annual earnings. Earnings include annual wage, income from businesses, faring, gardening, fishing, livestock, and retirement benefits. Instrument is a compulsory education law 1986. Controls are a quadratic age, gender, ethnic status, marital status, urban residency, province, and self-reported health status. (*): robust s.e.		0.0842 *** (0.00369)	0.0585 (0.0684)
26	Aydemir and Murat	Turkey	2002 - 2013 data of the Turkish Household	**Regression Discontinuity** Dependent variable is log hourly wages.	Men	0.050 *** (0.002)	0.005 (0.014)

	(2015)		Labour Force Surveys (HLFS), Men and Women.	Instrument is the extension of compulsory schooling from 5 to 8 years in 1997. Controls are year of birth, age, and urban status. (*): robust s.e.	Women	0.063 *** (0.003)	0.038 ** (0.015)

Notes: na – not available

Source: Author's compilation based on own observation, Card (1999), and Harmon, Oosterbeek, and Walker (2003)

APPENDIX II SAMPLE DESIGN (TAKEN FROM LABOUR FORCE SURVEY OFFICIAL REPORT 2012)

A Stratified Two - Stage Sampling was adopted for the survey. Provinces were constituted strata. The primary and secondary sampling units were enumeration areas (EAs) for municipal areas and non - municipal areas and private households / persons in the collective households respectively.

Stratification: Provinces were constituted strata. There were altogether 77 strata. Each stratum was divided into two parts according to the type of local administration, namely municipal areas and non - municipal areas.

Selection of primary sampling unit: The sample selection of enumeration areas was performed separately and independently in each part by using probability proportional to size - total number of households. The total sample enumeration areas were 5,970 from 127,460 EAs. The total number of sample enumeration areas selected for enumeration by region and type of local administration was as follows:

Selection of primary sampling unit (2012)

Region/Stratum	Total	Municipal Areas	Non-municipal Areas
Bangkok Metropolis	300	300	-
Central (Excluding Bangkok Metropolis)	1,902	900	1,002
North	1,278	630	648
Northeast	1,476	732	744
South	1,014	498	516
Total	5,970	3,060	2,910

Selection of secondary sampling unit: Private households were our ultimate sampling units. A new listing of private households was made for every sample enumeration areas to serve as the sampling frame. In each sample EAs, a systematic sample of private households were selected with the following sample size: Municipal areas: 16 sample households per EAs; Non - municipal areas: 12 sample households per EAs.

Before selecting sample private households in each sample EAs, the list of private households was rearranged by household' s size - member of the households. All collective households located within the sample areas were included in the sample and the persons in the collective household were systematically selected for the interviewing.

The total number of sample private households selected for enumeration by region and type of local administration was as follows:

Selection of secondary sampling unit (2012)

Region/Stratum	Total	Municipal Areas	Non-municipal Areas
Bangkok Metropolis	4,800	4,800	-
Central (Excluding Bangkok Metropolis)	26,424	14,400	12,024
North	17,856	10,080	7,776
Northeast	20,640	11,712	8,928
South	14,160	7,968	6,192
Total	83,880	48,960	34,920

APPENDIX III THE IDENTIFICATION FOR FUZZY REGRESSION DISCONTINUITY

The identification for fuzzy regression discontinuity is as following:

In the case of fuzzy regression discontinuity, treatment status or (actual) treatment indicator, D_i, is *not* a deterministic and discontinuous function of a covariate x_i. Instead, it utilizes discontinuities in the probability of treatment conditional on a covariate. In other words, treatment depends on whether $x_i \geq x_0$ or $x_i < x_0$ *and* endogenous factors:

$$P[D_i = 1|x_i] = \begin{cases} g_1(x_i), & x_i \geq x_0 \\ g_0(x_i), & x_i < x_0 \end{cases}, where\ \ g_1(x_i) \neq g_0(x_i).$$ [34]

$$\mathrm{E}[D_i|x_i] = P[D_i = 1|x_i] = g_0(x_i) + [g_1(x_i) - g_0(x_i)]T_i, where\ T_i = 1(x_i \geq x_0)$$

To elaborate:

If $x_i \geq x_0\ then\ T_i = 1$ then $P[D_i = 1|x_i] = g_1(X_i)$

If $x_i < x_0\ then\ T_i = 0$ then $P[D_i = 1|x_i] = g_0(X_i)$

The dummy variable, T_i, indicates the point where there is a discontinuity in $\mathrm{E}[D_i|x_i]$. Under the assumption of that $g_0(x_i)$ and $g_1(x_i)$ are *p*th-order polynomial functions, the fuzzy regression discontinuity can be mathematically illustrated as a simple Two-Stage least squares estimation (2SLS):

$$\mathrm{E}[D_i|x_i] = \gamma_{00} + \gamma_{01}x_i + \gamma_{02}x^2{}_i + \cdots + \gamma_{0p}X_i^p + [\pi + \gamma_1^* X_i + \gamma_2^* X_i^2 + \cdots + \gamma_p^* X_i^p]T_i$$

Where:

D_i = 1 if observation *i* receives treatment and,

0 if observation *i* does not receive treatment.

x_0 = the cut-off point.

With IV estimation, the first stage is as following:

$$\mathrm{E}[D_i|X_i] = \gamma_{00} + \gamma_{01}X_i + \gamma_{02}X_i^2 + \cdots + \gamma_{0p}X_i^p + \pi T_i + \gamma_1^* X_i T_i + \gamma_2^* X_i^2 T_i + \cdots + \gamma_p^* X_i^p T_i$$

[34] Most of mathematic expressions and equations are directly quoted from the book named, "Mostly Harmless Econometrics: An Empiricist's Companion" written by Angrist and Pischke (2009)

$$D_i = \gamma_0 + \gamma_1 I(x_i > x_0) + \eta_i$$

The second stage is as following:

$$y_i = \beta_0 + \beta_1 \widehat{D}_i + \delta(x_i) + \varepsilon_i$$

According to Hahn, Todd, and van der Klaauw (2011), fuzzy regression discontinuity estimates local average treatment effect (LATE), or the average treatment effect of the compliers, given that the standard assumptions of instrumental variable framework is met.

APPENDIX IV FULL EMPIRICAL RESULTS

First Stage: Estimated Effect of Compulsory Education Law on Education Attainment, Thailand, Ages 15-60, 1986-2012

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					(First Stage)				
					Dependent Variable: Number of Years of Schooling				
Compulsory Education	4.358***	6.385***	6.440***	6.395***	6.443***	6.413***	6.462***	6.057***	6.104***
	(0.713)	(0.0868)	(0.0884)	(0.105)	(0.106)	(0.105)	(0.105)	(0.0837)	(0.0834)
Fixed Effects:									
Region	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of Birth (Cohort)	Quartic	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region x Year of Birth (Cohort)	No	No	Yes	No	Yes	No	Yes	No	Yes
Additional Controls	None	None	None	Age Quadratic	Age Quadratic	Age Quadratic Gender	Age Quadratic Gender	Age Quadratic Gender Urban	Age Quadratic Gender Urban
Observations	1,307,988	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016
R-squared	0.091	0.123	0.125	0.137	0.139	0.138	0.140	0.192	0.194

Note: The dependent variables are number of years of schooling. Each regression includes controls for a birth cohort quartic polynomial or birth cohort dummies, regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (six years of compulsory education). Column (4) to (9) also include age controls: a quadratic polynomial and fixed effects where indicated. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. Robust standard errors in parentheses. ***, **, and * indicate p < 0.01, p < 0.05, and p < 0.1, respectively.

Source: Author's compilation based on LFS 1986-2012.

Reduced Form: Estimated Effect of Compulsory Education Law on Log Annual Earnings, Thailand, Ages 15-60, 1986-2012

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	(Reduced Form) Dependent Variable: Log Monthly Wages								
Compulsory Education	0.355***	0.523***	0.529***	0.547***	0.552***	0.537***	0.541***	0.476***	0.481***
	(0.0705)	(0.0326)	(0.0325)	(0.0291)	(0.0285)	(0.0298)	(0.0292)	(0.0255)	(0.0247)
Fixed Effects:									
Region	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of Birth (Cohort)	Quartic	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region x Year of Birth (Cohort)	No	No	Yes	No	Yes	No	Yes	No	Yes
Additional Controls	None	None	None	Age Quadratic	Age Quadratic	Age Quadratic Gender	Age Quadratic Gender	Age Quadratic Gender Urban	Age Quadratic Gender Urban
Observations	1,307,988	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016
R-squared	0.017	0.088	0.090	0.125	0.127	0.133	0.135	0.199	0.201

Note: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort quartic polynomial or birth cohort dummies, regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (six years of compulsory education). Column (4) to (9) also include age controls: a quadratic polynomial and fixed effects where indicated. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. Robust standard errors in parentheses. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively.

Source: Author’s compilation based on LFS 1986-2012.

OLS Returns to Schooling Estimates for Log Monthly Wages, 15-60 Years Old, 1986-2012

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					(OLS)				
	Dependent Variable: Log Monthly Wages								
Years of Schooling	0.113*** (0.00366)	0.111*** (0.00378)	0.111*** (0.00378)	0.112*** (0.00425)	0.112*** (0.00426)	0.112*** (0.00412)	0.112*** (0.00412)	0.109*** (0.00389)	0.109*** (0.00390)
Fixed Effects:									
Region	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of Birth (Cohort)	Quartic	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region x Year of Birth (Cohort)	No	No	Yes	No	Yes	No	Yes	No	Yes
Additional Controls	None	None	None	Age Quadratic	Age Quadratic	Age Quadratic Gender	Age Quadratic Gender	Age Quadratic Gender Urban	Age Quadratic Gender Urban
Observations	1,307,988	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016
R-squared	0.528	0.567	0.569	0.602	0.603	0.614	0.614	0.621	0.622

Note: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort quartic polynomial or birth cohort dummies, regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (six years of compulsory education). Column (4) to (9) also include age controls: a quadratic polynomial and fixed effects where indicated. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. Robust standard errors in parentheses. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively.

Source: Author's compilation based on LFS 1986-2012.

IV Returns to Schooling Estimates for Log Monthly Wages, 15-60 Years Old, 1986-2012

	(1)	(2)	(3)	(4)	(5) (IV)	(6)	(7)	(8)	(9)
	Dependent Variable: Log Monthly Wages								
Years of Schooling	0.0813***	0.0819***	0.0821***	0.0856***	0.0856***	0.0837***	0.0838***	0.0787***	0.0787***
	(0.00548)	(0.00424)	(0.00420)	(0.00354)	(0.00345)	(0.00363)	(0.00355)	(0.00352)	(0.00342)
Fixed Effects:									
Region	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of Birth (Cohort)	Quartic	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region x Year of Birth (Cohort)	No	No	Yes	No	Yes	No	Yes	No	Yes
Additional Controls	None	None	None	Age Quadratic	Age Quadratic	Age Quadratic Gender	Age Quadratic Gender	Age Quadratic Gender Urban	Age Quadratic Gender Urban
Observations	1,307,988	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016	1,307,016
R-squared	0.486	0.533	0.535	0.575	0.576	0.581	0.582	0.588	0.589

Note: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort quartic polynomial or birth cohort dummies, regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (six years of compulsory education). Column (4) to (9) also include age controls: a quadratic polynomial and fixed effects where indicated. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. Robust standard errors in parentheses. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively.

Source: Author's compilation based on LFS 1986-2012.

APPENDIX V FULL ESTIMATES FROM TWO-LEVEL DISAGGREGATED ANALYSIS

1. Gender

Dependent Variables	OLS	IV	Bias gap	Observations
Base model				
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Log monthly wages,	0.108***	0.0808***	0.0272	662,753
male	(0.00439)	(0.00374)		
Log monthly wages,	0.116***	0.0856***	0.0304	644,263
female	(0.00391)	(0.00367)		
Gender and Cohort				
Log monthly wages, male	0.122***	0.0827***	0.0393	413,455
cohort 1955-1970	(0.00340)	(0.00337)		
Log monthly wages, male	0.0965***	0.0807***	0.0158	512,457
cohort 1961-1985	(0.00410)	(0.00376)		
Log monthly wages, female	0.128***	0.0870***	0.041	399,067
cohort 1955-1970	(0.00359)	(0.00386)		
Log monthly wages, female	0.106***	0.0856***	0.0204	504,292
cohort 1961-1985	(0.00323)	(0.00367)		
Gender and Area				
Log monthly wages, urban male	0.102***	0.0776***	0.0244	424,146
	(0.00354)	(0.00192)		
Log monthly wages, rural male	0.101***	0.0687***	0.0323	238,607
	(0.00581)	(0.00657)		
Log monthly wages, urban female	0.113***	0.0871***	0.0259	432,471
	(0.00360)	(0.00210)		
Log monthly wages, rural female	0.106***	0.0668***	0.0392	211,792
	(0.00440)	(0.00613)		
Gender and Region				
Log monthly wages, male	0.0945***	0.0656***	0.0289	82,071
Bangkok and Metropolitan	(0.00214)	(0.00261)		
Log monthly wages, male	0.121***	0.101***	0.02	126,355
Northern area	(0.00458)	(0.00604)		

Log monthly wages, male	0.131***	0.0871***	0.0439	151,593
North eastern area	(0.00401)	(0.00425)		
Log monthly wages, male	0.0862***	0.0659***	0.0203	116,980
Southern area	(0.00503)	(0.00500)		
Log monthly wages, male	0.0959***	0.0750***	0.0209	186,181
Central area	(0.00462)	(0.00324)		
Log monthly wages, female	0.0968***	0.0796***	0.0172	80,328
Bangkok and Metropolitan	(0.00195)	(0.00118)		
Log monthly wages, female	0.125***	0.0938***	0.0312	129,881
Northern area	(0.00414)	(0.00459)		
Log monthly wages, female	0.148***	0.101***	0.047	146,544
North eastern area	(0.00366)	(0.0101)		
Log monthly wages, female	0.0998***	0.0774***	0.0224	105,201
Southern area	(0.00422)	(0.00273)		
Log monthly wages, female	0.0981***	0.0738***	0.0243	182,854
Central area	(0.00435)	(0.00232)		
Gender and Economic Sector				
Log monthly wages, male	0.0944***	0.0596***	0.0348	230,334
agricultural and elementary workers	(0.00488)	(0.00381)		
Log monthly wages, male	0.0990***	0.0687***	0.0303	138,554
non-agricultural manual workers	(0.00522)	(0.00284)		
Log monthly wages, male	0.0903***	0.0709***	0.0194	293,022
desk, service, and intellectual workers	(0.00290)	(0.00126)		
Log monthly wages, female	0.104***	0.0476***	0.0564	198,365
agricultural and elementary workers	(0.00437)	(0.00514)		
Log monthly wages, female	0.0831***	0.0675***	0.0156	99,720
non-agricultural manual workers	(0.00335)	(0.00286)		
Log monthly wages, female	0.109***	0.0850***	0.024	344,083
desk, service, and intellectual workers	(0.00329)	(0.00286)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author's compilation based on LFS 1986-2012.

2. Area of Residence

Dependent Variables	OLS	IV	Bias gap	Observations
Base model				
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Log monthly wages,	0.108***	0.0849***	0.0231	856,617
urban	(0.00369)	(0.00169)		
Log monthly wages,	0.105***	0.0709***	0.0341	450,399
rural	(0.0053)	(0.00622)		
Area and Gender				
Log monthly wages, urban male	0.102***	0.0776***	0.0244	424,146
	(0.00354)	(0.00192)		
Log monthly wages, rural male	0.101***	0.0687***	0.0323	238,607
	(0.00581)	(0.00657)		
Log monthly wages, urban female	0.113***	0.0871***	0.0259	432,471
	(0.00360)	(0.00210)		
Log monthly wages, rural female	0.106***	0.0668***	0.0392	211,792
	(0.00440)	(0.00613)		
Area and Cohort				
Log monthly wages, urban	0.120***	0.0858***	0.0342	537,768
cohort 1955-1970	(0.00292)	(0.00170)		
Log monthly wages, urban	0.0984***	0.0848***	0.0136	663,259
cohort 1961-1985	(0.00333)	(0.00175)		
Log monthly wages, rural	0.123***	0.0718***	0.0512	274,754
cohort 1955-1970	(0.00506)***	(0.00640)		
Log monthly wages, rural	0.0927***	0.0712***	0.0215	353,490
cohort 1961-1985	(0.00448)	(0.00609)		
Area and Region				
Log monthly wages, urban	0.0962***	0.0758***	0.0204	143,357
Bangkok and Metropolitan	(0.00198)	(0.00143)		
Log monthly wages, urban	0.121***	0.0997***	0.0213	162,960
Northern area	(0.00394)	(0.00235)		

Log monthly wages, urban	0.129***	0.0965***	0.0325	192,139
North eastern area	(0.00331)	(0.00336)		
Log monthly wages, urban	0.0912***	0.0726***	0.0186	133,456
Southern area	(0.00392)	(0.00245)		
Log monthly wages, urban	0.0957***	0.0773***	0.0184	225,565
Central area	(0.00419)	(0.00103)		
Log monthly wages, rural	0.0876***	0.0680***	0.0196	19,042
Bangkok and Metropolitan	(0.00291)	(0.00237)		
Log monthly wages, rural	0.115***	0.0818***	0.0332	93,276
Northern area	(0.00509)	(0.00766)		
Log monthly wages, rural	0.142***	0.0645***	0.0775	105,998
North eastern area	(0.00563)	(0.0143)		
Log monthly wages, rural	0.0838***	0.0660***	0.0178	88,725
Southern area	(0.00554)	(0.00481)		
Log monthly wages, rural	0.0938***	0.0695***	0.0243	143,470
Central area	(0.00509)	(0.00435)		
Area and Industry				
Log monthly wages, urban	0.109***	0.0708***	0.0382	171,004
agricultural and elementary workers	(0.00482)	(0.00346)		
Log monthly wages, urban	0.0971***	0.0750***	0.0221	158,304
non-agricultural manual workers	(0.00441)	(0.00139)		
Log monthly wages, urban	0.103***	0.0834***	0.0196	524,584
desk, service, and intellectual workers	(0.00305)	(0.000594)		
Log monthly wages, rural	0.0776***	0.0435***	0.0341	257,695
agricultural and elementary workers	(0.00418)	(0.00545)		
Log monthly wages, rural	0.0797***	0.0590***	0.0207	79,970
non-agricultural manual workers	(0.00451)	(0.00618)		
Log monthly wages, rural	0.0969***	0.0751***	0.0218	112,521
desk, service, and intellectual workers	(0.00414)	(0.00306)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate p < 0.01, p < 0.05, and p < 0.1, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author's compilation based on LFS 1986-2012.

3. Cohort

Dependent Variables	OLS	IV	Bias gap	Observations
Base model				
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Log monthly wages,	0.126***	0.0872***	0.0388	812,522
cohort 1955-1970	(0.00354)	(0.00347)		
Log monthly wages,	0.101***	0.0856***	0.0154	1,016,749
cohort 1961-1985	(0.00377)	(0.00354)		
Cohort and Gender				
Log monthly wages, male	0.122***	0.0827***	0.0393	413,455
cohort 1955-1970	(0.00340)	(0.00337)		
Log monthly wages, male	0.0965***	0.0807***	0.0158	512,457
cohort 1961-1985	(0.00410)	(0.00376)		
Log monthly wages, female	0.128***	0.0870***	0.041	399,067
cohort 1955-1970	(0.00359)	(0.00386)		
Log monthly wages, female	0.106***	0.0856***	0.0204	504,292
cohort 1961-1985	(0.00323)	(0.00367)		
Cohort and Area				
Log monthly wages, urban	0.120***	0.0858***	0.0342	537,768
cohort 1955-1970	(0.00292)	(0.00170)		
Log monthly wages, urban	0.0984***	0.0848***	0.0136	663,259
cohort 1961-1985	(0.00333)	(0.00175)		
Log monthly wages, rural	0.123***	0.0718***	0.0512	274,754
cohort 1955-1970	(0.00506)***	(0.00640)		
Log monthly wages, rural	0.0927***	0.0712***	0.0215	353,490
cohort 1961-1985	(0.00448)	(0.00609)		
Cohort and Region				
Log monthly wages, cohort 1955-1970	0.103***	0.0741***	0.0289	96,073
Bangkok and Metropolitan	(0.00154)	(0.000922)		
Log monthly wages, cohort 1955-1970	0.136***	0.100***	0.036	173,735
Northern area	(0.00381)	(0.00465)		

Log monthly wages, cohort 1955-1970	0.150***	0.0991***	0.0509	191,927
North eastern area	(0.00349)	(0.00656)		
Log monthly wages, cohort 1955-1970	0.109***	0.0750***	0.034	131,534
Southern area	(0.00428)	(0.00312)		
Log monthly wages, cohort 1955-1970	0.113***	0.0785***	0.0345	220,198
Central area	(0.00373)	(0.00199)		
Log monthly wages, cohort 1961-1985	0.0909***	0.0748***	0.0161	131,317
Bangkok and Metropolitan	(0.00201)	(0.00109)		
Log monthly wages, cohort 1961-1985	0.112***	0.0991***	0.0129	188,780
Northern area	(0.00380)	(0.00434)		
Log monthly wages, cohort 1961-1985	0.130***	0.0979***	0.0321	228,013
North eastern area	(0.00331)	(0.00669)		
Log monthly wages, cohort 1961-1985	0.0812***	0.0732***	0.008	177,346
Southern area	(0.00400)	(0.00349)		
Log monthly wages, cohort 1961-1985	0.0863***	0.0763***	0.01	291,794
Central area	(0.00406)	(0.00231)		
Cohort and Industry				
Log monthly wages, cohort 1955-1970	0.118***	0.0614***	0.0566	267,602
agricultural and elementary workers	(0.00488)	(0.00501)		
Log monthly wages, cohort 1955-1970	0.114***	0.0727***	0.0413	128,681
non-agricultural manual workers	(0.00390)	(0.00306)		
Log monthly wages, cohort 1955-1970	0.111***	0.0827***	0.0283	415,123
desk, service, and intellectual workers	(0.00266)	(0.000801)		
Log monthly wages, cohort 1961-1985	0.0902***	0.0569***	0.0333	330,622
agricultural and elementary workers	(0.00387)	(0.00412)		
Log monthly wages, cohort 1961-1985	0.0850***	0.0700***	0.015	199,898
non-agricultural manual workers	(0.00444)	(0.00298)		
Log monthly wages, cohort 1961-1985	0.0924***	0.0825***	0.0099	483,566
desk, service, and intellectual workers	(0.00288)	(0.000884)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author’s compilation based on LFS 1986-2012.

4. Region

Dependent Variables	OLS	IV	Bias gap	Observations
Base model				
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Log monthly wages,	0.0956***	0.0751***	0.0205	162,399
BKK	(0.00204)	(0.00116)		
Log monthly wages,	0.124***	0.0990***	0.025	256,236
North	(0.00439)	(0.00436)		
Log monthly wages,	0.140***	0.0977***	0.0423	298,137
Northeast	(0.0038)	(0.00683)		
Log monthly wages,	0.0928***	0.0733***	0.0195	222,181
South	(0.00473)	(0.00343)		
Log monthly wages,	0.0973***	0.0766***	0.0207	369,035
Centre	(0.00459)	(0.00225)		
Region and Gender				
Log monthly wages, male	0.0945***	0.0656***	0.0289	82,071
Bangkok and Metropolitan	(0.00214)	(0.00261)		
Log monthly wages, male	0.121***	0.101***	0.02	126,355
Northern area	(0.00458)	(0.00604)		
Log monthly wages, male	0.131***	0.0871***	0.0439	151,593
North eastern area	(0.00401)	(0.00425)		
Log monthly wages, male	0.0862***	0.0659***	0.0203	116,980
Southern area	(0.00503)	(0.00500)		
Log monthly wages, male	0.0959***	0.0750***	0.0209	186,181
Central area	(0.00462)	(0.00324)		
Log monthly wages, female	0.0968***	0.0796***	0.0172	80,328
Bangkok and Metropolitan	(0.00195)	(0.00118)		
Log monthly wages, female	0.125***	0.0938***	0.0312	129,881
Northern area	(0.00414)	(0.00459)		
Log monthly wages, female	0.148***	0.101***	0.047	146,544
North eastern area	(0.00366)	(0.0101)		
Log monthly wages, female	0.0998***	0.0774***	0.0224	105,201
Southern area	(0.00422)	(0.00273)		
Log monthly wages, female	0.0981***	0.0738***	0.0243	182,854
Central area	(0.00435)	(0.00232)		
Region and Cohort				
Log monthly wages, cohort 1955-1970	0.103***	0.0741***	0.0289	96,073
Bangkok and Metropolitan	(0.00154)	(0.000922)		
Log monthly wages, cohort 1955-1970	0.136***	0.100***	0.036	173,735
Northern area	(0.00381)	(0.00465)		
Log monthly wages, cohort 1955-1970	0.150***	0.0991***	0.0509	191,927
North eastern area	(0.00349)	(0.00656)		
Log monthly wages, cohort 1955-1970	0.109***	0.0750***	0.034	131,534
Southern area	(0.00428)	(0.00312)		
Log monthly wages, cohort 1955-1970	0.113***	0.0785***	0.0345	220,198
Central area	(0.00373)	(0.00199)		
Log monthly wages, cohort 1961-1985	0.0909***	0.0748***	0.0161	131,317
Bangkok and Metropolitan	(0.00201)	(0.00109)		
Log monthly wages, cohort 1961-1985	0.112***	0.0991***	0.0129	188,780
Northern area	(0.00380)	(0.00434)		
Log monthly wages, cohort 1961-1985	0.130***	0.0979***	0.0321	228,013
North eastern area	(0.00331)	(0.00669)		
Log monthly wages, cohort 1961-1985	0.0812***	0.0732***	0.008	177,346
Southern area	(0.00400)	(0.00349)		
Log monthly wages, cohort 1961-1985	0.0863***	0.0763***	0.01	291,794
Central area	(0.00406)	(0.00231)		
Region and Area				
Log monthly wages, urban	0.0962***	0.0758***	0.0204	143,357
Bangkok and Metropolitan	(0.00198)	(0.00143)		
Log monthly wages, urban	0.121***	0.0997***	0.0213	162,960
Northern area	(0.00394)	(0.00235)		

Log monthly wages, urban	0.129***	0.0965***	0.0325	192,139
North eastern area	(0.00331)	(0.00336)		
Log monthly wages, urban	0.0912***	0.0726***	0.0186	133,456
Southern area	(0.00392)	(0.00245)		
Log monthly wages, urban	0.0957***	0.0773***	0.0184	225,565
Central area	(0.00419)	(0.00103)		
Log monthly wages, rural	0.0876***	0.0680***	0.0196	19,042
Bangkok and Metropolitan	(0.00291)	(0.00237)		
Log monthly wages, rural	0.115***	0.0818***	0.0332	93,276
Northern area	(0.00509)	(0.00766)		
Log monthly wages, rural	0.142***	0.0645***	0.0775	105,998
North eastern area	(0.00563)	(0.0143)		
Log monthly wages, rural	0.0838***	0.0660***	0.0178	88,725
Southern area	(0.00554)	(0.00481)		
Log monthly wages, rural	0.0938***	0.0695***	0.0243	143,470
Central area	(0.00509)	(0.00435)		
Region and Industry				
Log monthly wages, Bangkok and Metropolitan	0.0916***	0.0590***	0.0326	18,206
agricultural and elementary workers	(0.00198)	(0.00341)		
Log monthly wages, Northern area	0.108***	0.0693***	0.0387	99,008
agricultural and elementary workers	(0.00486)	(0.00587)		
Log monthly wages, North eastern area	0.136***	0.0407***	0.0953	122,801
agricultural and elementary workers	(0.00528)	(0.00903)		
Log monthly wages, Southern area	0.0729***	0.0573***	0.0156	80,575
agricultural and elementary workers	(0.00485)	(0.00449)		
Log monthly wages, Central area	0.0884***	0.0628***	0.0256	108,148
agricultural and elementary workers	(0.00490)	(0.00412)		
Log monthly wages, Bangkok and Metropolitan	0.101***	0.0731***	0.0279	50,096
non-agricultural manual workers	(0.00372)	(0.00241)		
Log monthly wages, Northern area	0.0960***	0.0760***	0.02	34,490
non-agricultural manual workers	(0.00548)	(0.00330)		
Log monthly wages, North eastern area	0.0971***	0.0598***	0.0373	32,626
non-agricultural manual workers	(0.00499)	(0.00657)		
Log monthly wages, Southern area	0.0837***	0.0644***	0.0193	26,171
non-agricultural manual workers	(0.00520)	(0.00937)		
Log monthly wages, Central area	0.0889***	0.0693***	0.0196	95,064
non-agricultural manual workers	(0.00498)	(0.00417)		
Log monthly wages, Bangkok and Metropolitan	0.0931***	0.0791***	0.014	91,729
desk, service, and intellectual workers	(0.00174)	(0.00241)		
Log monthly wages, Northern area	0.114***	0.0944***	0.0196	122,639
desk, service, and intellectual workers	(0.00344)	(0.00278)		
Log monthly wages, North eastern area	0.113***	0.0929***	0.0201	142,589
desk, service, and intellectual workers	(0.00303)	(0.00255)		
Log monthly wages, Southern area	0.0899***	0.0702***	0.0197	115,376
desk, service, and intellectual workers	(0.00368)	(0.00183)		
Log monthly wages, Central area	0.0949***	0.0774***	0.0175	165,532
desk, service, and intellectual workers	(0.00375)	(0.00188)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.
Source: Author's compilation based on LFS 1986-2012.

5. Economic Sector

Dependent Variables	OLS	IV	Bias gap	Observations
Base model				
Log monthly wages,	0.112***	0.0856***	0.0264	1,307,016
all workers	(0.00425)	(0.00354)		
Log monthly wages,	0.101***	0.0590***	0.042	428,699
Agricultural sector	(0.00479)	(0.0044)		
Log monthly wages,	0.0936***	0.0706***	0.023	238,274
Manufacturing sector	(0.00468)	(0.00285)		
Log monthly wages,	0.102***	0.0825***	0.0195	637,105
Service sector	(0.00325)	(0.00078)		
Industry and Gender				
Log monthly wages, male	0.0944***	0.0596***	0.0348	230,334
agricultural and elementary workers	(0.00488)	(0.00381)		
Log monthly wages, male	0.0990***	0.0687***	0.0303	138,554
non-agricultural manual workers	(0.00522)	(0.00284)		
Log monthly wages, male	0.0903***	0.0709***	0.0194	293,022
desk, service, and intellectual workers	(0.00290)	(0.00126)		
Log monthly wages, female	0.104***	0.0476***	0.0564	198,365
agricultural and elementary workers	(0.00437)	(0.00514)		
Log monthly wages, female	0.0831***	0.0675***	0.0156	99,720
non-agricultural manual workers	(0.00335)	(0.00286)		
Log monthly wages, female	0.109***	0.0850***	0.024	344,083
desk, service, and intellectual workers	(0.00329)	(0.00286)		
Industry and Cohort				
Log monthly wages, cohort 1955-1970	0.118***	0.0614***	0.0566	267,602
agricultural and elementary workers	(0.00488)	(0.00501)		
Log monthly wages, cohort 1955-1970	0.114***	0.0727***	0.0413	128,681
non-agricultural manual workers	(0.00390)	(0.00306)		
Log monthly wages, cohort 1955-1970	0.111***	0.0827***	0.0283	415,123
desk, service, and intellectual workers	(0.00266)	(0.000801)		
Log monthly wages, cohort 1961-1985	0.0902***	0.0569***	0.0333	330,622
agricultural and elementary workers	(0.00387)	(0.00412)		
Log monthly wages, cohort 1961-1985	0.0850***	0.0700***	0.015	199,898
non-agricultural manual workers	(0.00444)	(0.00298)		
Log monthly wages, cohort 1961-1985	0.0924***	0.0825***	0.0099	483,566
desk, service, and intellectual workers	(0.00288)	(0.000884)		
Industry and Area				
Log monthly wages, urban	0.109***	0.0708***	0.0382	171,004
agricultural and elementary workers	(0.00482)	(0.00346)		
Log monthly wages, urban	0.0971***	0.0750***	0.0221	158,304
non-agricultural manual workers	(0.00441)	(0.00139)		
Log monthly wages, urban	0.103***	0.0834***	0.0196	524,584
desk, service, and intellectual workers	(0.00305)	(0.000594)		
Log monthly wages, rural	0.0776***	0.0435***	0.0341	257,695
agricultural and elementary workers	(0.00418)	(0.00545)		

Log monthly wages, rural	0.0797***	0.0590***	0.0207	79,970
non-agricultural manual workers	(0.00451)	(0.00618)		
Log monthly wages, rural	0.0969***	0.0751***	0.0218	112,521
desk, service, and intellectual workers	(0.00414)	(0.00306)		
Industry and Region				
Log monthly wages, Bangkok and Metropolitan	0.0916***	0.0590***	0.0326	18,206
agricultural and elementary workers	(0.00198)	(0.00341)		
Log monthly wages, Northern area	0.108***	0.0693***	0.0387	99,008
agricultural and elementary workers	(0.00486)	(0.00587)		
Log monthly wages, North eastern area	0.136***	0.0407***	0.0953	122,801
agricultural and elementary workers	(0.00528)	(0.00903)		
Log monthly wages, Southern area	0.0729***	0.0573***	0.0156	80,575
agricultural and elementary workers	(0.00485)	(0.00449)		
Log monthly wages, Central area	0.0884***	0.0628***	0.0256	108,148
agricultural and elementary workers	(0.00490)	(0.00412)		
Log monthly wages, Bangkok and Metropolitan	0.101***	0.0731***	0.0279	50,096
non-agricultural manual workers	(0.00372)	(0.00241)		
Log monthly wages, Northern area	0.0960***	0.0760***	0.02	34,490
non-agricultural manual workers	(0.00548)	(0.00330)		
Log monthly wages, North eastern area	0.0971***	0.0598***	0.0373	32,626
non-agricultural manual workers	(0.00499)	(0.00657)		
Log monthly wages, Southern area	0.0837***	0.0644***	0.0193	26,171
non-agricultural manual workers	(0.00520)	(0.00937)		
Log monthly wages, Central area	0.0889***	0.0693***	0.0196	95,064
non-agricultural manual workers	(0.00498)	(0.00417)		
Log monthly wages, Bangkok and Metropolitan	0.0931***	0.0791***	0.014	91,729
desk, service, and intellectual workers	(0.00174)	(0.00241)		
Log monthly wages, Northern area	0.114***	0.0944***	0.0196	122,639
desk, service, and intellectual workers	(0.00344)	(0.00278)		
Log monthly wages, North eastern area	0.113***	0.0929***	0.0201	142,589
desk, service, and intellectual workers	(0.00303)	(0.00255)		
Log monthly wages, Southern area	0.0899***	0.0702***	0.0197	115,376
desk, service, and intellectual workers	(0.00368)	(0.00183)		
Log monthly wages, Central area	0.0949***	0.0774***	0.0175	165,532
desk, service, and intellectual workers	(0.00375)	(0.00188)		

Notes: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort dummies (except for the models with explicit cohort variables), regional dummies (except for the models with explicit region variables), and an indicator whether a cohort faced a new compulsory education law (6 years of compulsory education). Moreover, each model also includes age controls: a quadratic polynomial. Each regression includes the sample of 15 to 60 years old from the 1986 through 2012 Labour Force Surveys. Data are first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort. ***, **, and * indicate $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively. Bias gap refers to the difference between OLS estimate and the IV estimate.

Source: Author's compilation based on LFS 1986-2012.