UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2019/105

Revisiting the returns to education during rapid structural and rural transformation in Thailand

A regression discontinuity approach

Upalat Korwatanasakul[1]

December 2019

Abstract: This paper estimates returns to schooling in Thailand, applying a regression discontinuity approach to the change in the compulsory schooling law in 1978. This law helped to enhance human capital investment on the eve of rapid structural transformation. The returns to schooling based on our instrumental variables estimation were around eight per cent, while ordinary least squares (OLS) overestimated such returns. Returns were higher in urban areas, service sectors, and underdeveloped northern regions. Our findings contrast sharply with recent studies exploiting similar institutional changes in developed countries, where OLS estimates *underestimate* returns to schooling, with the implication that former school dropouts tend to have higher returns than those who were already in school before the law changed. Ability bias is more likely to arise in developing countries, possibly because parents might be forced to keep only children with higher abilities in school, reinforcing inequality among children within the household.

Key words: returns to education, Mincer equation, ability bias, regression discontinuity, Thailand

JEL classification: I20, I21, I25, I28

Acknowledgements: The author expresses gratitude to the late Professor Nobuhiko Fuwa of Tokyo University for his valuable guidance, input, and comments during the writing of this paper.

---

[1] ASEAN-Japan Centre, Tokyo, Japan, korwatanasakul.upalat@gmail.com

This study has been prepared within the UNU-WIDER project Transforming economies—for better jobs.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-741-5

https://doi.org/10.35188/UNU-WIDER/2019/741-5

Typescript prepared by Merl Storr.

1 Introduction

The fundamental importance of human capital formation in the process of economic development is well understood. However, the quantitative magnitudes of the causal effects of education on earnings are still intensely debated in both developed and developing country contexts. Recent studies from developed countries have shown that endogeneity bias in conventional ordinary least squares (OLS) estimates is quite substantial, and that there is a great deal of heterogeneity in returns to education within populations. In developing countries, however, similar studies remain relatively scarce. Obtaining accurate estimates of returns to education is essential for policymaking. Without such estimates, governments are poorly informed as to how to allocate scarce resources among different types of capital (e.g., human versus physical capital), as well as how to allocate their education budgets efficiently (e.g., rural versus urban areas, different regions in a country, different levels of schooling, or primary versus higher education).

This paper applies a regression discontinuity approach to the change in Thailand's compulsory schooling law in 1978, which effectively extended compulsory schooling from four to six years. This methodology has had an increasing number of applications in developed countries but is rarely found in developing countries. We find that the compulsory schooling law played a role in enhancing human capital investment on the eve of the rapid structural transformation of the 1980s: the returns to schooling based on our instrumental variables (IV) estimation were around eight per cent, while OLS substantially overestimated (by 20 per cent) such returns; returns were higher in urban areas, in service (rather than agricultural) sectors, and surprisingly in underdeveloped northern regions. Our findings contrast with most recent studies exploiting similar institutional changes in developed countries, where OLS estimates tend to *underestimate* returns to schooling, with the implication that former school dropouts (whose behaviour is altered by compulsory schooling) tend to have higher returns than those who were already in school before the change in the law. Traditional 'ability bias' (which we confirm) may tend to arise only in developing (and not in developed) countries because parents are forced to keep only those of their (multiple) children who have higher abilities in school, thereby reinforcing (rather than compensating for) inequality among children within the household.

The rest of the paper is organized as follows. The next section provides a brief review of the relevant literature. Section 3 discusses the data, methodology, and identification strategy employed in this study. The empirical results and the interpretations of our estimation of returns to schooling are provided in section 4. Section 5 concludes.

2 Literature review

In our attempts to estimate the returns to education in Thailand, we follow the now classic approach developed by Mincer (1958, 1974):

$$\log y_i = \beta_0 + \beta_1 S_i + \beta_2 X_i + \beta_3 X_i^2 + e_i \quad [1]$$

where the log of individual earnings y_i is a linear function of the years of education an individual i has attained (S_i) and a quadratic function of the number of years the individual has worked after completing his/her education (X_i). e_i represents the disturbance term. The returns to education are measured by the coefficient β_1.

The Mincer specification has been found to fit data reasonably well. For example, Card (1999) shows that, based on the pooled samples of the 1994–96 March Current Population Surveys in the United States, the hourly wage-age profiles for men and women are reasonably well approximated by the Mincer specification; he concludes that the Mincerian 'human capital earnings function is alive and well' (Card 1999: 1809). As we will see in the next section, the Mincer equation appears also to fit our Thai data reasonably well.

Despite the simplicity of the Mincerian specification and the large number of empirical studies conducted in the past few decades, debates continue about the quantitative magnitudes of the causal effects of schooling on earnings in both developed and developing country contexts. Since years of schooling is an endogenous variable determined by the choice made by the household (including both parents and children), the association of schooling with earnings does not necessarily represent causal effects, but instead may also include the effects of other factors such as children's ability, heterogeneity in family backgrounds, and heterogeneity in school quality (e.g., Behrman 1999). In particular, a major focus in the empirical literature has been on the magnitude of the so-called ability bias: the tendency for OLS estimates of returns to schooling to be (presumably upwardly) biased due to the (presumably positive) correlation between schooling and student ability, since student ability is often unobserved and thus difficult to control for (e.g., Card 1999; Deere and Vesovic 2006; Schultz 1988; Willis 1986).

Interestingly, a number of recent studies, mostly from developed countries, find that ability bias may not be very serious.[1] In addition to the early sceptics about ability bias concerns (e.g., Becker 1964; Griliches 1977), studies using the incidence of identical twins as an instrumental variable tend to show similar results between OLS and IV estimations, suggesting that the magnitude of ability bias is likely to be small, if any (Card 1999). Furthermore, there has been an increasing number of empirical studies utilizing institutional features and changes in the law as natural experiments. Such 'revisionist' literature finds that OLS estimates of returns to schooling tend to underestimate, rather than overestimate (as the traditional ability bias arguments would predict), the true returns quite substantially (for literature reviews, see e.g., Card 1999; Heckman et al. 2006). Card (1999) concludes that the magnitude of such underestimation could be as large as 20 to 40 per cent.

The local average treatment effects (LATE) interpretations of these recent studies, based on the IV estimates exploiting changes in compulsory schooling, rather surprisingly suggest the possibility of negative rather than positive ability bias, where children whose schooling behaviour is affected by the change in the compulsory schooling law (i.e. those who were school dropouts before the law change) tend to have higher returns to schooling compared with those who were already in school before the law changed. Since these empirical findings run counter to the conventional ability bias story (i.e. that those with higher ability have higher returns to schooling, stay in school longer, and earn better), a number of explanations have been proposed to account for the findings. Exogenous constraints, such as credit constraints, may be one possibility, although some question the universal applicability of this explanation across the variety of country contexts where similar findings have been obtained (Oreopoulos 2006). Card (1999) develops a theoretical model of schooling choice which suggests that a negative correlation between returns to schooling and years of schooling can arise if ability differences are not 'too important' in the determination of schooling outcomes, and if the marginal return to schooling is decreasing. Others have argued that the

[1] Apart from the concern about ability bias, there has also been a parallel literature addressing other aspects of heterogeneity, such as family background and school quality. See Behrman (1999) for a review of the literature from developing countries. Behrman (1999) argues that the 'standard estimates', which do not address these concerns, tend to overestimate the impact of schooling attainment substantially, even by as much as 40 to 100 per cent.

attenuation bias in OLS estimation due to (classical) measurement errors in the schooling variable may account for some of the difference between OLS and IV estimates,[2] or that the bias due to discount rate heterogeneity (i.e. students who dropped out of school before the compulsory schooling law changed might include students with higher than average ability and higher discount rates) may offset the positive ability bias in OLS estimates (Card 1999; Lang 1993). Alternatively, Heckman et al. (2006) interpret the empirical finding of larger IV estimates than OLS estimates as suggesting that the ability space is multidimensional, rather than unidimensional as is typically assumed in the conventional literature. If there are different types of abilities, and if different levels of schooling are required for different types of jobs, individuals with different mixes of abilities and skills will sort across schooling levels in such a way that the best individuals at one schooling level will not do so well at other levels. In light of such possibilities, 'the idea that individuals with "higher ability" are more likely to enroll in school is no longer obvious' (Heckman et al. 2006: 390).[3]

In contrast with the continuing debates in developed country contexts, similar empirical studies exploiting compulsory schooling law changes in developing countries are still relatively scarce. It is possible that some of the explanations for the empirical findings on OLS and IV estimation results may not apply in developing country contexts; ability differences may be important in explaining schooling outcomes (Card's (1999) explanation), or the multidimensionality of the ability or skill space may not be as important in developing countries as in developed countries (Heckman et al.'s (2006) explanation). It is this lacuna in the literature that this paper intends to address.

3 Empirical methodology and data

3.1 Methodology and identification strategy

This paper follows closely the regression discontinuity approach of Oreopoulos (2006) and others that use the incidence of change in compulsory schooling laws as an IV for estimating returns to schooling. This paper applies this approach to the incidence of the 1978 Primary Education Act in Thailand. As can be seen from Figures 1 and 2, the timing of the act was immediately before the rapid growth in per-capita gross domestic product (GDP), which started in the mid-1980s (and lasted until the 1997 Asia crisis), and the industrialization episode of the 1980s.

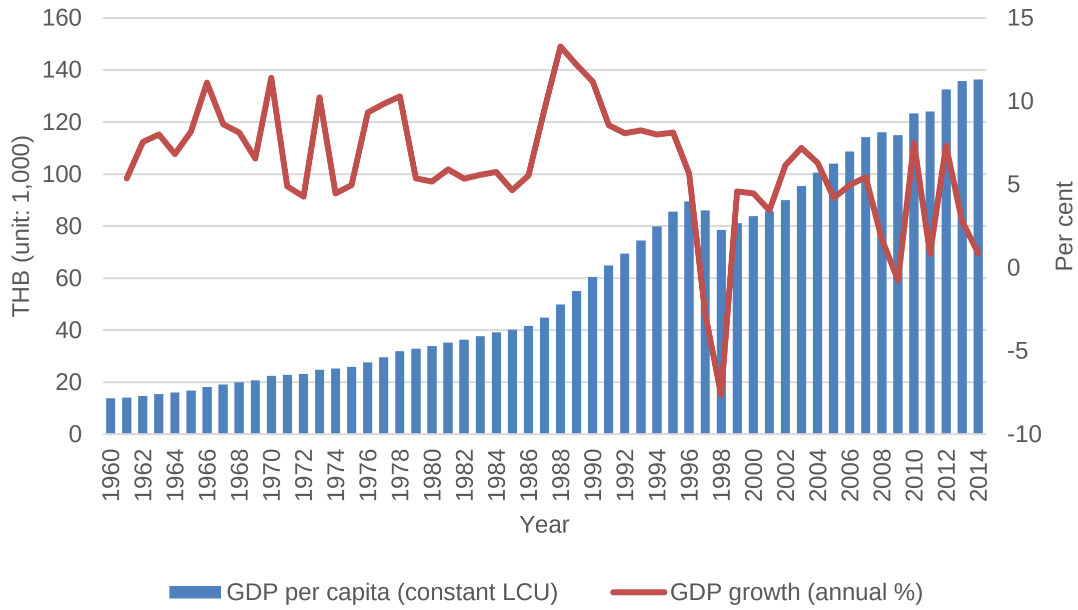
The 1978 Primary Education Act was the first education act with which every individual was required to comply. At that time, the whole educational structure was changed from a four-three-three-two structure to a six-three-three structure. Previously, the Thai education system had consisted of four years of lower primary education, three years of upper primary education, and five years of secondary education. The 1978 Primary Education Act reduced the total years of primary education to six, without division between the lower and upper levels, while secondary

[2] Card (1999) notes, however, that, based on what is known in the literature about the effects of random measurement errors, the magnitude of the downward bias of OLS estimates found in some studies is too large to be explained by the attenuation bias arising from classical measurement errors.

[3] As an example, if individuals with more schooling become teachers and those with lower schooling become plumbers, 'then the latter are better plumbers than the average teacher would be if he became a plumber' (Heckman et al. 2006: 389).

education remained the same. In addition, the government expanded compulsory education from four years to six years of primary education.

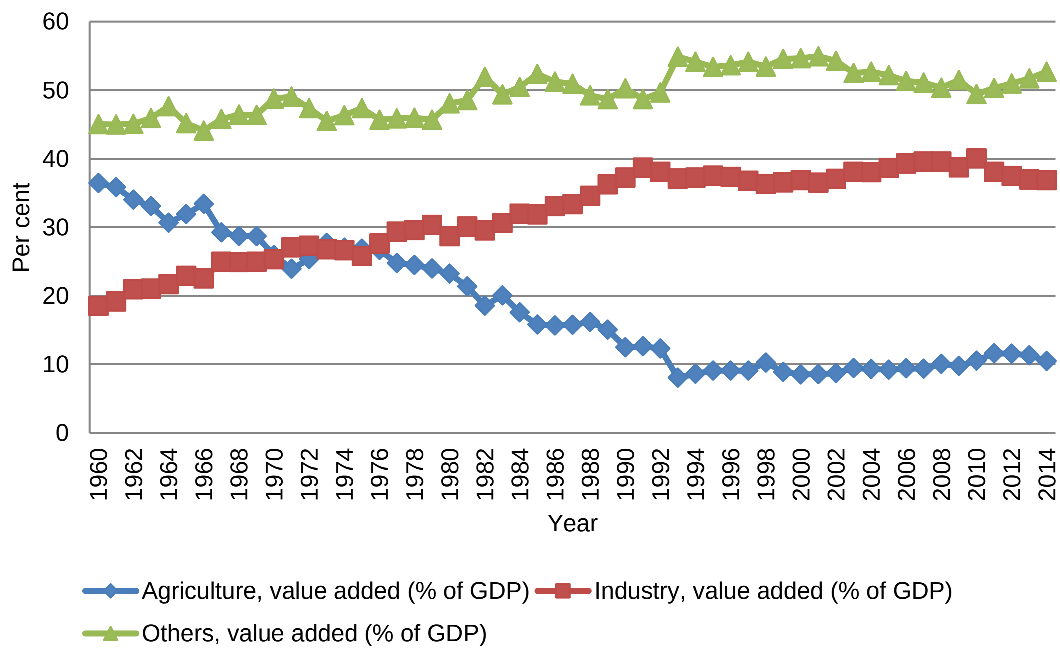
Figure 1: Real GDP growth and per-capita GDP in Thailand, 1960–2014



Note: 1 USD = 35 THB (as of 2016). LCU: local currency unit.

Source: author's compilation based on data from World Development Indicators for 2016.

Figure 2: Structural transformation: net output as percentage of GDP, 1960–2014



Source: author's compilation based on data from World Development Indicators for 2016.

With normal schooling progression (i.e. entering primary school at age six), the first birth cohort to be affected by the 1978 change to compulsory education was the cohort born in 1968. However, it was relatively commonly observed at the time that some students started enrolling in primary schools as late as age eight. As a result, we consider the cohort born in 1968 to be the first cohort affected by the 1978 law (Table 1). The law took effect immediately on 5 April 1978, a month before the start of the 1978 academic year. Therefore, all students enrolled in grade four in 1977 were required to move up to grade five in 1978. The possible age range of fourth-year students is between nine and eleven years old, which corresponds to the 1966–68 cohorts.

Table 1: Identification of first cohorts affected by the 1978 compulsory education law

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.

One additional complication for our current purposes is the fact that, during the process of expansion of compulsory education, the government provided a five-year adjustment period (1978 to 1982) to districts that were not ready for the compulsory education reform. Since the majority of schools had either four years of lower primary education or three years of upper primary education prior to the 1978 reform, these schools found it difficult to comply with the new compulsory education law immediately, and therefore needed extra time to build more classrooms or combine with other schools. The government nevertheless declared that by 1982 every student and every school had to comply with the 1978 Primary Education and Compulsory Education Acts. Taking the five-year adjustment period into account, we therefore exclude from the analysis the cohorts from 1966 to 1972.

The 1978 compulsory schooling law change in Thailand affected a large proportion of the population, requiring almost half the students in fourth-grade primary education to stay in school for two more years (until grade six, the final grade of primary education). As a result, similarly to the application by Oreopoulos (2006), the estimated LATE in this paper may arguably be closer to the population average treatment effect than similar studies that affect only relatively small (and arguably peculiar) fractions of the population.

3.2 Data

We use the Thai Labour Force Survey (LFS) conducted by the National Statistical Office for the years 1986 to 2012. The LFS is collected quarterly from about 80,000 randomly selected households for a total of about 200,000 observations per quarter, representing 0.1–0.5 per cent of the total Thai population.

The data set used in this study is constructed by pooling the 27 consecutive annual LFS. 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 the urban manufacturing sector and the rural agricultural sector. Nevertheless, they tend to migrate back to the rural agricultural sector during the rainy season in the third quarter of the year (Sussangkarn and Chalamwong 1996). Moreover, this study limits the sample to 157,390 wage workers aged 15 to 60 in the year of interview. The age restriction of 15 to 60 years is imposed because 15 years is the minimum legal age at which individuals can start working, and 60 years is the usual retirement age in Thailand. In addition, we also employ a birth cohort restriction 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 regression discontinuity estimation. 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.

3.3 Econometric specification

As discussed in the previous section, we construct a cohort panel using the 1986 to 2012 rounds of the Thai LFS. We compare the schooling and earnings outcomes between the cohorts who were covered by the 1978 compulsory schooling law change and the preceding cohorts who were not affected, by introducing into the conventional Mincer equation dummy variables indicating coverage by the law as the IV to control for endogenous years of schooling.

Our regression equation of main interest closely follows Oreopoulos (2006) and takes the form:

$$\log y_i = \gamma_0 + \gamma_1 \widehat{S_i} + \gamma_2 C_i^1 + \gamma_3 C_i^2 + \gamma_4 C_i^3 + \gamma_5 C_i^4 + \sum_{\mathrm{k}=16}^{60} \gamma_{6\mathrm{k}} \mathrm{A}_{\mathrm{ki}} + \sum_{\mathrm{l}=1}^{4} \gamma_{7\mathrm{l}} \mathrm{R}_{\mathrm{li}} + \vartheta_i \quad [2]$$

with the first-stage equation:

$$S_i = \pi_0 + \pi_1 F_i + \pi_2 C_i^1 + \pi_3 C_i^2 + \pi_4 C_i^3 + \pi_5 C_i^4 + \sum_{\mathrm{k}=16}^{60} \pi_{6\mathrm{k}} \mathrm{A}_{\mathrm{ki}} + \sum_{\mathrm{l}=1}^{4} \pi_{7\mathrm{l}} \mathrm{R}_{\mathrm{li}} + \varepsilon_i \quad [3]$$

where $\log y_i$ is the log of the monthly wages of individual *I*, S_i is the endogenous years of education of individual *i*, $\widehat{S_i}$ is the fitted value estimated from the first-stage least squares regression, and F_i represents a dummy variable to be used as the IV, indicating whether an individual had to comply with the 1978 compulsory education law (i.e. individuals who were born between 1966 and 1968 and had not left school at the time of the law change, as well as all subsequent birth year cohorts born from 1969 onwards). The control variables include a set of age (as a proxy for working experience) dummies (A_{ki}), quartic terms of birth cohort (C_i), and regional dummies (R_{li}). ϑ_i and ε_i (as well as θ_i, and e_i below) represent disturbance terms.

We also estimate the reduced form version:

$$\log y_i = \alpha_0 + \alpha_1 F_i + \alpha_2 C_i + \alpha_3 C_i^2 + \alpha_4 C_i^3 + \alpha_5 C_i^4 + \sum_{\mathrm{k}=16}^{60} \alpha_{6\mathrm{k}} \mathrm{A}_{\mathrm{ki}} + \sum_{\mathrm{l}=1}^{4} \alpha_{7\mathrm{l}} \mathrm{R}_{\mathrm{li}} + \theta_i \quad [4]$$

The IV estimation results from equation 2 will be compared with the results based on the OLS regression:

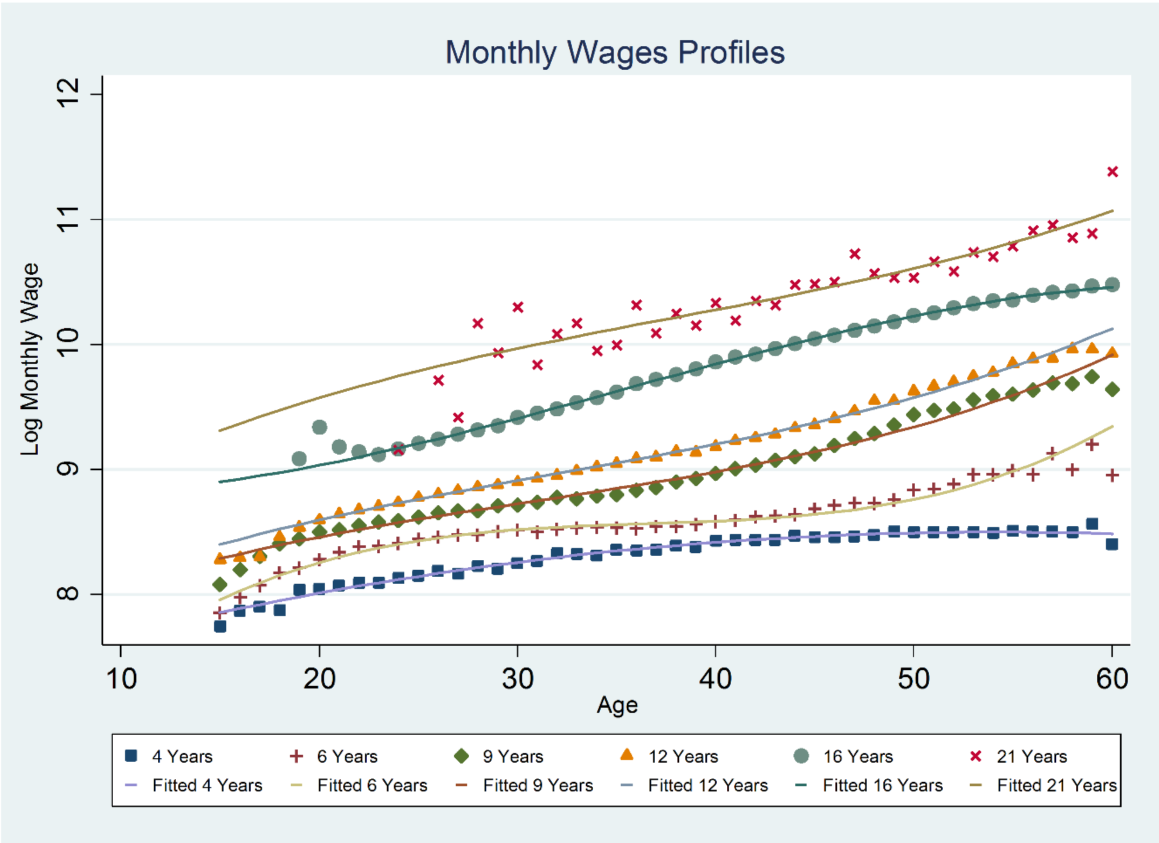
$$\log y_i = \beta_0 + \beta_1 S_i + \beta_2 C_i^1 + \beta_3 C_i^2 + \beta_4 C_i^3 + \beta_5 C_i^4 + \sum_{\mathrm{k}=16}^{60} \beta_{6\mathrm{k}} \mathrm{A}_{\mathrm{ki}} + \sum_{\mathrm{l}=1}^{4} \beta_{7\mathrm{l}} \mathrm{R}_{\mathrm{li}} + e_i \quad [5]$$

4 Empirical results

4.1 Some preliminary analysis of the data

Following Card (1999), we first examine age-wage profiles based on our Thai data, which leads us to a similar conclusion as Card's in that the estimates from the Mincer model approximate the actual age-wage profiles reasonably well. Figure 3 shows the age profiles of monthly wages by education level for Thai workers using pooled samples from the LFS for 1986 to 2012. The data represents the mean log monthly wage by age for individuals with four, six, nine, 12, 16, and 21 years of education. Four years of education refers to the minimum years of education required by the 1962 Compulsory Education Act, whereas six years of education represents the minimum compulsory education level enforced by the 1978 law. The other education years refer to the final year of each academic level: lower secondary level, upper secondary level, undergraduate level, and graduate level. Plotted lines along with the actual means are the fitted values obtained from the Mincer model, which includes only a quadratic term of age. Comparisons of the fitted and actual data suggest that age-earnings profiles for Thai workers are fairly smooth and reasonably well approximated. In contrast to age-wage profiles from the US, the problem in fitting the precise curvature appears less pronounced in the case of Thailand.

Figure 3: Age profiles of monthly wages, Thailand

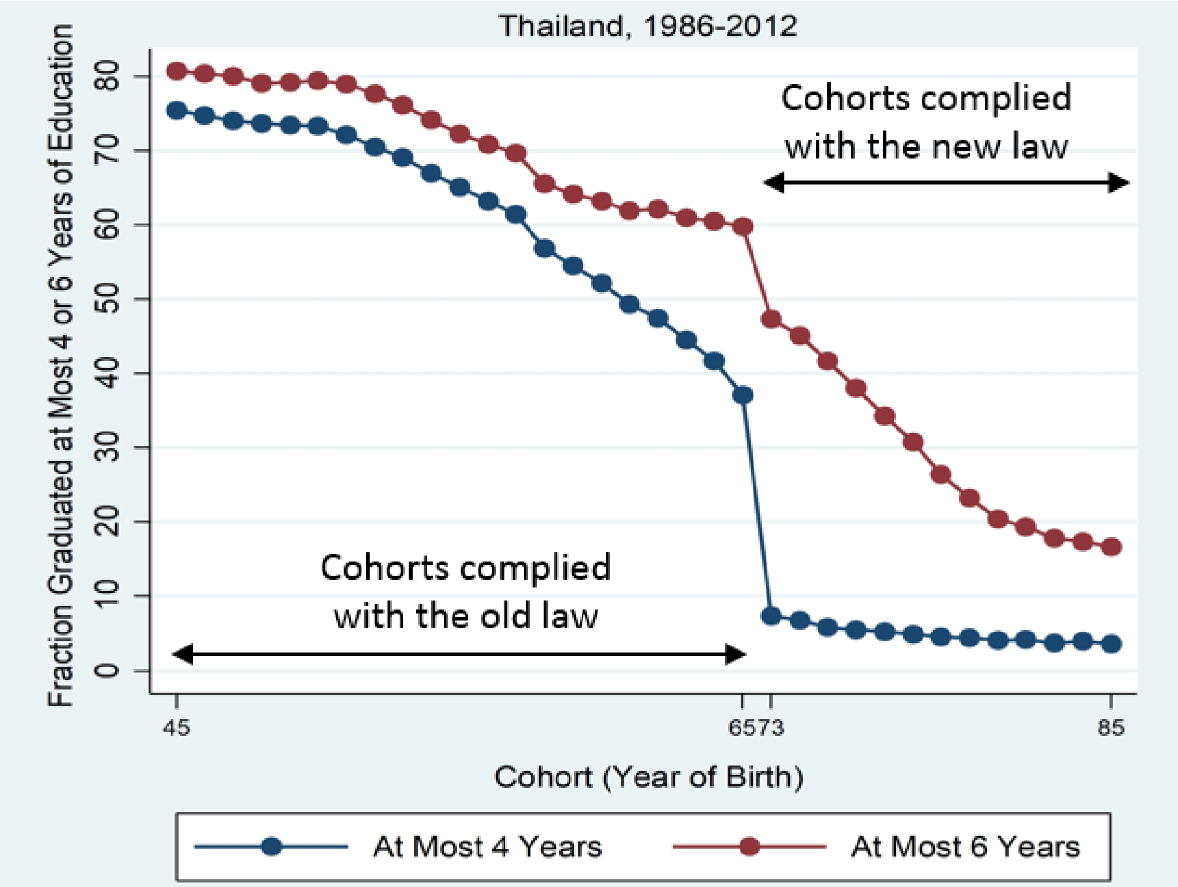


Note: The regression models include a linear education term and a quadratic in age.

Source: author's compilation based on LFS data for 1986–2012.

Figure 4 is a graphical representation of the effect of the change in the compulsory education law on the fraction of individuals graduating with at most four years of education. The lower line in Figure 4 shows the proportion of adults aged 15 to 60 who reported that their highest attained level of education was at most four years, while the upper line shows the proportion of adults aged 15 to 60 who reported that their highest attained level of education was at most six years. Both lines exhibit downward trends. The 1966 to 1972 cohorts were the first cohorts affected by the 1978 compulsory education law and were in the five-year adjustment period. Approximating the magnitude of the discontinuity, a sharp drop is observed in the fraction attaining at most four years of education, from 40 per cent to ten per cent.

Figure 4: Fraction attaining at most four and six years of education, 1986–2012



Source: author's compilation based on LFS data for 1986–2012.

4.2 First-stage regression results

Table 2 presents the first-stage regression results showing the effects of the change in the compulsory schooling law on educational attainment (equation 3). Each regression includes as regressors a birth cohort quartic polynomial, regional dummies, and a dummy variable indicating whether a cohort faced the new compulsory education law. Columns 3 to 5 also include age controls: a quadratic polynomial or age dummies where indicated. Each regression includes the sample aged 15 to 60 years from the LFS in 1986 to 2012. Following Oreopoulos (2006), individual-level observations are first aggregated into cell group means by cohort, age, sex, survey year, region, and industrial sector of employment, and are weighted by cell size. Regressions are clustered by birth cohort, region, and industrial sector of employment. The total number of cells is 157,087.

As shown in Table 2, the coefficients on the compulsory education dummy are statistically significant and robust across different specifications. The compulsory education law change, which extended the minimum years of schooling from four years to six years, led to roughly four additional years of schooling, corresponding to roughly twice the additional schooling required by the law. Some existing studies from developed countries have also found that the impact of compulsory schooling law changes goes beyond the additional years of schooling imposed by the law change (e.g., Oreopoulos 2003). While the quantitative magnitudes of the impact of compulsory schooling that we see in Thailand appear to be substantially larger than those found in the existing literature—which mostly comes from developed countries (typically ranging between 0.1 and 0.5 years of additional schooling), with a few from China (ranging between 0.8

and 1.2 years of additional schooling)—we should note that direct comparisons among such studies from different countries may not necessarily be warranted, since compulsory education is imposed somewhat differently (e.g., a minimum school leaving age, or a specified minimum compulsory level of education) in the different countries in the previous studies.

Table 2: First-stage regression results: estimated effects of compulsory education law on education attainment

	(1)	(2)	(3)	(4)	(5)
	First stage				
	Dependent variable: number of years of schooling				
Compulsory education	4.356***	4.294***	4.259***	4.270***	4.046***
	(0.392)	(0.391)	(0.365)	(0.364)	(0.313)
Fixed effects					
Regional controls	No	Yes	Yes	Yes	Yes
Birth cohort	Quartic	Quartic	Quartic	Quartic	Quartic
Additional controls	None	None	Age dummy	Age dummy Gender	Age dummy Gender Urban
Initial sample size	1,308,519	1,308,519	1,308,519	1,308,519	1,308,519
R-squared	0.091	0.104	0.128	0.129	0.184

Note: The dependent variables are the number of years of schooling. Each regression includes controls for a birth cohort quartic polynomial, regional dummies (except for the models with explicit region variables), and an indicator of whether a cohort faced the new compulsory education law (six years of compulsory education). Columns 3–5 also include age dummy variables. Each regression includes the sample aged 15-60 years from the 1986–2012 LFS. Data is first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort, region, and industrial sector of employment. 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 data for 1986–2012.

4.3 Reduced form

Table 3 shows the reduced form equation estimation of the effects of the compulsory education law change on monthly wages. We find that the change in compulsory schooling law led to an approximately 30 per cent increase in monthly wages. The relatively large reduced form effects are consistent with the relatively large effects on the years of schooling resulting from the change in compulsory schooling found in the first-stage regression results. They appear to be substantially larger compared with the effects of compulsory education laws on wages and earnings, especially in developed countries, as found in the existing literature with the major exception of Harmon and Walker (1995), who show similar effects of a change in compulsory education on earnings in the UK.

Table 3: Reduced form equation results: estimated effects of compulsory education law on log monthly wages

	(1)	(2)	(3)	(4)	(5)
	Reduced form				
	Dependent variable: log monthly wages				
Compulsory education	0.354***	0.343***	0.355***	0.348***	0.310***
	(0.0590)	(0.0559)	(0.0585)	(0.0592)	(0.0497)
Fixed effects					
Regional controls	No	Yes	Yes	Yes	Yes
Birth cohort	Quartic	Quartic	Quartic	Quartic	Quartic
Additional controls	None	None	Age dummy	Age dummy Gender	Age dummy Gender Urban
Initial sample size	1,308,519	1,308,519	1,308,519	1,308,519	1,308,519
R-squared	0.017	0.082	0.126	0.134	0.200

Note: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort quartic polynomial, regional dummies (except for the models with explicit region variables), and an indicator of whether a cohort faced the new compulsory education law (six years of compulsory education). Columns 3–5 also include age dummy variables. Each regression includes the sample aged 15–60 years from the 1986–2012 LFS. Data is first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort, region, and industrial sector of employment. 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 data for 1986–2012.

4.4 OLS and IV estimations

The OLS estimation results on returns to schooling are shown in Table 4. Each regression equation includes controls for birth cohort, regional dummies, and age. The rates of return to schooling based on the OLS estimation are approximately 11 per cent. Our OLS estimates from Thailand are somewhat higher than OLS estimates obtained in developed countries in the literature, which range roughly between eight and 10 per cent. Our OLS estimation results of 11 per cent are the same as those obtained by Warunsiri and McNown (2010) for Thailand.

Table 4: Returns to schooling estimates for log monthly wages (OLS)

	(1)	(2)	(3)	(4)	(5)
	OLS				
	Dependent variable: log monthly wages				
Years of schooling	0.113***	0.111***	0.112***	0.112***	0.109***
	(0.00184)	(0.00172)	(0.00186)	(0.00182)	(0.00165)
Fixed effects					
Regional controls	No	Yes	Yes	Yes	Yes
Birth cohort	Quartic	Quartic	Quartic	Quartic	Quartic
Additional controls	None	None	Age dummy	Age dummy Gender	Age dummy Gender Urban
Initial sample size	1,308,519	1,308,519	1,308,519	1,308,519	1,308,519
R-squared	0.527	0.567	0.603	0.614	0.621

Note: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort quartic polynomial, regional dummies (except for the models with explicit region variables), and an indicator of whether a cohort faced the new compulsory education law (six years of compulsory education). Columns 3–5 also include age dummy variables. Each regression includes the sample aged 15–60 years from the 1986–2012 LFS. Data is first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort, region, and industrial sector of employment. 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 data for 1986–2012.

The IV estimates of returns to schooling are shown in Table 5. We find that one additional year of schooling is associated with an approximately eight per cent increase in monthly wages, which is somewhat lower than the OLS estimates. The IV estimates from this study are less than, but somewhat consistent with, those found in Canada and the UK, which are approximately 10 per cent. While the OLS estimates of returns to schooling are the same in Warunsiri and McNown (2010) and our study, our IV estimates are lower than the estimate obtained by Warunsiri and McNown's cohort panel analysis (using cohort fixed effects) of 14 per cent, which is similar to estimates obtained from the US.[4]

Table 5: IV returns to schooling estimates for log monthly wages (IV estimation)

	(1)	(2)	(3)	(4)	(5)
			IV		
			Dependent variable: log monthly wages		
Years of schooling	0.0818***	0.0799***	0.0832***	0.0807***	0.0767***
	(0.00772)	(0.00680)	(0.00767)	(0.00790)	(0.00751)
Fixed effects					
Regional controls	No	Yes	Yes	Yes	Yes
Birth cohort	Quartic	Quartic	Quartic	Quartic	Quartic
Additional controls	None	None	Age dummy	Age dummy Gender	Age dummy Gender Urban
Initial sample size	1,308,519	1,308,519	1,308,519	1,308,519	1,308,519
R-squared	0.487	0.528	0.571	0.575	0.584

Note: The dependent variables are log monthly wages. Each regression includes controls for a birth cohort quartic polynomial, regional dummies (except for the models with explicit region variables), and an indicator of whether a cohort faced the new compulsory education law (six years of compulsory education). Columns 3–5 also include age dummy variables. Each regression includes the sample aged 15–60 years from the 1986–2012 LFS. Data is first aggregated into cell means and weighted by cell size. Regressions are clustered by birth cohort, region, and industrial sector of employment. 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 data for 1986–2012.

It is intriguing to find that our IV estimates are lower than our OLS estimates by the order of 20 per cent. There are a few other studies from developing countries (e.g., China and Turkey) that also suggest that OLS may overestimate returns to schooling, consistent with Behrman's (1999) view. In contrast, relatively recent studies based on compulsory schooling laws (mostly in developed countries) find that their IV estimates of returns to schooling are substantially higher than their OLS estimates. A number of explanations have been raised in the literature for the latter set of results. Among these, Card (1999) and Heckman et al. (2006) stand out in that they develop formal (although relatively simple) models to explore how such outcomes can arise. Card (1999) argues that a negative correlation between schooling and returns to schooling (and thus lower OLS estimates than IV estimates) may arise if ability differences are 'not too important' in the determination of the years of schooling.[5] This story appears to be plausible in explaining why positive ability bias is absent in developed countries while it may be relatively more important in

[4] Warunsiri and McNown (2010) conclude that the estimates from IV and panel fixed effects are similar in magnitude; therefore, the problem of endogeneity bias is fixed in their estimations. Unlike 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) argue that this IV adopted by Warunsiri and McNown may be correlated with individual ability due to the non-random nature of individuals and school allocation.

[5] Card (1999) also argues that while OLS attenuation bias due to classical measurement errors might partially account for the pattern, the quantitative magnitudes of the (negative) OLS bias are not likely to be measurement errors alone.

developing countries. Unless resource (financial) constraints are severe, parents in developed societies might make every effort to educate their children, regardless of their ability. Since fertility tends to be low in these countries, parents are unlikely to be under pressure to select only better-ability children to go to school. In contrast, in developing countries, where many households are resource-constrained and the number of children tends to be large, parents may not be able to keep all of their (multiple) children in school up to their perceived optimum. Hard pressed to keep a subset of their children in school while putting the rest to work, parents may try to keep only children with relatively better ability in school.[6] As a country grows richer, however, severe resource constraints are likely to be gradually lifted, and fertility is also likely to decline at the same time. As a result, parents are less likely to have to make such choices, and more likely to keep all of their (smaller number of) children in school, regardless of heterogeneity of ability. The relative importance of ability heterogeneity in determining schooling outcomes may decline, consistent with Card (1999).

Heckman et al. (2006) propose an alternative explanation for the absence of (positive) ability bias. They argue that ability is multidimensional: different types of ability or skill, and different levels of schooling, are required by different types of job or in different industries. According to this view, 'individuals sort themselves across schooling levels in such a way that the best individuals in one schooling level are the worst in the other, and vice versa' (Heckman et al. 2006: 374). In relatively industrialized and diversified economies, this story would be quite plausible. In less diversified and predominantly low-skilled economies, however, such possibilities are arguably less plausible. Based on the unidimensional skill/ability space view, on the other hand, the conventional positive ability bias in the determination of schooling may become quite important.

Thus, based on both Card's (1999) and Heckman et al.'s (2006) views of why positive ability bias may not be important in developed countries, the role of conventional positive ability bias in OLS estimates of returns to schooling can become relatively more important in developing country contexts, which is consistent with our empirical results.

4.5 Disaggregated analysis of returns to schooling

In addition to the overall estimates of returns to schooling, another important issue is potential heterogeneity in educational returns across individuals. This subsection presents initial results from our exploratory analysis of returns to schooling disaggregated by demographic and geographical aspects, including gender, birth cohort, area of residence (urban or rural), region, and industrial sector of employment.[7]

---

[6] Theoretically, it is not obvious that parents in poor households will invest more in the human capital of better endowed (higher-ability) children, thereby enhancing rather than compensating for inequality among their children in terms of endowment (Becker 1991). Behrman, Pollak, and Taubman (1982) show that whether parents compensate for or enhance sibling inequality depends on parental preferences (utility function) regarding the relative priority of ensuring equity among their children. Our empirical findings appear to be consistent with the possibility that parental preferences for equity among children are not strong.

[7] Simply splitting the sample into geographical and/or demographic subsamples could lead to the classic problems of self-selection. Schultz (1988) argues, for example, that stratifications of samples based on heterogeneous demographic characteristics lead to selection bias in estimating returns to education, especially in the case of developing countries, due to the prevalence of self-employment and non-wage labour, imbalanced economic development among different areas and regions, and gender-based segregation in occupational choices. Dahl (2002) is an example of directly addressing the self-selection problem due to migration, albeit in the context of the US, and finding the magnitudes of bias to be quite modest. We intend to explore these aspects in our future work.

We find that returns to schooling (based on the IV estimation) are similar between women and men, both at roughly eight per cent (8.3 per cent for females and 7.9 per cent for males). This appears in line with the conventional view that gender disparity is much less pronounced in South-East Asia compared with, say, South or East Asia (e.g., Atkinson and Errington 1990). We also disaggregated the sample between (relatively) earlier (1955 to 1970) and more recent (1961 to 1985) birth cohorts: the returns are slightly higher for the earlier birth cohort than for the later birth cohort (8.6 per cent versus 8.2 per cent).[8] This is rather surprising, since during rapid structural transformation and economic development, returns to schooling may increase over time, and thus be higher among later cohorts than earlier ones. On the other hand, not surprisingly, we find that returns to schooling in urban areas are substantially higher (8.3 per cent) than in rural areas (6.8 per cent). While the difference between urban and rural returns appears to be sizeable, it is not immediately clear to what extent this magnitude (or even the sign) may be causal, due to selective rural-to-urban migration. In addition, consistent with the urban-rural gap in returns to schooling is the finding that returns are higher in non-agricultural sectors than in the agricultural sector. The estimated returns to schooling are 5.8 per cent, 6.8 per cent, and 8.1 per cent in the agricultural, industrial, and service sectors respectively. The relatively lower returns in the agricultural sector seem consistent with the conventional view on the process of economic development, while the higher returns in the service sector than the industrial sector are somewhat puzzling in light of the rapid industrialization of the 1980s (Figure 2). As in the case of the urban-rural gap, estimates based on sectoral disaggregation are potentially subject to self-selection bias, however, and thus causal inferences may not be warranted. We also attempted disaggregation across geographical regions (i.e. Bangkok, north, north-east, south, and centre). Rather surprisingly again, returns to schooling appear to be relatively low in the Bangkok area at 7.4 per cent (second lowest, next only to the south) and relatively high in the relatively underdeveloped regions of the north (9.3 per cent) and north-east (9.7 per cent). This pattern is quite puzzling. There would be no question that relatively higher return opportunities were located in the Bangkok area than in the northern area. Furthermore, while the self-selection problem is a potential issue, the conventional wisdom would predict the opposite direction in bias, where those with higher returns would migrate to the Bangkok area. While there may be a possibility that a rapid increase in the labour supply in Bangkok due to rural-to-urban migration might have depressed wages (and thus returns to schooling) during the period under analysis, the extent of the magnitude still appears to be quite large.

According to Table 6, the overall results are similar to the main results reported earlier, in the sense that the OLS estimates of returns to schooling are higher than the IV estimates. There appear to be some variations in terms of the magnitude of the difference between OLS and IV estimates, which may suggest the severity of the endogeneity biases. The differences appear relatively more pronounced for the older age cohort (12.5 per cent based on OLS, and 8.6 per cent based on IV) rather than the younger cohort, for rural areas (10.4 per cent based on OLS, and 6.8 per cent based on IV), for the north region (14.0 per cent based on OLS, and 9.3 per cent based on IV), and for the agricultural sector (10.0 per cent based on OLS, and 5.8 per cent based on IV). Consistent with our interpretation of the main results, there appears to be a tendency that the extent of the ability bias in OLS estimates is larger in geographical areas or economic sectors where agents (households) are relatively more resource-constrained.

[8] The two cohorts compared here overlap, because we need to include the birth year cohort in the neighbourhood of the compulsory law change (i.e. the 1966 cohort) in order to apply our IV estimation.

Table 6: 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***	0.0832***	0.0288	1,308,519		
all workers	(0.00186)	(0.00767)				
Log monthly wages,	0.108***	0.0790***	0.029	663,501	Female > male	Female > male
male	(0.00185)	(0.00932)				
Log monthly wages,	0.116***	0.0831***	0.0329	645,018		
female	(0.00190)	(0.00718)				
Log monthly wages,	0.125***	0.0860***	0.039	813,981	Old > young	Old > young
cohort 1955–70	(0.00205)	(0.00549)				
Log monthly wages,	0.101***	0.0816***	0.0194	1,017,586		
cohort 1961–85	(0.00183)	(0.00623)				
Log monthly wages,	0.108***	0.0834***	0.0246	857,828	Urban > rural	Rural > urban
urban	(0.00159)	(0.00546)				
Log monthly wages,	0.104***	0.0680***	0.036	450,691		
rural	(0.00225)	(0.00907)				
Log monthly wages,	0.0953***	0.0737***	0.0216	162,399	North-east, north > others	North-east, north > others
Bangkok	(0.00145)	(0.00391)				
Log monthly wages,	0.124***	0.0965***	0.0275	256,447		
north	(0.00309)	(0.0114)				
Log monthly wages,	0.140***	0.0925***	0.0475	298,457		
north-east	(0.00379)	(0.0293)				
Log monthly wages,	0.0926***	0.0709***	0.0217	222,181		
south	(0.00309)	(0.00959)				
Log monthly wages,	0.0972***	0.0748***	0.0224	369,035		
centre	(0.00295)	(0.00708)				
Log monthly wages,	0.100***	0.0583***	0.0417	428,987	Service > manufacturing > agriculture	Agriculture > manufacturing > service
agricultural sector	(0.00299)	(0.00536)				
Log monthly wages,	0.0936***	0.0682***	0.0254	238,514		
manufacturing sector	(0.00252)	(0.00307)				
Log monthly wages,	0.102***	0.0812***	0.0208	638,080		
service sector	(0.00184)	(0.00225)				

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

Source: author's compilation based on LFS for 1986–2012.

5 Conclusion

This paper estimates returns to schooling in Thailand, using a regression discontinuity approach applied to the change in compulsory schooling law in 1978. We find that the compulsory schooling law played a role in enhancing human capital investment on the eve of the rapid structural transformation of the 1980s. The returns to schooling based on our IV estimation were around eight per cent, while OLS somewhat overestimates (by 20 per cent) such returns; returns were higher in urban areas, in service (rather than agricultural) sectors, and surprisingly in the relatively underdeveloped northern region. Our findings are in sharp contrast with most recent studies exploiting similar institutional changes in developed countries, where OLS estimates tend to *underestimate* returns to schooling, with the implication that former school dropouts (whose behaviour was altered by the new compulsory schooling law) tend to have *higher* returns than those who were already in school before the law changed. Ability bias (which we confirm) is more likely to arise in developing (but not so much in developed) countries, because parents may be forced to keep only those of their (multiple) children who have higher abilities in school, thereby reinforcing (rather than compensating for) inequality among children within the household.

References

Atkinson, J.M., and S. Errington (eds) (1990). *Power and Difference: Gender in Island Southeast Asia*. Stanford: Stanford University Press.

Becker, G. (1964). *Human Capital*. New York: Columbia University Press.

Becker, G. (1991). *A Treatise on the Family*. Cambridge, MA: Harvard University Press.

Behrman, J.R. (1999). 'Labor Markets in Developing Countries'. In O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics, Vol. 3B*. Amsterdam: Elsevier.

Behrman, J.R., R. Pollak, and P. Taubman (1982). 'Parental Preferences and Provision for Progeny'. *Journal of Political Economy*, 90(1): 52–73.

Card, D. (1999). 'The Causal Effect of Education on Earnings'. In O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics, Vol. 3A*. Amsterdam: Elsevier.

Dahl, G. (2002). 'Mobility and the Returns to Education: Testing a Roy Model with Multiple Markets'. *Econometrica*, 70(6): 2367–420.

Deere, D.R., and J. Vesovic (2006). 'Educational Wage Premiums and the US Income Distribution: A Survey'. In E.A. Hanushek and F. Welch (eds), *Handbook of the Economics of Education, Vol. 1*. Amsterdam: Elsevier.

Griliches, Z. (1977). 'Estimating the Returns to Schooling: Some Econometric Problems'. *Econometrica*, 45: 1–22.

Harmon, C., and I. Walker (1995). 'Estimates of the Economic Return to Schooling for the United Kingdom'. *American Economic Review*, 85(5): 1278–86.

Heckman, J.J., L.J. Lochner, and P.E. Todd (2006). 'Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond'. In E.A. Hanushek and F. Welch (eds), *Handbook of the Economics of Education, Vol. 1*. Amsterdam: Elsevier.

Lang, K. (1993). 'Ability Bias, Discount Rate Bias, and the Return to Education'. Unpublished discussion paper. Boston, MA: Boston University, Department of Economics.

Meghir, C., and S. Rivkin (2011). 'Econometric Methods for Research in Education'. In E.A. Hanushek, S. Machin, and L. Woessmann (eds), *Handbook of the Economics of Education, Vol. 3*. Amsterdam: Elsevier.

Mincer, J. (1958). 'Investment in Human Capital and Personal Income Distribution'. *Journal of Political Economy*, 66(4): 281–302.

Mincer, J. (1974). *Schooling Experience and Earnings*. Cambridge, MA: NBER.

Oreopoulos, P. (2003). 'Do Dropouts Drop Out Too Soon? International Evidence from Changes in School-Leaving Laws'. Working Paper 10155. Cambridge, MA: NBER.

Oreopoulos, P. (2006). 'Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter'. *American Economic Review*, 96(1): 152–75.

Schultz, T.P. (1988). 'Education Investment and Returns'. In H. Chenery and T.N. Srinivasan (eds), *Handbook in Development Economics, Vol. 1*. Amsterdam: Elsevier.

Sussangkarn, C., and Y. Chalamwong (1996). 'Thailand Development Strategies and Their Impacts on Labour Markets and Migration'. In D. O'Connor and L. Farsakh (eds), *Development Strategy, Employment, and Migration*. Paris: OECD.

Warunsiri, S., and R. McNown (2010). 'The Returns to Education in Thailand: A Pseudo-Panel Approach'. *World Development*, 38(11): 1616–25.

Willis, R.J. (1986). 'Human Capital and the Rise and Fall of Families: Comment'. *Journal of Labor Economics*, 4(3): 40–47.