

**HOUSEHOLD EDUCATION EXPENDITURE,
SCHOOL EFFECTS, AND TEACHER
ABSENTEEISM IN MALAYSIA**

SARIMAH BINTI SURIANSHAH

Thesis submitted to the University of Nottingham
for the Degree of Doctor of Philosophy

July, 2019

Abstract

This thesis consists of three independent topics that mainly contribute to the literature on the economics of education. It employs two different datasets, the Households Expenditure Survey (HES) and the Trends in International Mathematics and Science Study (TIMSS). It considers a few methodologies to address the topics. The first topic is a comprehensive study on gender gaps in household education expenditure. It shows tentative evidence of gender gaps in the allocation of household education expenditure. The school effects on student attainment by gender is the second topic. One finding worth noting is that larger classes are positively and significantly associated with high male students' attainment in particular. Two behaviour perspectives are discussed as well: the first is how students' absence from school affects teachers' behaviour, and the second is how teachers' behaviour might affect students' behaviour as well as school performance. As expected, we find that there is significant positive teacher-student behaviour linkage in Malaysian schools. This thesis initially contributes to the survey of the *Jawatankuasa Kajian Dasar Pendidikan Negara (JKDPN)* 2019 under the Malaysia Ministry of Education by providing recommendations to improve the Malaysian education system.

Acknowledgements

First of all, I praise to Allah, for 'this'.

I would like to express my sincere gratitude to my supervisors, Professor Gianni De Fraja and Associate Professor Sarah Bridges who support and guide me a lot in various aspects; to my parents who have prayed and supported me endlessly; to my other family members including my positive friends for their warm support for me; to the institutions especially the Ministry of Education Malaysia, the University of Malaysia Sabah and the University of Nottingham who have gave me the opportunities to be in this unexpected journey; to the Department of Statistics Malaysia for providing me with one of the datasets; and lastly for those support and pray for me either directly or indirectly.

Contents

1	Introduction	1
2	Gender Differences in Household Education Expenditure in Malaysia	5
2.1	Introduction	5
2.2	Background	8
2.2.1	Expenditure on education in Malaysia	8
2.2.2	Gender gaps in student enrolment in Malaysia	9
2.3	Selected literature in gender differences and determinants of household education expenditure	10
2.4	Data	12
2.5	Empirical strategy	18
2.6	Empirical results	24
2.6.1	The determinants of household education expenditure . . .	34
2.6.2	Robustness tests	39
2.7	Conclusion	45
	Appendix A	47
3	School Effects on Student Attainment in Malaysia	57
3.1	Introduction	57
3.2	Background on the Malaysia Education Blueprint 2013-2025: Preschool to Post-Secondary Education	60
3.3	Selected literature on the effect of school factors on student attainment	62
3.4	Empirical strategy	64

3.5	Data and descriptive statistics	67
3.6	Empirical results	75
3.6.1	School effects on student attainment by school location (urban and rural)	88
3.6.2	School effects on student attainment by school size	93
3.6.3	Sensitivity analysis	96
3.7	Conclusion	100
	Appendix B	101
4	Teacher Absenteeism and Student Attainment in Malaysia	116
4.1	Introduction	116
4.2	Background	119
4.2.1	Teacher attendance and leave systems in Malaysia	119
4.2.2	The District Transformation Program	119
4.3	Selected literature on absenteeism: Evidence from teaching and other industries	120
4.4	Empirical strategy	123
4.5	Data and descriptive statistics	126
4.6	Empirical results	130
4.6.1	Teacher absenteeism effect on student attainment	141
4.7	Conclusion	143
	Appendix C	144
5	Conclusion	145
5.1	Summary of results and implications	145
5.2	Future suggestions	147
5.3	Recommendations for Malaysian development	148
	Bibliography	150

List of Figures

2.1	Percentage of student enrolment in Malaysia by education level	9
2.2	Ratio male to female who have enrolled in school for the selected sample	14
2.3	Share of education expenditure by ratio of age-gender and household expenditure on insurance	29
2.4	Share of education expenditure by ratio of age-gender and household expenditure on transportation by road	31
2.5	Relationship between household heads' educational attainment and ratios of children aged 13-17	35
2.6	Relationship between types of family and ratios of children aged 5-12	38
3.1	Cumulative distribution of students' test scores in mathematics by gender	70
3.2	Cumulative distribution of students' test scores in science by gender	71
3.3	Distribution of class size by gender	74
3.4	The relationship between teacher's qualification, age, and students' attainment	82
4.1	The effect of community surrounding school over school size on teacher absenteeism	136
4.2	Probability of teacher motivation (high)	137
4.3	Percent teacher absenteeism by school size	138
4.4	Percent teacher absenteeism by mentoring (often) and school size	139
4.5	Percent teacher absenteeism by teacher evaluation and school size	139

List of Tables

2.1	Summary statistics for the full sample by households whether or not report education expenditure	16
2.2	Estimates of household education expenditure using the hurdle method	25
2.3	Estimates of household education expenditure using the Heckman selection method	26
2.4	First-stage of the Heckman selection method	27
2.5	Estimates of household education expenditure in rural areas . . .	28
2.6	Estimates of household education expenditure in Malay ethnicity	37
2.7	Estimated results of Heckman selection model including a few specifications	40
2.8	Estimates of household education expenditure for households in rural areas of West and East Malaysia	41
2.9	Summary estimates of household education expenditure model . .	44
3.1	Descriptive statistics	72
3.2	Estimates of school effects on student attainment in mathematics	76
3.3	Estimates of school effects on student attainment in science	79
3.4	Estimates of school effects on student attainment in mathematics including clustering and sampling weight	86
3.5	IV estimates of school effects on student attainment in mathematics by school location	89
3.6	IV estimates of school effects on student attainment in science by school location	90

3.7	Estimates of school effects on student attainment by dummy of school size (more than 974 students)	94
3.8	Estimates of school effects on student attainment including dummy class size	97
3.9	Heterogeneity in class size effects	98
4.1	Descriptive statistics	129
4.2	Marginal effects of the teacher absenteeism model	131
4.3	Instrumental variable estimates of teacher absenteeism model . . .	132
4.4	Effects of teacher absenteeism on student attainment	141

Acronyms

B40 Bottom 40%

CIs Confidence Interval

DTP District Transformation Program

GDP Gross Domestic Product

HES Household Expenditure Survey

IT Information technology

IV Instrumental variable

MEB Malaysia Education Blueprint

MOE Ministry of Education Malaysia

MYR Malaysian Ringgit

OLS Ordinary least squares

SIPartner+ School Improvement Partner

SISC+ School Improvement Specialist Coach

SKMM Malaysian Communications and Multimedia Commission

TIMSS Trends in International Mathematics and Science Study

TVET Technical and Vocational Education and Training

UNICEF United Nations Children's Fund

2SLS Two-stage least squares

Chapter 1

Introduction

There is still an ongoing debate among the educationists and economists as to how to improve human capital and economic development. Among earlier researchers in the discipline, [Mincer \(1984\)](#), [Becker \(1994\)](#), [Schultz \(1993\)](#), and [Psacharopoulos \(1994\)](#) highlighted the significant effect of investment in education on economic growth. Recently, a crucial role has been suggested for educational equity and quality ([Hanushek, 2013](#); [Castelló-Climent and Hidalgo-Cabrillana, 2012](#); [Hanushek and Woessmann, 2012](#); [Breton, 2011](#); [Castelló-Climent, 2010](#); [Hanushek and Woessmann, 2008](#)). For instance, [Castelló-Climent and Hidalgo-Cabrillana \(2012\)](#) state that two advantages of higher quality is that: (i) it makes education more accessible, and (ii) encourages an individuals' investment in education.

The Malaysian education system has a long standing issue of inequality in student achievement, and many education reforms have been implemented to lessen the problem. This thesis focuses on three areas that policymakers have focused on to tackle these issues. First, it investigates the incidence of gender gaps in household education expenditure in Malaysia. Second, it examines the effectiveness of recent school policy, framed by the current Malaysian education policy, the 2013-2025 Malaysia Education Blueprint (MEB)¹ on student attainment. Lastly, it examines

¹See subsection 3.2 for details.

the effects student absenteeism and the District Transformation Program (DTP)² have had on teacher absenteeism in Malaysian schools.

Malaysia is unusual compared to many other countries in terms of its student attainment. In general, girls outperform boys at all levels of schooling ([World Bank, 2017](#)). For instance, the MEB reported that in some higher institutions, approximately 70% of total student enrolments were female. One potential reason for the huge gap could be due to the differences in parental spending on their children's education ([Wan, 2018](#)). Chapter 2 provides up-to-date evidence of gender differences in the allocation of education expenditure, and tries to explain this gap. It also discusses other determinants of household decision making in education expenditure. An examination of the gender differences and determinants of household education expenditure may help the government and policymakers to diversify their strategies for the national education budget. This chapter employs hurdle method and Heckman selection method to address significant missing data in the Household Expenditure Survey (HES) dataset. We consider this to be an additional novelty of the chapter in Malaysian context. We find that there are significant gender differentials in the provision of household education expenditure, within households, especially the advantaged ones, favour females compared to males' higher education. Factors such as cultural influence, females' incentives in the labour market, and girls' modesty and safety aspects might explain this finding ([Kenayathulla, 2016](#); [Ismail, 2015](#); [Jelas et al., 2014](#); [Herz and Sperling, 2004](#); [Nasrudin and Nor, 2013](#)).

There is a lack of extensive and empirical studies which have examined the effect of each of the initiatives within the MEB policy on student attainment. Recently, [Aroff \(2014\)](#), [Majid \(2011\)](#), and [Sulaiman et al. \(2017\)](#) among others, have been interested in examining the effect of particular education reforms in the MEB such as the Higher Order Thinking Skills (HOTS) and the School-Based Assessment (SBA) on school outcomes. Examining the MEB effect on student outcomes is

²See subsection 4.2.2 for details.

important as it consists of comprehensive measures for improving the quality of government, schools, and teachers in Malaysia. Thus, Chapter 3 attempts to present the degree to which these measures have been achieved through its effect on student attainment by gender. Specifically, the chapter looks into the impact of teacher incentives, school resources, teacher hiring, principals' role, and parental role in schooling on student outcomes. Findings from the investigation may become a guide for the government or policymakers in making strategic decisions for any future educational plans such as to prepare Malaysian students to be more competitive in international assessments. Chapter 3 provides a novel set of results for Malaysia context when it accounts for class size endogeneity using the instrumental variable (IV) method. The availability of the Trend in International Mathematics and Science Study (TIMSS) dataset is one of the components facilitating the assessment. This is because the dataset provides comprehensive information on student scores in the eighth-grade including several questions that may represent the strategies within the MEB. It also has a suitable proxy to be used as an instrument to address the potential endogeneity problem of the education production model. Findings show a tentative evidence of class size and IT staff effects and a positive effect of teacher hiring, school resources, and parental involvement in schools effects on student attainment ([Hattie, 2005](#); [Woessmann, 2016](#); [Lazear, 2001](#); [Fajolu et al., 2016](#); [Gunnarsson et al., 2009](#)). More specifically, we find the negative impact of class size only when we estimate the model focusing on rural areas and small schools.

One of the initiatives within the MEB policy that has been conducted to improve the equity of education in Malaysia the implementation of the DTP. The annual report of the DTP shows that teacher attendance is a significant factor to lessen the inequality in student attainment ([Ministry of Education Malaysia, 2015](#)). However, the worrying number of disciplinary problems among Malaysian students as reported by the United Nations Children's Fund (UNICEF) in 2016 may significantly affect teacher attendance rates in schools. Although teacher absenteeism

problem in Malaysia might be isolated (for instance, [Vanar \(2018\)](#), “[Hundreds of Teachers](#)” ([2011](#)), [Rahim \(2018\)](#)), if schools fail to act immediately, it is likely that the disciplinary problem in the schools will worsen and spread to other schools. Therefore, Chapter 4, which also utilizes the TIMSS data investigates the effect of student absenteeism and the DTP initiatives on teacher absenteeism. It ends with an investigation of the impact of teacher absenteeism on students’ attainment in school. Similar to Chapter 3, an originality in this chapter is that it addresses a potential reverse causality problem in the estimated model. Besides, to the best of my knowledge, the chapter is the first research on worker absenteeism in the education sector in Malaysia. Findings from the analysis are important to the government and policymakers in deciding on the continuation of the DTP program in the Wave 4 2018 phase and subsequent phases. In addition, it is important to improve quality of education in Malaysia in general. As expected, we find evidence of a positive relationship between student absenteeism and teacher absenteeism. This finding is consistent with those of [Banerjee et al. \(2012\)](#) and [Irvine’s \(1986\)](#) findings. We also find that schools which implement DTP initiatives teaching mentoring, which is about 1 to 3 times a week may reduce the rate of teacher absenteeism by approximately 20 percent. Finally, schools that suffer from teacher absenteeism have lower student attainment.

The next chapters provide a detailed explanation of the chapters that have been explained above. The main findings of each chapter, some recommendations for further research and policy implications are summarised in the final chapter of this thesis.

Chapter 2

Gender Differences in Household Education Expenditure in Malaysia

2.1 Introduction

This study examines evidence on and the determinants of gender differences in household education expenditure in Malaysia. A large number of studies show that females outperform men in terms of school enrolment in Malaysia and many other countries ([Wan, 2018](#); [Jackman and Morrain-webb, 2019](#); [Tienxhi, 2017](#); [Parker et al., 2018](#); [Buchmann et al., 2008](#); [Buchmann and Diprete, 2006](#); [Clark et al., 2008](#); [Majzub and Rais, 2010](#)). For example, in 2011, the Malaysian female student enrolment in post-secondary¹ and tertiary levels in some institutions were about 64% compared with 34% for male student enrolment ([Ministry of Education Malaysia, 2013a,c](#)). The differences in school systems², labour market opportunities, and household preferences for investment in children's education might explain these disparities. However, there are very few studies that explore the gender differences in household education expenditure in Malaysia, and a recent study by [Wan \(2018\)](#) suggests that the disparity in household education expenditure is a potential reason for the male-female gap in student enrolment in Malaysia.

¹Education level before the tertiary education that is either Form 6 (which equivalent to GCE A-level in the UK), pre-university, Matriculation, or certificate level.

²For instance, as discussed by [Kim and Lee \(2010\)](#).

Examining the gender differences and determinants of household education expenditure is crucial for the government and policymakers when making educational spending decisions. In addition, an increase in male education in Malaysia³, in particular, may improve households' social and economic development (Becker, 1994; Bilenkisi et al., 2015; Behrman et al., 2017; Qian and Smyth, 2011; Botha, 2010), reduce child mortality (Cochrane, 1979; Breierova and Duflo, 2004; Vallières et al., 2013; Sichande et al., 2014), and improve economic growth (Chung, 2003; Barro, 2013; Anderson et al., 2003; Glewwe and Jacoby, 2004). Besides, it may help Malaysia to achieve the education policy aspirations (e.g. the Malaysia Education Blueprint 2013-2025 and the Eleventh Malaysia Plan (11MP) Education) and the United Nations Sustainable Development Goals, to promote equity in education.

This study uses an up-to-date and extensive dataset from the Malaysian Household Expenditure Survey (HES) 2014⁴. The dataset is suitable because it provides rich information about both the head of the household and its other members, including the household education expenditures. This study first estimates the household education expenditure model using the two-step hurdle method. The method is chosen because it may help to prevent potential bias estimates given that the HES dataset has a significant number of households that do not report their education expenditure (missing)⁵. Rubin (1976) and Enders (2010) among others show the importance of handling the missing observations by examining patterns to yield probable inferences in estimations. The first step of the hurdle method is to examine whether households enrol their child into school or not, and the second step is to examine the amount to be spent, given that households have decided to enrol their children into school. This study also estimates the household education expenditure model including those households with no reported education expenditure but have enrolled children in school using

³ Given that male students performing worse than female students and they may become household heads.

⁴ A previous study by Kenayathulla (2016) used the Malaysian Household Expenditure Survey, 2004 dataset.

⁵ See Section 2.5 for further explanation and Kingdon (2005).

the Heckman selection approach⁶. The analyses are performed on the Malaysian sample as a whole and by region to suggest appropriate policy responses. Lastly, several robustness checks are performed to support the baseline results.

This study finds significant gender differences in expenditure allocation favouring girls, particularly those in the secondary age group 13-17, with approximately 10% points expenditure higher than their boy counterparts. This finding is in line with [Kenayathulla's \(2016\)](#) findings where in some regions it shows that females outweighed males in terms of household educational spending. However, this study also finds pro-male gaps (i.e. males' expenses are twice as higher as females' expenses) for children aged 5-12 within the Malaysian households as a whole and a subsample of rural areas. These results are robust when we analyse the household education expenditure function regardless of age group, focusing on the sample of households with one child only, and the subsample of households in more detailed regions in Malaysia.

Other than child gender, household characteristics also found to have significant effects in determining the allocation of the children's educational spending. Specifically, high socioeconomic status households, those households headed by graduates and living in urban areas, non-Malay households, and nuclear types of households tend to spend greater amounts on their children's education compared to their counterparts such as Malays and extended types of family households. These findings correspond to studies by [Tilak \(2002\)](#), [Ebaidalla \(2017\)](#), [Donkoh and Amikuzuno \(2011\)](#), and [Tansel \(2002\)](#), among others.

This study is structured as follows. In section 2.2 an overview of household education expenditures in Malaysia and gender gaps in student outcomes in Malaysia is provided. Section 2.3 reviews selected literature on household education expenditure. Section 2.4 describes the dataset including its limitation and summary statistics. Section 2.5 outlines the empirical strategy. Section 2.6 discusses the

⁶Identification strategy for this approach is explained in Section 2.5.

results and concluding remarks are presented in Section 2.7.

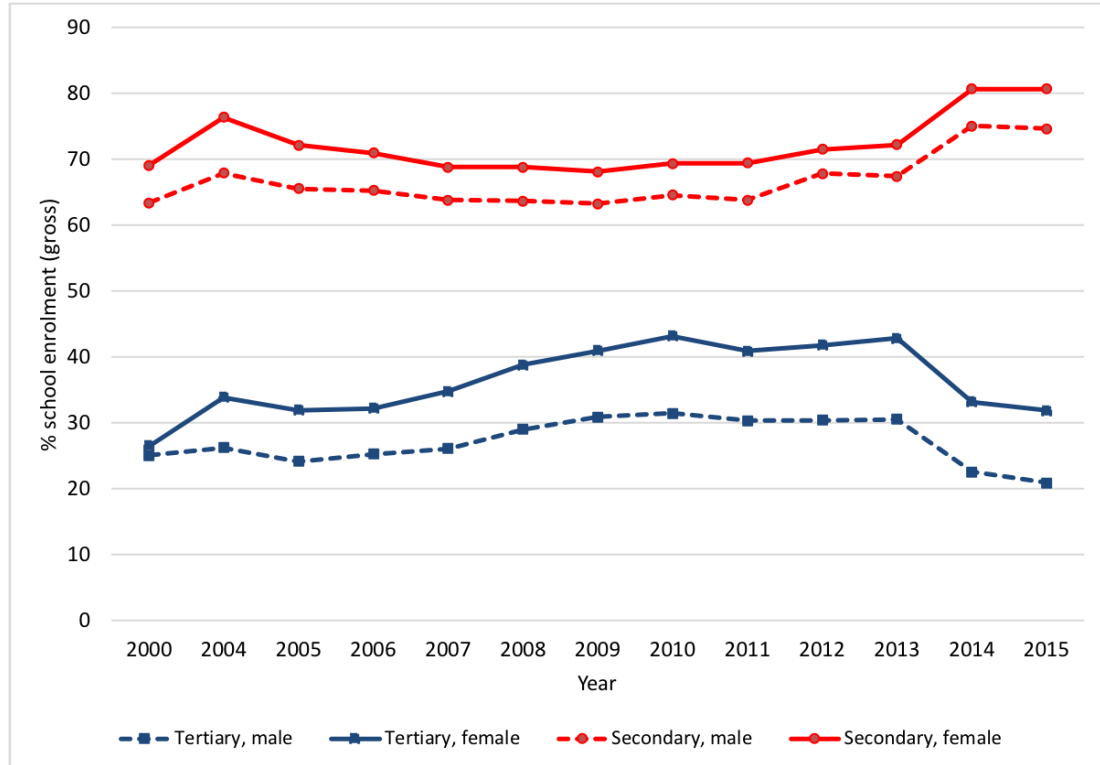
2.2 Background

2.2.1 Expenditure on education in Malaysia

Education in Malaysia at public schools is generally free, with six years of free primary education, which is compulsory education level, and five years of free secondary education. Nevertheless, there are other costs, which consist of contribution fees according to the requirement of schools are payable such as activity fees (e.g. co-curricular, seminar, sports day, and subjects week), membership fees (e.g. the Parent-Teacher Association, cooperative, and monthly class fees), learning materials fees (e.g. exercise books and textbooks), and out of school tuition classes fees ([“Ibu bapa mesti jelas”](#), 2007; [“Ibu bapa masih perlu”](#), 2007; [“Usah berkira”](#), 2008).

At the pre-university level, only small fees are required, with some institutions provide allowances and a level with free of charge, for example, the college community, the Industrial Training Institute (ITI) at the certificate level, and the Centre for Instructors and Advanced Skills Training (CIAST) at the certificate level ([StudyMalaysia.com](#), 2016). However, at the university level (i.e. higher education), students are required to pay fees accordingly. For example, local students who enrolled in public universities rather than private universities only need to pay about 10% of the full fees ([Yun](#), 2017).

2.2.2 Gender gaps in student enrolment in Malaysia



Sources: World Development Indicators and author's calculation.

Notes: Gross enrolment ratio is the ratio of total enrolment, regardless of age, to the population of the age group that officially corresponds to the level of education shown. No significant difference between boys and girls enrolment in the primary education, thus omitted for brevity.

Figure 2.1: Percentage of student enrolment in Malaysia by education level

Figure 2.1 shows the evidence of gender differences in school enrolment in Malaysia. It shows that there is some evidence of gender differences with female enrolment exceeds male enrolment in both education levels over the years. For example, the difference in tertiary education level, in general, have increased from 5 to 10% over a decade before decrease significantly in recent years⁷.

⁷Othman (2014) reported that an introduction of the School-Based Assessment (SBA) (i.e. a student assessment approach that requires teachers to evaluate by their own their students' performance) and the Higher Order Thinking Skills (HOTS) (i.e. a school reforms that need the teaching and learning process to be conducted in high-level) were among the reasons for a sharp decrease in the number of student enrolment in recent years.

2.3 Selected literature in gender differences and determinants of household education expenditure

Many studies find evidence of significant gender differences in the allocation of household education expenditure. Studies in Cote d'Ivoire ([Deaton, 1987](#)), Pakistan ([Aslam and Kingdon, 2008](#)), China ([Song et al., 2006](#); [Burgess and Zhuang, 2000](#)), Nepal ([Khanal, 2018](#)), and India ([Azam and Kingdon, 2013](#); [Saha, 2013](#); [Kaul, 2018](#)) show that households are more likely to spend money on males' than females' education. In contrast, studies in Malaysia by [Kenayathulla \(2016\)](#) and in Thailand by [Wongmonta and Glewwe \(2017\)](#) found that females were preferred to males in the distribution of the household educational spending. They stated that these findings are due to culture changes, where it is becoming more common for daughters who are working to support their elderly parents. Besides, in Malaysia, an increase in females' incentives in the workplace would encourage households to allocate more funds for girls' than boys' education ([Kenayathulla, 2016, 2013](#)).

By considering households' region, [Kenayathulla \(2016\)](#) found that Malaysian households in urban areas in a developed region tend to allocate the education expenditures on girls' education compared to boys' education at age 5-9. Households in the urban areas (rural areas) in less developed regions prefer to allocate their education expenses on boys' (girls') education in the 5-9 age ranges. Similarly, in Thailand, [Wongmonta and Glewwe \(2017\)](#) show that households in the rural areas favoured spending on females over males' education, and vice versa in urban areas, for children in the younger age group (6-11 years). These findings show that the evidence of gender gaps in household education expenditure is inconsistent, and further studies should provide an update and consider the topic extensively as suggested by [Aikman and Rao \(2012\)](#) and [Zimmermann \(2012\)](#). They show that evidence on discrimination in the allocation of education expenditure varies depending on the statistical method, expenditure measure, or the direct (e.g.

materials inputs in relation to children's education) and indirect inputs (e.g. economic, political, and social life) used in the estimation.

There are many studies showing that household education expenditure is positively related to the household's socioeconomic status, including parental education and income ([Donkoh and Amikuzuno, 2011](#); [Ebaidalla, 2017](#); [Qian and Smyth, 2011](#); [Tansel, 1997](#); [Glewwe and Jacoby, 2004](#); [Andreou, 2012](#); [Sulaiman et al., 2012](#); [Vu, 2012](#); [Kim and Lee, 2010](#)). A large effect of the mother's education over the father's education was found in rural China ([Song et al., 2006](#); [Brown, 2006](#)), Ghana ([Sackey, 2007](#)), Turkey ([Tansel, 2002](#)), and Malaysia ([Lillard and Willis, 1994](#)). Specifically, they mainly found that mothers tend to spend on daughters' education, whereas fathers tend to spend on sons' education. However, [Tansel \(1997\)](#) shows that the effect of the father's education was greater compared to the mother's education in Ghana and Cote d'Ivoire. Regarding household income, unexpectedly, [Acar et al. \(2016\)](#) and [Chi and Qian \(2016\)](#) found that only the middle- and low-income families in Turkey and China respectively were willing to spend more money on their children's education. Besides, [Brown \(2006\)](#) found a significant positive effect of both the low and high-income families on their children's human capital development in China. Possibly, these findings have endogeneity problems, as [Donkoh and Amikuzuno \(2011\)](#) have suggested; it is not only current income, but wealth or home possessions such as number of books and other home resources, which also may affect households' preference to invest in their children's education.

Other than household socioeconomic background, households' decisions to allocate expenditure to their children's education also may be affected by the number of family members in the household. Increase in household size has negative effects on children's educational outcomes in the US ([Cáceres-Delpiano, 2006](#)), in Ouagadougou ([Lachaud et al., 2014](#)), in Vietnam ([Dang and Rogers, 2015](#)), in Cyprus ([Andreou, 2012](#)), and in Malaysia ([Sudha, 1997](#); [Pong, 1997](#)). In contrast, a recent study by [Kenayathulla \(2016\)](#) shows a significant and positive

effect of household size on household education expenditure in Malaysia. Her finding has two potential explanations: first, based on economies of scale, a large household size may allow households to share household public goods (e.g. food and clothing) while being able to spend more on education, and second, an increase in the number of school-age children in a household might directly increase the households' total education spending.

To summarise, findings on the existence of gender differences and determinants of household education expenditure differ across countries as well as regions. Thus, this study will provide further evidence about gender differentials in the Malaysia context. Unlike [Kenayathulla's \(2016\)](#) and [Sulaiman et al.'s \(2017\)](#) studies, this study considers the age based on school age groups on each education level and the effects of other household factors including parental characteristics on household education expenditure. The dataset and empirical strategy used to address these questions are explained in the next sections.

2.4 Data

The Malaysian Household Expenditure Survey (HES) is a cross-sectional survey dataset conducted by the Department of Statistics Malaysia. The survey was first conducted in the year 1957/1958 and is carried out at five-year intervals. The two-stage stratified random sampling design was used in the HES, which covered all 13 states⁸ (including three federal territories) and urban and rural strata in Malaysia⁹. The available HES 2014 dataset contains 14,838 randomly selected households, which is 30% of the total sample of the HES dataset¹⁰. It provides rich information about the background of both the head of household¹¹

⁸See Figure A2.1 in Appendix A for Malaysia map.

⁹See [Department of Statistics Malaysia, 2014](#), p. 176 for further details on the HES sampling design. The rural areas are defined as the areas with a population of less than 10,000.

¹⁰The total sample in the HES dataset is 49,862 households ([Department of Statistics Malaysia, 2014](#)). But, due to the confidential data factor, we were only provided with 30% of the sample.

¹¹Head of household is defined as any members whether male or female who is considered as head of household by other members and must be an income recipient and age 15 years and over ([Department of Statistics Malaysia, 2014](#)).

and its other members¹² such as age, marital status, ethnicity, education level, and activity status (e.g. employed, unemployed, student, etc.) including the household total income, total education expenditure, and other types of expenditure¹³. The total household education expenditure consists of household expenditure on pre-primary and primary education, secondary education, post-secondary education, and tertiary education. In detail, the education expenditure items include the household expenses on school fees (government or private), examination fees, tuition fees, membership fees, and fees for higher institution learning (public, private-local or private-overseas).

However, the dataset has some limitations: there is i) no information about education expenditure on each child, it ii) does not distinguish between households with zero education expenditure, if any, and households who refuse to answer the education expenditure question, and iii) data on each child is recorded only once, not over years, making it difficult for further research in examining the implication of this study, for instance, on children's return to schooling. Thus, to address the first limitation, we use household-level data in the estimation and allocate household expenditure on education equally across children in the different age groups. However, we note that expenditure per student does increase with the age of the child and create three samples by age group. A further explanation for the age groups is discussed below. For the second limitation, the two-step hurdle model and Heckman selection model approaches are employed, which will be discussed in the next section.

¹²Wife or husband of the head of household, which we refer to as spouse; unmarried children of the head of household; married children of the head of household; daughter or son-in-law; grandchildren; father or mother of the head of household or spouse of the head of household; grandfather or grandmother of the head of household or spouse of the head of household; siblings of the head of household or spouse of the head of household; other relatives of the head of household or spouse of the head of household; maid; and non-related relatives of the head of household or spouse of the head of household.

¹³Other types of expenditure included in the survey are: food and non-alcoholic beverages; alcoholic beverages and tobacco; clothing and footwear; housing, water, electricity, gas and other fuels; furnishings, household equipment and routine household maintenance; medical products, appliances and equipment; transport; communication; recreation services and culture; restaurants and hotels; and miscellaneous goods and services.

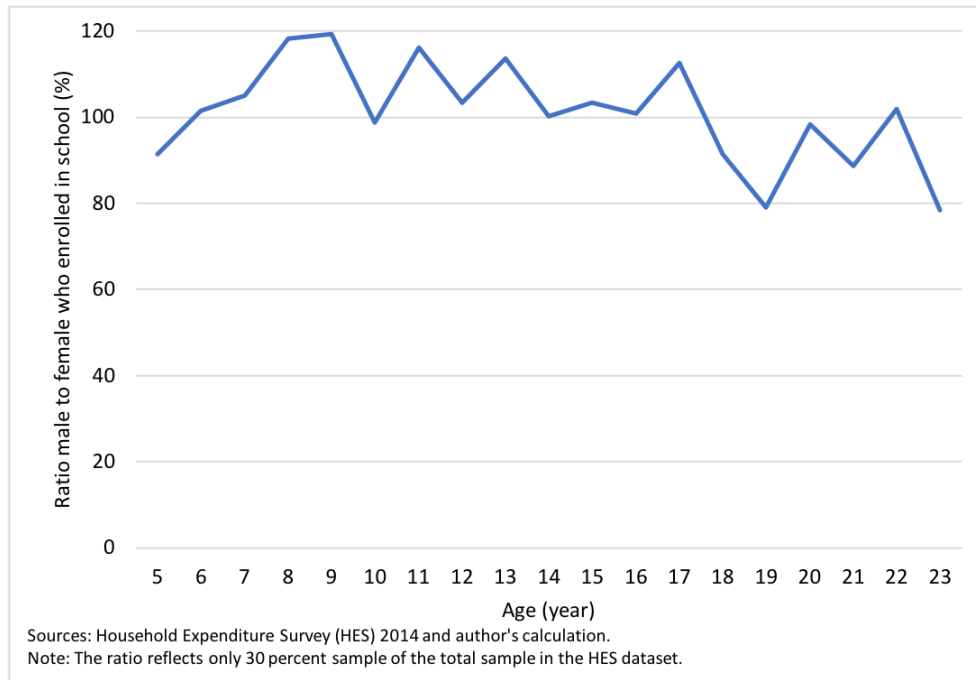


Figure 2.2: Ratio male to female who have enrolled in school for the selected sample

We have restricted the sample to households that have at least a child or member¹⁴ between 5 and 23 years of age, which reduces the sample to 8,332¹⁵. The lower limit of the sample (age 5) was chosen to represent the minimum age for children who enter pre-primary education, while the upper bound (age 23) was chosen to limit the extent to which children, in general, have finished the first-degree study and will start to live independently. To check potential sample selection bias in the selected age, Figure 2.2 shows the pattern of the ratio of studying males to females by age. The figure illustrates that there are more boys than girls who are studying in the primary and secondary education levels (7 to 17 years), while there are more girls than boys who are studying in the higher education level (18 years and above). In comparison to Figure 2.1 of Section 2.2, we could say that males are over-represented in the HES sample relative to the population.

¹⁴This study considered other members such as grandchildren, sibling of the head of household or spouse, other related members, and non-related members that possibly received the household education expenditure.

¹⁵This study excludes sample of households with no children (29.73%), households with children less than 5 years old and/or more than 23 years old (about 20% each), households with unmarried head of households who support his or her parents and/or siblings (12%), households with no family members (5%), and households with more than one spouse (0.03%).

Briefly, among potential reasons for the high enrolment of males at the primary education level is that parents might regard their sons as future caregivers for the family or those who will contribute to the family income. Thus, parents are more likely to invest in the sons' compared to daughters' education. Another potential explanation for a fewer number of girls enrolled in the primary education is related to the parental choice to obtain an alternative education for their children such as homeschooling for their daughters in particular. As of 2015, it was reported that approximately 5,000 children were homeschooled in Malaysia ([Thevadass, 2018](#)). [Kim-Soon et al. \(2015\)](#) show that social and school factors are among the reasons for parents in Malaysia to favour homeschooling for their children. However, boys' high interest to work rather than pursue higher education might be a potential reason for them to leave school earlier than girls, as discussed by [Jelas et al. \(2014\)](#) and [Ismail \(2015\)](#).

From the selected main sample, we then construct three separate samples of the household which are households with at least a child or member in the following age ranges: i) 5-12 years, ii) 13-17 years, and iii) 18-23 years, on the assumption that households allocate a different amount of expenditure based on children's age level (see [Deaton, 1997](#), p. 223). These samples are constructed based on the school age level in the pre-primary and primary, secondary, and tertiary education in Malaysia. However, to note that for each age sample the estimation includes a variable to capture if the household has children in other age groups.

Table 2.1: Summary statistics for the full sample by households whether or not report education expenditure

Variables	Monthly total household education expenditure (1=Observed)				Monthly total household education expenditure (0=Not observed)			
	Mean	Standard deviation	Minimum	Maximum	Mean	Standard deviation	Minimum	Maximum
School enrolment decision (1 =Yes)	0.992	0.088	0	1	0.847	0.360	0	1
Monthly total household education expenditure (MYR)	96.995	194.898	0	5427	0	0	0	0
Monthly total household expenditure (MYR)	3791.969	2618.052	521	40094	3272.785	2207.336	523	45586
Monthly gross household total income (MYR)	6197.162	5304.883	723	66680	5646.920	5057.199	529	84567
Parental characteristics								
Father's education (base: No formal)								
Primary	0.082	0.274	0	1	0.126	0.332	0	1
Secondary	0.527	0.499	0	1	0.488	0.500	0	1
Tertiary	0.181	0.385	0	1	0.131	0.337	0	1
Mother's education (base: No formal)								
Primary	0.089	0.285	0	1	0.151	0.358	0	1
Secondary	0.576	0.494	0	1	0.520	0.500	0	1
Tertiary	0.161	0.367	0	1	0.107	0.309	0	1
Father's age	42.857	7.367	25	65	44.118	8.541	24	67
Mother's age	40.566	7.041	24	67	42.001	8.067	23	68
Father missing	0.199	0.399	0	1	0.228	0.420	0	1
Mother missing	0.153	0.360	0	1	0.182	0.386	0	1
Household characteristics								
Types of family (base: Nuclear family)								
Single-parent family	0.036	0.186	0	1	0.049	0.215	0	1
Extended family	0.275	0.447	0	1	0.292	0.455	0	1
Others	0.001	0.035	0	1	0.003	0.052	0	1
Female headed	0.088	0.283	0	1	0.104	0.305	0	1
Rural	0.318	0.466	0	1	0.355	0.479	0	1
West Malaysia	0.748	0.434	0	1	0.657	0.475	0	1
Ethnic Malays	0.756	0.429	0	1	0.754	0.431	0	1
Non-biological parents	0.141	0.348	0	1	0.163	0.369	0	1
Number of households	4378				3954			

Notes: Total household education expenditure consists of household expenditure on pre-primary and primary education, secondary education, post-secondary education, and tertiary education. Total household expenditure consists of expenditure on: education; food and non-alcoholic beverages; alcoholic beverages and tobacco; clothing and footwear; housing, water, electricity, gas and other fuels; furnishings, household equipment and routine household maintenance; medical products, appliances and equipment; transport; communication; recreation services and culture; restaurants and hotels; and miscellaneous goods and services. The non-biological parents refer to parents for households who have not fulfilled the parent-child age restrictions where the age difference between the household head (and/or spouse) and oldest child of the household is in the range of 18-45 years, and the age difference between the household head (and/or spouse) and youngest child of the household is in the range of 20-45 years (see Figures A2.2 and A2.3 in Appendix A for graphical illustration of the restrictions). We suspect them to be those who have married early (e.g. at the age of less than 20 years), married late (e.g. at the age of more than 45 years), married to an older or a younger woman, parents of the married child of the household with children, or individuals with data entry errors. Rural areas are defined as the areas with a population of less than 10,000. MYR refers to Malaysian currency, Malaysian Ringgit. This full sample refers to sample of households with school age children between 5 and 23 years.

Table 2.1 reports summary statistics for the full sample for households who have reported their education expenditure. We also show the summary statistics for households who do not report their education expenditure for general information purposes only. On average, households spend on education approximately 2.7% of their total spending. The sample consists of approximately 10% households with no child enrolled in school, and to note, half of the households with studying child do not report their education expenses. Based on the sample of school age group (see Table A2.1 in Appendix A), the share of education expenditure in the households with 13-17 year-old children was slightly higher than the households with children in other age groups (5-12 and 18-23 years).

Table 2.1 also reports the households' socioeconomic characteristics. A majority of the households (households with biological parents) have secondary education, are categorised as nuclear types of family, lived in urban areas, and Malays ethnicity. The sample also consists of 15% of households with non-biological parents. The non-biological parents refer to parents for households who have not fulfilled the parent-child age restrictions where the age difference between the household head (and/or spouse) and oldest child of the household is in the range of 18-45 years, and the age difference between the household head (and/or spouse) and youngest child of the household is in the range of 20-45 years (see Figures A2.2 and A2.3 in Appendix A for graphical illustration of the restrictions). We suspect them to be those who have married early (e.g. at the age of less than 20 years), married late (e.g. at the age of more than 45 years), married to an older or a younger woman (a potential explanation for the households with parents aged 65 and above), parents of the married child of the household with children (extended families including grandparents, parents, and children), or individuals with data entry errors.

2.5 Empirical strategy

This chapter considers two approaches to examine the evidence on, and the determinants of, gender differences in household education expenditure: the hurdle model and the Heckman selection model. The hurdle model estimates carefully the household education expenditure model in two stages, decision to spend stage and how much to spend stage, excluding the missing observations, while the Heckman selection model estimates the household education expenditure model taking into account the missing observations. In each model the Working-Leser expenditure specification is employed.

Unlike many previous studies¹⁶, this study straightforwardly estimates the household education expenditure model using the two-step hurdle method as proposed by [Kingdon \(2005\)](#). We do so to prevent potential bias estimates in the household education expenditure model given that the HES dataset consists of a significant number of households with missing education expenditure. However, we also estimate the model using the standard Working-Leser model as a robustness check. [Kingdon \(2005\)](#) has discussed that there are two possible channels through which pro-male bias may occur in expenditure on any particular community. The first channel is via zero purchases for daughters (nonenrolment for girls) and positive purchases for sons. The second channel is via lower expenditure on daughters than on sons, conditional on positive education expenditure for both daughters and sons. [Kingdon \(2005\)](#) has raised a concern if gender bias operates mainly through the first channel but that, conditional on enrolment, expenditure on girls is similar or exceeds that on boys. “Averaging across the two mechanisms, there may not be significant gender bias but, via the nonenrolment mechanism only, there may be strong” ([Kingdon, 2005](#), p. 410), as what happened in [Subramanian and Deaton’s \(1991\)](#) study, among others.

¹⁶For example, [Wongmonta and Glewwe \(2017\)](#) use the Engle Curve Ordinary Least Squares and Tobit approach, [Saha \(2013\)](#) uses the Oaxaca-Blinder decomposition approach, [Qian and Smyth \(2011\)](#) apply the Tobit model, and [Song et al. \(2006\)](#) and [Subramanian and Deaton \(1991\)](#) use the conventional Engle Curve OLS method.

The hurdle model is a count model that permits the probability of zero (e.g. nonenrolment of children) and positive data are constrained not to be identical (Mullahy, 1986). The model was first suggested by Cragg (1971). Based on Cragg, 1971, p. 831, the hurdle model may apply on count data, where the response variable is a nonnegative integer, with the following situations; a decision first has to be made, for example, to consider a change or not, then a decision on the amount of the change is taken. The first decision might be represented by the probit model, and the second by a standard regression model. Hence, in this study, the first step is to estimate whether households enrol their child into school or not, and the second step is to estimate how much expenses are spent on children's education conditional on households with children enrolled in school only. In both equations, we employ the standard Working-Leser Engle expenditure specification that was introduced by Working (1943). The Engle expenditure specification is chosen due to data limitation that requires the household-level data is employed instead of child individual-level data in the estimation. Zimmermann (2012) has shown that the approach will allow the use of household-level data to examine the evidence of gender bias within households. The approach will consider the share of particular goods of the household total expenditure rather than the actual expenditure on the particular goods as the dependent variable of the model. It has been widely used in literature on expenditure such as by Hoddinott and Haddad (1995), Kenayathulla (2016), Acar et al. (2016), Song et al. (2006), Dai et al. (2012), and Aslam and Kingdon (2008), and has the following general expression:

$$(p_i q_i)^h = f_i(x, a, z, u) \quad (2.1)$$

where $p_i q_i$ is expenditure on good h in household i , x is household total expenditure, a is household demographic composition, z is other household characteristics, and u is an error term. A few functional forms of the model have been discussed such as in Leser (1963) and Holcomb et al. (1995), and this study follows the model

proposed by Deaton (1987). Deaton's (1987) model includes a vector of household age-gender categories that will explain whether girls get less expenditure than boys based on age group.

The hurdle model can be expressed as follows: equation (2.2) is the probit model, with a dependent variable equivalent to one if households have children enrolled in school and zero if households do not have children enrolled in school. Equation (2.3) is the model of household education expenditure that shows how much expenditure a household will allocate on their children's education, given that the household has at least one child who is enrolled in school.

$$Pr(W_i = 1|x_i) = f(x_i'\beta) \quad (2.2)$$

$$\ln W_i = \alpha + \delta \ln(Y_i / N_i) + \sum_{k=1}^m \lambda_k (N_{k,i} / N_i) + \gamma D_i + \varepsilon_i \quad (2.3)$$

where $Pr(W_i = 1|x_i)$ is the probability that household i has decided to spend, given x is a vector of explanatory variables, f is cumulative distribution function of the standard normal distribution, and β are the parameters to be estimated typically by maximum likelihood. Using the maximum likelihood estimator, equation (2.2) will yield coefficients β which maximize the log-likelihood function $f(x_i'\beta)$ ¹⁷. But for interpretation, we then compute the marginal effects of the probit estimates that explain the change in $Pr(W_i = 1)$ to a one-unit change in variable of interest, holding other explanatory variables at their means. For equation (2.3), the $\ln W_i$ is the log share of education expenditure in household i of the household total expenditure, Y_i is the monthly household total expenditure, N_i is the household size, $N_{k,i}$ is the number of children or members¹⁸ in the

¹⁷ $L(x_i'\beta) = \prod_{k=1}^n (f(x_i'\beta)^{y_i} [1 - f(x_i'\beta)]^{(1-y_i)})$, where x' denotes a 1 x K row vector of regressor values for observation i , and β are a K x 1 column vector of regression coefficients.

¹⁸To note, this study considered other members who are enrolled in schools that may be non-biological children to the household head. They may be the grandchildren, sibling of the head of household or spouse, other related members, and non-related members that possibly receive the household education expenditure.

k th age-gender group, D_i is the vector of parental and household characteristics such as parental education, parental age, types of family, household residential region, and ethnic group. ε is an error term, and α , δ , λ , and γ are parameters to be estimated. This study utilises the share of household education expenditure in logarithmic form in order to express the change in the share relative to the initial values. In addition, since different households may have different items of education expenditure, logarithmic transformation would enable the estimated results to be compared.

The λ_k is the coefficient of interest in this study that will explain the percentage change of an addition of one child in the k th age-gender group on the household's share of education expenditure. The share of household education expenditure is measured by the ratio of total household expenditure in education, which is the sum of the household's total expenditure on pre-primary and primary education, secondary education, post-secondary education, and tertiary education, to the total household expenditure. Three age groups are considered in this study that correspond to the school age level in Malaysia. Specifically, each group takes into account three measures. For instance, in the sample of households with at least a child aged 5 to 12 years, the ratio of males aged 5-12, ratio of females aged 5-12, and others, which refers to children in ages other than 5-12 years old are considered. To examine whether there are gender differences in household education expenditure in each age group, for example, whether children aged 5-12 are treated differently by gender, we refer to an F-test for equality of coefficients of the 5-12 age group by gender in addition to the direct interpretation of respective coefficients. [Snedecor and Cochran \(1989\)](#) claim that the F-test is a standard method to examine whether the variance of two sample means is different or not.

Next, this study estimates the household education expenditure model including those households that have decided to enrol a child in school but not provide information on education expenditure. This measure is taken mainly to prevent possible loss of data in the estimation given that half of the households in the

sample do not report their education expenditure¹⁹ (Rubin, 1976; Enders, 2010). Furthermore, addressing the missing observations problem may check potential bias estimates in the hurdle model. Conceptually, we regard them as non-missing because as mentioned in Section 2.2, even though education in Malaysia is free at the public primary and secondary education, there must be some expenses on schooling incurred by the families. The households may refuse to answer the education expenditure question due to several reasons such as the households only spend a small amount of money on education because they are categorised as low-income families, or the head of household is not a permanent income recipient. We employ the Heckman selection model (Heckman, 1976, 1979) after examining that the missing observations could be classified as missing at non-random (see Figures A2.4 and A2.5 in Appendix A for evidence). The missing at non-random status may arise when the missing values on a variable are related to the values of the variable itself, even after controlling for other variables. None of the figures show persistency that the probability of missing data on education expenditure is related to household socioeconomic backgrounds. Thus, we expect that variables of whether the household receives a low or high amount of school financial aid, and whether or not a household is headed by a self-employed individual would become selected variables to correct the missing observations. This study utilises the households who are not in below 40% (B40) income group status²⁰ to represent those low school aid recipient households. They are expected to be those who are less likely to receive school financial aid due to a few education endowments such as the Tuition Aid Scheme, the Poor Student's Trust Fund, the Supplementary Food Scheme, and the National Higher Education Fund Corporation (*PTPTN*) that are mainly provided for students that come from the B40 income group status homes (Ministry of Education Malaysia, 2019). However, excellent students that

¹⁹See Table 2.1 of Section 2.4.

²⁰Households in Malaysia are categorised into three main group of income: the top 20% (T20) that refers to households with monthly income of MYR8,319 and above, the middle 40% (M40) that refers to households with monthly income of between MYR3,860 and MYR8,319, and the bottom 40% (B40) that refers to households with monthly income of below MYR3,860. The cut-off of the income group is based on the Report on Household Income and Basic Amenities 2014 by the Department of Statistics Malaysia.

come from the middle- and top-income status home also have chances to receive some of this aid. Next, households may not report their education expenses due to a financial instability problem. Households with self-employed household heads are those that we expect tend to encounter this situation due to unstable earnings that they have which might depend on weather, customer behaviour, and so on.

The Heckman selection model can be derived as follows. Consider the following general selection rule:

$$z_i = \mu k_i + u_i \quad (2.4)$$

where z_i is the dependent variable of whether household i has reported education expenditure or not, k_i is the vector of the selected variables, and u_i is the error term.

Let the population function for equation (2.3) be written as follows:

$$E(W_i|X_i) = X_i\Theta \quad (2.5)$$

where X_i is the vector of the explanatory variables. In the assumption that the conditional expectation of ε of equation (2.3) is nonzero, then the selected sample regression function²¹ can be written as follows:

$$E(W_i|X_i, z_i > 0) = X_i\Theta + E(\varepsilon_i|u_i > -\mu k_i) \quad (2.6)$$

where the regression function depends on X_i and k_i .

Lastly, for further analysis, the household education expenditure model by households' region and interaction terms between the explanatory variables are estimated to identify the determinants of the household education expenditure comprehen-

²¹See [Heckman \(1979\)](#) for further details.

sively. Our intuition to estimate the household education expenditure by region is to examine to what extent the government incentives²² may help to reduce households' burden in education spending, in general, and to reduce inequity in the allocation of households' education expenditure by gender within the region, if any.

2.6 Empirical results

This section first presents the evidence of gender differences in education expenditure within the three school-age groups using the Malaysia sample as a whole. Then it shows the estimated results of the education expenditure by region and other determinants of household education expenditure. Findings of the robustness checks for the baseline models are presented at the end of this section.

²²For example, the Tuition Aid Scheme, the Poor Students' Trust Fund, the Supplementary Food Programme and the Early Education Scholarship, which implemented under the Ministry of Education Malaysia, and the *Tuisyen Rakyat@KKLW* programme, which conducted by the Ministry of Rural and Regional Development Malaysia.

Table 2.2: Estimates of household education expenditure using the hurdle method

Dependent variables:	5-12 years		13-17 years		18-23 years	
Education expenditure (1 = Observed); Education expenditure (log)	Probit model (1)	OLS model (2)	Probit model (3)	OLS model (4)	Probit model (5)	OLS model (6)
Ratio of males aged 5-12	0.033*** (0.010)	0.589** (0.263)				
Ratio of females aged 5-12	0.048*** (0.012)	0.214 (0.259)				
Ratio of males aged 13-17			0.007 (0.006)	0.904*** (0.332)		
Ratio of females aged 13-17			0.014** (0.007)	1.151*** (0.328)		
Ratio of males aged 18-23					-0.356*** (0.043)	-0.270 (0.456)
Ratio of females aged 18-23					-0.348*** (0.045)	0.404 (0.441)
F-test for equality by gender						
5-12	3.07**	2.72*				
13-17			1.48	0.91		
18-23					0.06	3.28*
Number of households	5340	3133	3727	2598	3647	1609
Adjusted R-squared		0.13		0.15		0.13

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level. Education expenditure refers to the total household education expenditure.

The ratio of males (females) is defined as the relative number of males (females) in each age group to the total members in the household. Omitted variable in each age group is others, which refer to children in other ages.

Table 2.2 shows the estimated results of the coefficient of interest of this study, ratio of age-gender categories, on the household education expenditure using the hurdle method²³. The coefficient differences test, an F-test that indicates differences between boys and girls if any, is provided at the second bottom section of the table. Overall, we found a low or not significant difference between males' and females' education expenditure. Only in the younger age range, 5-12 years, is the presence of boys is slightly associated with a less household educational spending compared to the presence of girls. In general, this weak effect might be related to the male or female ratios in a household. If households have a very low ratio of female at the primary age group, they probably have less expenditure on education. However, we also expect that significant missing observations in the full sample contribute to this weak result and lead us to employ the Heckman selection approach in addressing the problem.

²³To note, this study also estimates the household education expenditure model using the standard Engle-Curve method but found no significant results (see Table A2.3 in Appendix A).

Chapter 2. Gender Differences in Household Education Expenditure in Malaysia

Table 2.3: Estimates of household education expenditure using the Heckman selection method

Dependent variable:	5-12 years		13-17 years		18-23 years	
Share of education expenditure (log)	(1)		(2)		(3)	
Ratio of males aged 5-12	0.793***	(0.151)				
Ratio of females aged 5-12	0.391***	(0.138)				
Ratio of males aged 13-17			0.865***	(0.161)		
Ratio of females aged 13-17			1.217***	(0.167)		
Ratio of males aged 18-23					-0.200	(0.379)
Ratio of females aged 18-23					0.399	(0.271)
Father's education (base: No formal)						
Primary	0.576***	(0.149)	0.246*	(0.147)	0.389***	(0.137)
Secondary	0.497***	(0.156)	0.143	(0.147)	0.159	(0.138)
Tertiary	0.527***	(0.155)	0.294*	(0.166)	0.216	(0.153)
Mother's education (base: No formal)						
Primary	0.476***	(0.096)	0.158	(0.122)	0.258**	(0.123)
Secondary	0.562***	(0.104)	0.580***	(0.141)	0.330**	(0.134)
Tertiary	1.027***	(0.123)	1.094***	(0.163)	0.821***	(0.157)
Father's age	-0.026	(0.029)	0.205***	(0.046)	0.058	(0.050)
Mother's age	0.049*	(0.026)	-0.070*	(0.041)	-0.108**	(0.049)
Father's age squared	0.001	(0.000)	-0.002***	(0.000)	-0.001	(0.001)
Mother's age squared	-0.001**	(0.000)	0.001	(0.000)	0.001**	(0.001)
Father missing	0.515***	(0.179)	0.260	(0.192)	0.298	(0.200)
Mother missing	0.083	(0.152)	0.123	(0.197)	-0.016	(0.228)
Types of family (base: Nuclear family)						
Single-parent family	-0.157*	(0.091)	-0.142	(0.122)	0.109	(0.129)
Extended family	-0.199***	(0.032)	-0.095**	(0.039)	-0.173***	(0.039)
Others	-2.304***	(0.374)	-1.613***	(0.470)	-1.183***	(0.329)
Female headed	0.418***	(0.078)	0.114	(0.085)	0.024	(0.106)
Rural	-0.303***	(0.027)	-0.253***	(0.035)	-0.227***	(0.036)
West Malaysia	0.770***	(0.068)	0.979***	(0.068)	0.649***	(0.060)
Ethnic Malays	-0.620***	(0.031)	-0.865***	(0.045)	-0.684***	(0.046)
Non-biological parents	0.390**	(0.167)	0.156	(0.204)	0.031	(0.235)
Constant	-7.290***	(0.679)	-9.843***	(1.058)	-4.147***	(1.145)
F-test for equality by gender						
5-12	15.22***					
13-17			6.51***			
18-23					8.85***	
Number of observations	5191		4477		3054	

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level. Share of education expenditure is the ratio of total household education expenditure to the total household expenditure. The ratio of males (females) is defined as the relative number of males (females) in each age group to the total members in the household. Omitted variable in each age group is others, which refer to children in other ages. The non-biological parents refer to parents for households who have not fulfilled the parent-child age restrictions (see Figures A2.2 and A2.3 in Appendix A for graphical illustration of the restrictions). See the Data Section of this study for more details of the variables.

The estimated results of the household education expenditure using the Heckman selection method is shown in Table 2.3. The first stage estimates of the method are presented in Table 2.4, which employs two selected variables, dummies of school aid recipient household and self-employed head of household, as instruments to address missing observations in the main sample. Table 2.4 shows some evidence that selection bias exists since the coefficient of the inverse Mills ratios (λ) is statistically significant, and the ρ , which is the correlation between errors in the selection and the outcome equation, is non-zero. Thus, we can state that the estimated results in Table 2.2 yields biased estimates. All coefficients of the selected variables also show statistically significant effects and indicate that the expectations of this study are supported²⁴.

Table 2.4: First-stage of the Heckman selection method

Dependent variable: Education	5-12 years		13-17 years		18-23 years	
expenditure (1 = Observed)	(1)		(2)		(3)	
School-aid recipient (1 = Low)	0.047***	(0.017)	0.076***	(0.018)	0.013	(0.022)
Self-employed head of household	-0.108***	(0.018)	-0.079***	(0.019)	-0.127***	(0.022)
Lambda	0.81*	(0.41)	1.83***	(0.51)	-0.28	(0.45)
Censored observations	2058		1879		1445	
Rho	0.50		0.89		-0.19	
Chi-squared	1301.10***		1052.00***		1111.11***	

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level. Education expenditure refers to the total household education expenditure.

Other explanatory variables are omitted for brevity.

The evidence of gender differentials in household education expenditure is shown by the significance value of the F-test for equality by gender in each column of Table 2.3. Overall, it shows that there are significant differences in educational spending between boys and girls in the primary and secondary age groups. Specifically, households preferred to spend money on boys' education about double of the girls' education expenditure at the 5-12 years group. However, households tend to spend 10% points less on boys' education aged 13-17 as opposed on their girl counterparts if their ratio increased by 0.1 unit. The explanation for other explanatory variables in Table 2.3 is provided in the next subsection.

²⁴As discussed in the empirical strategy's section, this study expects that the probability of a household to report education expenditure is high if the household is less likely to receive school aids and not headed by a self-employed individual.

Table 2.5: Estimates of household education expenditure in rural areas

Dependent variable:	5-12 years		13-17 years		18-23 years	
Share of education expenditure (log)	OLS model	Heckman model	OLS model	Heckman model	OLS model	Heckman model
	(1)	(2)	(3)	(4)	(5)	(6)
Ratio of males aged 5-12	0.477 (0.466)	0.499*** (0.187)				
Ratio of females aged 5-12	-0.266 (0.485)	-0.262 (0.202)				
Ratio of males aged 13-17			0.878 (0.557)	2.731*** (1.049)		
Ratio of females aged 13-17			1.653*** (0.550)	3.637*** (1.089)		
Ratio of males aged 18-23					-0.752 (0.852)	-1.564** (0.660)
Ratio of females aged 18-23					0.727 (0.803)	0.255 (0.465)
F-test for equality by gender						
5-12	3.50*	20.53***				
13-17			3.23*	2.60*		
18-23					4.55**	21.27***
Lambda		-0.03 (0.69)		5.02** (2.02)		1.27 (1.05)
Number of observations	985	1684	867	1546	521	1031
Adjusted R-squared	0.16		0.19		0.12	
Censored observations		699		679		510
Rho		-0.02		1.00		0.73
Chi-squared		922.93***		114.27***		354.53***

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level. Share of education expenditure is the ratio of total household education expenditure to the total household expenditure. The ratio of males (females) is defined as the relative number of males (females) in each age group to the total members in the household. Omitted variable in each age group is others, which refer to children in other ages. See Table A2.4 in Appendix A for full results.

Regarding household region (Table 2.5), we only find significant gender differentials in the allocation of educational spending, with high spending on the girls compared to the boys' education (about double) for those aged 13-17 in the rural areas (see first the F-test and Lambda results in Table 2.5). There are no significant results of gender differences in education expenditure found in the urban areas, as presented in Table A2.5 in Appendix A. The significant effect on the share of household education expenditure in the presence of girls than boys aged 13-17 years in the rural areas is supported by [Jelas et al. \(2014\)](#), [Jelas and Dahan \(2010\)](#), and [Patel \(2014\)](#). They show that boys in the secondary age group are more likely to drop out of school than girls due to financial constraints, among other reasons. These findings suggest for parents and schools to take further interventions to engage the disadvantaged boys in school in particular.

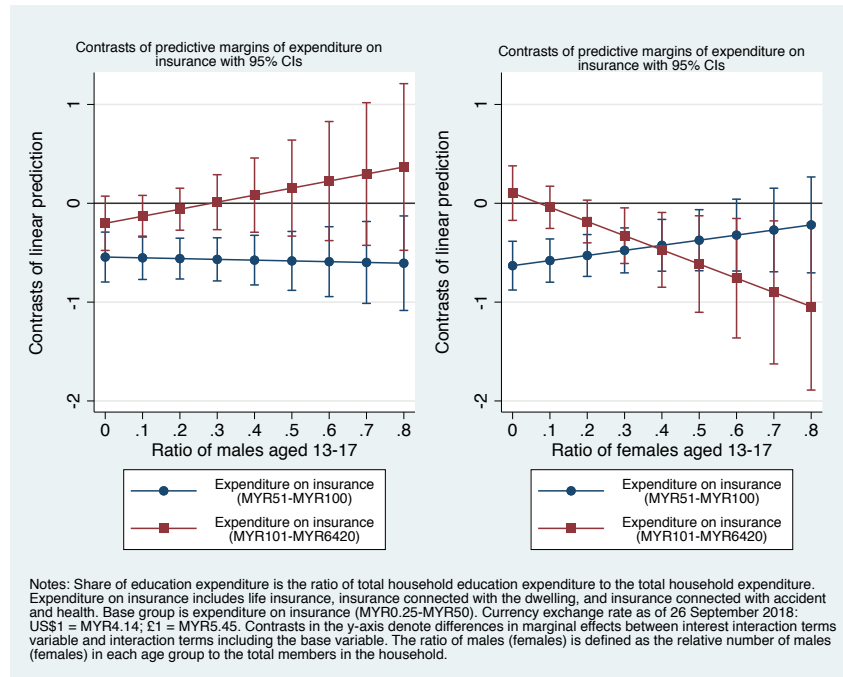


Figure 2.3: Share of education expenditure by ratio of age-gender and household expenditure on insurance

There is a number of potential reasons for higher educational spending on girls' than boys' education in Malaysia. First, it may be due to stakes and cultural influence in the household. Parents may be willing to spend more on daughters' educational resources than sons as a future economic incentive for the parents (Kenayathulla, 2016), as well as for the daughters. Additionally, the filial piety culture, especially among the Chinese Malaysian population, might encourage children to provide better resources for their elderly parents. Hence, this study expects that parents favour daughters' education in aspiration for better future resources. This expectation is based on findings by Ismail (2015) and Jelas et al. (2014), which show that females tend to achieve better outcomes as they have shown great attitudes, aspirations, and hard work regarding education. Figure 2.3 shows the evidence of differences (contrasts) in marginal effects between households with high expenditure on insurance and low expenditure on insurance on the household share on education expenditure as the ratio of children aged 13-17 years increases. As expected, as parents expect better returns on education from their

daughters, they have spent less on insurance expenses²⁵ and possibly prefer to spend more on their daughters' education. Next, better education for females will provide better socioeconomic security for themselves, as suggested by [Alderman and King \(1998\)](#). For example, regarding the culture of the allocation of the share of inheritance in the Muslim households, typically where the household has both male and female children, a daughter's share is equal to one-half of a son's share ([Muslim Women's League, 1995](#)). This culture may force parents to concentrate their investment in females compared to males' education in order to maintain the social and economic future of their daughters.

Second, an increase in women's facilities and welfare in the workplace potentially encourages households to spend more on females' than males' higher education in aspiration to have better returns on education²⁶, given that females show great performance in schooling ([Ismail, 2015](#); [Jelas et al., 2014](#)). [Alderman and King \(1998\)](#) show that children's broad market opportunities may increase parents' willingness to spend on the benefited children's education. [Kenayathulla \(2016\)](#) has pointed out a few incentives introduced by the government for women in the workforce, for example, the child-care centres, paid maternity leave, and flexible working hours. These incentives may be stimulator for why households are willing to spend more on girls' than boys' education.

²⁵The expenditure on insurance refers to expenditure on life insurance, insurance connected with the dwelling, and insurance connected with accident and health.

²⁶The returns on education might refer to the earnings of a person after she or he has finished investing in human capital ([Becker and Chiswick, 1996](#)).

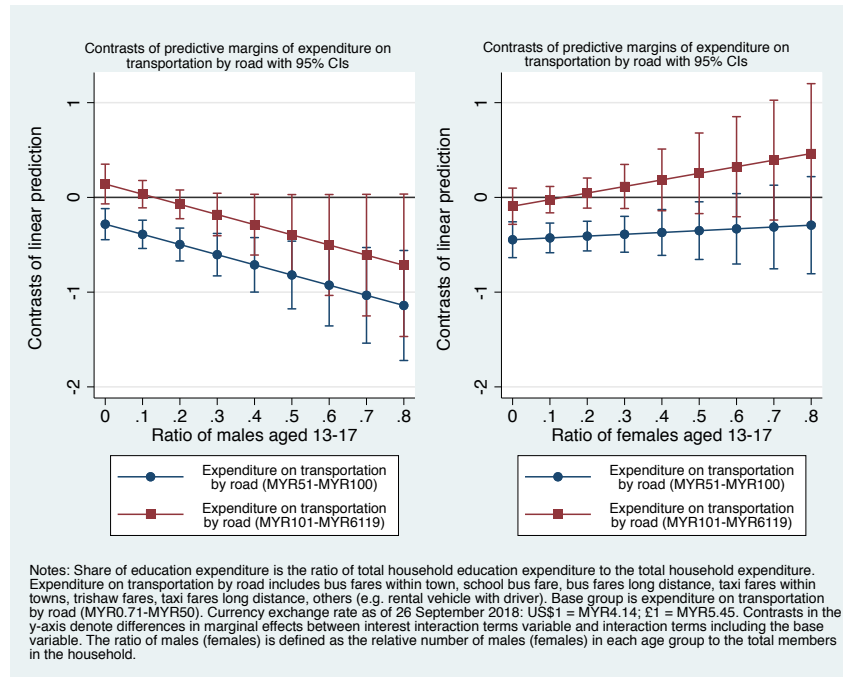


Figure 2.4: Share of education expenditure by ratio of age-gender and household expenditure on transportation by road

Lastly, higher education spending on girls than on boys may be caused by the indirect costs for girls such as in terms of modesty and safety ([Herz and Sperling, 2004](#); [Kenayathulla, 2016](#)). Parents may be spending more on girls' modesty aspect such as when they begin wearing headscarves between 9 to 16 years, which is also a cultural and religious belief for Muslim households. In addition, additional cost in transportation to school such as school bus fees may be applied for girls rather than boys. Specifically, the extra cost applies to those girls who lived far from school and their parents are unable to accompany them to school. Figure 2.4 shows the effect of household expenditure on transportation by road, which may consist of school bus fare, on the share of household education expenditure as the ratio of children aged 13-17 years increases. As expected, the figure shows that increases in the ratio of females aged 13-17 will increase household spending on transportation by road. This finding is corresponding to [Nasrudin and Nor's \(2013\)](#) study which shows that due to safety, the number of girls who walk to school was lower than boys in one of the large districts in Malaysia. Thus, in

order to reduce households' expenses on their children's transportation in general, this study suggests for the government to provide subsidies for transportation to students who do not live close to the school in particular and implement safety programs such as Safe Routes to School²⁷, as discussed by [McDonald and Aalborg \(2009\)](#).

Regarding the pro-male gap in education expenditure in the primary school age group (5-12), it suggests that parents may be more likely to send their sons to tuition classes in order to boost their sons' interest in schooling, in assumption that there are more challenges in the early childhood education for boys compared to girls. These challenges may include the boys' lack of interest in attending school and their having disciplinary problems. This explanation is in line with [Björklund and Salvanes \(2011\)](#) and [Heckman and Mosso's \(2014\)](#) suggestions concerning the importance of children's development of educational investment in the early years. On the other hand, these days, parents might be interested in providing homeschooling for their daughters. [Kim-Soon et al. \(2015\)](#) have shown that an inefficient school system, social issues among students, and an adverse school environment in some Malaysian schools are among potential reasons why parents chose the alternative channel of education for their daughters.

However, for sons' higher education, due to family socioeconomic constraints, households, especially in the rural areas, may prefer to enrol their sons into non-formal education beyond secondary school. For example, parents may be interested in sending their sons to skill-based training institutions²⁸ that required a low fee or were free of charge and the students potentially have a high demand upon graduation ([The World Bank, 2015](#); [Sani, 2018](#); [Vu et al., 2012](#)). In addition, the government incentives to transform the Technical and Vocational Education

²⁷This is a program that encourages children to walk and bicycle by making the modes of transportation safe and appealing.

²⁸For example, the community college, the Skills Training Institute of Youth and Sports (*Institut Latihan Kemahiran Belia dan Sukan (ILKBS)*), certificate level in the Industrial Training Institute (ITI), and the Centre for Instructors and Advanced Skills Training (CIAST).

and Training (TVET) programs²⁹ also may boost households' preference to send their sons to the skills institutions. These explanations correspond to Wan's (2018) findings. He shows that approximately 65% of the student population in community colleges, which is one of the main skills institutions in Malaysia, are males. It is also closely linked with the statistical discrimination theory³⁰, where in this study, the sons' role or ability to be more likely to get a job in the informal sector³¹ or the skilled-based industry had urged the risk-averse families to send their sons to vocational or skill institutions. Lastly, a potential explanation for the low education expenses for boys is due to the opportunity cost (Psacharopoulos and Woodhall, 1985). For instance, males may choose to work such as continuing the family business or working in the informal sector instead of enrol in higher institutions, given a high rate of unemployment among the graduates in Malaysia³² (Hanapi and Nordin, 2014; Omar and Rajoo, 2016; Dass, 2018). In addition, generally their opportunity to participate in the informal sectors is higher compared to females.

²⁹A program (i.e. available from the upper secondary education level) that aims to equip students with skills before they participate in the labour market even they do not perform well in the Malaysian Certificate of Education examination, which equivalent to the O-level examination, at Form 5 or Grade 10. A total of RM4.9 billion has been allocated to implement the TVET plans in 2018, and the budget increased by more than twice from 2.1 billion in 2014 (StudyMalaysia, 2016; NST Team, 2017).

³⁰According to Fang and Moro (2011), "statistical discrimination generally refers to the phenomenon of a decision-maker using observable characteristics of individuals as a proxy for unobservable, but outcome-relevant characteristics. The observable characteristics are easily recognizable physical traits which are used in society to broadly categorize demographic groups by race, ethnicity, or gender. But, sometimes the group characteristics can also be endogenously chosen, such as club membership or language." (p. 134).

³¹In Malaysia, in 2017, 61.5% workers in the informal sector were males (Department of Statistic Malaysia, 2018). The informal sector refers to three industries which are: i) manufacturing, ii) construction, and iii) services (e.g. wholesale and retail trade).

³²For example, in 2015, the youth graduate unemployment rate in Malaysia was approximately 15.3%, while the unemployment rate of youth without tertiary education was 9.8% (Ibrahim and Mahyuddin, 2016). In this sentence, youth refers to the population, labour force, and employed and unemployed workers between the ages of 15 to 24. Additionally, it was reported that there is a substantial increase in the female compared to male labour force participation rate in 2016 by 7.5% from 46.8% in 2010 (Department of Statistics Malaysia, 2017a).

2.6.1 The determinants of household education expenditure

Table 2.3 reports a few noteworthy findings of the other explanatory variables³³. As expected, this study found that high socioeconomic status households tend to allocate more educational resources for their children compared to the low socioeconomic families. For instance, this study found that households with highly educated mothers tend to have a high amount of education expenditure compared to households with low-educated mothers. This study also found that, in general, highly educated mothers tend to allocate a high education spending for the children compared to highly educated fathers. This finding is consistent with a study by [Lillard and Willis \(1994\)](#) in Malaysia that shows a positive and significant impact of highly educated mothers on household education expenditure.

Regarding another household socioeconomic measure that is household residential location, we found that households who lived in the less developed regions that are the rural and East Malaysia areas (compared with West Malaysia³⁴) preferred to spend less on their children's education expenses. This finding appears to reflect the need of the continuation of the government's education provisions such as on the Tuition Aid Scheme program that provides free tuition classes on students in the areas. However, further investigation regarding the impact of both regions by children gender and age group are suggested to provide more appropriate policy responses for the education provisions.

³³The findings are unchanged after the household education function is estimated by region. See Table A2.4 in Appendix A.

³⁴[Naidu \(2008\)](#) shows that the west coast of Peninsular Malaysia remains much more developed than the east of Malaysia. This disparity can be explained clearly through the infrastructure development in that region.

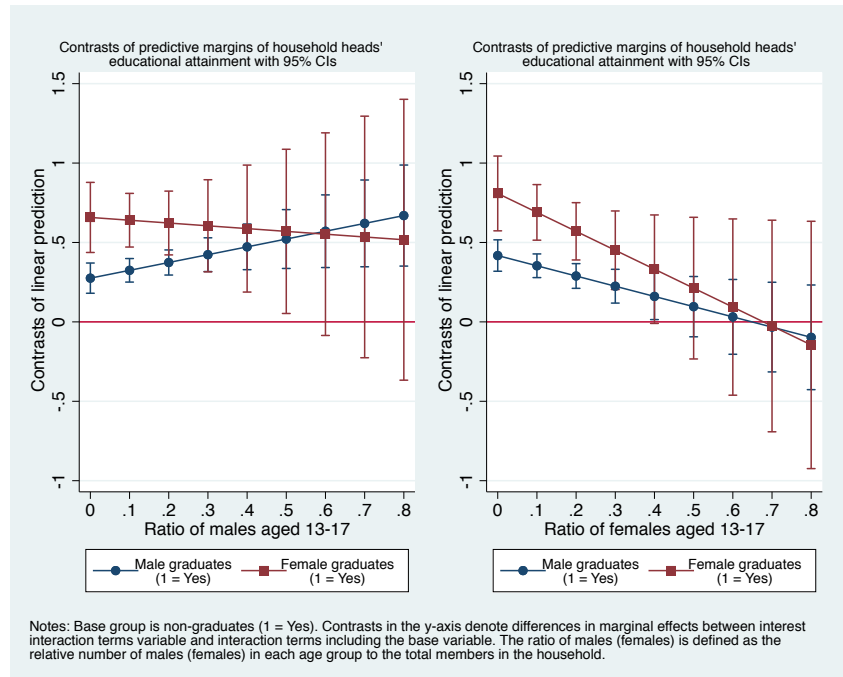


Figure 2.5: Relationship between household heads' educational attainment and ratios of children aged 13-17

To the best of my knowledge, no study in Malaysia examines whether the existence of gender differences in household education expenditure varies by household socioeconomic background. Song et al. (2006) suggest that the interactions between parental education and gender difference measures can explain sex-specific intergenerational effects. Figure 2.5 shows the interaction term effects between household heads' educational attainment and the ratio of children aged 13-17 by gender³⁵. The figure seems to indicate that there is a weak intergenerational socioeconomic status effect of the head of households' educational attainment on their members' educational attainment. Specifically, it shows that there is a positive effect of highly educated male heads of household as the ratio of males aged 13-17 increases. On the other hand, unexpectedly, the positive effect of the highly educated female heads of household decreases as the ratio of females aged 13-17 in the households increases. A potential explanation for these findings is

³⁵This study found a less significant effect of the interaction between parental education level and the gender discriminations. Thus, it employs the heads' education level, which applied to all type of household, as a proxy for the household socioeconomic status.

due to economies of scale. As the number of secondary-educated females in the households increases, households' willingness to invest in their children's secondary education resources decreases. Probably, the daughters can share similar resources such as reading materials as they have been sent to the same school. In addition, sibling gender composition and sibship density in the households also might affect girls' education resources, as discussed by [Powell et al. \(1995\)](#), [Lee \(2009\)](#), and [Hauser and Kuo \(1998\)](#) among others. [Lee \(2009\)](#) shows that Japanese girls with college-educated brothers have higher education resources than their peers without brothers. Thus, it is suggested for further study to examine whether sibship gender composition or density (e.g. the number of siblings closely spaced) does affect parents' decisions to invest in their children education.

Another noteworthy finding from Table 2.3 is that the majority of ethnic households in Malaysia, which are Malays, were found to spend at least 60% points less on education expenditure compared to their non-Malay counterparts (i.e. Chinese and Indian). The finding is supported by [Pong \(1993\)](#) and [Agadjanian and Liew \(2005\)](#) that shows a potential reason for less educational expenses from Malays than non-Malay households are due to Malaysian education preferential policies. An example is regarding sponsorship and quotas in higher institutions, which are commonly in favour of the majority ethnicity. This finding implies that a high expenditure on education by the ethnic minority may be why Malay students still lag behind non-Malay students in education attainment. To propose more appropriate responses on the preferential policy, we examine whether there is evidence of gender gaps in household education expenditure within the two ethnicities.

Table 2.6: Estimates of household education expenditure in Malay ethnicity

Dependent variable:	5-12 years		13-17 years		18-23 years	
Share of education expenditure (log)	OLS model	Heckman model	OLS model	Heckman model	OLS model	Heckman model
	(1)	(2)	(3)	(4)	(5)	(6)
Ratio of males aged 5-12	0.651** (0.297)	1.228*** (0.201)				
Ratio of females aged 5-12	0.214 (0.292)	0.819*** (0.201)				
Ratio of males aged 13-17			0.611 (0.386)	0.964** (0.395)		
Ratio of females aged 13-17			0.783** (0.374)	1.484*** (0.436)		
Ratio of males aged 18-23					-0.238 (0.529)	0.082 (0.452)
Ratio of females aged 18-23					0.536 (0.509)	0.703** (0.335)
F-test for equality by gender						
5-12	2.93*	9.19***				
13-17			0.34	2.45		
18-23					3.21*	7.46***
Lambda		1.94*** (0.52)		4.30*** (1.28)		-0.63 (0.58)
Number of observations	2413	4025	1948	3408	1212	2319
Adjusted R-squared	0.11		0.12		0.11	
Censored observations		1612		1460		1107
Rho		0.93		1.00		-0.40
Chi-squared		447.13***		141.91***		684.19***

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level. Share of education expenditure is the ratio of total household education expenditure to the total household expenditure. The ratio of males (females) is defined as the relative number of males (females) in each age group to the total members in the household. Omitted variable in each age group is others, which refer to children in other ages.

Table 2.6 shows the evidence of gender gaps in household education expenditure within the Malay households with a statistically significant F-test for the sample of households with children aged 5-12 years (column 2) in particular. This study found no significant evidence in the non-Malay households, and thus the results are presented in Appendix A, Table A2.6. Specifically, Table 2.6 shows that there are significant gender differentials in the allocation of education expenditure between girls and boys at the primary school age group (5-12 years), whereas no evidence of gender gaps in education expenditure is found in the other age groups. These findings suggest that no preferential policies by gender, which mainly applied to higher education students, happens among the Malay households. Besides, it shows that a pro-male gap in education expenditure for the primary school age children (5-12 years), as reported in Tables 2.3 and 2.5, is associated with Malays

households.

Lastly, both parents' age and family types show some significant impact on household education expenditure (Table 2.3). Specifically, we found tentative evidence that as the age of father and mother increases, the share of household education expenditure will significantly increase in the households with younger age groups (5-12 and 13-17 years). Nevertheless, the younger and older parents show a low investment in education, and these findings are consistent with [Tansel and Bircan's \(2006\)](#) study which shows that the age of the household head increases at a decreasing rate. Possibly, parents at a younger age have financial constraints as they have just entered the labour market and are thus restricted from spending more on their children's early education. However, for some parents who are older and economically inactive, potentially their children can seek financial assistance to finance their education fees.

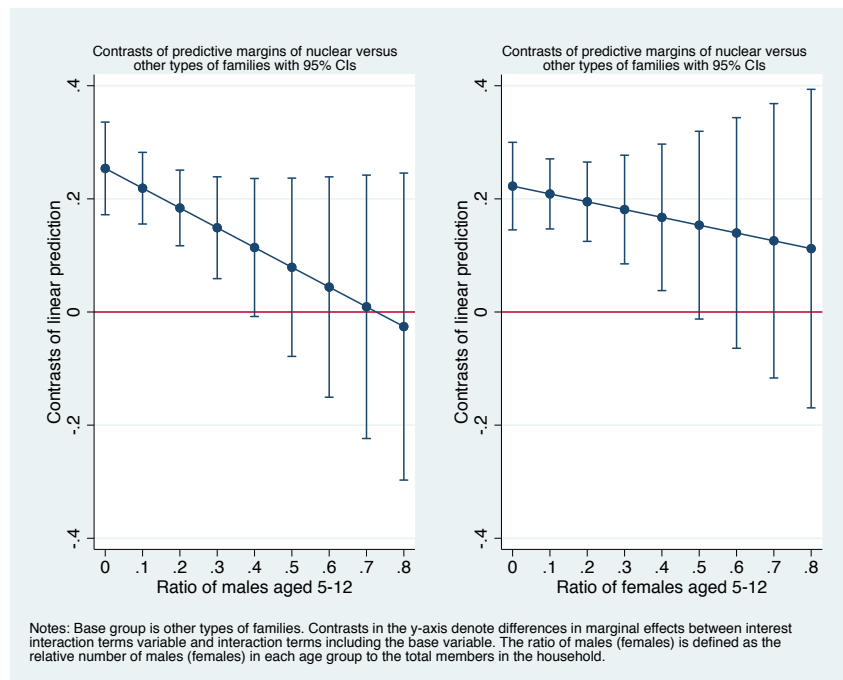


Figure 2.6: Relationship between types of family and ratios of children aged 5-12

For the family types, it is found that the nuclear families relative to other types of families tend to allocate high education expenditure on daughters as opposed to sons' education. Figure 2.6 shows evidence that the effect of the types of families (nuclear versus others) diminishes, which is closer to zero, and becomes insignificant as the ratio of males aged 5-12 years increases. To note, here we report the contrasts or differences in marginal effects between interest interaction terms variable, the nuclear types of family and the ratio of males (or females) aged 5-12 years, and interaction terms effect including the base variable. These findings correspond to [Kazianga and Wahhaj's \(2017\)](#) findings which show that nuclear family households (i.e. have strong familial ties) are more likely to achieve Pareto efficient³⁶ allocation of productive resources and consumption than other types of families (e.g. the extended family households).

2.6.2 Robustness tests

In this subsection, the estimated results of the household education expenditure function include two specifications: i) another household composition measure that is the female ratio in general and ii) the sample of households with one child only, given that this study does not have information on education expenditure for each child, are first presented. It then presents further findings of the gender differences in household education expenditure in two main regions in Malaysia, East and West Malaysia. We are interested in providing further evidence of gender differences in both regions by areas (urban-rural) since they have prevalent differences in economic development such as regarding infrastructures facilities (see studies by [Naidu \(2008\)](#), [Yusof and Bhattasali \(2008\)](#), and [Lim \(1986\)](#) among others), school aids ([Ministry of Education Malaysia, 2019](#)), and so on.

³⁶Pareto efficiency is a state where one individual is better off without making someone worse off ([Barr, 2012](#)).

Chapter 2. Gender Differences in Household Education Expenditure in Malaysia

Table 2.7: Estimated results of Heckman selection model including a few specifications

Dependent variable: Share of	5-12 years		13-17 years		18-23 years	
education expenditure (log)	(1)		(2)		(3)	
<i>a)</i>						
Ratio of female to total members of household	0.118	(0.073)	0.233**	(0.096)	0.284**	(0.121)
<i>b) Estimates for households with one child only</i>						
Ratio of males aged 5-12	−3.835***	(1.405)				
Ratio of females aged 5-12	−4.198***	(1.407)				
Ratio of males aged 13-17			−1.750***	(0.474)		
Ratio of females aged 13-17			−1.324***	(0.452)		
Ratio of males aged 18-23					−5.083**	(2.466)
Ratio of females aged 18-23					−4.536*	(2.322)
F-test for equality by gender						
5-12		1.00				
13-17				4.23**		
18-23						1.62

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level. Share of education expenditure is the ratio of total household education expenditure to the total household expenditure. Other variables are omitted for brevity. The ratio of males (females) is defined as the relative number of males (females) in each age group to the total members in the household. Omitted variable in each age group is others, which refer to children in other ages.

Overall, the estimated results in Table 2.7 support the baseline results of this study. For instance, Table 2.7 a) shows that the share of education expenditure increases as the ratio of females to household size increases specifically in the sample of households with children aged 13-17 and 18-23 years. This study also found similar evidence of gender gaps in the education expenditure for children aged 13-17 when the sample of households with one child only are considered.

Table 2.8: Estimates of household education expenditure for households in rural areas of West and East Malaysia

Dependent variables:	5-12 years				13-17 years				18-23 years			
Share of education	Rural West		Rural East		Rural West		Rural East		Rural West		Rural East	
expenditure (log); Education expenditure (1 = Observed)	OLS model (1)	Heckman model (2)	OLS model (3)	Heckman model (4)	OLS model (5)	Heckman model (6)	OLS model (7)	Heckman model (8)	OLS model (9)	Heckman model (10)	OLS model (11)	Heckman model (12)
Ratio of males aged 5-12	0.799 (0.557)	-0.551 (1.263)	-0.685 (0.799)	-0.560 (0.353)								
Ratio of females aged 5-12	-0.295 (0.614)	-0.806 (0.739)	-0.746 (0.785)	-0.575* (0.329)								
Ratio of males aged 13-17					1.011 (0.720)	2.439* (1.358)	0.631 (0.856)	2.664** (1.044)				
Ratio of females aged 13-17					1.665** (0.684)	2.533** (1.117)	1.677* (0.918)	4.624*** (1.346)				
Ratio of males aged 18-23									-0.712 (1.046)	-0.501 (0.544)	-0.377 (1.501)	-3.230** (1.590)
Ratio of females aged 18-23									1.046 (0.939)	1.089** (0.420)	-0.282 (1.557)	-2.335 (1.464)
Ethnic Malays	-0.760*** (0.228)	-0.588** (0.229)	-0.891** (0.377)	-0.799*** (0.142)	-0.938*** (0.250)	-0.839*** (0.303)	-1.170*** (0.370)	-1.040*** (0.238)	-0.733 (0.450)	-0.745*** (0.191)	-1.850*** (0.529)	-1.692*** (0.407)
<i>First-stage selection model</i>												
School aid recipient (1= Low)		-0.026 (0.037)		0.060 (0.048)		0.080** (0.038)		0.123** (0.050)		0.007 (0.043)		0.180*** (0.061)
Self-employed head of household		-0.060* (0.035)		-0.154*** (0.044)		-0.021 (0.036)		0.165*** (0.047)		-0.131*** (0.042)		0.197*** (0.062)
F-test for equality by gender	4.78**	0.09	0.01	0.03	1.46	0.01	2.11	7.89***	4.45**	15.62***	0.01	0.58
Lambda		-4.00 (3.36)		0.21 (0.61)		5.45 (3.40)		3.20*** (1.17)		-0.65 (0.84)		3.81*** (1.27)
Number of observations	628	1022	357	662	557	970	310	576	349	691	172	340
Adjusted R-squared	0.10		0.17		0.11		0.14		0.06		0.16	
Censored observations		394		305		413		266		342		168
Rho		-1.00		0.15		1.00		1.00		-0.43		1.00
Chi-squared		54.08***		261.70***		36.26***		81.64***		121.76***		47.18***

Continued on next page

Table 2.8 – continued from previous page

Dependent variables:	5-12 years				13-17 years				18-23 years			
Share of education	Rural West		Rural East		Rural West		Rural East		Rural West		Rural East	
expenditure (log);	OLS	Heckman	OLS	Heckman	OLS	Heckman	OLS	Heckman	OLS	Heckman	OLS	Heckman
Education expenditure	model	model	model	model	model	model	model	model	model	model	model	model
(1 = Observed)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. *** Significant at the 1% level. Education expenditure refers to the total household education expenditure. Share of education expenditure is the ratio of total household education expenditure to the total household expenditure. Other variables are omitted for brevity. The ratio of males (females) is defined as the relative number of males (females) in each age group to the total members in the household. Omitted variable in each age group is others, which refer to children in other ages.

Table 2.8 reports the evidence of gender differences in education expenditure for households in rural areas³⁷ of two major parts of Malaysia, East and West Malaysia. As expected, we found noticeable evidence especially in the rural areas of East Malaysia with the education expenditure on females is approximately twice as high as that on males at the secondary age. A potential reason for this is due to a low development in the regions in areas such as road infrastructure³⁸ (Naidu, 2008; Yusof and Bhattasali, 2008; Lim, 1986), which causes households in the areas to be unable to pay the costly transportation fees for their children. In addition, due to financial instability, sons probably choose to work to contribute to the family income while daughters have the opportunity to further education.

³⁷The rural areas are defined as the areas with a population of less than 10,000.

³⁸For example, in 1965, the total length of railway tracks in the Peninsular Malaysia (i.e. West Malaysia) was 1,600 km and increased to 1,667 km in 2005, whereas, the length of railway in Sabah (i.e. East Malaysia) remains at 131 km over the years (Naidu, 2008).

The following table summarises the significant results that we have discussed above to form the next section, Conclusion.

Table 2.9: Summary estimates of household education expenditure model

Dependent variables: Education expenditure (1=observed); Share of education expenditure (log)	Hurdle method		Heckman Selection method		
	5-12 years		5-12 years	13-17 years	18-23 years
	Probit (1)	OLS	(2)	(3)	(4)
A) Full sample					
Ratio of males aged 5-12	0.033*** (0.010)	0.589** (0.263)	0.793*** (0.151)		
Ratio of females aged 5-12	0.048*** (0.012)	0.214 (0.259)	0.391*** (0.138)		
Ratio of males aged 13-17				0.865*** (0.161)	
Ratio of females aged 13-17				1.217*** (0.167)	
Father's education (base: No formal)					
Primary			0.576*** (0.149)	0.246* (0.147)	0.389*** (0.137)
Secondary			0.497*** (0.156)	0.143 (0.147)	0.159 (0.138)
Tertiary			0.527*** (0.155)	0.294* (0.166)	0.216 (0.153)
Mother's education (base: No formal)					
Primary			0.476*** (0.096)	0.158 (0.122)	0.258** (0.123)
Secondary			0.562*** (0.104)	0.580*** (0.141)	0.330** (0.134)
Tertiary			1.027*** (0.123)	1.094*** (0.163)	0.821*** (0.157)
Rural			-0.303*** (0.027)	-0.253*** (0.035)	-0.227*** (0.036)
West Malaysia			0.770*** (0.068)	0.979*** (0.068)	0.649*** (0.060)
B) Region (rural areas)					
Ratio of males aged 13-17				2.731*** (1.049)	
Ratio of females aged 13-17				3.637*** (1.089)	
C) Ethnicity (Malay)					
Ratio of males aged 5-12			1.228*** (0.201)		
Ratio of females aged 5-12			0.819*** (0.201)		
F-test for equality by gender					
A) Full sample					
5-12	3.07**	2.72*	15.22***		
13-17				6.51***	
B) Region (rural areas)					
13-17				2.60*	
C) Ethnicity (Malay)					
5-12			9.19***		

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level. Share of education expenditure is the ratio of total household education expenditure to the total household expenditure. The ratio of males (females) is defined as the relative number of males (females) in each age group to the total members in the household. Omitted variable in each age group is others, which refer to children in other ages.

2.7 Conclusion

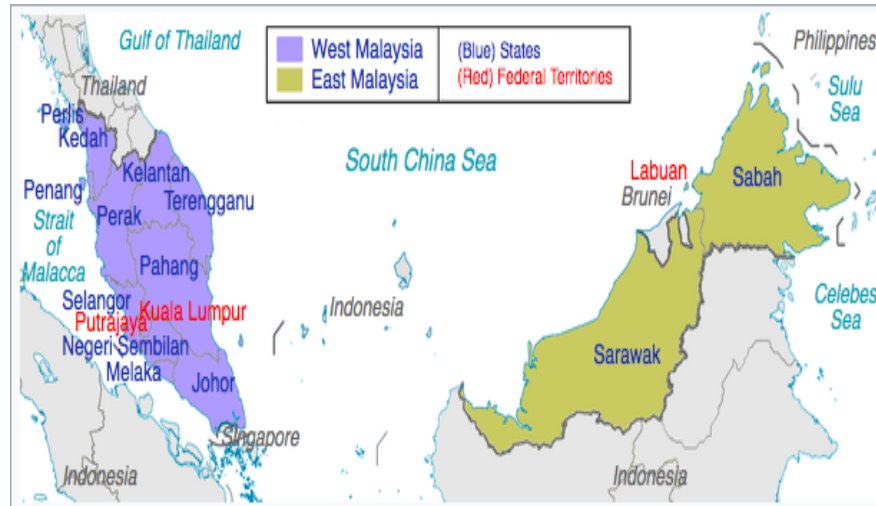
This study first examines whether there exist gender differences in the allocation of household education expenditure in Malaysia as a whole and by region. It then investigates the determinants of the household education expenditure. We found pro-male gaps in household education expenditure for primary school age children (5-12 years) within the Malaysian households as a whole, in the rural areas and Malay ethnicity. However, it is found that households prefer to invest in females compared to males' education in the post-primary school ages (13-17 years). These findings correspond to the report by the [Ministry of Education Malaysia \(2013a\)](#) on a large proportion of females compared to males' enrolment in the post-primary education levels in particular, over the years. Regarding the determinants of household education expenditure, we found consistent evidence with many previous studies such as [Tansel \(2002\)](#) and [Behrman et al. \(2017\)](#) that show that those advantaged households tend to spend more money on their children's education compared to their counterparts.

Findings from this study provide several guidelines for the government or policymakers to intervene in the education expenditure decision-making process appropriately. For instance, they can encourage women to participate in the technical streams that are increasingly demanded in the workplace as suggested by [The World Bank \(2015\)](#), [Sani \(2018\)](#), and [Vu et al. \(2012\)](#). On the other hand, they can assign more education provision for men to attract them to enrol in the non-technical fields of study where women are in the majority³⁹. Our findings also suggest that an implementation of policies to reduce the incidence of gender differences in household education expenditure should focus on the households that come from low socioeconomic status (i.e. headed by non-graduates and living in rural areas), Malay ethnicity, and non-nuclear families.

³⁹[Wan \(2018\)](#) shows that the proportion of female students enrolment in Malaysian higher education across all fields of study, except for the technical streams such as engineering, manufacturing, and construction categories, was approximately 60% of the total student enrolments.

Another implication of this study is an expectation of a lack of impact of household investment on daughters' returns to education. Even though we found some evidence of pro-female gaps in household education expenditure for children in older ages, in reality, Malaysia still has a significant pro-male pay gap in the workplace. For instance, in 2016, the mean monthly salaries and wages of male employees were higher at MYR2,500 compared to female employees, MYR2,398 ([Department of Statistics Malaysia, 2017b](#)). It implies that Malaysian women's contribution to the workplace is less valuable than men when it is assumed that they are better educated. [Barro \(2013\)](#) argues that many countries follow discriminatory practices that allow males to be in the top positions in most firms. Thus, it is not surprising that most females earned no more than males. This circumstance contradicts [Kingdon \(2005\)](#), [Zimmermann \(2018\)](#), and [Wongmonta and Glewwe's \(2017\)](#) suggestion that parents' willingness to invest in their children's education is closely linked to children's returns to education. Therefore, the government or firms in Malaysia are encouraged to offer promotion to women employees accordingly to provide returns that are expected by their parents and give them the opportunity to contribute to the improvement of human capital in Malaysia. To support this expectation, further studies are suggested to examine the impact of gender differentials in education expenditure on gender differentials in the returns to education.

Appendix A



Source: [Wikipedia contributors \(2018\)](#).

Figure A2.1: The map of Malaysia

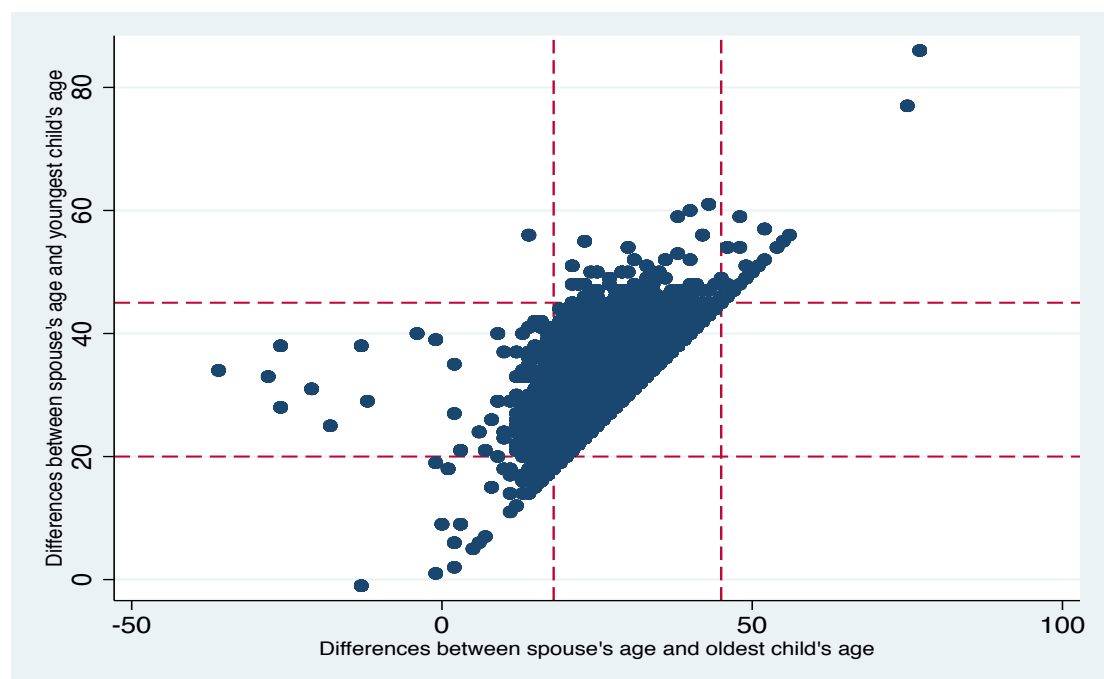


Figure A2.2: Differences between spouse's age and child's age

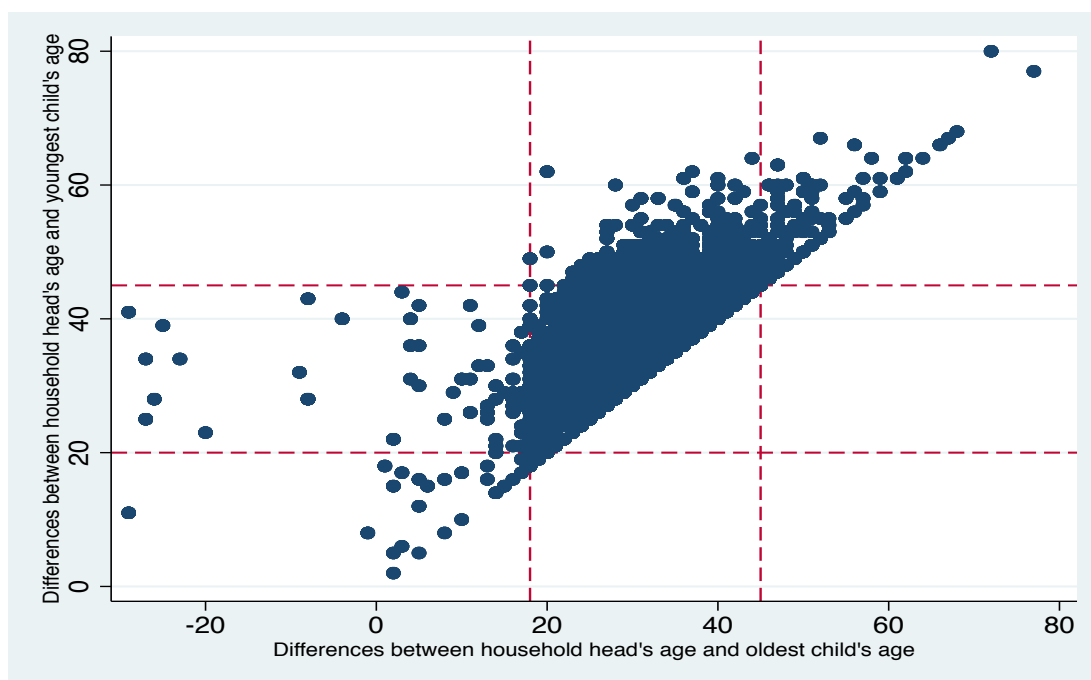


Figure A2.3: Differences between household head's age and child's age

Table A2.1: Summary statistics by school age group for households with children enrolled in school only

Variables	5-12 years				13-17 years				18-23 years			
	Mean	Standard deviation	Min-imum	Max-imum	Mean	Standard deviation	Min-imum	Max-imum	Mean	Standard deviation	Min-imum	Max-imum
Monthly total household education expenditure (MYR)	92.84	132.80	0.17	2675	96.39	151.61	0.17	5427	93.14	139.99	0.17	5427
Monthly total household expenditure (MYR)	3467.57	2336.12	521	37185	3498.33	2440.45	521	45586	3622.86	2435.45	534	45586
Monthly total household education expenditure (1 = Observed)	0.408	0.491	0	1	0.423	0.494	0	1	0.523	0.499	0	1
Parental characteristics												
Father's education (base: No formal)												
Primary	0.080	0.271	0	1	0.101	0.301	0	1	0.131	0.338	0	1
Secondary	0.513	0.500	0	1	0.507	0.500	0	1	0.479	0.500	0	1
Tertiary	0.177	0.381	0	1	0.135	0.342	0	1	0.103	0.304	0	1
Mother's education (base: No formal)												
Primary	0.091	0.287	0	1	0.118	0.322	0	1	0.157	0.364	0	1
Secondary	0.557	0.497	0	1	0.556	0.497	0	1	0.517	0.500	0	1
Tertiary	0.155	0.362	0	1	0.105	0.306	0	1	0.078	0.268	0	1
Father's age	41.139	6.572	24	65	44.383	6.958	24	64	46.494	8.332	25	67
Mother's age	38.747	6.258	23	67	42.242	6.210	24	63	44.716	7.495	23	68
Father missing	0.214	0.410	0	1	0.240	0.427	0	1	0.259	0.438	0	1
Mother missing	0.172	0.377	0	1	0.190	0.393	0	1	0.207	0.405	0	1
Household characteristics												
Types of family (base: Nuclear family)												
Single-parent family	0.032	0.175	0	1	0.047	0.212	0	1	0.046	0.210	0	1
Extended family	0.296	0.456	0	1	0.250	0.433	0	1	0.307	0.461	0	1
Others	0.002	0.039	0	1	0.002	0.039	0	1	0.004	0.059	0	1
Female headed	0.082	0.274	0	1	0.099	0.299	0	1	0.107	0.309	0	1
Rural	0.345	0.475	0	1	0.362	0.481	0	1	0.362	0.481	0	1
West Malaysia	0.693	0.461	0	1	0.703	0.457	0	1	0.701	0.458	0	1
Ethnic Malays	0.787	0.409	0	1	0.777	0.416	0	1	0.773	0.419	0	1
Non-biological parents	0.160	0.367	0	1	0.177	0.382	0	1	0.192	0.394	0	1
Number of households	5191				4477				3054			

Notes: Total household education expenditure consists of household expenditure on pre-primary and primary education, secondary education, post-secondary education, and tertiary education. Total household expenditure consists of expenditure on: education; food and non-alcoholic beverages; alcoholic beverages and tobacco; clothing and footwear, housing, water, electricity, gas and other fuels; furnishings, household equipment and routine household maintenance; medical products, appliances and equipment; transport; communication; recreation services and culture; restaurants and hotels; and miscellaneous goods and services. The non-biological parents refer to parents for households who have not fulfilled the parent-child age restrictions where the age difference between the household head (and/or spouse) and oldest child of the household is in the range of 18-45 years, and the age difference between the household head (and/or spouse) and youngest child of the household is in the range of 20-45 years (see Figures A2.2 and A2.3 in Appendix A for graphical illustration of the restrictions). We suspect them to be those who have married early (e.g. at the age of less than 20 years), married late (e.g. at the age of more than 45 years), married to an older or a younger woman, parents of the married child of the household with children, or individuals with data entry errors. Rural areas are defined as the areas with a population of less than 10,000.

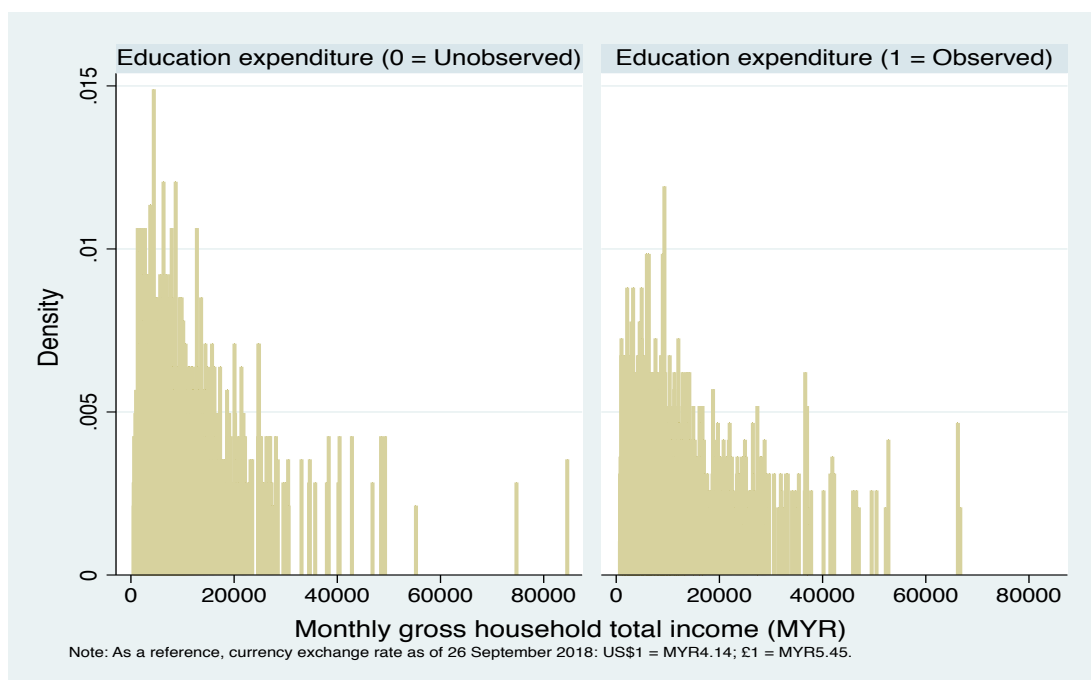


Figure A2.4: The distribution of household total income on the dummy of household education expenditure

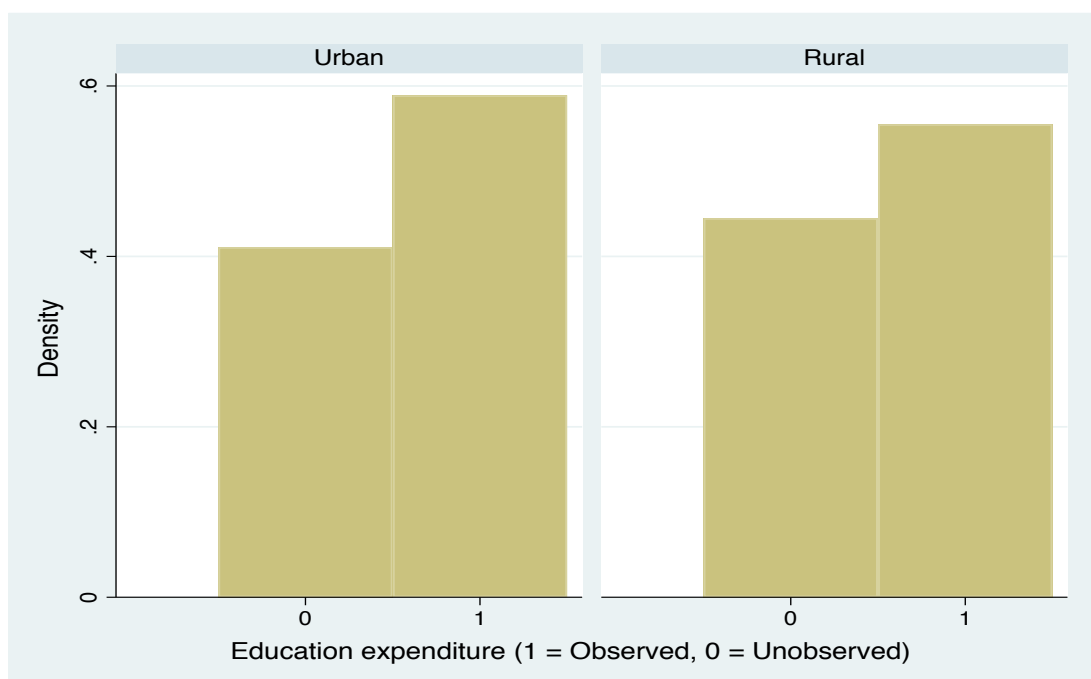


Figure A2.5: Graph of household education expenditure by household residential location

Chapter 2. Gender Differences in Household Education Expenditure in Malaysia

Table A2.2: Estimates of household education expenditure using the hurdle method

Dependent variables:	5-12 years		13-17 years		18-23 years	
Education expenditure (1 = Observed); Share of education expenditure (log)	Probit model (1)	OLS model (2)	Probit model (3)	OLS model (4)	Probit model (5)	OLS model (6)
Ratio of males aged 5-12	0.033*** (0.010)	0.589** (0.263)				
Ratio of females aged 5-12	0.048*** (0.012)	0.214 (0.259)				
Ratio of males aged 13-17			0.007 (0.006)	0.904*** (0.332)		
Ratio of females aged 13-17			0.014** (0.007)	1.151*** (0.328)		
Ratio of males aged 18-23					-0.356*** (0.043)	-0.270 (0.456)
Ratio of females aged 18-23					-0.348*** (0.045)	0.404 (0.441)
Log expenditure per person	0.007*** (0.002)	-0.250*** (0.062)	0.004** (0.002)	-0.254*** (0.066)	-0.058*** (0.009)	-0.250*** (0.086)
Father's education (base: No formal)						
Primary	0.020*** (0.006)	0.467* (0.261)	0.001 (0.003)	0.246 (0.275)	0.038 (0.028)	0.407 (0.270)
Secondary	0.017*** (0.005)	0.401 (0.258)	0.007** (0.003)	0.093 (0.271)	0.068** (0.028)	0.234 (0.270)
Tertiary	0.017*** (0.006)	0.521* (0.273)	0.000 (^a)	0.186 (0.290)	0.151*** (0.035)	0.377 (0.297)
Mother's education (base: No formal)						
Primary	0.001 (0.005)	0.503** (0.206)	0.006** (0.003)	0.045 (0.216)	-0.005 (0.023)	0.346 (0.254)
Secondary	0.011** (0.005)	0.535*** (0.188)	0.007** (0.003)	0.344* (0.208)	0.041* (0.023)	0.469* (0.250)
Tertiary	-0.004 (0.006)	1.048*** (0.209)	0.000 (^a)	0.883*** (0.237)	0.105*** (0.033)	1.063*** (0.289)
Father's age	0.000 (0.002)	-0.052 (0.060)	0.000 (0.002)	0.173* (0.092)	0.026** (0.013)	0.056 (0.127)
Mother's age	0.008*** (0.002)	0.065 (0.055)	0.001 (0.001)	-0.058 (0.082)	0.004 (0.012)	-0.099 (0.120)
Father's age squared	0.000 (0.000)	0.001 (0.001)	0.000 (0.000)	-0.002* (0.001)	0.000** (0.000)	-0.001 (0.001)
Mother's age squared	0.000*** (0.000)	-0.001 (0.001)	0.000 (0.000)	0.001 (0.001)	0.000 (0.000)	0.001 (0.001)
Father missing	0.047*** (0.011)	0.466 (0.351)	-0.005 (0.005)	0.176 (0.373)	0.057 (0.052)	0.373 (0.443)
Mother missing	-0.001 (0.009)	0.153 (0.349)	-0.001 (0.004)	0.026 (0.384)	-0.150*** (0.043)	0.163 (0.510)
Types of family (base: Nuclear family)						
Single-parent family	-0.005 (0.007)	-0.171 (0.200)	0.002 (0.003)	0.079 (0.216)	0.034 (0.025)	0.110 (0.282)
Extended family	0.005** (0.002)	-0.251*** (0.073)	0.002 (0.001)	-0.134 (0.085)	-0.033*** (0.012)	-0.202** (0.102)
Others	0.000 (^a)	-2.159*** (0.389)	0.000 (^a)	-1.164*** (0.334)	-0.022 (0.055)	-1.286*** (0.367)
Female headed	-0.005 (0.006)	0.382** (0.158)	0.001 (0.002)	0.111 (0.159)	-0.023 (0.025)	-0.034 (0.192)
Rural	0.001 (0.002)	-0.369*** (0.066)	0.000 (0.001)	-0.338*** (0.072)	-0.016 (0.010)	-0.284*** (0.093)
West Malaysia	0.006*** (0.002)	0.677*** (0.069)	0.004*** (0.001)	0.792*** (0.077)	0.059*** (0.010)	0.691*** (0.096)
Ethnic Malays	0.016***	-0.676***	0.008***	-0.848***	0.022**	-0.764***

Continued on next page

Chapter 2. Gender Differences in Household Education Expenditure in Malaysia

Table A2.2: – continued from previous page

Dependent variables:	5-12 years		13-17 years		18-23 years	
Education expenditure (1 = Observed); Share of education expenditure (log)	Probit model (1)	OLS model (2)	Probit model (3)	OLS model (4)	Probit model (5)	OLS model (6)
	(0.002)	(0.076)	(0.002)	(0.082)	(0.011)	(0.105)
Non-biological parents	−0.023** (0.012)	0.273 (0.405)	0.005 (0.006)	0.209 (0.447)	0.103** (0.049)	−0.029 (0.570)
F-test for equality by gender						
5–12	3.07**	2.72*				
13–17			1.48	0.91		
18–23					0.06	3.28*
Number of households	5340	3133	3727	2598	3647	1609
Adjusted R-squared		0.13		0.15		0.13

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level. Education expenditure refers to the total household education expenditure.

Share of education expenditure is the ratio of total household education expenditure to the total household expenditure. The ratio of males (females) is defined as the relative number of males (females) in each age group to the total members in the household. Omitted variable in each age group is others, which refer to children in other ages. The non-biological parents refer to parents for households who have not fulfilled the parent-child age restrictions where the age difference between the household head (and/or spouse) and oldest child of the household is in the range of 18-45 years, and the age difference between the household head (and/or spouse) and youngest child of the household is in the range of 20-45 years (see Figures A2.2 and A2.3 in Appendix A for graphical illustration of the restrictions). We suspect them to be those who have married early (e.g. at the age of less than 20 years), married late (e.g. at the age of more than 45 years), married to an older or a younger woman, parents of the married child of the household with children, or individuals with data entry errors. ^a is omitted value.

Table A2.3: Estimates of household education expenditure using the Engle-Curve method

Dependent variable: Share of education expenditure (log)	5-12 years		13-17 years		18-23 years	
	(1)		(2)		(3)	
Ratio of males aged 5-12	0.589**	(0.263)				
Ratio of females aged 5-12	0.214	(0.259)				
Ratio of males aged 13-17			0.904***	(0.332)		
Ratio of females aged 13-17			1.151***	(0.328)		
Ratio of males aged 18-23					−0.270	(0.456)
Ratio of females aged 18-23					0.404	(0.441)
Log expenditure per person	−0.250***	(0.062)	−0.254***	(0.066)	−0.250***	(0.086)
Father's education (base: No formal)						
Primary	0.467*	(0.261)	0.246	(0.275)	0.407	(0.270)
Secondary	0.401	(0.258)	0.093	(0.271)	0.234	(0.270)
Tertiary	0.521*	(0.273)	0.186	(0.290)	0.377	(0.297)
Mother's education (base: No formal)						
Primary	0.503**	(0.206)	0.045	(0.216)	0.346	(0.254)
Secondary	0.535***	(0.188)	0.344*	(0.208)	0.469*	(0.250)
Tertiary	1.048***	(0.209)	0.883***	(0.237)	1.063***	(0.289)
Father's age	−0.052	(0.060)	0.173*	(0.092)	0.056	(0.127)
Mother's age	0.065	(0.055)	−0.058	(0.082)	−0.099	(0.120)
Father's age squared	0.001	(0.001)	−0.002*	(0.001)	−0.001	(0.001)
Mother's age squared	−0.001	(0.001)	0.001	(0.001)	0.001	(0.001)
Father missing	0.466	(0.351)	0.176	(0.373)	0.373	(0.443)
Mother missing	0.153	(0.349)	0.026	(0.384)	0.163	(0.510)

Continued on next page

Chapter 2. Gender Differences in Household Education Expenditure in Malaysia

Table A2.3: – continued from previous page

Dependent variable: Share of	5-12 years		13-17 years		18-23 years	
education expenditure (log)	(1)		(2)		(3)	
Types of family						
(base: Nuclear family)						
Single-parent family	−0.171	(0.200)	0.079	(0.216)	0.110	(0.282)
Extended family	−0.251***	(0.073)	−0.134	(0.085)	−0.202**	(0.102)
Others	−2.159***	(0.389)	−1.164***	(0.334)	−1.286***	(0.367)
Female headed	0.382**	(0.158)	0.111	(0.159)	−0.034	(0.192)
Rural	−0.369***	(0.066)	−0.338***	(0.072)	−0.284***	(0.093)
West Malaysia	0.677***	(0.069)	0.792***	(0.077)	0.691***	(0.096)
Ethnic Malays	−0.676***	(0.076)	−0.848***	(0.082)	−0.764***	(0.105)
Non-biological parents	0.273	(0.405)	0.209	(0.447)	−0.029	(0.570)
Constant	−4.673***	(1.032)	−6.478***	(1.941)	−3.109	(1.949)
F-test for equality by gender						
5–12	2.72*					
13–17			0.91			
18–23					3.28*	
Number of observations	3133		2598		1609	
Adjusted R-squared	0.13		0.15		0.13	

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level. Share of education expenditure is the ratio of total household education expenditure to the total household expenditure. The ratio of males (females) is defined as the relative number of males (females) in each age group to the total members in the household. Omitted variable in each age group is others, which refer to children in other ages. The non-biological parents refer to parents for households who have not fulfilled the parent-child age restrictions where the age difference between the household head (and/or spouse) and oldest child of the household is in the range of 18-45 years, and the age difference between the household head (and/or spouse) and youngest child of the household is in the range of 20-45 years (see Figures A2.2 and A2.3 in Appendix A for graphical illustration of the restrictions). We suspect them to be those who have married early (e.g. at the age of less than 20 years), married late (e.g. at the age of more than 45 years), married to an older or a younger woman, parents of the married child of the household with children, or individuals with data entry errors. No or less significant gender differences in each age group shows the relevant to use an alternative approach such as the hurdle method.

Table A2.4: Estimates of household education expenditure in rural areas

	5-12 years		13-17 years		18-23 years	
	OLS model (1)	Heckman model (2)	OLS model (3)	Heckman model (4)	OLS model (5)	Heckman model (6)
<i>Corrected model</i>						
Dependent variable: Share of education expenditure (log)						
Ratio of males aged 5-12	0.477 (0.466)	0.499*** (0.187)				
Ratio of females aged 5-12	−0.266 (0.485)	−0.262 (0.202)				
Ratio of males aged 13-17			0.878 (0.557)	2.731*** (1.049)		
Ratio of females aged 13-17			1.653*** (0.550)	3.637*** (1.089)		
Ratio of males aged 18-23					−0.752 (0.852)	−1.564** (0.660)

Continued on next page

Chapter 2. Gender Differences in Household Education Expenditure in Malaysia

Table A2.4: – continued from previous page

	5-12 years		13-17 years		18-23 years	
	OLS model (1)	Heckman model (2)	OLS model (3)	Heckman model (4)	OLS model (5)	Heckman model (6)
Ratio of females aged 18-23					0.727 (0.803)	0.255 (0.465)
Father's education (base: No formal)						
Primary	0.137 (0.297)	0.101 (0.153)	-0.058 (0.253)	-0.868* (0.527)	0.439 (0.324)	0.295 (0.206)
Secondary	0.064 (0.297)	-0.035 (0.175)	-0.150 (0.248)	-0.284 (0.429)	0.336 (0.331)	0.405* (0.214)
Tertiary	-0.146 (0.352)	-0.378** (0.174)	-0.531 (0.340)	-0.499 (0.519)	-0.079 (0.449)	-0.145 (0.247)
Mother's education (base: No formal)						
Primary	-0.026 (0.260)	-0.026 (0.126)	-0.299 (0.260)	0.169 (0.436)	-0.038 (0.343)	0.107 (0.227)
Secondary	0.044 (0.246)	-0.027 (0.139)	-0.122 (0.275)	0.584 (0.498)	-0.091 (0.361)	0.012 (0.233)
Tertiary	1.105*** (0.316)	0.827*** (0.151)	1.166*** (0.362)	2.014*** (0.655)	0.819* (0.462)	0.889*** (0.312)
Father's age	0.039 (0.097)	0.033 (0.059)	0.207 (0.128)	0.299* (0.164)	0.036 (0.200)	0.218 (0.163)
Mother's age	0.013 (0.077)	0.019 (0.042)	-0.020 (0.110)	-0.310* (0.179)	-0.098 (0.181)	-0.243* (0.136)
Father's age squared	0.000 (0.001)	0.000 (0.001)	-0.002 (0.001)	-0.003* (0.002)	0.000 (0.002)	-0.002 (0.002)
Mother's age squared	0.000 (0.001)	-0.001 (0.001)	0.000 (0.001)	0.003 (0.002)	0.001 (0.002)	0.003* (0.001)
Father missing	0.758* (0.446)	0.595** (0.265)	-0.042 (0.473)	0.530 (0.704)	0.283 (0.614)	0.552 (0.409)
Mother missing	-0.088 (0.435)	-0.157 (0.221)	-0.976 (0.628)	-1.001 (0.740)	0.208 (0.730)	0.308 (0.465)
Log expenditure per person	-0.399*** (0.105)		-0.408*** (0.107)		-0.273* (0.146)	
Types of family (base: Nuclear family)						
Single-parent family	-0.610** (0.305)	-0.604*** (0.139)	0.122 (0.382)	-0.635 (0.521)	-0.211 (0.481)	-0.318 (0.231)
Extended family	-0.466*** (0.111)	-0.423*** (0.043)	-0.240* (0.129)	-0.205 (0.150)	-0.372** (0.164)	-0.324*** (0.065)
Female headed	0.210 (0.222)	0.250* (0.147)	0.247 (0.216)	-0.345 (0.424)	-0.004 (0.271)	-0.135 (0.205)
West Malaysia	0.805*** (0.100)	0.762*** (0.076)	0.945*** (0.112)	1.148*** (0.166)	0.733*** (0.149)	0.781*** (0.073)
Ethnic Malays	-0.793*** (0.202)	-0.699*** (0.074)	-1.005*** (0.208)	-0.781*** (0.227)	-1.005*** (0.354)	-0.771*** (0.188)
Non-biological parents	-0.429 (0.547)	-0.346 (0.291)	0.798 (0.798)	0.056 (0.918)	-0.446 (0.839)	-0.725 (0.531)
Constant	-4.190*** (1.594)	-6.419*** (1.183)	-6.570** (2.728)	-8.398*** (3.076)	-1.911 (2.806)	-5.320*** (1.959)
<i>First-stage selection model</i>						
Dependent variable: Education expenditure (1 = Observed)						
School aid recipient (1= Low)		0.023 (0.029)		0.103*** (0.030)		0.073** (0.035)
Self-employed head of household		-0.090*** (0.027)		0.047* (0.028)		-0.042 (0.034)
F-test for equality by gender						

Continued on next page

Chapter 2. Gender Differences in Household Education Expenditure in Malaysia

Table A2.4: – continued from previous page

	5-12 years		13-17 years		18-23 years	
	OLS model (1)	Heckman model (2)	OLS model (3)	Heckman model (4)	OLS model (5)	Heckman model (6)
5-12	3.50*	20.53***				
13-17			3.23*	2.60*		
18-23					4.55**	21.27***
Lambda		-0.03 (0.69)		5.02** (2.02)		1.27 (1.05)
Number of observations	985	1684	867	1546	521	1031
Adjusted R-squared	0.16		0.19		0.12	
Censored observations		699		679		510
Rho		-0.02		1.00		0.73
Chi-squared		922.93***		114.27***		354.53***

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level. Education expenditure refers to the total household education expenditure. Share of education expenditure is the ratio of total household education expenditure to the total household expenditure. The ratio of males (females) is defined as the relative number of males (females) in each age group to the total members in the household. Omitted variable in each age group is others, which refer to children in other ages. The non-biological parents refer to parents for households who have not fulfilled the parent-child age restrictions where the age difference between the household head (and/or spouse) and oldest child of the household is in the range of 18-45 years, and the age difference between the household head (and/or spouse) and youngest child of the household is in the range of 20-45 years (see Figures A2.2 and A2.3 in Appendix A for graphical illustration of the restrictions). We suspect them to be those who have married early (e.g. at the age of less than 20 years), married late (e.g. at the age of more than 45 years), married to an older or a younger woman, parents of the married child of the household with children, or individuals with data entry errors. The school aid recipient variable refers to the households who are not categorised in the bottom 40% (B40) income household (i.e. households with monthly income of below MYR3,860). Other explanatory variables for the first-stage selection model are omitted for brevity. The unmarried-headed family is omitted because of no observation.

Table A2.5: Estimates of household education expenditure in urban areas

Dependent variable:	5-12 years		13-17 years		18-23 years	
Share of education expenditure (log)	OLS model (1)	Heckman model (2)	OLS model (3)	Heckman model (4)	OLS model (5)	Heckman model (6)
Ratio of males aged 5-12	0.568* (0.322)	0.810*** (0.210)				
Ratio of females aged 5-12	0.336 (0.308)	0.521*** (0.176)				
Ratio of males aged 13-17			0.997** (0.413)	0.742*** (0.197)		
Ratio of females aged 13-17			0.976** (0.408)	0.819*** (0.189)		
Ratio of males aged 18-23					-0.092 (0.542)	-0.096 (0.372)
Ratio of females aged 18-23					0.319 (0.537)	0.263 (0.292)
F-test for equality by gender						
5-12	0.71	4.79**				
13-17			0.00	0.24		
18-23					0.86	2.82*

Continued on next page

Chapter 2. Gender Differences in Household Education Expenditure in Malaysia

Table A2.5: – continued from previous page

Dependent variable:	5-12 years		13-17 years		18-23 years	
Share of education expenditure (log)	OLS model (1)	Heckman model (2)	OLS model (3)	Heckman model (4)	OLS model (5)	Heckman model (6)
Lambda		0.69 (0.46)		1.32*** (0.41)		-0.17 (0.38)
Number of observations	2148	3507	1731	2931	1088	2023
Adjusted R-squared	0.08		0.10		0.10	
Censored observations		1359		1200		935
Rho		0.42		0.71		-0.11
Chi-squared		500.24***		541.77***		518.32***

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level. Share of education expenditure is the ratio of total household education expenditure to the total household expenditure. Other variables are omitted for brevity. The ratio of males (females) is defined as the relative number of males (females) in each age group to the total members in the household. Omitted variable in each age group is others, which refer to children in other ages.

Table A2.6: Estimates of household education expenditure in non-Malay ethnicity

Dependent variable:	5-12 years		13-17 years		18-23 years	
Share of education expenditure (log)	OLS model (1)	Heckman model (2)	OLS model (3)	Heckman model (4)	OLS model (5)	Heckman model (6)
Ratio of males aged 5-12	0.336 (0.567)	0.575* (0.345)				
Ratio of females aged 5-12	0.081 (0.559)	-0.035 (0.293)				
Ratio of males aged 13-17			2.081*** (0.640)	1.780*** (0.337)		
Ratio of females aged 13-17			2.521*** (0.661)	2.213*** (0.348)		
Ratio of males aged 18-23					-0.617 (0.903)	-1.754** (0.690)
Ratio of females aged 18-23					-0.221 (0.877)	-0.861* (0.495)
F-test for equality by gender						
5-12	0.26*	3.13*				
13-17			0.73	2.42		
18-23					0.30	4.15**
Lambda		1.65* (0.98)		1.80*** (0.57)		1.27* (0.72)
Number of observations	720	1166	650	1069	397	735
Adjusted R-squared	0.06		0.06		0.06	
Censored observations		446		419		338
Rho		0.81		0.86		0.68
Chi-squared		133.19***		168.13***		103.64***

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level. Share of education expenditure is the ratio of total household education expenditure to the total household expenditure. Other variables are omitted for brevity. The ratio of males (females) is defined as the relative number of males (females) in each age group to the total members in the household. Omitted variable in each age group is others, which refer to children in other ages.

Chapter 3

School Effects on Student Attainment in Malaysia

3.1 Introduction

This study examines the extent to which initiatives (which we refer to as school factors) outlined in the current Malaysian education policy, the Malaysia Education Blueprint 2013-2025 (MEB), affect student attainment in Malaysia. We examine the school factors based on student gender as an attempt to address the issue of gender gaps in student attainment. In comparison to other high-middle income countries such as Iran, Lebanon, and Romania, the gender gap in Malaysia regarding student attainment in particular in the international TIMSS assessment is widening while its attainment falls dramatically¹. On average, in 2011, gender differences in mathematics and science achievement, with the difference favouring girls, was 449 versus 430 and 434 versus 419 respectively. In detail, for mathematics assessment, 71% females compared to 59.7% males had achieved at least a minimum proficiency level². The disparity has been consistent over the years and is of great concern to the Ministry of Education Malaysia (MOE).

To the best of my knowledge, no study on Malaysia comprehensively explores the effects of school factors within the MEB on student attainment. Studies by

¹See Figure A3.1 in Appendix B for more details on average scores in Malaysia compared with other similar high-middle income countries.

²The minimum proficiency level refers to Low International Benchmark where in mathematics, refers to the students who have some knowledge of whole numbers and basic graphs ([Ministry of Education Malaysia, 2017](#)).

Hanushek and Luque (2003), Glewwe and Muralidharan (2016), and Hanushek (1997) have shown the importance of taking into account the effectiveness of direct policy implications on student outcomes in order to make strategic decisions in education. Additionally, studies that examine the effect of school factors by student gender are scarce. Many studies in Malaysia highlight the importance of student-family factors on student attainment, for example, Jelas and Dahan (2010), Jelas et al. (2014), Ismail (2015), Saw (2016), and Lillard and Willis (1994). Recently, however, researchers are interested in studying the effects of school reforms such as the School-Based Assessment (SBA) and the Higher-Order Thinking Skills (HOTS) on student performance (Sulaiman et al., 2017; Ghazali, 2016; Saido et al., 2015; Majid, 2011; Malakolunthu and Hoon, 2010). Thus, this study improves on existing work in the areas. Hanushek (2013), Castelló-Climent (2010), Klasen (2002), and Klasen and Lamanna (2009) show that investigating school effects on student attainment is important in enhancing school quality, human capital, and eventually, economic growth of a country.

This study uses eighth-grade students' test scores in mathematics and science in the international assessment of the Trends in International Mathematics and Science Study (TIMSS) dataset as the measures of student attainment. The students' test scores refer to the students' average of five plausible values, where the plausible value is defined as the range of competence that a student might rationally have, given the student's item responses (Wu, 2005). The TIMSS dataset consists of randomly selected schools from all 13 states in Malaysia. It also provides information on various school factors associated with the MEB policy and students' family background, which makes the data well suited for examining the educational production model in Malaysia. We examine this model using standard ordinary least squares (OLS) and instrumental variable (IV) approaches. Our intuition to use the IV method is to address the potential reverse causality effect between class size and student scores due to school sorting effect as argued by other authors and as explained in detail in Section 3.4 below. We use the

eight-grade and schools size as instruments in the method, which we expect to affect class size directly and student attainment indirectly. Lastly, we include several specifications in the estimation to support the baseline results and provide the appropriate policy responses. These include the estimation using a sub-sample of school location, a sub-sample of school size, other proxies of the variables of interest, and estimation using the weighted least squares method which recognises the stratification and clustering of the TIMSS data and includes sampling weights in the estimation so that the education production estimates are representative of the Malaysian school population.

The results presented in this research provide tentative evidence of class size and IT staff effects, in particular. Specifically, we find that class size and IT staff have a positive effect on student attainment in most of our models estimate. However, we find a negative evidence when we estimate the model focusing on rural areas and small schools only. Even though the class size result is not consistent with many theoretical and empirical studies on class size, it is supported by [Hattie \(2005\)](#), [Woessmann \(2016\)](#), and [Lazear \(2001\)](#) among others. They show that the positive effect of large class size is related to factors such as school system and peer effects in the classroom. For example, [Lazear \(2001\)](#) shows that a student's previous experience, in assumption he or she enrolled in the same school, able to help the student to cope with such situation of large number of students in a classroom. However, other school factors such as teacher hiring, school resources, and parental involvement in schools show an expected positive effect on student performance. Overall, these results are robust when we estimate the education attainment model using alternative specifications as mentioned above.

This paper is structured as follows. The next section provides further information about the Malaysia Education Blueprint 2013-2025 (MEB). Section 3.3 reviews selected literature on the effects of school factors on student attainment. Section 3.4 outlines the empirical strategy used to estimate the education production function. Section 3.5 describes the dataset and provides descriptive statistics.

Section 3.6 discusses the results, and concluding remarks are presented in Section 3.7.

3.2 Background on the Malaysia Education Blueprint 2013-2025: Preschool to Post-Secondary Education

The Malaysia Education Blueprint 2013-2025: Preschool to Post-Secondary Education was introduced in October 2011 to replace the previous educational policy of the Malaysia Education Blueprint 2006-2010. The blueprint is a comprehensive policy that was formulated based on extensive research from various organisations and individuals related to the development of education in Malaysia.

According to the [Ministry of Education Malaysia, 2013a](#), p. E-9–E-10, the blueprint has five system aspirations or challenges: i) Access (100 percent enrolment across all education level), ii) Quality (top third of countries in international assessments in 15 years³), iii) Equity (50 percent reduction in achievement gaps), iv) Unity (an education system that gives children shared values and experiences by embracing diversity), and v) Efficiency (a system which maximises student outcomes within current budget). To achieve these aspirations, eleven transformation measures or in this study, we refer to as school factors had been formed to drive the Malaysian education system into a knowledge-based economy by 2020. The eleven transformations are: 1) provide equal access to quality education of international standard; 2) ensure every child is proficient in Malay language and English language; 3) develop values-driven Malaysians; 4) transform teaching into the profession of choice; 5) ensure high performing school leaders in every school; 6) empower State Education Department (*JPNs*), District Education Office (*PPDs*), and schools to customise solutions based on needs; 7) leverage ICT to scale up quality learning across Malaysia; 8) transform ministry delivery capabilities and capacity; 9) partner with parents, community and private sector at scale; 10)

³Malaysia ranked 10 out of 50 in TIMSS assessment 2003. The aspiration may correspond to this achievement.

maximise student outcomes for every *ringgit* (Malaysia currency); and 11) increase transparency for direct public accountability.

Only seven out of the eleven transformations were considered in this chapter due to data unavailability, which excludes items 2), 3), 8), and 11). For transformation item 1), this study aims to address sub-transformation: raise quality of all preschools and encourage universal enrolment by 2020. For item 4), this study aims to examine the effectiveness for two sub-transformations: raise the entry bar for teachers from 2013 to be amongst top 30% of graduates and upgrade the quality of continuous professional development (CPD) from 2013. For item 5), this study aims to address the sub-transformation: roll out a New Principal Career Package in waves with greater support and sharper accountability for improving student outcomes. Both items 6) and 7) related to school resources where we aim to address sub-transformations: allow greater school-based management and autonomy for schools that meet a minimum performance criteria; ensure 100% of schools meet basic infrastructure requirements by 2015; and provide internet access and virtual learning environments via 1BestariNet for all 10,000 schools by 2013. For transformation 9), this chapter aims to examine the effectiveness sub-transformations: equip every parent to support their child's learning through a parent engagement toolkit and access to their child's in-school progress and invite every PIBG (refers to Parent and Teacher Association) to provide input on contextualisation of curriculum and teacher quality from 2016. Lastly, for transformation 10), this chapter aims to address sub-transformation: capture efficiency opportunities, with funding reallocated to the most critical areas such as teacher training and upskilling.

3.3 Selected literature on the effect of school factors on student attainment

The evidence of the effect of school factors on student attainment is mixed. Among possible explanations for the mixed evidence are differences in the school system and development of respective countries as argued in studies by [Hanushek \(2013\)](#), [Lee and Barro \(2001\)](#), and [Castelló-Climent \(2010\)](#).

For the school factor of school resources, a positive effect of school resources (e.g. instructional materials, pupil-teacher ratios, infrastructure) on student achievements is shown by [Lee and Barro \(2001\)](#), [Zhao and Glewwe \(2010\)](#), [Glewwe and Jacoby \(1994\)](#), and [Card and Krueger \(1996\)](#) in China, Ghana, and North and South Carolina among others. However, a number of studies such as [Hanushek \(1997\)](#), [Houtenville and Conway \(2008\)](#), [Woessmann \(2003\)](#), and [Jeong et al. \(2017\)](#) show a less important impact of school resources on student achievements. For instance, [Houtenville and Conway \(2008\)](#) and [Jeong et al. \(2017\)](#) each suggest that instead of school resources, family factors in the US and fiscal decentralisation⁴ in Korea are important factors in improving student performance. While [Hanushek \(1997\)](#) argues that the lack of school resources' impact on student achievement is closely related to the school's resources measure.

A large number of studies highlight the importance of smaller class sizes on student attainment, for example, in Columbia ([Breton, 2014](#)), in Minnesota, US ([Cho et al., 2012](#)) and in some universities in the east coast of the US ([Monks and Schmidt, 2011](#)). A similar finding is also seen in cross-country studies by [Lee and Barro \(2001\)](#) and [Krueger \(1999\)](#) among others. However, the incentives are found to be not significant in improving student attainment such as in studies by [Dieterle \(2015\)](#), [Galton and Pell \(2012\)](#), [Konstantopoulos \(2008\)](#), and [Hoxby \(2000\)](#). The differences in teacher quality, shifts in teaching methods, and student

⁴The study refers fiscal decentralisation as the allocation of tax and spending powers to lower levels of government.

characteristics (e.g. ability and socioeconomic backgrounds) are found to be among the reasons for their negative finding. [Lazear \(2001\)](#) and [Woessmann and West \(2006\)](#) show that the mixed effect of class size is due to factors such as classroom congestion effects, peers effects, and teacher characteristics (e.g. teacher salary), and suggests further study to address the class size effect.

Other school factors such as teacher incentives, a high level of parental involvement in the school, and a high level of the principal's role in the school also show a positive effect on student attainment. For example, [Gunnarsson et al. \(2009\)](#), [Woessmann \(2005\)](#), [Dhuey and Smith \(2014\)](#), and [Bloom et al. \(2015\)](#) show that a high parental involvement in their children's schooling, teacher incentives, and the principal's leadership have a significant and positive effect on student performance in their respective studies in eight Latin American countries (Argentina, Bolivia, Brazil, Chile, Colombia, the Dominican Republic, Honduras, and Peru), two East Asian countries (Korea and Singapore), British Columbia, the UK, Sweden, Canada, and the US. [Brewer \(1993\)](#), [Bloom et al. \(2015\)](#), [Louis et al. \(2010\)](#), and [Eberts and Stone \(1988\)](#) show that principals have a significant impact on student outcomes through their roles such as in selecting teachers, setting the school's goals, leadership, and conflict resolution. However, [Coelli and Green \(2012\)](#) and [Gertler et al. \(2012\)](#) show that factors such as time and school surroundings may influence the impact of both the principal and parents' roles in schools. [Gertler et al. \(2012\)](#) show that the school factor of high parental involvement in school has a negative effect on students' attainment especially for those living in remote areas, while [Coelli and Green \(2012\)](#) show that principals will take time to maximise their efforts at a school.

Different from previous literature, this study attempts to examine the determinants of student attainment by focusing on the effects of current educational policy in Malaysia by including a broad number of school factors corresponding to the educational policy to examine the effectiveness of each of the factors in improving Malaysian students' attainment comprehensively. In addition, this

study contributes to the literature by showing the effect of hiring new teachers on student attainment, which is lacking attention in a Malaysian context in particular. Results from the estimate may help to address issues such as teacher shortage in major cities and a high rate of graduate unemployment in Malaysia ([Bernama, 2018](#); [Ibrahim and Mahyuddin, 2016](#)).

3.4 Empirical strategy

This study begins with an estimation of the education production function using the simple ordinary least squares (OLS) approach to examine the education production function of the effect of school factors on student attainment. Then, the instrumental variable (IV) approach is employed to address potential endogeneity problems in the OLS estimates. The sources of this problem will be discussed below.

In general, the education production function for each observation student i from school s can be written as follows:

$$T = \alpha_0 + \alpha_j X_j + \delta D + \mu \quad (3.1)$$

where T is the measure of student attainment, student i 's test scores in mathematics or science in school s ; X_j for $j = 1, \dots, p$ is the vector of school factors including class size, school resources (i.e. instructional materials, infrastructure, IT staff, and computer availability), teacher incentives, teacher hiring, principal's role in attending leadership activities, and parental role in supporting their child's performance; D is the vector of control variables such as student i 's family background and school characteristics; and μ is an error term. The student test scores refer to student average plausible values⁵, which is the range of competences that a student might have, given the student's item responses in an assessment ([Wu,](#)

⁵See Section 3.5 for details on this measure.

2005). We examine equation (3.1) using standard OLS separately by student gender. From the estimation, we can straightforwardly determine whether the presence of the school factors may alter the student achievement gap by gender or not. In addition to the standard regression, we consider two specifications: 1) the estimates using robust clustering and 2) the estimates using weighted sampling and robust clustering. The first specification regards that the observations are not independent and thus the model estimates are clustered by school identification (see next section, Section 3.5 for details on the TIMSS sampling method) while the second specification allows the model estimates to be interpreted nationally. However, we expect that these specifications produce unchanged results with our standard regression due to the small number of schools (180) and students (5,733) considered in the TIMSS 2011 assessment for Malaysia compared to the 2,271 schools and 2,296,034 students enrolment at secondary level in the year in Malaysia ([Ministry of Education Malaysia, 2013c](#)).

Our alternative way to examine the education production model in detail is by estimating the model using a sub-sample. Two sub-sample categories are considered, one by school location (urban-rural) and one by school size dummy. The resulting estimates, in general, will show the effectiveness of government initiatives in allocating education funds to conduct the school factors or initiatives within the MEB policy according to the sub-samples.

Next, we are concerned about the potential endogeneity problem of class size effect. This is because there might be an issue of reverse causality between class size and student attainment in equation (3.1). In Malaysia, a student's performance is the main factor to determine his or her class (school sorting effect) ([Tuah, 2017](#); "[Tiada lagi pengasingan](#)", 2017). [West and Woessmann \(2006\)](#) and [Lazear \(2001\)](#) among others claim that the sorting practices may lead to an endogeneity bias in the class-size effect. Thus, to address the potential endogeneity problem, this study uses the instrumental variable (IV) approach with two instruments which are the eighth-grade size and school size. The eighth-grade size refers to

the number of student enrolments in the eighth-grade, and the school size refers to the total student enrolments in a school. We expect that these instruments are good as they do not directly link with other determinants of this study's dependent variables, student test scores in mathematics and science. Studies by [West and Woessmann \(2006\)](#), [Gary-Bobo and Mahjoub \(2013\)](#), and [Akerhielm \(1995\)](#) among others have shown the validity of the instruments.

The instrumental variable (IV) model with two-stage least squares (2SLS) estimator can be written as follows:

$$T = \alpha_0 + \alpha_1 \hat{X}_1 + \alpha_2 X_2 + \alpha_3 X_3 + \dots + \alpha_p X_p + \delta D + \varepsilon_{i,s} \quad (3.2)$$

where X_1 is an endogenous variable, class size, and $\hat{X}_1$ is the predicted result from the following first-stage estimation of equation (3.3).

$$\hat{X}_1 = \beta_0 + \lambda_k Z_k + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_p X_p + \delta D + \varepsilon \quad (3.3)$$

where Z_k for $k = 1, \dots, q$ are the instruments for the class size, that is the eighth-grade size and school size. To check the validity of this method, we first look at the significant result of [Wooldridge's \(1995\)](#) robust score test, which shows whether class size is an endogenous variable or not. In other words, [Wooldridge's \(1995\)](#) score test that refers to the endogeneity test will tell whether the baseline OLS model yields bias estimates or not. We then identify the significant results of the F-statistic of the first-stage regression of equation (3.3), which affirms whether the selected instruments are not weak. A not weak instrument is the instrument that is highly correlated with the endogenous variable and satisfies [Stock et al.'s \(2002\)](#) suggestion of an F-statistic that should exceed ten to be reliable when there is one endogenous regressor. For an analysis with more than one instrument, we also check on the overidentifying restrictions test of [Wooldridge's](#) robust score test. This tests whether the instruments are uncorrelated with the structural error

term in equation (3.1). It also identifies whether our IV model is misspecified and if one or more of the excluded exogenous variables should, in fact, be included in the structural equation.

Overall, the key assumption for the appropriateness of the IV estimates is that the instruments are highly correlated with the endogenous variables but uncorrelated with other determinants of the dependent variable of the education production model. The term uncorrelated with other determinants of the dependent variables is like saying $Cov(\varepsilon, Z_k) = 0$, is the exclusion restriction of the IV model ([Angrist and Pischke, 2009](#)).

We end this research with some sensitivity analysis including the estimates of the education production model using other proxies of the variables of interest and interaction terms between class size and students' socioeconomic background.

3.5 Data and descriptive statistics

The Trends in International Mathematics and Science Study (TIMSS) dataset is a dataset sourced from a survey organised by the International Association for the Evaluation of Educational Achievement (IEA), an independent educational institution. In 2011, a total of 5,733 Malaysian students who were in the eighth-grade in 2010 from 180 schools had participated in the survey. The TIMSS dataset includes schools from all 13 states in Malaysia and consists of comprehensive information obtained from the students, teachers, and school principal in each participating school, making the data well suited for examining the educational production model in Malaysia. Other than because it is a rich dataset, we are interested in using the TIMSS dataset due to its assessment contents that is curriculum-based rather than application-based, which is consistent with the Malaysian education system. Hence, this study assumes that the measures of student attainment, students' scores in the assessment, are less likely to produce bias estimates that might be sourced from heterogeneity in student ability.

The two-stage random sampling design was used to conduct the TIMSS survey (see [TIMSS & PIRLS \(2013\)](#) for further details on the sampling design). Briefly, the design begins with selecting a random sample of schools in Malaysia and then within each of these schools, one class in the eighth-grade is chosen randomly. Every student in the class was then assigned randomly with only one booklet of an assessment, and the TIMSS used all available data from the student's responses to the assessment, including the student's background and school characteristics to impute the student's test scores or "plausible values". According to [Wu \(2005\)](#), the plausible values refer to the range of competence that a student might have given the student's item responses in an assessment. The TIMSS chose to use the imputation process due to the complicated matrix-sampling design that has been used in the assessment.

This study uses the average of students' five plausible values in mathematics or science as a measure of students' attainment. It considers some school factors⁶ that have been outlined within the Malaysia Education Blueprint 2013-2025 (MEB). These include the class size, school instructional materials, infrastructure, IT staff, computers per student, teacher incentives, teacher hiring, principal's attendance in leadership activities, and parental involvement in school. All of the selected explanatory variables are categorical variables, except the class size and computer availability. This study also includes a dummy for missing variables for each of the categorical variables, which takes the value of one if the observation is missing and zero otherwise to reduce the loss of information due to missing data. For the continuous variables, we use the mean imputation method.

In detail, class size is measured by the number of students in the eighth-grade class. To note, although there is only one eighth grade class per school involved in the TIMSS survey, in general, each school in Malaysia has more than one eighth grade class. School resources such as instructional materials, infrastructure,

⁶Other school factors such as moral, language, and administration factors are not included due to data limitations.

and IT staff are carefully generated using the principal component analysis. The rotated component matrix results for the analysis are presented in the Appendix B in Table A3.1. We use the analysis to reduce the number of variables that may be correlated and regrouped them into a new variable accordingly. In the analysis, we specify three components to be generated, and the components are formed according to the characteristics of the selected variables that it has. Another school resource proxy, computers per student, refers to the ratio of the number of computers for eighth-grade students to the number of students in the eighth grade.

School factors related to teachers such as teacher incentives refer to whether the school grants incentives such as bonuses, housing, and smaller classes to recruit new teachers. Teacher hiring, which is measured from a question about “how difficult was it to fill teaching vacancies”, is categorised directly from the options of the question into three categories: no vacancy, easy teacher hiring, and difficult teacher hiring.

School factors related to principals’ role and parents’ role in schools, respectively, were generated from the school principal questionnaire about “how much time have you spent on participating in professional development activities specifically for school principals” and the school’s level of “parental support for student achievement”. The principal attendance in leadership activities is categorised into two levels: high, which refers to a lot of time, and low, which refers to no time or some time. While the parental support for student attainment, which is used as a proxy of parental involvement in the school, is coded directly from the question into three levels: high, medium, and low.

Other than school factors, this study includes students’ family background and school characteristics as control variables in the estimation. In detail, for family background, we consider student’s parental educational level in five categories: i) some primary or did not go to school, ii) secondary, iii) tertiary, iv) child does

not know, and v) missing, as a measure of students' household socioeconomic background. We also consider students' home possessions⁷ availability as other proxies of students' socioeconomic status as suggested by McKenzie (2005) and Chowa et al. (2013). Similarly, the principal component analysis was applied as a tool to categorise home possessions into several types (see Table A3.2 in Appendix B). As a result, other than the number of books at home (i.e. one or more bookcases), the home possessions measure consists of reading resources, technology resources, and if there is a study room with each of them are categorised into high and low levels.

For school characteristics, we include school location whether urban or rural, a dummy of whether schools have more than 50 percent of students come from economically disadvantaged homes, and a dummy of instructional time⁸ based on if they have six and more hours per day or not.

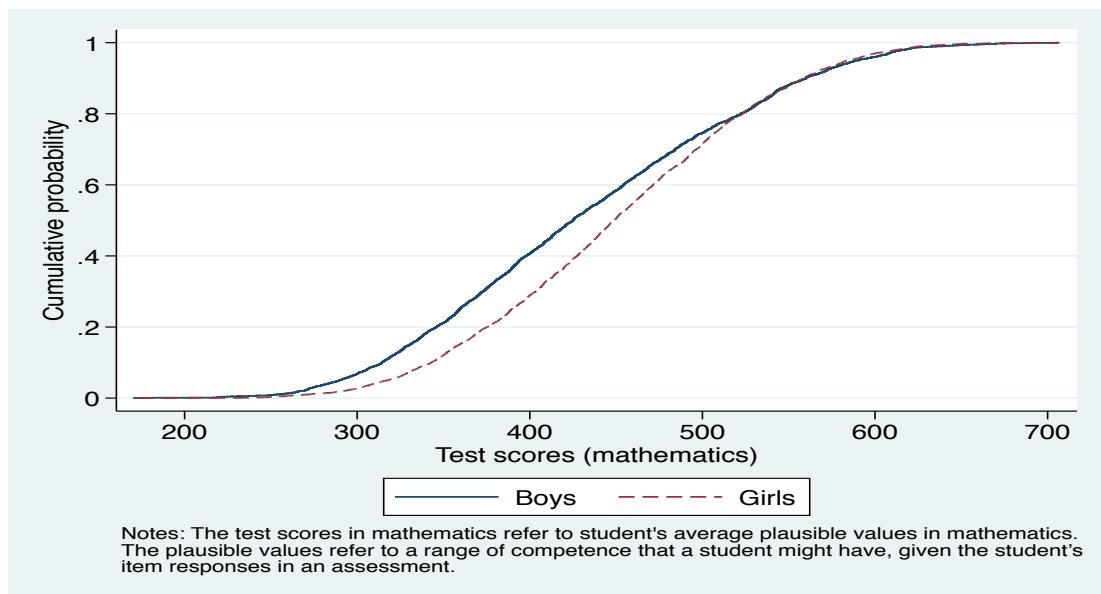


Figure 3.1: Cumulative distribution of students' test scores in mathematics by gender

⁷These include the computer, study desk/table for your use, books of your own (do not count your school books), your own room, internet connection, calculator, dictionary, reference book, and video.

⁸The instructional time refers to the engaged time between students and teachers in school.

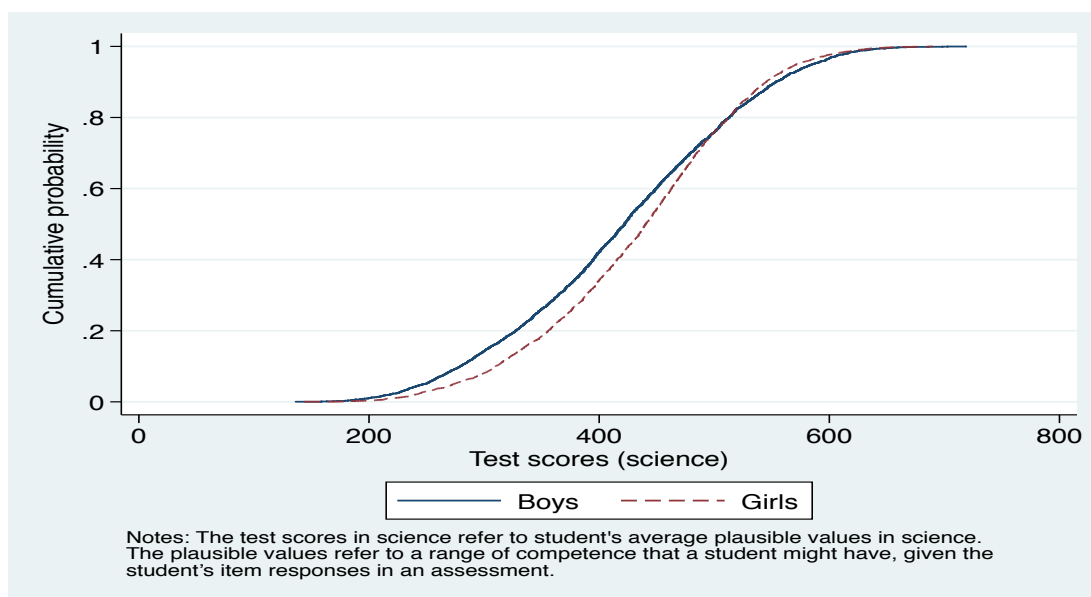


Figure 3.2: Cumulative distribution of students' test scores in science by gender

Figures 3.1 and 3.2 show the cumulative distribution function of students' test scores in mathematics and science. The figures illustrate that there is no difference between the number of girls and boys who score more than 500, while more boys than girls score below 500 in both subjects.

Table 3.1: Descriptive statistics

Variables	Boys		Girls		Test of difference (p-value)
	Mean	Standard deviation	Mean	Standard deviation	
	(1)	(2)	(3)	(4)	(5)
Test scores (mathematics)	431.75	93.46	449.58	82.10	0.00
Test scores (science)	419.85	102.61	433.63	90.30	0.00
Class size	34.50	6.71	35.39	6.37	0.00
Class size missing (1 = yes)	0.02	0.16	0.02	0.15	0.52
Instructional materials (base: a lot)					
Instructional materials (a little)	0.57	0.49	0.56	0.50	0.585
Infrastructure (base: a lot)					
Infrastructure (some)	0.31	0.46	0.27	0.44	0.00
Infrastructure (little)	0.28	0.45	0.31	0.46	0.00
IT staff (base: a lot)					
IT staff (a little)	0.45	0.50	0.49	0.50	0.00
IT staff missing	0.01	0.10	0.00	0.00	0.00
Computers per student	0.11	0.09	0.11	0.09	0.13
Computers missing (1 = yes)	0.02	0.15	0.03	0.16	0.52
Teacher incentives (1 = yes)	0.27	0.44	0.27	0.45	0.63
Teacher incentives missing (1 = yes)	0.01	0.08	0.00	0.06	0.22
Teacher hiring (base: difficult)					
Teacher hiring (easy)	0.48	0.50	0.55	0.50	0.00
No teacher hiring	0.41	0.49	0.36	0.48	0.00
Teacher hiring missing	0.01	0.11	0.01	0.08	0.03
Principal attendance in leadership activities (base: low)					
Principal attendance in leadership activities (high)	0.41	0.49	0.42	0.49	0.59
Principal attendance in leadership activities missing	0.00	0.06	0.01	0.07	0.19
Parental involvement in school (base: low)					
Parental involvement in school (high)	0.29	0.45	0.33	0.47	0.00
Parental involvement in school (medium)	0.53	0.50	0.50	0.50	0.01
Family background					
Mother's education (base: no formal or primary)					
Secondary	0.39	0.49	0.47	0.50	0.00
Tertiary	0.19	0.39	0.20	0.40	0.39
Child's does not know	0.33	0.47	0.21	0.41	0.00
Missing	0.01	0.10	0.01	0.08	0.10
Father's education (base: no formal or primary)					
Secondary	0.35	0.48	0.43	0.50	0.00
Tertiary	0.22	0.41	0.22	0.42	0.79
Child's does not know	0.36	0.48	0.26	0.44	0.00
Missing	0.01	0.10	0.01	0.09	0.90
<i>Home possessions</i>					
One and more bookcases	0.28	0.45	0.36	0.48	0.00
Books availability missing	0.00	0.06	0.00	0.04	0.15
Reading resources (low)	0.34	0.47	0.17	0.37	0.00
Reading resources missing	0.01	0.10	0.01	0.10	0.91
Technology resources (low)	0.63	0.48	0.65	0.48	0.12
Technology resources missing	0.01	0.12	0.02	0.14	0.12
Study room (low)	0.44	0.50	0.42	0.49	0.15
Study room missing	0.01	0.08	0.01	0.10	0.37

Continued on next page

Table 3.1 – continued from previous page

Variables	Boys		Girls		Test of
	Mean	Standard deviation	Mean	Standard deviation	difference (p-value)
	(1)	(2)	(3)	(4)	(5)
School characteristics					
Students come from disadvantaged homes (> 50%)	0.41	0.49	0.41	0.49	0.48
Students come from disadvantaged homes missing	0.02	0.13	0.03	0.17	0.00
Rural	0.60	0.49	0.62	0.48	0.11
Instructional time (6 and more hours per day)	0.47	0.50	0.50	0.50	0.03
Instructional time missing	0.15	0.35	0.14	0.35	0.58
Number of observations	2815		2918		

Notes: *Significant at the 10% level. **Significant at the 5% level. *** Significant at the 1% level. The test scores refer to the average of students' five plausible values in the test. The plausible values refer to the range of competence that a student might have, given the student's item responses in an assessment. The test of difference refers to the t-test, which is the test of mean difference of continuous variables by gender, and the Chi-squared test, which is the test of mean differences between the categorical variables and students' gender.

Table 3.1 shows the statistical evidence of the difference in means of student test scores by gender for the TIMSS sample. On average, we found that girls achieve statistically significantly higher test scores than boys. Table 3.1 also shows the descriptive statistics of variables of interest of this study, which is school factors, including control variables of students' family background and school characteristics. A brief comparison of the TIMSS sample and population in Malaysia, in 2011, secondary school in Malaysia had an average school size of 974 pupils per school in comparison to 1,449 in the TIMSS sample. Regarding average class size, there are 28 pupils per class compared to around 35 students in the TIMSS sample. 49.27% (1,119) secondary schools from the total number of secondary schools in Malaysia (2,271) located in rural areas in contrast to 60% in the TIMSS sample. Lastly, about 2% (47 out of 2,271) secondary school in Malaysia is categorised as single-sex schools as compared to 5% in the TIMSS sample. Table 3.1 presents the descriptive statistics of the variables of interest by gender to examine if there is a relationship between the corresponding variable and students' gender. The resulted relationship tests are summarised in column (5) of Table 3.1, which consists of a t-test and Chi-squared test. The t-test refers

to the test of mean differences between the continuous variables and students' gender, while the Chi-squared test refers to the test of the relationship between the categorical variables and students' gender.

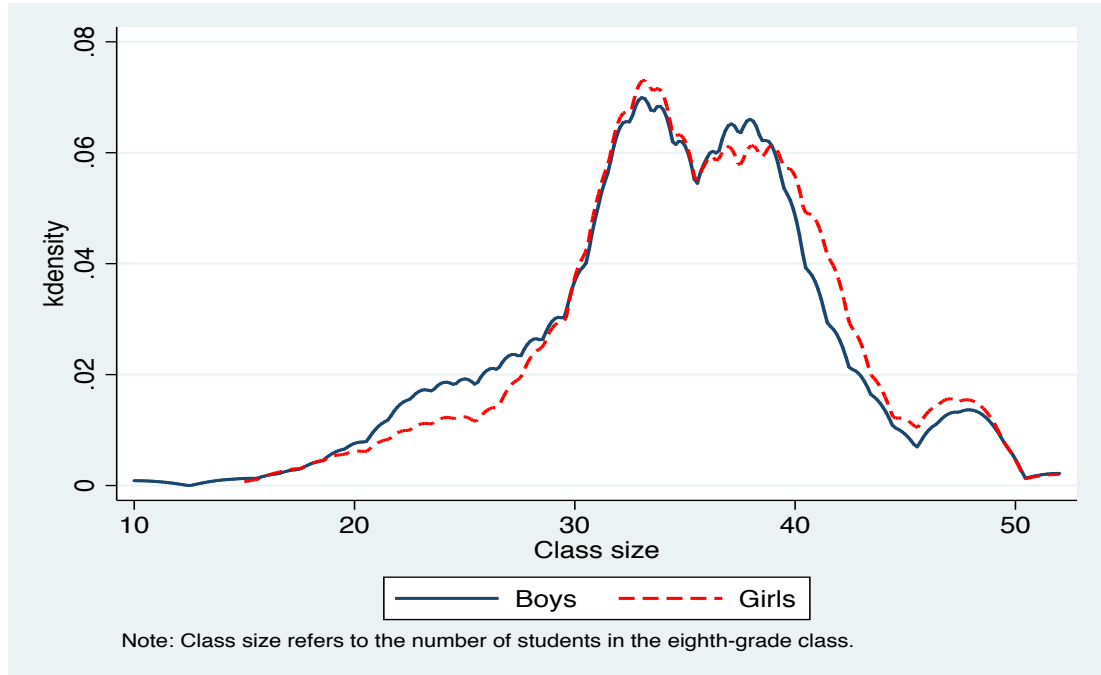


Figure 3.3: Distribution of class size by gender

Table 3.1 shows that, on average, the number of students in a class is 35 students, with girls are more likely to be in a large class (see Figure 3.3 for graphical illustration). Approximately 50% of the students enrolled in the schools with a high supply of school resources such as instructional materials (e.g. textbooks, papers, and pencils), infrastructure (e.g. school buildings, lighting systems, and classrooms), and IT staff. About half of the students enrolled in schools that do not have difficulty in hiring new teachers, have a low principal attendance in leadership activities, and have a medium level of parental involvement in school. Regarding the mean difference of the variable of interest by student's gender, most of the variables, except instructional materials, teacher incentives, and school principals' role, show statistically significant results, indicating the appropriateness of this research to analyse the education production model by student gender.

For students' family background, a large proportion of the students either reported that they have secondary education parents (approximately 40%) or they do not know their parents' educational attainment (20-30%). A majority of the students, especially boys, have less than one bookcase at home (i.e. less than 26 books excluding magazines, newspapers, or school books). However, only one-third or less of the students have reported a low number of reading resources, referring to their own books (except school books), references books, and a dictionary. For other home possessions, approximately 60% of the students have low technology resources, and half of the students do not have their own study room. Overall, only parental education and reading resources show a significant difference based on student gender.

In regards to students' school characteristics, a majority of them enrolled in schools that have 50 percent and fewer students that come from disadvantaged homes, located in urban areas, and have instructional time of six and more hours per day.

3.6 Empirical results

In this section, we first present the estimated effects of the school factors on student attainment in the full sample (Tables 3.2 and 3.3) with dependent variables of students' test scores in mathematics and science using two methods, OLS and IV. We then report the estimated results of the school factors on student attainment in sub-samples of school location (urban and rural) and school size dummy (whether a school has student enrolment of more than 974, the average school size in Malaysia in 2011⁹, or not). We end this section with the reports of the sensitivity analysis of the baseline models using alternative specifications.

⁹See a report by the [Ministry of Education Malaysia \(2013c\)](#).

Table 3.2: Estimates of school effects on student attainment in mathematics

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models		IV 2SLS models			
			instrument = eight-grade size		instruments = eight-grade size and school size	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)
Class size	3.122*** (0.227)	1.916*** (0.221)	2.021*** (0.529)	1.796*** (0.425)	2.499*** (0.508)	2.345*** (0.403)
Class size squared	-0.007 (0.020)	0.020 (0.025)				
Instructional materials (base: a lot)						
Instructional materials (a little)	-29.102*** (3.823)	-18.367*** (3.445)	-30.157*** (3.793)	-18.295*** (3.346)	-29.735*** (3.789)	-17.583*** (3.351)
Infrastructures (base: a lot)						
Infrastructure (some)	-14.254*** (4.049)	-11.587*** (3.938)	-12.709*** (4.028)	-11.562*** (3.908)	-13.346*** (4.034)	-12.220*** (3.924)
Infrastructure (little)	-17.023*** (5.264)	-20.414*** (4.620)	-16.502*** (5.087)	-20.620*** (4.549)	-16.648*** (5.096)	-20.617*** (4.565)
IT staff (base: a lot)						
IT staff (a little)	21.442*** (3.598)	19.257*** (3.395)	21.992*** (3.542)	19.417*** (3.437)	21.700*** (3.536)	18.706*** (3.433)
Computers per student	17.568 (17.388)	7.927 (15.165)	0.461 (18.956)	7.129 (16.402)	7.769 (18.787)	16.603 (16.342)
Teacher incentives (1 = yes)	2.861 (3.065)	2.759 (2.775)	1.519 (3.100)	2.877 (2.753)	2.067 (3.080)	2.971 (2.754)
Teacher hiring (base: difficult)						
Teacher hiring (easy)	17.030*** (5.306)	-1.473 (4.686)	19.039*** (5.430)	-1.561 (4.674)	18.239*** (5.350)	-1.920 (4.633)
No teacher hiring	19.736*** (5.246)	3.279 (4.597)	22.800*** (5.537)	3.401 (4.651)	21.522*** (5.430)	2.356 (4.599)
Principal attendance in leadership activities (base: low)						
Principal attendance in leadership activities (high)	-2.661 (3.226)	2.399 (2.726)	-1.414 (3.216)	2.237 (2.667)	-1.924 (3.207)	1.919 (2.667)
Parental involvement in school (base: low)						
Parental involvement in school (high)	41.867*** (4.991)	43.408*** (4.221)	40.112*** (5.045)	43.594*** (4.210)	40.830*** (5.041)	44.060*** (4.210)
Parental involvement in school (medium)	27.552*** (3.944)	28.605*** (3.589)	27.089*** (3.914)	28.788*** (3.558)	27.288*** (3.923)	28.574*** (3.564)

Continued on next page

Table 3.2 – continued from previous page

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models		IV 2SLS models			
			instrument = eight-grade size		instruments = eight-grade size and school size	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)
Constant	421.105*** (10.012)	434.083*** (8.050)	423.234*** (9.883)	435.414*** (7.993)	422.121*** (9.840)	433.600*** (7.975)
<i>First-stage estimates</i>						
Dependent variable: Class size						
Eight-grade size			0.023*** (0.001)	0.024*** (0.001)	0.014*** (0.002)	0.013*** (0.001)
School size (enrolment)					0.002*** (0.000)	0.003*** (0.000)
Number of observations	2815	2918	2815	2918	2815	2918
Adjusted R-squared	0.396	0.373	0.391	0.373	0.395	0.372
χ^2 statistic (p-value) of endogeneity test			0.018	0.775	0.153	0.228
First-stage F-statistics			455.956***	525.401***	278.255***	414.667***
χ^2 statistic (p-value) of overidentifying restrictions test					0.000	0.000

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. *** Significant at the 1% level. The plausible values refer to the range of competence that a student might have, given the student's item responses in an assessment. See Appendix B Table A3.3 for full results including students' family background and school characteristics.

Table 3.2 shows the estimated results of equation (3.1) using students' scores in mathematics. The results using students' scores in science are, in general, very similar except the teacher incentives and computers per student effects (see Table 3.3). Both tables show the estimated results using OLS and IV 2SLS methods by student gender. The IV method includes the models that use the instruments of eight-grade size and both eight-grade size and school size in the estimations. The post-estimation results for the IV method are reported in the bottom section of each table. These post-estimation results include the test of endogeneity, first-stage regression statistics, and test of overidentifying restrictions. As we specified a robust VCE¹⁰ as the variance estimator in each estimation, we then chose Wooldridge's score test as the endogeneity test and the overidentifying restrictions test. An advantage of Wooldridge's score test is that it can tolerate heteroskedastic and autocorrelated errors in the model estimates.

Overall, the IV results indicate that there is some evidence of an endogeneity problem in the OLS estimates, especially in the estimation using students' test scores in science. The evidence is supported by statistically significant results of the endogeneity test, that is Wooldridge's score test with a null hypothesis that the class size variable is exogenous. The instruments employed in the models of interest also have shown reliable or not-weak results with statistically significant results of the F-statistics of the first-stage regression and Wooldridge's robust score test of overidentifying restrictions. The significant results of the IV post-estimation indicate that the exclusion restriction of the method is satisfied.

¹⁰Robust to some types of misspecification to ensure that the observations are independent, which is one of the features of the estimator.

Table 3.3: Estimates of school effects on student attainment in science

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models		IV 2SLS models			
			instrument = eight-grade size		instruments = eight-grade size and school size	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)
Class size	3.771*** (0.253)	2.076*** (0.239)	0.329 (0.619)	0.335 (0.487)	0.609 (0.601)	0.468 (0.473)
Class size squared	-0.022 (0.023)	0.002 (0.026)				
Instructional materials (base: a lot) Instructional materials (a little)	-33.978*** (4.209)	-22.893*** (3.538)	-37.280*** (4.321)	-25.129*** (3.549)	-37.033*** (4.300)	-24.957*** (3.547)
Infrastructure (base: a lot) Infrastructure (some)	-12.454*** (4.332)	-11.126*** (3.943)	-7.619* (4.340)	-9.050** (3.927)	-7.992* (4.326)	-9.209** (3.919)
Infrastructure (little)	-17.167*** (5.932)	-23.049*** (4.681)	-15.525*** (5.741)	-23.077*** (4.589)	-15.611*** (5.732)	-23.076*** (4.587)
IT staff (base: a lot) IT staff (a little)	24.964*** (4.108)	26.270*** (3.584)	26.674*** (4.173)	28.523*** (3.679)	26.502*** (4.154)	28.351*** (3.670)
Computers per student	58.444*** (17.859)	46.889*** (15.454)	4.954 (20.436)	16.966 (17.306)	9.239 (20.220)	19.252 (17.190)
Teacher incentives (1 = yes)	8.024** (3.476)	4.917* (2.955)	3.823 (3.590)	4.632 (2.940)	4.145 (3.566)	4.655 (2.936)
Teacher hiring (base: difficult) Teacher hiring (easy)	20.348*** (6.071)	2.712 (5.527)	26.641*** (6.524)	3.837 (5.642)	26.171*** (6.464)	3.750 (5.623)
No teacher hiring	24.552*** (5.994)	8.104 (5.399)	34.138*** (6.607)	11.408** (5.572)	33.388*** (6.532)	11.156** (5.548)
Principal attendance in leadership activities (base: low) Principal attendance in leadership activities (high)	-2.542 (3.481)	3.240 (2.789)	1.362 (3.577)	4.228 (2.770)	1.063 (3.561)	4.151 (2.768)
Parental involvement in school (base: low) Parental involvement in school (high)	40.915*** (5.510)	51.559*** (4.522)	35.419*** (5.655)	50.106*** (4.554)	35.840*** (5.642)	50.219*** (4.550)
Parental involvement in school (medium)	26.971*** (4.510)	36.470*** (3.803)	25.525*** (4.572)	37.162*** (3.823)	25.641*** (4.561)	37.110*** (3.814)

Continued on next page

Table 3.3 – continued from previous page

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models		IV 2SLS models			
			instrument = eight-grade size		instruments = eight-grade size and school size	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)
Constant	376.587*** (11.500)	375.348*** (9.312)	383.203*** (11.786)	381.181*** (9.317)	382.551*** (11.731)	380.743*** (9.310)
<i>First-stage estimates</i>						
Dependent variable: Class size						
Eight-grade size			0.023*** (0.001)	0.024*** (0.001)	0.014*** (0.002)	0.013*** (0.001)
School size (enrolment)					0.002*** (0.000)	0.003*** (0.000)
Number of observations	2815	2918	2815	2918	2815	2918
Adjusted R-squared	0.393	0.416	0.350	0.404	0.357	0.406
χ^2 statistic (p-value) of endogeneity test			0.000	0.000	0.000	0.000
First-stage F-statistics			455.956***	525.401***	278.255***	414.667***
χ^2 statistic (p-value) of overidentifying restrictions test					0.031	0.311

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. *** Significant at the 1% level. The plausible values refer to the range of competence that a student might have, given the student's item responses in an assessment. See Appendix B Table A3.4 for full results including students' family background and school characteristics.

Both tables (Tables 3.2 and 3.3) show that class size has a positive and statistically significant effect on students' test scores, with the test scores approximately twice as high for boys as for girls when class size increases by one point. This finding is inconsistent with previous studies such as [Glewwe and Jacoby \(1994\)](#), [Cho et al. \(2012\)](#), [Lee and Barro \(2001\)](#), and [Monks and Schmidt \(2011\)](#) who have examined the effects of the corresponding countries' policy of class size reduction. However, it is in line with [Dieterle \(2015\)](#), [Galton and Pell \(2012\)](#), [Breton \(2014\)](#), [Konstantopoulos \(2008\)](#), and [Hoxby's \(2000\)](#) findings of a positive class size effect. The results, to some extent, are consistent even after the causality effect between the class size and student attainment has been addressed. However, it should be emphasised that the positive effect of the class size on student attainment is relatively small.

Although the results are inconsistent with the theory and empirical work in class size effect, we found evidence of a negative effect of the class size squared but not statistically significant. The negative effect of the class size squared indicates that students' test scores first increase with class size but at a decreasing rate (diminishing marginal returns), as students put more efforts in the learning process or as students are able to cope with the situation probably starting from his or her previous grade (in the assumption that education is a cumulative process). Then, at some point (223 for maths and 171 for science), the students' scores stop growing, reach the optimal level, and then start to decrease as class size becomes larger. The decrement is probably in line with [Lazear's \(2001\)](#) theory of class size, where if a classroom gets congested, students' and teachers' focus on learning and teaching in the classroom is limited. This situation may reduce students' motivation or ability to learn and ultimately affect their performance.

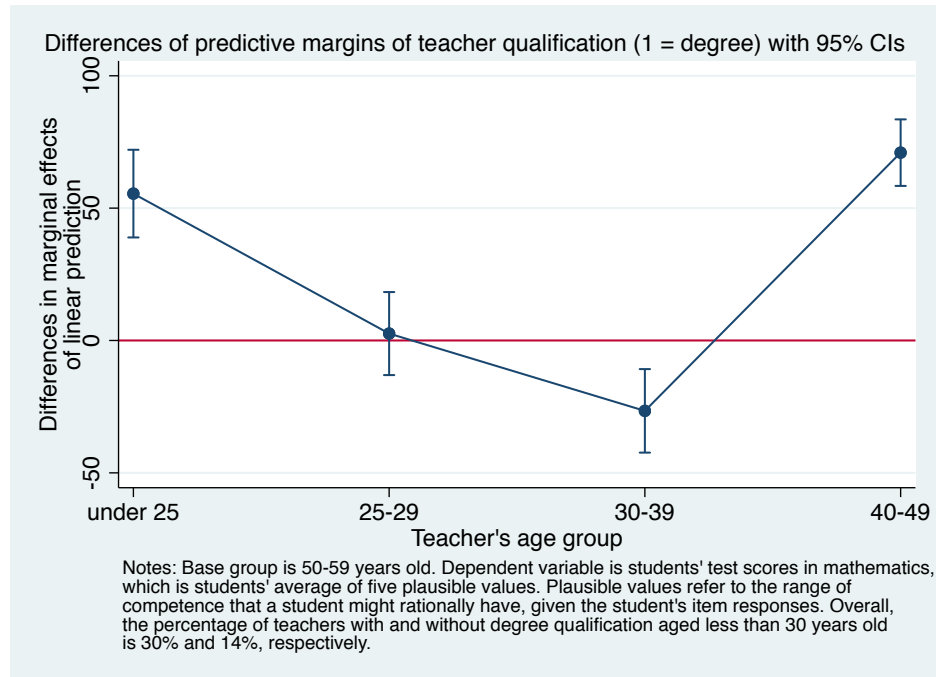


Figure 3.4: The relationship between teacher's qualification, age, and students' attainment

According to [Hattie \(2005\)](#), a potential explanation for the positive effect of large class size is that students in the class are probably taught by highly capable teachers that can handle a large number of students in a class. Figure 3.4 shows the evidence of a positive relationship between teachers' qualifications and students' attainment (where we add teacher characteristics in the baseline model), with students tending to perform better if they are taught by degree-qualified teachers either younger or older. In addition, a positive surrounding in the classroom may help the low achiever students, in particular, to perform better with the help of the high achiever students for the classroom, supporting the idea of positive peer effects ([Blatchford et al., 2011](#); [Duflo et al., 2011](#); [Lazear, 2001](#)). Lastly, due to resource constraints, the opportunity cost for schools to have smaller class sizes may be higher compared to other options such as providing LCD projector for each of the large class sizes, which results in better student performance ([Lazear, 2001](#); [Krugger, 2003](#)).

Regarding school resources, Tables 3.2 and 3.3 show that an adequate supply of instructional materials (e.g. textbooks, papers, and pencils) and infrastructure (e.g. school buildings, lighting systems, and classrooms) may improve students' test scores. These findings are consistent with studies by [Demir et al. \(2010\)](#), [Lee and Barro \(2001\)](#), and [Zhao and Glewwe \(2010\)](#) among others.

Another noteworthy finding from Tables 3.2 and 3.3 is that both IT related resources, IT staff resources (a little), which refers to technologically competent staff, and computers per student show positive and statistically significant effects on student attainment. However, we only find a significant effect of computers per student when we use the students' test scores in science as a dependent variable. For instance, Table 3.3 shows that an increase in one unit of a computer per student will increase students' scores in science about 58 points, with a standard deviation of 18. However, unexpectedly, the supply of IT staff at "a little" level compared to "a lot" level has a large positive effect on students' performance. A potential explanation for the weak findings of the IT staff resources at a high level is that there might be other related resources that contribute to the improvement in students outcome such as the provision of netbooks and access to the internet. In 2010, the Malaysian Communications and Multimedia Commission (*SKMM*) had allocated one million 1Malaysia Netbooks with ready broadband access to students from low-income families in particular ([Malaysian Communications and Multimedia Commission, 2010](#)). The allocation possibly indirectly affected the benefit of the provision of IT competent staff in schools on student outcomes. This could be considered a crowding out effect, where the government or the *SKMM* initiatives to allocate netbooks among the students has reduced the usage of IT staff and computers in schools. In this respect, the allocation of netbooks among students that might increase students' interest in learning allows the impact of those IT resources to be lowered, offsetting the government's spending on those same causes. Among others, the benefit of the availability of the computers and technology at an unremarkable level on educational outcomes is in line with

studies by [Bulman and Fairlie \(2016\)](#) and [Fuchs and Wößmann \(2004\)](#).

Similar to IT resources effects, teacher incentives show a weak effect on student attainment. Only when we employ student attainment in science, do teacher incentives have a statistically significant effect on students' scores. In particular, Table 3.3 shows that male students may score 8 points more, while female students may score 5 points more if schools reported that they are willing to provide teacher incentives such as housing, bonuses, and smaller classes in school. For teacher hiring, which refers to the level of schools' difficulty in hiring teachers, both Tables 3.2 and 3.3 show that easiness in teacher hiring has a significant and positive effect on male student attainment. Ease of hiring of teachers may increase male students' test scores by approximately 20 points. This finding appears to indicate that the Ministry's plan in the MEB policy to tighten teacher entry requirements is less effective in enhancing student achievement.

Lastly, we found that principals' role in attending leadership activities is not effective in enhancing students' attainment. This finding is potentially related to the school characteristics such as having limited resources that hinder school principals from conducting the goals that they set after attending the professional leadership activities. In addition, the finding may be in line with [Coelli and Green's \(2012\)](#) argument that show that principals will take time to maximise their efforts at a school. However, we found a significant and strong effect of parents' role in school on students' attainment. In particular, Tables 3.2 and 3.3 show that in the presence of a high level of parental involvement in school, female students' scores will increase about 45 points, while male students' scores will increase about 35 points. The substantial effect of the parental involvement in school corresponding with studies such as by [Fajoku et al. \(2016\)](#) and [Gunnarsson et al. \(2009\)](#) among others.

Overall, we found that students may obtain better scores in mathematics and science if schools have an adequate supply of resources (in particular, "a lot" of

instructional materials and infrastructure while at “a little” level for IT staff), ease in hiring teachers, and a high level of parental involvement in schools. In regards to an attempt to reduce the gender gap between male and female students attainment with boys performing worse than girls, this study suggests boys should, in particular, be in schools with larger classes, a high tendency to easily hire new teachers, and that have an adequate supply of school resources. These overall findings are nearly similar with the results of mean difference test in Table 3.1, except instructional materials effect, thus affirm the above discussion.

Table 3.4: Estimates of school effects on student attainment in mathematics including clustering and sampling weight

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models				IV 2SLS models			
	Yes No		Yes Yes		Yes No		Yes Yes	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)	Boys (7)	Girls (8)
Class size	3.122*** (0.639)	1.916*** (0.585)	3.017*** (0.654)	1.828*** (0.580)	2.021 (1.529)	1.796 (1.147)	1.954 (1.609)	1.725 (1.179)
Class size squared	-0.007 (0.054)	0.020 (0.064)	-0.007 (0.054)	0.020 (0.067)				
Instructional materials (base: a lot)								
Instructional materials (a little)	-29.102*** (11.145)	-18.367* (9.519)	-29.288** (11.256)	-18.794* (9.549)	-30.157*** (10.946)	-18.295** (9.075)	-30.477*** (10.948)	-18.709** (9.053)
Infrastructure (base: a lot)								
Infrastructure (some)	-14.254 (12.859)	-11.587 (11.407)	-12.209 (12.805)	-9.808 (11.476)	-12.709 (12.608)	-11.562 (11.302)	-10.393 (12.563)	-9.840 (11.369)
Infrastructure (little)	-17.023 (16.179)	-20.414 (13.945)	-17.668 (16.280)	-20.201 (13.927)	-16.502 (15.654)	-20.620 (13.690)	-17.002 (15.662)	-20.372 (13.685)
IT staff (base: a lot)								
IT staff (a little)	21.442** (10.475)	19.257** (9.238)	19.627* (10.438)	18.082* (9.213)	21.992** (10.236)	19.417** (9.308)	20.285** (10.187)	18.193* (9.357)
Computers per student	17.568 (52.855)	7.927 (42.683)	22.527 (50.710)	11.885 (42.703)	0.461 (58.488)	7.129 (47.060)	4.639 (57.205)	11.869 (48.755)
Teacher incentives (1 = yes)	2.861 (8.559)	2.759 (7.464)	2.267 (8.670)	2.242 (7.476)	1.519 (8.597)	2.877 (7.355)	0.877 (8.700)	2.333 (7.361)
Teacher hiring (base: difficult)								
Teacher hiring (easy)	17.030 (16.684)	-1.473 (11.744)	20.887 (16.789)	0.516 (12.153)	19.039 (17.545)	-1.561 (11.637)	22.784 (17.816)	0.441 (12.042)
No teacher hiring	19.736 (16.685)	3.279 (11.428)	23.665 (16.994)	5.192 (11.857)	22.800 (17.980)	3.401 (11.506)	26.638 (18.474)	5.290 (11.947)
Principal attendance in leadership activities (base: low)								
Principal attendance in leadership activities (high)	-2.661 (9.762)	2.399 (7.545)	-1.728 (9.828)	3.077 (7.437)	-1.414 (9.646)	2.237 (7.249)	-0.330 (9.801)	2.876 (7.175)
Parental involvement in school (base: low)								
Parental involvement in school (high)	41.867*** (15.256)	43.408*** (12.551)	41.811*** (15.127)	43.774*** (12.406)	40.112*** (15.456)	43.594*** (12.538)	40.194*** (15.332)	43.984*** (12.443)

Continued on next page

Table 3.4 – continued from previous page

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models				IV 2SLS models			
	Yes No		Yes Yes		Yes No		Yes Yes	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)	Boys (7)	Girls (8)
Parental involvement in school (medium)	27.552** (11.289)	28.605*** (10.513)	27.776** (11.101)	28.972*** (10.415)	27.089** (11.051)	28.788*** (10.367)	27.433** (10.876)	29.179*** (10.319)
<i>First-stage estimates</i>								
Dependent variable: Class size								
Eight-grade size					0.023*** (0.005)	0.024*** (0.004)	0.022** (0.005)	0.024*** (0.005)
Number of observations	2815	2918	2815	2918	2815	2918	2815	2918
Adjusted R-squared	0.396	0.373	0.399	0.372	0.391	0.373	0.394	0.372
χ^2 statistic (p-value) of endogeneity test					0.398	0.916	0.445	0.947
First-stage F-statistics					24.971***	29.814***	21.787***	27.086***

Notes: Standard errors in parentheses. Standard error clustered at the school level. *Significant at the 10% level. **Significant at the 5% level. *** Significant at the 1% level. The plausible values refer to the range of competence that a student might have, given the student's item responses in an assessment. Variables missing, students' family background and school characteristics are controlled but not reported for brevity.

As robustness checks for the findings, we specify the education production model including clustering and sampling weight. Our aim is to examine the effectiveness of the school factors explicitly and nationally. Table 3.4 shows the estimated school effects on student scores in mathematics in the presence of the specifications. The estimated results on student scores in science are similar and thus presented in Appendix B (Table A3.5) for simplicity.

As we take account of expected correlated error terms within schools, columns (1), (2), (5), and (6) show unchanged results in comparison to our baseline model in Table 3.2 even though the school effects including the post-estimation tests of the IV method have a weak significance level. Other columns, which have controlled for the sampling weight of total students, have also reported the same magnitude results. However, again, note that the significance level of the variables of interest diminishes. For instance, the effects of school resources (infrastructure) and teacher hiring become insignificant at a 10 percent significant level on student attainment.

3.6.1 School effects on student attainment by school location (urban and rural)

This subsection attempts to present the education production model based on school location (urban and rural areas) to support our discussion for the baseline models and provide appropriate policy responses in respective areas. This subsection also aims to determine to what extent the initiatives within the MEB may help to reduce student achievement gaps based on school location and gender in Malaysia.

Table 3.5: IV estimates of school effects on student attainment in mathematics by school location

Dependent variable: Test scores (i.e. average of 5 plausible values)	Urban		Rural	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)
Class size	-0.892 (1.056)	2.117** (1.065)	4.009*** (0.653)	2.222*** (0.463)
Class size squared				
Instructional materials (base: a lot)				
Instructional materials (a little)	-45.478*** (5.993)	-29.602*** (6.404)	-32.671*** (5.107)	-18.406*** (4.072)
Infrastructure (base: a lot)				
Infrastructure (some)	-27.178*** (5.694)	-28.094*** (6.543)	1.621 (6.254)	0.050 (5.145)
Infrastructure (little)	-45.941*** (9.746)	-41.555*** (9.191)	3.683 (7.049)	-10.247* (5.648)
IT staff (base: a lot)				
IT staff (a little)	37.120*** (6.348)	34.463*** (8.792)	18.590*** (4.629)	14.418*** (4.146)
Computers per student	10.759 (24.081)	-19.394 (22.157)	-23.364 (32.462)	12.324 (27.953)
Teacher incentives (1 = yes)	17.121*** (6.370)	10.028* (5.329)	-2.808 (3.820)	3.001 (3.355)
Teacher hiring (base: difficult)				
Teacher hiring (easy)	8.226 (7.289)	-4.624 (7.621)	21.870*** (8.220)	6.275 (6.532)
No teacher hiring	24.257*** (7.567)	-2.457 (7.010)	19.273** (8.254)	10.975 (6.836)
Principal attendance in leadership activities (base: low)				
Principal attendance in leadership activities (high)	2.785 (4.809)	12.919*** (4.877)	-12.415*** (4.548)	-5.435 (3.681)
Parental involvement in school (base: low)				
Parental involvement in school (high)	-5.754 (8.735)	28.071*** (7.768)	62.865*** (6.293)	52.518*** (5.128)
Parental involvement in school (medium)	-1.767 (6.908)	9.642* (5.817)	47.105*** (5.117)	41.111*** (4.359)
Constant	458.249*** (14.924)	459.789*** (14.361)	396.966*** (14.148)	411.665*** (10.486)
Number of observations	1114	1095	1701	1823
χ^2 statistic (p-value) of endogeneity test	0.002	0.399	0.816	0.820
First-stage F-statistics	151.380***	151.609***	319.352***	577.887***

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level.

*** Significant at the 1% level. The plausible values refer to the range of competence that a student might have, given the student's item responses in an assessment. Variables missing, students' family background, and school characteristics are controlled but not reported for brevity. See Table A3.6 in Appendix B for results using other methods.

Table 3.6: IV estimates of school effects on student attainment in science by school location

Dependent variable: Test scores (i.e. average of 5 plausible values)	Urban		Rural	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)
Class size	-4.532*** (1.275)	-1.583 (1.200)	2.405*** (0.788)	1.105** (0.562)
Instructional materials (base: a lot)				
Instructional materials (a little)	-62.082*** (7.310)	-43.095*** (6.820)	-31.358*** (5.526)	-24.131*** (4.374)
Infrastructure (base: a lot)				
Infrastructure (some)	-10.121 (6.737)	-18.466*** (6.872)	-0.244 (6.496)	4.312 (5.411)
Infrastructure (little)	-19.897* (10.893)	-36.851*** (8.379)	-5.836 (8.021)	-11.959* (6.353)
IT staff (base: a lot)				
IT staff (a little)	40.423*** (8.135)	52.260*** (8.725)	21.219*** (5.361)	20.986*** (4.553)
Computers per student	6.371 (28.599)	-36.652 (24.467)	-15.643 (37.061)	21.739 (28.842)
Teacher incentives (1 = yes)	15.067** (7.681)	13.798** (5.824)	-1.506 (4.445)	3.292 (3.628)
Teacher hiring (base: difficult)				
Teacher hiring (easy)	21.292** (9.131)	-4.079 (8.971)	20.371** (10.359)	9.240 (8.446)
No teacher hiring	57.867*** (9.269)	9.611 (7.964)	14.195 (10.433)	14.526* (8.799)
Principal attendance in leadership activities (base: low)				
Principal attendance in leadership activities (high)	7.466 (5.696)	8.521* (4.950)	-5.714 (5.107)	-0.008 (3.895)
Parental involvement in school (base: low)				
Parental involvement in school (high)	-14.289 (10.356)	41.835*** (7.794)	56.216*** (7.079)	55.270*** (5.730)
Parental involvement in school (medium)	2.600 (8.431)	23.736*** (6.292)	43.603*** (5.948)	46.986*** (4.845)
Constant	390.633*** (18.526)	398.443*** (16.364)	380.615*** (16.312)	367.698*** (12.441)
<i>First-stage estimates</i>				
Dependent variable: Class size				
Eight-grade size	0.020*** (0.002)	0.021*** (0.002)	0.026*** (0.001)	0.027** (0.001)
Number of observations	1114	1095	1701	1823
χ^2 statistic of endogeneity test	0.000	0.000	0.000	0.007
First-stage F-statistics	150.477***	152.038***	338.495***	595.215***

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level.

*** Significant at the 1% level. The plausible values refer to the range of competence that a student might have, given the student's item responses in an assessment. Variables missing, students' family background, and school characteristics are controlled but not reported for brevity. See Table A3.7 in Appendix B for results using other methods. The resulted estimates including clustering and sampling weight specifications are the same as with our baseline model that specified the specifications.

The estimated school effects on student attainment in mathematics and science by school location, urban and rural areas, are shown in Tables 3.5 and 3.6, respectively. In this study, rural areas refer to the medium size city or large town, small town or village, and remote rural. The urban and suburban areas are referred to as the urban area with an estimated population of 10,000 or more people based on indicators by the [Department of Statistics Malaysia \(2014\)](#).

Only Table A3.6 in Appendix B and the Table 3.6) show almost the same findings as in the full sample, Tables 3.2 and 3.3, with class size, school resources, and parental involvement in school show a significant impact on students' outcomes. In addition, we find that the impact of class size changes when we estimate the variable on students' scores in urban schools. In particular, male students in the urban schools tend to perform better as class size decreases, with a ten-unit decrease in class size will increasing boys' scores in science by about 50 points. This finding is consistent with many previous studies and economies of scale, where an institution becomes more efficient if the cost in purchasing capital such as hiring more workers could be minimised.

All school resource factors (institutional materials, infrastructure, and IT staff) show an unchanged impact as in the full sample. Specifically, Tables 3.5 and 3.6 show some evidence that fewer school resources, except the IT staff, have significantly reduced male students' attainment in urban compared to rural areas. The less significant impact of a low level of school resources in rural areas may be due to the government's initiatives in providing free tuition classes and netbooks for students especially those from the rural areas. On the other hand, we assume that students in urban areas are less likely to be affected with a low supply of IT staff because they have a better experience in using technology and may need less help from the IT staff than students in rural areas. Among others, their frequency of using a computer at home might help them to perform better compared to their peers in rural areas. Lastly, a substitution effect possibly explains the unexpected findings, for example, if schools do not need IT staff, then they might choose

to spend that money on better resources such as student exercise books, new classrooms, and so on.

Next, we discuss the impact of teacher incentives, ease of hiring teachers, and principals and parental roles in school on student attainment. Both teacher incentives and teacher hiring show significant and positive effects on male student attainment in urban areas in particular. However, Table 3.6 shows that if schools reported that they find hiring new teachers easy, only male students' achievement in science is improved. In the presence of a high level of principals' role in the schools and parental roles in schools, respectively, male students in rural schools and urban schools tend to perform worse. In addition, male students in rural areas also tend to achieve low test scores if they are attending schools with a high parental involvement. These results may be explained by the reverse causality effect, where the impact of both factors is affected by student outcomes. For instance, even though school principals have encouraged teachers to provide extra tuition classes in order to improve students' performance, students, especially those with a long commute to the area may not make a full effort. In addition, male students' lack of interest in schooling as argued by [Jelas et al. \(2014\)](#) and [Ismail \(2015\)](#) may cause the impact of this effort to be limited for the enhancement of student achievement.

Overall, an increase in the number of students in a classroom in rural areas and school resources regarding instructional materials and infrastructure in urban areas appear to be possible ways in reducing the gender gap between male and female students' in science in particular. Additionally, to reduce the student achievement gap by school location regardless of students' gender, very often parental involvement in the school is recommended.

3.6.2 School effects on student attainment by school size

Some researchers argued that large schools are more cost-effective ([Hanushek, 2003](#); [Jepsen and Rivkin, 2009](#); [Dee and West, 2011](#)). This is because large schools may have the ability to save on costs such as on school resources through bulk purchasing power. In addition, resources specialization and innovative curriculum come at a lower cost in large schools. However, a report by [Hanover Research \(2015\)](#) shows that the benefits of a large school size are only up to certain point¹¹ and that the smallest and largest schools have the highest cost to operate. Thus, in this subsection, we are interested in examining the direct and indirect costs associated with the large schools through the impact of this study's school factors of interest on student attainment based on school size. School size refers to the total number of students enrolled in a school. We interact the school size dummy, whether a school has more than average school enrolment, 974 students enrolled in the given year¹², and school factors in the MEB. We use the school size dummy to provide a guideline for policymakers on whether to increase (e.g. through school consolidation) or decrease school size quotas in Malaysian schools based on the threshold.

¹¹For instance, [Monk \(1987\)](#) shows that the benefit of larger enrolment in a school is diminished once the school's enrolment exceeds 400 students.

¹²The threshold is based on reported average school size in Malaysia from 2011 (see a report by the [Ministry of Education Malaysia \(2013c\)](#)).

Table 3.7: Estimates of school effects on student attainment by dummy of school size (more than 974 students)

Dependent variables: Test scores (i.e. average of 5 plausible values)	Mathematics				Science			
	0 = No		1 = Yes		0 = No		1 = Yes	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)	Boys (7)	Girls (8)
Class size	4.001*** (0.923)	1.819** (0.878)	3.596*** (0.249)	2.221*** (0.257)	5.858*** (1.044)	1.941** (0.974)	4.103*** (0.273)	2.354*** (0.275)
Instructional materials (base: a lot)								
Instructional materials (a little)	-12.745 (9.320)	-18.130** (8.140)	-24.353*** (4.167)	-15.146*** (4.154)	-22.064** (11.044)	-11.766 (8.993)	-28.420*** (4.617)	-20.836*** (4.200)
Infrastructure (base: a lot)								
Infrastructure (some)	-25.406*** (8.939)	-23.586*** (7.287)	-7.025 (4.636)	-1.884 (4.462)	-22.272** (10.086)	-10.449 (8.316)	-8.540* (4.985)	-3.210 (4.451)
Infrastructure (little)	19.787* (10.324)	4.133 (8.069)	-27.517*** (6.308)	-27.115*** (5.956)	32.177*** (11.691)	15.095* (8.971)	-29.683*** (7.040)	-32.754*** (5.943)
IT staff (base: a lot)								
IT staff (a little)	-14.925* (9.030)	-2.991 (7.432)	24.439*** (4.260)	22.661*** (4.011)	-31.948*** (9.618)	-18.249** (8.452)	35.382*** (4.882)	34.646*** (4.272)
Computers per student	-23.202 (34.439)	46.416* (24.932)	5.866 (27.952)	-43.309* (23.223)	15.688 (38.642)	50.038* (27.024)	78.496*** (27.676)	13.164 (23.506)
Teacher incentives (1 = yes)	32.812*** (6.485)	18.303*** (6.629)	-3.577 (3.912)	-1.074 (3.557)	41.996*** (7.586)	22.345*** (7.095)	-1.816 (4.269)	-0.520 (3.751)
Teacher hiring (base: difficult)								
Teacher hiring (easy)	4.781 (10.991)	0.940 (9.875)	11.656* (6.733)	-13.147* (6.726)	-8.784 (13.480)	20.658* (11.815)	26.106*** (7.327)	-7.351 (7.718)
No teacher hiring	1.863 (10.654)	-6.356 (9.885)	19.980*** (6.701)	-2.991 (6.686)	-9.267 (12.884)	-1.367 (11.663)	34.901*** (7.284)	4.933 (7.630)
Parental involvement in school (base: low)								
Parental involvement in school (high)	16.296* (8.521)	16.828*** (6.274)	60.328*** (6.184)	59.917*** (5.209)	29.607*** (10.211)	19.232*** (6.808)	55.522*** (6.841)	69.869*** (5.543)
Parental involvement in school (medium)	11.300* (6.689)	28.255*** (5.861)	37.460*** (4.926)	34.843*** (4.342)	3.017 (7.778)	29.560*** (6.357)	36.399*** (5.643)	45.409*** (4.541)
Number of observations	725	760	2090	2158	725	760	2090	2158
Adjusted R-squared	0.545	0.482	0.391	0.385	0.513	0.513	0.398	0.438

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. *** Significant at the 1% level. The plausible values refer to the range

of competence that a student might have, given the student's item responses in an assessment. Variables missing, students' family background, and school characteristics are controlled but not reported for brevity. We found no significant effect of principal's attendance in leadership activities and less efficient IV estimates, thus they are not reported for brevity.

Table 3.7 reports the estimated results of equation (3.1), including interaction terms between the school size dummy and school factors, and we provide the evidence separately by a dummy of school size for clarity purposes. We can see that the effect of class size based on school size, in general, is almost similar to the class size effect in the baseline estimates, Tables 3.2 and 3.3, with male students appear to benefit more in large class sizes compared to female students if there is an enrolment of 974 students or less in the school. A potential explanation for this is that *ceteris paribus*, the proportion of the number of students to teachers in the school is low which result in teachers are more likely to be efficient in delivery and students performing better. The increment in the proportion also may be lower than that of a larger size schools.

As expected, Table 3.7 shows that a substantial supply of instructional materials and infrastructure are more likely to benefit students who are attending large sized schools (more than 974 students enrolled in school). A potential explanation for these findings is that large sized schools are able to provide more resources (economies of scale), then students have the opportunity to use the abundant resources in their learning. For the supply of IT staff, we found a similar result as with our baseline models, Tables 3.2 and 3.3, especially for schools with a large capacity. On the other hand, schools with a small capacity (i.e. low number of students enrolled) show that the impact of the shortage in IT staff in the school will decrease students' scores in science in particular. These findings give another potential explanations of why IT staff at "a little" level in the baseline models is contradict with many previous findings. Besides, peer effects may contribute to the positive effect of a shortage in IT staff in the large schools. For instance, large schools probably have some students who have a high ability in IT that may help their peers who are less skilled. However, we highlight here that these explanations may be weak as we have no suitable instruments to address potential unobserved effects of the IT staff (for example, the IT staffs' level of expertise and their intervention approach in helping the students) on student attainment.

Looking at the computers per students, the ease of the hiring of teachers, and a high level of parental involvement in school, we found that these school factors are more likely to benefit boys in large schools. In contrast, male students may benefit from the presence of teacher incentives only in small-sized schools. The positive relationship between parental involvement in school and student attainment, in large schools in particular, is inconsistent with a study by [Dee et al. \(2007\)](#) that show that high parental involvement in school is positively related to student attainment in small schools, in general. However, to note, the effect of parental involvement in small school sizes is still positive although it is less than the effect in large sized schools (more than 974 students). Overall, the positive impacts of the school factors on student attainment in large schools indicate the economically efficient characteristics of the implementation of corresponding school factors on student outcomes.

3.6.3 Sensitivity analysis

Overall, findings from this study show pronounced differences in comparison to many previous studies. However, we have provided some potential explanations associated with the respective results, and some of them are not supported by statistical arguments due to data limitations. Thus, to support the appropriateness of the explanations that we have been made, in this subsection, we examine the baseline model of this study, including alternative specifications. These include class size dummy, where we replace the discrete variable of class size into a dummy variable according to average class size in Malaysia in the given year¹³ and exclude the control variables of students' family background and school characteristics. Lastly, we present evidence for the class size effects on students' attainment based on their socioeconomic backgrounds as an alternative measure to address the potential bias problem in the class size estimates.

¹³See [Ministry of Education Malaysia \(2013c\)](#).

Table 3.8: Estimates of school effects on student attainment including dummy class size

Dependent variables: Test scores (i.e. average of 5 plausible values) Family background controls School characteristics controls	Mathematics				Science			
	No		Yes		No		Yes	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)	Boys (7)	Girls (8)
Dummy class size (more than 28 students =1)	71.365*** (4.785)	43.190*** (5.252)	55.099*** (4.335)	31.385*** (4.585)	80.092*** (5.447)	42.903*** (5.854)	64.705*** (4.889)	33.252*** (4.971)
Instructional materials (base: a lot)								
Instructional materials (a little)	-35.275*** (4.281)	-30.745*** (3.784)	-25.538*** (3.835)	-17.968*** (3.483)	-38.387*** (4.695)	-34.004*** (3.981)	-30.057*** (4.259)	-22.750*** (3.599)
Infrastructure (base: a lot)								
Infrastructure (some)	-25.544*** (4.492)	-17.634*** (4.264)	-17.491*** (4.012)	-11.824*** (3.902)	-20.836*** (4.773)	-16.284*** (4.380)	-15.952*** (4.310)	-11.207*** (3.917)
Infrastructure (little)	-28.624*** (5.903)	-22.333*** (5.170)	-18.946*** (5.198)	-20.300*** (4.626)	-26.366*** (6.615)	-26.460*** (5.412)	-19.019*** (5.854)	-22.729*** (4.667)
IT staff (base: a lot)								
IT staff (a little)	32.516*** (4.200)	27.129*** (3.910)	24.296*** (3.657)	20.354*** (3.450)	35.667*** (4.665)	34.379*** (4.205)	28.131*** (4.167)	27.487*** (3.622)
Computers per student	18.331 (19.192)	15.292 (16.359)	-3.516 (17.358)	-7.157 (15.137)	78.348*** (19.966)	65.584*** (16.536)	31.494* (17.565)	28.877* (15.117)
Teacher incentives (1 = yes)	4.279 (3.569)	6.423** (3.144)	1.649 (3.098)	3.373 (2.781)	10.526*** (4.017)	10.261*** (3.406)	6.317* (3.487)	5.425* (2.949)
Teacher hiring (base: difficult)								
Teacher hiring (easy)	10.611* (6.002)	-13.659** (5.334)	20.201*** (5.250)	-0.531 (4.748)	17.937*** (6.833)	-7.219 (6.421)	24.586*** (5.996)	3.902 (5.551)
No teacher hiring	17.967*** (5.853)	-3.074 (5.148)	28.384*** (5.212)	6.928 (4.633)	25.869*** (6.673)	1.574 (6.175)	35.236*** (5.932)	12.161** (5.404)
Principal attendance in leadership activities (base: low)								
Principal attendance in leadership activities (high)	2.560 (3.453)	7.635** (2.984)	-3.582 (3.210)	0.634 (2.740)	4.231 (3.752)	11.411*** (3.184)	-3.362 (3.475)	1.622 (2.807)
Parental involvement in school (base: low)								
Parental involvement in school (high)	71.805*** (5.315)	69.814*** (4.434)	44.846*** (4.922)	44.507*** (4.173)	71.183*** (5.816)	80.965*** (4.905)	44.044*** (5.495)	52.404*** (4.508)
Parental involvement in school (medium)	42.036*** (4.374)	39.772*** (3.957)	26.964*** (3.832)	29.935*** (3.527)	43.363*** (4.885)	50.977*** (4.424)	26.228*** (4.451)	37.765*** (3.800)
Adjusted R-squared	0.231	0.191	0.396	0.369	0.233	0.217	0.389	0.412

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level.

Table 3.8 shows that most of the school factors present consistent results with our baseline models. A class size dummy also shows unchanged impacts: boys who have attended a classroom with more than 28 students would expect to achieve scores almost twice that of girls' scores. However, noteworthy, the class size effect decreases substantially as we take into account students' family and school characteristics. It indicates the rationale of our next strategy, which is to quantify the class size effect based on students' socioeconomic background.

Table 3.9: Heterogeneity in class size effects

Dependent variables: Test scores (i.e. average of 5 plausible values)	Mathematics		Science	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)
Class size	2.587** (1.024)	3.171*** (0.960)	3.806*** (1.166)	2.550** (1.092)
<i>Interaction terms</i>				
Class size × Mother's education (base: no formal or primary)				
Secondary	−0.540 (0.815)	−0.846 (0.688)	−1.963** (0.984)	−1.510* (0.774)
Tertiary	−0.904 (1.028)	−1.276 (0.875)	−3.094** (1.210)	−2.009** (0.978)
Child's does not know	−0.513 (0.881)	−1.288 (0.790)	−2.049* (1.052)	−1.174 (0.885)
Father's education (base: no formal or primary)				
Secondary	−0.004 (0.847)	−0.789 (0.782)	0.796 (0.945)	0.602 (0.906)
Tertiary	−1.463 (0.967)	−2.431** (0.979)	−0.050 (1.103)	−0.855 (1.099)
Child's does not know	0.053 (0.900)	0.325 (0.836)	1.017 (0.996)	1.165 (0.979)
<i>Home possessions</i>				
One and more bookcases	−0.048 (0.538)	−0.604 (0.471)	−0.879 (0.581)	−0.732 (0.494)
Reading resources (low)	1.399*** (0.459)	1.523*** (0.528)	2.291*** (0.495)	1.772*** (0.566)
Technology resources (low)	−0.508 (0.461)	−0.395 (0.481)	−0.876* (0.502)	0.009 (0.531)
Study room (low)	0.050 (0.424)	−0.183 (0.430)	0.301 (0.455)	−0.348 (0.447)
Rural	1.891*** (0.450)	0.930** (0.468)	1.930*** (0.490)	0.713 (0.501)
Number of observations	2815	2918	2815	2918
Adjusted R-squared	0.403	0.384	0.406	0.426

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level. The plausible values refer to the range of competence that a student might have, given the student's item responses in an assessment. Missing observations, main effects of the student characteristics, other school factors and school characteristics are controlled. See Appendix B Table A3.8 for full results.

The interaction-terms effects between class size and students' socioeconomic backgrounds in addition to our baseline model, equation (3.1), are presented in Table 3.9. We can see that the increase in students' scores in a large class size differs for high and low socioeconomic status students. We found consistent findings with studies by Lazear (2001), Fajoku et al. (2016), Gunnarsson et al. (2009), and Martins and Veiga (2010) among others, of the positive effect of class size reduction for education outcomes of advantaged students. Lazear (2001) shows that students that come from the advantaged home have parents who tend to be actively involved in their children's schooling apart from their ability to send their children to the best schools with adequate resources.

However, in accordance with the main findings of this study, disadvantaged students who have low educated parents, a low supply of reading resources at home, and have attended rural schools appear to benefit more from large class sizes than their advantaged peers. Factors such as financial constraints in low-income families may influence children's motivation to aspire for higher education even though they are allocated to large classes (a reverse causality effect). In addition, the government initiatives to provide scholarships for excellent students in higher education level may boost students' motivation to excel in the classroom. However, it should be emphasised that, while statistically significant, the difference in these heterogeneous class size effects based on students' socioeconomic status are relatively small in magnitude. This implies that the initiatives from education policymakers among others to address the achievement gap problem by student socioeconomic background, such as tailoring class size to a different groups of students, would be less effective.

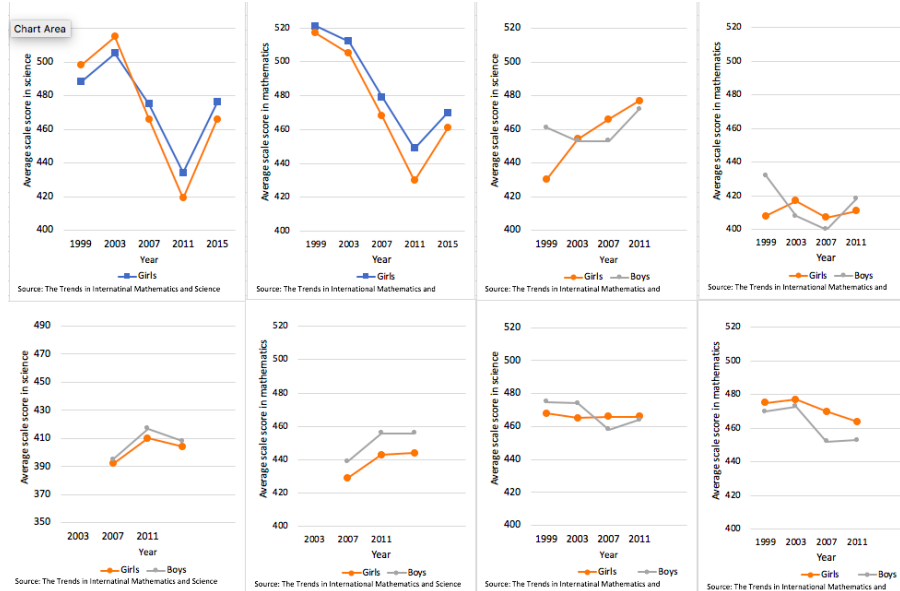
3.7 Conclusion

Using data on eighth-grade Malaysian students from the Trends in International Mathematics and Science Study (TIMSS), we examined the relationship between initiatives within the MEB policy, which in this study we refer to as school factors and student attainment by gender.

Although we found a strong positive effect of school factors, especially school resources (i.e. instructional materials and infrastructure) and parental involvement in schools, we found tentative evidence of the class size and IT staff on student attainment. Specifically, we find that male students in urban schools and small school sizes tend to perform worse if schools have large class sizes and “a little” level of IT staff. Factors such as differences in practices in schools and peer effects might explain these results, as argued by [Hattie \(2005\)](#), [Woessmann \(2016\)](#), and [Lazear \(2001\)](#) among others. While most of our education model, including the model of class size by student socioeconomic background, show a positive effect of the two factors on student attainment. These findings imply that strategic interventions in implementing both factors should be emphasised. For further study, an education production model that controls for the type of schools is suggested. Also, the TIMSS over-samples large schools (or class sizes) and rural schools should be addressed. The intuition is to lessen the effect of different school practices or autonomy between the school types or sample that we expect may lead to inconclusive evidence of the two factors.

In addition to class size, teacher hiring and school resources in regards to instructional materials and infrastructure appear to become effective initiatives to alleviate the issue of the student achievement gap by gender in Malaysia. Other than that, parental engagement in their child’s schools was also found to help in lessening the gender gap in student achievement if we take into account students’ school location and school size. The findings suggest that future education plans should consist of those aspects in order to improve school outcomes in Malaysia.

Appendix B



Note: Author's calculation.

Figure A3.1: The Malaysian performance in TIMSS assessment by gender and year

Table A3.1: Rotated component matrix for school resources			
Variables	Component		
	1	2	3
Instructional materials (e.g. textbooks)	0.690	-0.057	-0.054
Supplies (e.g. papers, pencils)	0.662	-0.058	-0.028
School buildings and grounds	0.220	0.381	0.137
Heating/cooling and lighting systems	-0.076	0.858	-0.100
Instructional space (e.g. classrooms)	0.176	0.328	0.270
Technologically competent staff	-0.032	-0.063	0.946
Extraction method: Principal Component.			
Rotation method: Varimax with Kaiser Normalization.			

Table A3.2: Rotated component matrix for home possessions			
Variables	Component		
	1	2	3
Computer	0.022	0.623	-0.060
Study desk/table for your use	0.210	0.056	0.490
Books of your very own (do not count your school books)	0.521	-0.054	-0.064
Your own room	-0.052	-0.038	0.839
Internet connection	0.001	0.632	-0.053
Calculator	0.403	0.145	-0.043
Dictionary	0.468	-0.040	0.075
Reference book	0.538	0.007	-0.043
Video	-0.101	0.427	0.190
Extraction method: Principal Component.			
Rotation method: Varimax with Kaiser Normalization.			

Table A3.3: Estimates of school factors on student attainment in mathematics

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models		IV 2SLS models			
			instrument = eight-grade size		instruments = eight-grade size and school size	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)
Class size	3.122*** (0.227)	1.916*** (0.221)	2.021*** (0.529)	1.796*** (0.425)	2.499*** (0.508)	2.345*** (0.403)
Class size squared	-0.007 (0.020)	0.020 (0.025)				
Instructional materials (base: a lot) Instructional materials (a little)	-29.102*** (3.823)	-18.367*** (3.445)	-30.157*** (3.793)	-18.295*** (3.346)	-29.735*** (3.789)	-17.583*** (3.351)
Infrastructures (base: a lot) Infrastructure (some)	-14.254*** (4.049)	-11.587*** (3.938)	-12.709*** (4.028)	-11.562*** (3.908)	-13.346*** (4.034)	-12.220*** (3.924)
Infrastructure (little)	-17.023*** (5.264)	-20.414*** (4.620)	-16.502*** (5.087)	-20.620*** (4.549)	-16.648*** (5.096)	-20.617*** (4.565)
IT staff (base: a lot) IT staff (a little)	21.442*** (3.598)	19.257*** (3.395)	21.992*** (3.542)	19.417*** (3.437)	21.700*** (3.536)	18.706*** (3.433)
Computers per student	17.568 (17.388)	7.927 (15.165)	0.461 (18.956)	7.129 (16.402)	7.769 (18.787)	16.603 (16.342)
Teacher incentives (1 = yes)	2.861 (3.065)	2.759 (2.775)	1.519 (3.100)	2.877 (2.753)	2.067 (3.080)	2.971 (2.754)
Teacher hiring (base: difficult) Teacher hiring (easy)	17.030*** (5.306)	-1.473 (4.686)	19.039*** (5.430)	-1.561 (4.674)	18.239*** (5.350)	-1.920 (4.633)
No teacher hiring	19.736*** (5.246)	3.279 (4.597)	22.800*** (5.537)	3.401 (4.651)	21.522*** (5.430)	2.356 (4.599)
Principal attendance in leadership activities (base: low) Principal attendance in leadership activities (high)	-2.661 (3.226)	2.399 (2.726)	-1.414 (3.216)	2.237 (2.667)	-1.924 (3.207)	1.919 (2.667)
Parental involvement in school (base: low) Parental involvement in school (high)	41.867*** (4.991)	43.408*** (4.221)	40.112*** (5.045)	43.594*** (4.210)	40.830*** (5.041)	44.060*** (4.210)

Continued on next page

Table A3.3: – continued from previous page

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models		IV 2SLS models			
			instrument = eight-grade size		instruments = eight-grade size and school size	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)
Parental involvement in school (medium)	27.552*** (3.944)	28.605*** (3.589)	27.089*** (3.914)	28.788*** (3.558)	27.288*** (3.923)	28.574*** (3.564)
Family background						
Mother's education (base: no formal or primary)						
Secondary	4.335 (5.640)	10.104** (4.195)	3.714 (5.677)	10.169** (4.170)	3.974 (5.634)	9.744** (4.140)
Tertiary	13.840** (6.807)	20.846*** (5.237)	13.315* (6.800)	21.070*** (5.195)	13.541** (6.772)	20.781*** (5.182)
Child's does not know	−0.979 (5.970)	2.245 (5.139)	−0.829 (5.999)	2.449 (5.103)	−0.890 (5.962)	1.926 (5.079)
Father's education (base: no formal or primary)						
Secondary	2.621 (6.632)	−2.470 (5.168)	3.440 (6.642)	−2.645 (5.134)	3.108 (6.610)	−2.393 (5.103)
Tertiary	10.480 (7.333)	−0.279 (5.995)	11.938 (7.329)	−0.570 (5.936)	11.314 (7.308)	−0.312 (5.918)
Child's does not know	0.828 (6.759)	−8.275 (5.740)	1.444 (6.763)	−8.535 (5.697)	1.181 (6.733)	−8.198 (5.670)
<i>Home possessions</i>						
One and more bookcases	28.275*** (3.381)	30.660*** (2.741)	29.002*** (3.393)	30.794*** (2.721)	28.698*** (3.376)	30.714*** (2.725)
Reading resources (low)	−28.355*** (3.136)	−32.891*** (3.507)	−29.529*** (3.173)	−32.877*** (3.495)	−29.028*** (3.157)	−32.351*** (3.473)
Technology resources (low)	−42.043*** (3.303)	−32.774*** (2.889)	−43.507*** (3.384)	−32.976*** (2.914)	−42.865*** (3.371)	−32.313*** (2.911)
Study room (low)	7.875*** (2.859)	8.248*** (2.570)	8.096*** (2.856)	8.406*** (2.553)	7.996*** (2.847)	8.243*** (2.554)
School characteristics						
Students come from disadvantaged homes (> 50%)	−15.695*** (3.179)	−15.044*** (2.881)	−17.019*** (3.193)	−15.253*** (2.969)	−16.409*** (3.180)	−14.419*** (2.955)
Rural	−5.295* (3.177)	−4.671* (2.825)	−5.949* (3.156)	−4.597 (2.817)	−5.694* (3.152)	−4.241 (2.823)

Continued on next page

Table A3.3: – continued from previous page

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models		IV 2SLS models			
			instrument = eight-grade size		instruments = eight-grade size and school size	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)
Instructional time (6 and more hours per day)	16.490*** (3.486)	11.256*** (3.024)	14.425*** (3.561)	10.623*** (3.028)	15.353*** (3.541)	11.536*** (3.010)
Constant	421.105*** (10.012)	434.083*** (8.050)	423.234*** (9.883)	435.414*** (7.993)	422.121*** (9.840)	433.600*** (7.975)
<i>First-stage estimates</i>						
Dependent variable: Class size						
Eight-grade size			0.023*** (0.001)	0.024*** (0.001)	0.014*** (0.002)	0.013*** (0.001)
School size (enrolment)					0.002*** (0.000)	0.003*** (0.000)
Number of observations	2815	2918	2815	2918	2815	2918
Adjusted R-squared	0.396	0.373	0.391	0.373	0.395	0.372
χ^2 statistic (p-value) of endogeneity test			0.018	0.775	0.153	0.228
First-stage F-statistics			455.956***	525.401***	278.255***	414.667***
χ^2 statistic (p-value) of overidentifying restrictions test					0.000	0.000

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. *** Significant at the 1% level. The plausible values refer to the range of competence that a student might have, given the student's item responses in an assessment. Variables missing are controlled but omitted for brevity.

Table A3.4: Estimates of school effects on student attainment in science

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models		IV 2SLS models			
			instrument = eight-grade size		instruments = eight-grade size and school size	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)
Class size	3.771*** (0.253)	2.076*** (0.239)	0.329 (0.619)	0.335 (0.487)	0.609 (0.601)	0.468 (0.473)
Class size squared	-0.022 (0.023)	0.002 (0.026)				
Instructional materials (base: a lot) Instructional materials (a little)	-33.978*** (4.209)	-22.893*** (3.538)	-37.280*** (4.321)	-25.129*** (3.549)	-37.033*** (4.300)	-24.957*** (3.547)
Infrastructure (base: a lot) Infrastructure (some)	-12.454*** (4.332)	-11.126*** (3.943)	-7.619* (4.340)	-9.050** (3.927)	-7.992* (4.326)	-9.209** (3.919)
Infrastructure (little)	-17.167*** (5.932)	-23.049*** (4.681)	-15.525*** (5.741)	-23.077*** (4.589)	-15.611*** (5.732)	-23.076*** (4.587)
IT staff (base: a lot) IT staff (a little)	24.964*** (4.108)	26.270*** (3.584)	26.674*** (4.173)	28.523*** (3.679)	26.502*** (4.154)	28.351*** (3.670)
Computers per student	58.444*** (17.859)	46.889*** (15.454)	4.954 (20.436)	16.966 (17.306)	9.239 (20.220)	19.252 (17.190)
Teacher incentives (1 = yes)	8.024** (3.476)	4.917* (2.955)	3.823 (3.590)	4.632 (2.940)	4.145 (3.566)	4.655 (2.936)
Teacher hiring (base: difficult) Teacher hiring (easy)	20.348*** (6.071)	2.712 (5.527)	26.641*** (6.524)	3.837 (5.642)	26.171*** (6.464)	3.750 (5.623)
No teacher hiring	24.552*** (5.994)	8.104 (5.399)	34.138*** (6.607)	11.408** (5.572)	33.388*** (6.532)	11.156** (5.548)
Principal attendance in leadership activities (base: low) Principal attendance in leadership activities (high)	-2.542 (3.481)	3.240 (2.789)	1.362 (3.577)	4.228 (2.770)	1.063 (3.561)	4.151 (2.768)
Parental involvement in school (base: low) Parental involvement in school (high)	40.915*** (5.510)	51.559*** (4.522)	35.419*** (5.655)	50.106*** (4.554)	35.840*** (5.642)	50.219*** (4.550)

Continued on next page

Table A3.4: – continued from previous page

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models		IV 2SLS models			
			instrument = eight-grade size		instruments = eight-grade size and school size	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)
Parental involvement in school (medium)	26.971*** (4.510)	36.470*** (3.803)	25.525*** (4.572)	37.162*** (3.823)	25.641*** (4.561)	37.110*** (3.814)
Family background						
Mother's education (base: no formal or primary)						
Secondary	3.852 (7.020)	3.289 (4.754)	1.909 (7.547)	4.634 (4.914)	2.061 (7.487)	4.532 (4.898)
Tertiary	19.955** (8.106)	26.263*** (5.721)	18.315** (8.526)	27.193*** (5.806)	18.447** (8.470)	27.123*** (5.791)
Child's does not know	-5.903 (7.399)	-9.539* (5.653)	-5.433 (7.829)	-7.876 (5.809)	-5.469 (7.775)	-8.002 (5.789)
Father's education (base: no formal or primary)						
Secondary	9.269 (7.743)	9.760 (6.019)	11.832 (8.138)	8.951 (6.124)	11.637 (8.083)	9.012 (6.108)
Tertiary	30.765*** (8.425)	22.256*** (6.788)	35.324*** (8.778)	21.419*** (6.838)	34.958*** (8.720)	21.481*** (6.824)
Child's does not know	10.683 (7.936)	4.980 (6.595)	12.609 (8.287)	3.897 (6.695)	12.455 (8.235)	3.978 (6.678)
<i>Home possessions</i>						
One and more bookcases	27.437*** (3.663)	29.793*** (2.924)	29.712*** (3.730)	30.057*** (2.933)	29.534*** (3.710)	30.038*** (2.928)
Reading resources (low)	-49.195*** (3.514)	-57.544*** (3.797)	-52.869*** (3.714)	-59.202*** (3.876)	-52.574*** (3.688)	-59.075*** (3.861)
Technology resources (low)	-14.946*** (3.612)	-7.620** (3.109)	-19.521*** (3.843)	-9.726*** (3.170)	-19.145*** (3.817)	-9.566*** (3.165)
Study room (low)	7.046** (3.116)	5.864** (2.713)	7.735** (3.215)	6.390** (2.716)	7.677** (3.198)	6.351** (2.710)
School characteristics						
Students come from disadvantaged homes (> 50%)	-12.742*** (3.519)	-14.376*** (3.005)	-16.874*** (3.637)	-17.025*** (3.090)	-16.516*** (3.610)	-16.823*** (3.068)
Rural	-4.010 (3.504)	-1.057 (2.903)	-6.059* (3.564)	-2.171 (2.925)	-5.909* (3.552)	-2.085 (2.926)

Continued on next page

Table A3.4: – continued from previous page

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models		IV 2SLS models			
			instrument = eight-grade size		instruments = eight-grade size and school size	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)
Instructional time (6 and more hours per day)	21.792*** (3.846)	18.537*** (3.294)	15.344*** (4.060)	15.604*** (3.329)	15.888*** (4.028)	15.824*** (3.308)
Constant	376.587*** (11.500)	375.348*** (9.312)	383.203*** (11.786)	381.181*** (9.317)	382.551*** (11.731)	380.743*** (9.310)
<i>First-stage estimates</i>						
Dependent variable: Class size						
Eight-grade size			0.023*** (0.001)	0.024*** (0.001)	0.014*** (0.002)	0.013*** (0.001)
School size (enrolment)					0.002*** (0.000)	0.003*** (0.000)
Number of observations	2815	2918	2815	2918	2815	2918
Adjusted R-squared	0.393	0.416	0.350	0.404	0.357	0.406
χ^2 statistic (p-value) of endogeneity test			0.000	0.000	0.000	0.000
First-stage F-statistics			455.956***	525.401***	278.255***	414.667***
χ^2 statistic (p-value) of overidentifying restrictions test					0.031	0.311

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. *** Significant at the 1% level. The plausible values refer to the range of competence that a student might have, given the student's item responses in an assessment. Variables missing are controlled but omitted for brevity.

Table A3.5: Estimates of school effects on student attainment in science including clustering and sampling weight specifications

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models				IV 2SLS models			
	Yes No		Yes Yes		Yes No		Yes Yes	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)	Boys (7)	Girls (8)
Class size	3.771*** (0.714)	2.076*** (0.653)	3.734*** (0.728)	2.040*** (0.650)	0.329 (1.917)	0.335 (1.442)	0.267 (2.026)	0.264 (1.464)
Class size squared	-0.022 (0.059)	0.002 (0.069)	-0.024 (0.060)	-0.001 (0.071)				
Instructional materials (base: a lot)								
Instructional materials (a little)	-33.978*** (12.335)	-22.893** (9.826)	-33.486*** (12.398)	-22.806** (9.651)	-37.280*** (13.018)	-25.129** (9.848)	-37.389*** (12.974)	-25.407*** (9.540)
Infrastructure (base: a lot)								
Infrastructure (some)	-12.454 (13.507)	-11.126 (11.107)	-9.432 (13.455)	-8.772 (11.230)	-7.619 (13.324)	-9.050 (10.909)	-3.484 (13.288)	-6.595 (10.997)
Infrastructure (little)	-17.167 (18.665)	-23.049* (13.847)	-16.114 (18.881)	-21.158 (13.858)	-15.525 (18.242)	-23.077* (13.401)	-13.895 (18.325)	-20.891 (13.236)
IT staff (base: a lot)								
IT staff (a little)	24.964** (12.305)	26.270** (10.454)	20.641* (12.169)	23.329** (10.238)	26.674** (12.675)	28.523*** (10.781)	22.753* (12.549)	26.005** (10.542)
Computers per student	58.444 (54.010)	46.889 (43.623)	61.339 (53.163)	48.113 (43.169)	4.954 (64.684)	16.966 (50.584)	2.855 (65.029)	14.556 (51.078)
Teacher incentives (1 = yes)	8.024 (10.262)	4.917 (8.388)	8.286 (10.369)	5.395 (8.316)	3.823 (10.600)	4.632 (8.196)	3.732 (10.850)	5.157 (8.194)
Teacher hiring (base: difficult)								
Teacher hiring (easy)	20.348 (19.263)	2.712 (15.443)	22.981 (19.458)	3.515 (15.962)	26.641 (21.674)	3.837 (15.716)	29.205 (22.230)	4.519 (16.445)
No teacher hiring	24.552 (19.413)	8.104 (15.019)	27.302 (19.750)	8.427 (15.567)	34.138 (22.160)	11.408 (15.398)	37.028 (22.954)	11.757 (16.223)
Principal attendance in leadership activities (base: low)								
Principal attendance in leadership activities (high)	-2.542 (10.569)	3.240 (7.930)	-1.240 (10.668)	3.990 (7.885)	1.362 (10.893)	4.228 (7.847)	3.342 (11.025)	5.175 (7.822)
Parental involvement in school (base: low)								
Parental involvement in school (high)	40.915** (16.663)	51.559*** (13.712)	41.026** (16.487)	51.335*** (13.666)	35.419** (17.219)	50.106*** (13.761)	35.731** (17.058)	49.345*** (13.721)

Continued on next page

Table A3.5: – continued from previous page

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS models				IV 2SLS models			
	Yes No		Yes Yes		Yes No		Yes Yes	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)	Boys (7)	Girls (8)
Parental involvement in school (medium)	26.971** (13.438)	36.470*** (11.033)	27.935** (13.213)	35.812*** (11.014)	25.525* (13.572)	37.162*** (10.993)	26.813** (13.571)	36.119*** (11.018)
<i>First-stage estimates</i>								
Dependent variable: Class size								
Eight-grade size					0.023*** (0.005)	0.024*** (0.004)	0.022** (0.005)	0.024*** (0.005)
Number of observations	2815	2918	2815	2918	2815	2918	2815	2918
Adjusted R-squared	0.393	0.416	0.397	0.416	0.350	0.404	0.353	0.403
χ^2 statistic (p-value) of endogeneity test					0.013	0.117	0.019	0.122
First-stage F-statistics					24.971***	29.814***	21.787***	27.086***

Notes: Standard errors in parentheses. Standard error clustered at the school level. *Significant at the 10% level. **Significant at the 5% level. *** Significant at the 1% level. The plausible values refer to the range of competence that a student might have, given the student's item responses in an assessment. Variables missing, students' family background and school characteristics are controlled but not reported for brevity.

Table A3.6: Estimates of school factors on student attainment in mathematics by school location

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS				OLS including clustering and sampling weight			
	Urban		Rural		Urban		Rural	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)	Boys (7)	Girls (8)
Class size	2.206*** (0.381)	1.424*** (0.409)	4.129*** (0.291)	2.259*** (0.267)	2.392*** (0.880)	1.528 (0.943)	3.922*** (0.792)	2.131*** (0.717)
Class size squared	0.027 (0.050)	0.080 (0.048)	-0.027 (0.025)	0.010 (0.029)	0.036 (0.134)	0.086 (0.124)	-0.030 (0.067)	0.006 (0.076)
Instructional materials (base: a lot)								
Instructional materials (a little)	-41.454*** (5.709)	-30.068*** (6.252)	-31.922*** (5.212)	-18.473*** (4.172)	-45.009*** (12.571)	-31.015** (13.239)	-31.127** (15.528)	-18.688 (11.809)
Infrastructure (base: a lot)								
Infrastructure (some)	-33.940*** (5.615)	-27.259*** (6.245)	0.818 (6.214)	0.104 (5.193)	-31.385** (13.711)	-25.821 (16.136)	3.447 (20.201)	1.656 (14.005)
Infrastructure (little)	-52.489*** (11.019)	-40.965*** (8.867)	2.641 (7.234)	-10.150* (5.729)	-51.612 (34.425)	-38.391 (25.377)	3.278 (22.871)	-9.481 (15.611)
IT staff (base: a lot)								
IT staff (a little)	32.893*** (8.108)	35.050*** (7.376)	18.719*** (4.679)	14.516*** (4.182)	31.435 (21.638)	31.542* (18.074)	15.593 (13.850)	13.378 (10.604)
Computers per student	31.786 (21.742)	-27.333 (20.856)	-20.661 (29.677)	12.188 (25.536)	45.672 (54.925)	-18.352 (47.033)	-30.197 (85.112)	15.021 (80.532)
Teacher incentives (1 = yes)	24.196*** (6.370)	7.055 (5.697)	-2.662 (3.852)	2.954 (3.364)	26.534 (18.307)	6.396 (15.354)	-4.428 (10.601)	1.861 (9.037)
Teacher hiring (base: difficult)								
Teacher hiring (easy)	7.687 (7.098)	-2.317 (7.967)	20.270** (8.211)	6.205 (6.564)	13.518 (19.543)	0.171 (18.980)	23.296 (24.754)	8.711 (17.964)
No teacher hiring	15.053** (7.065)	0.406 (7.021)	17.793** (8.037)	10.886 (6.845)	17.835 (20.320)	0.558 (15.848)	21.867 (24.156)	14.613 (18.403)
Principal attendance in leadership activities (base: low)								
Principal attendance in leadership activities (high)	3.527 (4.983)	12.933*** (4.702)	-12.690*** (4.515)	-5.371 (3.703)	6.717 (13.303)	16.095 (12.139)	-12.838 (13.500)	-5.595 (10.005)
Parental involvement in school (base: low)								
Parental involvement in school (high)	-0.180 (8.492)	26.065*** (7.360)	63.616*** (6.338)	52.397*** (5.225)	-1.022 (23.083)	27.012 (20.063)	62.188*** (19.153)	50.852*** (14.617)
Parental involvement in school (medium)	-4.871 (7.086)	9.630 (5.886)	47.410*** (4.990)	41.060*** (4.405)	-6.807 (18.964)	8.935 (16.561)	46.312*** (12.977)	40.164*** (11.673)

Continued on next page

Table A3.6: – continued from previous page

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS				OLS including clustering and sampling weight			
	Urban		Rural		Urban		Rural	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)	Boys (7)	Girls (8)
<i>First-stage estimates</i>								
Dependent variable: Class size								
Eight-grade size	0.019*** (0.002)	0.021*** (0.002)	0.026** (0.001)	0.027*** (0.001)				
Number of observations	1114	1095	1701	1823	1114	1095	1701	1823
Adjusted R-squared	0.422	0.338	0.421	0.401	0.439	0.338	0.416	0.394

Notes: Standard errors in parentheses. Standard error clustered at the school level. *Significant at the 10% level. **Significant at the 5% level. *** Significant at the 1% level. The plausible values refer to the range of competence that a student might have, given the student's item responses in an assessment. Variables missing, students' family background and school characteristics are controlled but not reported for brevity. There are fewer significant coefficients in the IV estimates, so we do not specify it using clustering and sampling weight specifications.

Table A3.7: Estimates of school factors on student attainment in science by school location

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS				IV 2SLS including clustering and sampling weight			
	Urban		Rural		Urban		Rural	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)	Boys (7)	Girls (8)
Class size	2.506*** (0.419)	2.014*** (0.429)	4.966*** (0.321)	2.365*** (0.294)	-3.880 (3.661)	-1.547 (3.189)	1.993 (2.770)	0.899 (1.856)
Class size squared	-0.036 (0.054)	0.029 (0.048)	-0.035 (0.027)	0.002 (0.032)				
Instructional materials (base: a lot) Instructional materials (a little)	-54.466*** (6.428)	-36.889*** (6.396)	-27.834*** (5.528)	-22.071*** (4.361)	-66.180*** (21.038)	-45.881*** (15.473)	-30.446* (16.539)	-24.182** (11.818)
Infrastructure (base: a lot) Infrastructure (some)	-26.430*** (6.231)	-28.102*** (6.211)	-5.060 (6.481)	2.397 (5.406)	-4.084 (18.574)	-15.301 (17.952)	4.355 (21.095)	7.103 (14.431)
Infrastructure (little)	-38.914*** (11.830)	-36.106*** (8.668)	-7.170 (8.142)	-12.240* (6.371)	-15.175 (34.378)	-32.640 (21.162)	-3.272 (26.067)	-9.306 (17.984)
IT staff (base: a lot) IT staff (a little)	37.623*** (8.980)	36.588*** (7.338)	20.854*** (5.312)	20.909*** (4.571)	35.764 (22.476)	50.231** (20.927)	17.022 (16.710)	18.129 (12.841)
Computers per student	53.024** (21.428)	-1.693 (20.601)	32.181 (33.214)	51.305** (25.698)	22.036 (83.175)	-34.462 (57.079)	-39.952 (121.467)	16.791 (96.765)
Teacher incentives (1 = yes)	34.536*** (6.835)	18.272*** (5.871)	-0.457 (4.435)	2.131 (3.647)	17.136 (23.467)	12.704 (15.443)	-2.528 (13.117)	3.924 (10.142)
Teacher hiring (base: difficult) Teacher hiring (easy)	20.454*** (7.709)	-2.971 (8.605)	12.177 (9.768)	8.245 (8.301)	25.670 (30.131)	-4.722 (24.042)	22.466 (32.794)	11.056 (25.009)
No teacher hiring	36.255*** (7.487)	4.923 (7.609)	4.912 (9.578)	12.483 (8.578)	58.962* (30.534)	7.193 (19.874)	18.200 (33.276)	17.269 (25.754)
Principal attendance in leadership activities (base: low) Principal attendance in leadership activities (high)	7.347 (5.251)	13.423*** (4.674)	-10.919** (4.974)	-2.024 (3.877)	10.650 (17.872)	10.901 (13.007)	-4.422 (15.453)	0.450 (10.981)
Parental involvement in school (base: low) Parental involvement in school (high)	-2.115 (9.382)	43.517*** (7.590)	61.490*** (7.126)	55.762*** (5.769)	-11.526 (31.785)	42.251** (20.595)	52.893** (21.330)	53.314*** (17.248)
Parental involvement in school (medium)	-5.229 (7.950)	18.879*** (6.012)	48.359*** (5.823)	46.954*** (4.874)	4.880 (25.276)	24.440 (16.411)	40.997** (17.178)	44.695*** (13.563)

Continued on next page

Table A3.7: – continued from previous page

Dependent variable: Test scores (i.e. average of 5 plausible values)	OLS				IV 2SLS including clustering and sampling weight			
	Urban		Rural		Urban		Rural	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)	Boys (5)	Girls (6)	Boys (7)	Girls (8)
Constant	389.293*** (17.353)	387.582*** (15.511)	377.875*** (16.299)	362.549*** (12.526)	377.745*** (42.112)	399.412*** (33.244)	380.601*** (40.905)	368.177*** (30.307)
<i>First-stage estimates</i>								
Dependent variable: Class size								
Eight-grade size					0.020*** (0.007)	0.021*** (0.006)	0.024** (0.006)	0.026*** (0.005)
Number of observations	1114	1095	1701	1823	1114	1095	1701	1823
Adjusted R-squared	0.466	0.424	0.384	0.420	0.361	0.389	0.345	0.409
χ^2 statistic (p-value) of endogeneity test					0.008	0.200	0.163	0.336
First-stage F-statistics					8.313***	11.681***	20.616***	32.253***

Notes: Standard errors in parentheses. Standard error clustered at the school level. *Significant at the 10% level. **Significant at the 5% level. *** Significant at the 1% level. The plausible values refer to the range of competence that a student might have, given the student's item responses in an assessment. Variables missing, students' family background and school characteristics are controlled but not reported for brevity.

Table A3.8: Heterogeneity in class size effects

Dependent variables: Test scores (i.e. average of 5 plausible values)	Mathematics		Science	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)
Class size	2.587** (1.024)	3.171*** (0.960)	3.806*** (1.166)	2.550** (1.092)
Family background				
Mother's education (base: no formal or primary)				
Secondary	3.770 (5.685)	8.685** (4.129)	1.244 (6.901)	1.856 (4.634)
Tertiary	12.810* (6.900)	19.063*** (5.297)	17.576** (8.036)	24.513*** (5.768)
Child's does not know	-1.061 (5.977)	0.813 (5.131)	-7.756 (7.262)	-11.434** (5.596)
Father's education (base: no formal or primary)				
Secondary	2.368 (6.835)	-1.071 (5.191)	9.626 (7.834)	11.038* (6.074)
Tertiary	10.886 (7.573)	3.494 (6.157)	31.325*** (8.586)	25.527*** (7.002)
Child's does not know	-0.142 (6.930)	-7.147 (5.787)	10.092 (7.997)	6.406 (6.673)
<i>Home possessions</i>				
One and more bookcases	27.997*** (3.394)	31.227*** (2.791)	27.585*** (3.671)	30.402*** (2.989)
Reading resources (low)	-27.163*** (3.136)	-30.991*** (3.497)	-46.921*** (3.488)	-56.067*** (3.793)
Technology resources (low)	-41.336*** (3.309)	-31.579*** (2.988)	-14.758*** (3.588)	-7.098** (3.257)
Study room (low)	7.937*** (2.842)	8.784*** (2.593)	7.510** (3.077)	6.341** (2.739)
Rural	-4.775 (3.207)	-3.878 (2.957)	-2.868 (3.495)	0.285 (3.078)
<i>Interaction terms</i>				
Class size × Mother's education (base: no formal or primary)				
Secondary	-0.540 (0.815)	-0.846 (0.688)	-1.963** (0.984)	-1.510* (0.774)
Tertiary	-0.904 (1.028)	-1.276 (0.875)	-3.094** (1.210)	-2.009** (0.978)
Child's does not know	-0.513 (0.881)	-1.288 (0.790)	-2.049* (1.052)	-1.174 (0.885)
Father's education (base: no formal or primary)				
Secondary	-0.004 (0.847)	-0.789 (0.782)	0.796 (0.945)	0.602 (0.906)
Tertiary	-1.463 (0.967)	-2.431** (0.979)	-0.050 (1.103)	-0.855 (1.099)
Child's does not know	0.053 (0.900)	0.325 (0.836)	1.017 (0.996)	1.165 (0.979)
<i>Home possessions</i>				
One and more bookcases	-0.048 (0.538)	-0.604 (0.471)	-0.879 (0.581)	-0.732 (0.494)
Reading resources (low)	1.399*** (0.459)	1.523*** (0.528)	2.291*** (0.495)	1.772*** (0.566)
Technology resources (low)	-0.508 (0.461)	-0.395 (0.481)	-0.876* (0.502)	0.009 (0.531)
Study room (low)	0.050 (0.424)	-0.183 (0.430)	0.301 (0.455)	-0.348 (0.447)

Continued on next page

Table A3.8: – continued from previous page

Dependent variables: Test scores (i.e. average of 5 plausible values)	Mathematics		Science	
	Boys (1)	Girls (2)	Boys (3)	Girls (4)
Rural	1.891*** (0.450)	0.930** (0.468)	1.930*** (0.490)	0.713 (0.501)
Other factors				
Instructional materials (base: a lot)				
Instructional materials (a little)	–29.282*** (3.781)	–18.763*** (3.377)	–34.896*** (4.130)	–23.446*** (3.506)
Infrastructure (base: a lot)				
Infrastructure (some)	–15.049*** (4.078)	–13.487*** (3.939)	–13.871*** (4.332)	–13.617*** (3.963)
Infrastructure (little)	–16.507*** (5.090)	–20.910*** (4.493)	–16.581*** (5.685)	–24.161*** (4.644)
IT staff (base: a lot)				
IT staff (a little)	22.497*** (3.551)	19.955*** (3.403)	26.585*** (4.025)	27.295*** (3.611)
Computers per student	7.887 (17.447)	3.814 (15.098)	45.343*** (17.478)	40.445*** (15.619)
Teacher incentives (1 = yes)	1.509 (3.059)	1.656 (2.781)	6.096* (3.461)	3.645 (2.977)
Teacher hiring (base: difficult)				
Teacher hiring (easy)	15.900*** (5.263)	–3.277 (4.604)	19.151*** (6.025)	0.342 (5.443)
No teacher hiring	19.803*** (5.219)	2.586 (4.493)	24.060*** (5.983)	7.025 (5.293)
Principal attendance in leadership activities (base: low)				
Principal attendance in leadership activities (high)	–2.882 (3.208)	1.159 (2.695)	–2.817 (3.446)	1.827 (2.752)
Parental involvement in school (base: low)				
Parental involvement in school (high)	41.901*** (5.026)	44.643*** (4.231)	41.369*** (5.529)	52.261*** (4.541)
Parental involvement in school (medium)	29.104*** (3.996)	29.533*** (3.651)	29.167*** (4.540)	37.359*** (3.868)
School characteristics				
Students come from disadvantaged homes (> 50%)	–15.424*** (3.162)	–15.897*** (2.886)	–13.330*** (3.479)	–15.504*** (3.000)
Instructional time (6 and more hours per day)	14.708*** (3.490)	8.901*** (2.989)	20.332*** (3.852)	16.767*** (3.235)
Constant	422.827*** (10.037)	436.905*** (7.985)	379.369*** (11.464)	378.960*** (9.318)
Number of observations	2815	2918	2815	2918
Adjusted R-squared	0.403	0.384	0.406	0.426

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level.

*** Significant at the 1% level. The plausible values refer to the range of competence that a student might have given the student's item responses in an assessment. Variables missing are omitted for brevity.

Chapter 4

Teacher Absenteeism and Student Attainment in Malaysia

4.1 Introduction

This study aims to investigate the effect of students' behaviour with respect to their absence from school on teachers' absenteeism. In 2016, the United Nations Children's Fund (UNICEF) reported that approximately 80% of students in Malaysia suffered from some form of bullying that might significantly affect their willingness to come to school. Research in this area is rare. Previous studies tend to focus on examining the effect of student absenteeism has on attainment. See, for example, [Gershenson et al. \(2017\)](#), [Durden and Ellis \(1995\)](#), [Marburger \(2001\)](#), [Gottfried \(2010\)](#), [Miller et al. \(2008\)](#), [Massingham and Herrington \(2006\)](#), [Yahaya et al. \(2010\)](#), and [Aucejo and Romano \(2016\)](#) among others.

Malaysia has a few reported problems of teacher absenteeism (for instance, [Vanar \(2018\)](#)). A survey by the School Inspectorate and Quality Assurance reported that the number of teachers who were found to be absent in three months from at least 11 schools increased by 34% to 157 in May 2010 ("[Hundreds of Teachers](#)", [2011](#)). Although this problem might be isolated, if schools fail to act immediately, it is likely that the disciplinary problem in the schools will worsen and eventually reduce student attainment in the schools and encourage absenteeism in other schools. According to the chairman of Parent Action Group for Education Malaysia [Rahim \(2018\)](#), the teacher absenteeism problem is prevalent in schools with low

community monitoring and parental involvement. However, a report from the Malaysian education policy, the Malaysian Education Blueprint 2013-2025, shows that implementation of one of its initiatives called the District Transformation Program (DTP)¹ has a positive impact on teacher attendance rate and perhaps on student attainment in school. Thus, this study examines to what extent the DTP program has affected teacher absenteeism in Malaysian schools. Examining the determinants of teacher absenteeism is important, as the problem may disrupt a classroom's learning process and consequently, lowering student attainment (Glewwe et al., 2012; Tavares, 2015; Herrmann and Rockoff, 2012; Cucchiella et al., 2014).

This study uses data from the 2011 Trends in International Mathematics and Science Study (TIMSS) to address this issue. This rich dataset contains information on the frequency of teacher and student absenteeism in 180 schools from all 13 states in Malaysia. The dataset also contains comprehensive measures of initiatives that have been implemented within the DTP program such as teacher mentoring, monitoring, and the school's level of resources and development activities, which makes it possible for this study to examine the extent of how this program affects teacher absenteeism in Malaysia.

This study expects that there might be a reverse causality in the estimates of the probit teacher absenteeism model. This is because not only might student absenteeism affect teacher absenteeism, teacher absenteeism also might affect student absenteeism, even though the case may be not serious (Ahmad and Salleh, 1997, p. 66). Thus, to address the causality problem, we use an instrumental variable (IV) approach to estimate the teacher absenteeism model. We instrument student absenteeism using schools' levels of eighth-grade student disciplinary problems and if schools have a majority of students from low socioeconomic status, which we argue directly affects eighth-grade student absenteeism and the teacher absenteeism at school level indirectly. In general, it is possible that

¹More details are presented in the next section.

teacher absenteeism does affect student absenteeism but [Clotfelter et al. \(2006\)](#) and [Román and Valenzuela \(2015\)](#) show that problem in teacher-student matching may result in bias estimates. This explanation further supports that the instruments, which also refer to the eighth-grade students' behaviour, in general, is acceptable and directly related to the student rather than teacher absenteeism. We end this chapter with an estimation of the impact of teacher absenteeism on student attainment using the simple ordinary least squares (OLS) method.

We find evidence of a positive relationship between student absenteeism and teacher absenteeism. This effect of student absenteeism doubles in size after endogeneity issue have been addressed. This positive relationship between student absenteeism and teacher absenteeism is consistent with those of [Banerjee et al. \(2012\)](#) and [Irvine's \(1986\)](#) teacher-student interaction findings. Teaching mentoring is another important factor in reducing teacher absenteeism in schools. Other school factors such as a high level of teacher evaluation, a high level of parental involvement in school, and small school size have also shown a significant impact on lowering teacher absenteeism. Finally, consistent with a report in the "[Tackling the dropout](#)" (2006) by *The Star Online Malaysia*, this chapter finds that schools that suffer from teacher absenteeism have lower student attainment.

This chapter is structured as follows. Section 4.2 provides an overview of the teacher attendance and leave systems in Malaysia and the District Transformation Program. Section 4.3 outlines the literature on absenteeism in teaching and other industries. Section 4.4 outlines the empirical strategy. Section 4.5 describes the TIMSS dataset and reports some descriptive statistics. Section 4.6 discusses the results and concluding remarks are presented in Section 4.7.

4.2 Background

4.2.1 Teacher attendance and leave systems in Malaysia

The system to monitor teacher's attendance in Malaysian government schools was set by the Ministry of Education (MOE). The system has gone through some changes, from a manual logbook system to a punch card system and most recently, a biometric system. A biometric system is a system that records the teachers' attendance using their thumbprint. To improve teacher quality in school, the teacher attendance system also includes an electronic system called e-hadir, which ensures that school teacher attendance is updated to the MOE daily.

Aside from school holidays teachers in Malaysia are entitled to ten days of special leave (*Cuti Rehat Khas (CRK)*) per year. This includes leave for emergency matters, religious ceremonies, and other events with the approval of the head of the department. These days cannot be carried forward to the next year, although any unused leave will be paid upon retirement (maximum of 150 days). The number of accumulated leave days does not increase with the amount of work experience, which may be one of the contributing factors behind problem of teacher absenteeism in Malaysia.

4.2.2 The District Transformation Program

The District Transformation Program (DTP) is an educational equity programs in Malaysia. It was introduced in 2013 under the Malaysian Education Blueprint 2013-2020². Its main objective is to reduce gaps in student achievement in terms of gender, socioeconomic status, and between urban and rural schools in Malaysia.

The program was structured over two waves. The first wave was a pilot survey implemented in two states in Malaysia that have many gaps in student achievement, while the second wave is a national survey. The [Ministry of Education Malaysia](#)

²See Chapter 3 Section 3.2 for information on this policy.

(2013b) claims that the program has shown that teacher attendance is an important factor in improving student achievement.

The main initiatives that were introduced as part of the DTP include mentoring, monitoring through performance dialogue, and school autonomy in term of resources and development activities. Two categories of mentoring were introduced in the program: i) the SISC+ (School Improvement Specialist Coach) and ii) the SIPartner+ (School Improvement Partner). The SISC+ is a process of mentoring from selected senior or excellent teachers to new teachers in particular. While the SIPartner+ is a process of mentoring from selected senior or excellent principals to new school principals in particular. The main role of the selected mentor teachers and principals is to assist the new teachers and principals in the teaching and leadership process. For teacher monitoring, teachers' performances are monitored regularly through a new approach called performance dialogue via a video conference between the MOE and selected schools. The approach was chosen to save time and costs of monitoring, especially for rural schools with low infrastructure facilities. Lastly, for school autonomy, the District Education Office (*PPDs*) in each district is given the autonomy to allocate resources and conduct development activities in schools in their respective districts.

4.3 Selected literature on absenteeism: Evidence from teaching and other industries

Several countries have experimented with programs that aim to improve teacher attendance such as incentive programs, sick leave schemes, and performance pay. A study in one of the New York, US districts found that implementation of an attendance incentive plan increases the number of teachers with perfect attendance from 8 percent to 34 percent (Jacobson, 1989). Dee and Wyckoff (2015) report a different program of teacher evaluation called IMPACT in the District of Columbia Public Schools, US. They found that financial incentives have

boosted the performance of high performing teachers more than low performing teachers. Similarly, the rate of teacher absenteeism decreases in the presence of a monitoring project that uses cameras in the rural Udaipur district, India ([Banerjee and Duflo, 2006](#)). [Banerjee and Duflo's \(2006\)](#) study suggest that rationalising job descriptions may be a crucial step to improve the teacher attendance rate.

There are mixed effects of school incentives whether monetary or non-monetary on teacher absenteeism. The non-monetary incentives can range from providing teacher training opportunities, improving school resources, and increasing teachers leave days. Review papers by [Galiani and Perez-Truglia \(2013\)](#) and [Rogers and Vegas \(2009\)](#) find that teacher incentive schemes have a positive impact in reducing absenteeism and improving student achievement in developing countries, especially in India, Indonesia, Uganda, Ecuador and Zambia. In India, a combination of financial incentives and monitoring has significantly reduced the teacher absence rate ([Duflo et al., 2012](#)). While, the combination of the leave policies and monetary incentives has had a positive effect in reducing the rate of teacher absenteeism in North Carolina ([Clotfelter et al., 2009](#); [Scott et al., 2007](#)) and New York State ([Ehrenberg et al., 1991](#)).

On the other hand, some studies show that teacher incentives did not significantly reduce the rate of teacher absence; for example, teacher absences in India ([Kremer et al., 2005](#)), Dutch Secondary Schools ([Waterreus, 2003](#)), and in Israel and Kenya ([Glewwe and Kremer, 2006](#)). [Glewwe and Kremer \(2006\)](#) argued that one potential explanation for the negative effects in Israel and Kenya come from an inefficient allocation of funds in schools in those regions. According to [Waterreus \(2003\)](#), the promotion of teachers is not associated with lower teacher absenteeism due to the potential selection effect where teachers who were in the highest salary scales are selected as the sample of the study. Incentives also show no or little effect on teacher absenteeism in California ([Goldschmidt and Phelps, 2010](#)) or in the elementary and middle schools of Florida ([Harris and Sass, 2011](#)). [Goldschmidt and Phelps \(2010\)](#) and [Guerrero et al. \(2013\)](#) argue that the effect of incentives in

terms of development training only lasted for a short period of time, and further studies should examine the long-term effects of the program.

We can conceptually divide school interventions into direct (such as teacher monitoring) and indirect (such as involving the community in the schools) interventions. Many previous studies show that in the presence of close monitoring of teachers from the principals or communities, the potential of high teacher absenteeism can be reduced (Duflo et al., 2012; Rogers and Vegas, 2009; Suryahadi and Sambodho, 2013; Guerrero et al., 2013). Indirect interventions in terms of teacher evaluation and involving parents in schools also have a positive effect in reducing teacher absenteeism (Jacob and Walsh, 2011; Rogers and Vegas, 2009; Guerrero et al., 2013; Kremer et al., 2005). Kremer et al. (2005) note that parental involvement in schools may become an alternative form of teacher monitoring. Apart from interventions through monitoring, an early intervention such as the mentoring of teachers is also a crucial school policy. Kemmis et al. (2014) suggest that the practices of mentoring new teachers in a school can act as supervision, support, and collaborative self-development.

Mentoring also has a significant impact on workers' absenteeism in other industries. For instance, Brown et al. (1999) found that employee-sharing practices may reduce absenteeism in several industries in French firms. Additionally, previous research has identified the important effects of incentives on worker absenteeism in other industries. Financial incentives have significantly reduced worker absenteeism in the Dutch manufacturing industry (Hassink and Koning, 2009) and the Indian health industry (Chaudhury et al., 2006). Chaudhury et al. (2006) show that high salaries combined with effective monitoring may boost workers' attendance in the health sector. Similarly, Bryson and Dale-Olsen (2017) argue that the high level of absenteeism in Norway is related to the high state sick pay in that region. In addition, different types of incentives such as the income replacement program for short term sickness and income taxes may affect the worker absenteeism problem in Swedish companies (Johansoon and Palme, 2002). Other types of incentives

such as smaller firm sizes are also found to significantly reduce German workers' absenteeism rate (Winkelmann, 1999; Barmby and Stephen, 2000).

In conclusion, many studies have examined the effectiveness of worker attendance programmes, and this study attempts to fill the literature on worker absenteeism in the context of the teaching sector in Malaysia. This study also aims to focus on initiatives that have been implemented in the DTP, including other school policies to extensively examine the program's effects on teacher absenteeism. The effect teacher absenteeism has on student attainment will be further investigated, given that some studies have shown adverse effects of teacher absenteeism on student attainment (Clotfelter et al., 2009; Duflo et al., 2012; Miller et al., 2008). The empirical strategy used to tackle these objectives is explained in the next section.

4.4 Empirical strategy

In this section, the strategy used to examine the determinants of teacher absenteeism is explained. A model of student attainment, examining the impact of teacher absenteeism on student attainment is also presented.

First, this study estimates the teacher absenteeism model for each observation school s of the following probit form:

$$Pr(Y = 1|X_j) = F(X'_j\alpha_j) \quad (4.1)$$

where $Pr(Y = 1|X_j)$ is the probability that school s has a high level of teacher absenteeism given X_j . X_j for $j = 1, \dots, p$ are the vector of factors influencing absenteeism such as student absenteeism and initiatives that have been implemented in the District Transformation Program (DTP). α_j is the coefficient estimates from the probit regression that maximise the log likelihood of the cumulative distribution function of the standard normal distribution, $F(X'_j\alpha_j)$ ³. We present

³ $L(X'_j\alpha) = \prod_{k=1}^n (F(X'_j\alpha)^{y_j} [1 - F(X'_j\alpha)]^{(1-y_j)})$, where X' denotes a $1 \times K$ row vector of regressor values

the results in terms of marginal effects. The marginal effects will explain the change in the probability that schools have a high teacher absenteeism problem in response to a one unit change in the explanatory variable of equation (4.1), holding other explanatory variables at their means.

This study considers two specifications of the marginal effects. First, the marginal effects are evaluated at the mean values of X_j , holding other explanatory variables at their respective means. Second, the marginal effects evaluated at certain values of X_j , holding other explanatory variables at their respective means. For example, this study would like to examine whether the probability of schools reporting teacher absenteeism will increase if the school size is increased by 500. This will place our findings into context, and provides some insights as to whether the intervention would be beneficial in terms of the effect it is likely to have on absenteeism.

Second, this study employs an identification strategy to address the potential endogeneity problem in equation (4.1). This is because, not only might student absenteeism affect teacher absenteeism, teacher absenteeism also might affect student absenteeism in schools, although these are infrequent cases ([Ahmad and Salleh, 1997](#), p. 66). Thus, to solve the problem by following [Angrist and Pischke's \(2009\)](#) suggestion, we introduce instruments in equation (4.1) and estimate the instrumental variable (IV) model using the two-stage least squares (2SLS) estimator with a second stage regression as follows:

$$Y = \alpha_0 + \alpha_1 \hat{X}_1 + \alpha_2 X_2 + \alpha_3 X_3 + \dots + \alpha_p X_p + \varepsilon \quad (4.2)$$

where X_1 is an endogenous variable, student absenteeism, and $\hat{X}_1$ is the predicted result from the following first-stage regression:

for observation j , and α are a $K \times 1$ column vector of regression coefficients.

$$\hat{X}_1 = \beta + \lambda_k Z_k + \beta_2 X_2 + \beta_3 X_3 + \beta_p X_p + \varepsilon \quad (4.3)$$

where Z_k for $k = 1, \dots, q$ are the instruments for the student absenteeism, dummy variables indicating whether the school has a high level of bullying, classroom disturbance, and students that come from disadvantaged homes. The key assumption for this method is that the instruments, Z , which refer to eighth-grade student attribute-level data, are correlated with the causal variable of interest, student absenteeism X_1 , but uncorrelated with any other explanatory variables at school-level data in equation (4.3). The term uncorrelated with other explanatory variables is like saying $Cov(\varepsilon, Z_k) = 0$, which is also the exclusion restriction of the IV model ([Angrist and Pischke, 2009](#)).

Statistically, to check the validity of the IV estimation, we look at the significance of the Wu-Hausman ([Wu, 1974](#); [Hausman, 1978](#)) F-statistics, the F-statistic of the first-stage regression of equation (4.3), and the [Sargan's \(1958\)](#) χ^2 statistics. The Wu-Hausman F-statistic shows the validity of the causal variable of interest, whether it is endogenous or not, and hence ensures whether the baseline model of equation (4.1) yields biased estimates or not. Both the F-statistic of the first-stage regression and the [Sargan's \(1958\)](#) χ^2 statistic indicate the appropriateness of the instruments.

We expect that the instruments will influence the students' willingness to attend school. UNICEF has reported that approximately 80% of students in Malaysia have suffered from some form of bullying, and such incidents tend to happen in the classroom ([UNICEF Malaysia Communication, 2009](#)). In addition, [Yahaya et al. \(2009\)](#) and [Ahmad and Salleh \(1997\)](#) show that disciplinary problems are mainly attributable to the influence of the students' peer group. These explanations reflect the fact that bullying could be the main reasons why Malaysian students tend to play truant. Other than the disciplinary problems, [Yahaya et al. \(2010\)](#) and [Ahmad and Salleh \(1997\)](#) show that low socioeconomic status also might

affect truancy, with students from a low socioeconomic background being more likely to be truant. Eventually, a low attendance rate of students in school leads to a lack of teachers' motivation to attend school.

Next, a model of student attainment to identify to what extent a high level of teacher absenteeism has an impact on student attainment in schools is estimated. We use the standard ordinary least squares (OLS) method with the following model:

$$T_{i,s} = \alpha + S_{i,s}\beta + D_{i,s}\delta + \mu_{i,s} \quad (4.4)$$

where $T_{i,s}$ is the measure of student attainment which includes the student i 's test scores (i.e. average plausible values⁴) in mathematics and science in school s . S is the vector of teacher absenteeism. D is the vector of control variables of students' background and school characteristics as discussed in detail in Chapter 3, and μ is an error term. We are not controlling student absenteeism variable as we expect that both teacher and student absenteeism are inter-related. In addition, this study only focuses on handling the teacher absenteeism determinants and its effect on student outcome.

4.5 Data and descriptive statistics

This study uses data from the Trends in International Mathematics and Science Study (TIMSS), 2011, to estimate the determinants of teacher absenteeism. The dataset is an international assessment survey conducted by an independent educational institution, the International Association for the Evaluation of Educational Achievement (IEA). The dataset consists of a random sample of 180 schools and 5,733 eighth-grade students in all states in Malaysia. The dataset is unique as it consists of data on various types of school disciplinary problems such as

⁴The plausible values refer to the range of competence that a student might have given the student's item responses in an assessment (Wu, 2005).

absenteeism in school. It includes information answered by school principals about the level of teacher absenteeism, which can be: i) not a problem, ii) minor problem, iii) moderate problem, or iv) serious problem, which becomes the key variable in this study. After examining the distribution of the teacher absenteeism data, we then categorised it into a dummy variable of teacher absenteeism which equals one if a school has high teacher absenteeism (consisting of schools with a minor, moderate, or serious problem of absenteeism) and zero otherwise.

The TIMSS dataset also includes measures of student absenteeism and initiatives conducted within the District Transformation Program (DTP) such as teacher mentoring, teacher monitoring, school resources, and school development activities. To note, both teacher and student absenteeism refer to the school rather than the class level. Information on other school factors such as teacher evaluations, parental involvement in school, and teacher incentives is also available in the dataset. Finally, the dataset provides some information on school characteristics such as location, the number of communities surrounding the school, and school size.

All of the variables are acquired from questions answered by the principal of the schools, and they have mainly been categorised into dummy variables, except teacher mentoring and school resources, after considering their distribution. In detail, student absenteeism refers to either schools that have a high (i.e. moderate and serious problem) or low (i.e. minor or not a problem) level of student absenteeism. The mentoring variable is measured from a question answered by a representative teacher in each school about the teacher's frequency in sharing experiences with other teachers, categorised into three groups: i) mentoring (very often), if teachers have almost daily sharing sessions, ii) mentoring (often), if teachers have weekly sharing sessions, and iii) mentoring (seldom), if teachers have monthly or almost never have sharing sessions. For teacher monitoring, this study uses a proxy of whether school principals have a lot of time or no time in monitoring teachers' implementation of the school's educational goals in

their teaching. For school resources, this study uses a proxy of schools' level of availability of instructional materials (e.g. textbooks), recorded at three levels: high, medium, and low. For the level of development activities, the survey refers to whether school principals participate a lot or some time in the professional development activities in the school.

Table 4.1: Descriptive statistics

	Teacher absenteeism (high = 1)	Teacher absenteeism (high = 0)	Test of difference
	Mean (1)	Mean (2)	(3)
Student absenteeism (high)	0.395	0.214	-2.572*
District Transformation Program			
Mentoring (base: seldom)			
Mentoring (very often)	0.132	0.133	0.021
Mentoring (often)	0.184	0.316	2.031*
High teacher monitoring	0.697	0.776	1.151
Resources (base: low)			
Resources (high)	0.368	0.459	1.206
Resources (medium)	0.289	0.214	-1.124
High level of development activities	0.421	0.398	-0.305
Other factors			
Teacher evaluation (yes)	0.671	0.929	4.276***
High level of parental involvement in schooling	0.447	0.704	3.479***
Incentives (yes)	0.289	0.276	-0.202
School characteristics			
School size (in thousands)	1.519	1.303	-2.179*
Rural	0.434	0.459	0.327
Community surrounding school (base: 15,000 and less people)			
Community surrounding school (>100,000 people)	0.158	0.204	0.787
Community surrounding school (15,001-100,000 people)	0.487	0.480	-0.094
Number of observations	76	98	

Notes: *Significant at the 10% level. **Significant at the 5% level. *** Significant at the 1% level. The test of difference refers to the t-statistics of the test of difference in means between schools with a high teacher absenteeism problem and schools with a low teacher absenteeism problem.

Table 4.1 reports summary statistics for the key variables. This study separately examines the mean of each variable for schools that have high teacher absenteeism and low teacher absenteeism, respectively. The tested difference between two means of each variable is presented in column (3). Overall, this study found significant mean differences in schools that have high student absenteeism, implemented mentoring often, high teacher evaluation, and high parental involvement in the school.

4.6 Empirical results

Next we present an overview of our results. The estimated results of the teacher absenteeism model using the probit method are first presented. It is followed by a presentation of the teacher absenteeism model using the IV method. Lastly, the impact of high teacher absenteeism on student attainment shown in this study is presented.

Table 4.2 reports the marginal effects of the probit regression of equation (4.1). The results are presented in three models: 1) the District Transformation Program model, 2) the model that excludes the insignificant factors of Model 1, and 3) model 1 including other school factors, with each of the models estimated with and without control variables. We also include a student absenteeism factor in each of the models, as we are interested in identifying its impact on teacher absenteeism in the presence of the school factors of interest.

Table 4.2: Marginal effects of the teacher absenteeism model

Dependent variable:	Model 1		Model 2		Model 3	
Teacher absenteeism (high)	(1)	(2)	(3)	(4)	(5)	(6)
Student absenteeism (high)	0.234*** (0.088)	0.272*** (0.090)	0.187** (0.091)	0.207** (0.093)	0.202** (0.096)	0.238** (0.098)
District Transformation Program						
Mentoring (base: seldom)						
Mentoring (very often)	-0.057 (0.119)	-0.062 (0.121)	-0.045 (0.119)	-0.044 (0.125)	-0.027 (0.120)	-0.032 (0.128)
Mentoring (often)	-0.201** (0.093)	-0.191** (0.096)	-0.248*** (0.094)	-0.245** (0.096)	-0.248** (0.098)	-0.244** (0.100)
High teacher monitoring	-0.131* (0.092)	-0.115 (0.094)	-0.036 (0.097)	-0.022 (0.099)	-0.058 (0.099)	-0.044 (0.101)
Resources (base: low)						
Resources (high)	-0.054 (0.092)	-0.053 (0.093)			-0.164* (0.097)	-0.171* (0.101)
Resources (medium)	0.064 (0.102)	0.077 (0.102)			-0.012 (0.107)	0.011 (0.110)
High level of development activities	0.088 (0.084)	0.112* (0.084)			0.108 (0.089)	0.141* (0.092)
Other factors						
Teacher evaluation (yes)			-0.422*** (0.118)	-0.439*** (0.124)	-0.460*** (0.119)	-0.488*** (0.127)
High level of parental involvement in school			-0.232*** (0.085)	-0.248*** (0.088)	-0.258*** (0.087)	-0.289*** (0.090)
Incentives (yes)					0.012 (0.093)	0.034 (0.097)
School characteristics						
School size (in thousands)		0.194*** (0.063)		0.179** (0.071)		0.211*** (0.073)
Rural		-0.044 (0.093)		-0.019 (0.092)		-0.009 (0.093)
Community surrounding school (base: 15,000 and less people)						
Community surrounding school (>100,000 people)		-0.151 (0.141)		-0.226* (0.142)		-0.210* (0.147)
Community surrounding school (15,001-100,000 people)		-0.080 (0.098)		-0.066 (0.101)		-0.043 (0.101)
Number of observations	174	174	174	174	174	174

Notes: Standard errors in parentheses. *Significant at the 20% level. **Significant at the 5% level. ***Significant at the 1% level.

Table 4.2 shows that regardless of model student absenteeism has a statistically significant positive effect on teacher absenteeism. Specifically, it shows that an increase in student absenteeism increases the probability of teacher absenteeism by approximately 20 points. The positive effect of student absenteeism is similar to Banerjee et al.'s (2012) and Irvine's (1986) teacher-student interaction findings.

They show that in the presence of student absenteeism problems in school, teachers' preference for going to school decreases. The positive relationship between student and teacher absenteeism is also consistent with Laibson and List's (2015) argument on how an individual's behaviour is closely related to others' behaviour.

Table 4.3: Instrumental variable estimates of teacher absenteeism model

Dependent variable:	Model 1		Model 2		Model 3	
Teacher absenteeism (high)	(1)	(2)	(3)	(4)	(5)	(6)
Student absenteeism (high)	0.507*** (0.161)	0.497*** (0.154)	0.374** (0.157)	0.375** (0.154)	0.351** (0.153)	0.348** (0.148)
District Transformation Program						
Mentoring (base: seldom)						
Mentoring (very often)	-0.045 (0.114)	-0.057 (0.112)	-0.034 (0.105)	-0.038 (0.104)	-0.019 (0.103)	-0.029 (0.102)
Mentoring (often)	-0.193** (0.088)	-0.177** (0.085)	-0.196** (0.081)	-0.188** (0.080)	-0.196** (0.080)	-0.185** (0.079)
High teacher monitoring	-0.125 (0.086)	-0.106 (0.084)	-0.040 (0.080)	-0.020 (0.079)	-0.056 (0.080)	-0.039 (0.078)
Resources (base: low)						
Resources (high)	-0.041 (0.087)	-0.042 (0.085)			-0.118 (0.081)	-0.110 (0.079)
Resources (medium)	0.051 (0.101)	0.066 (0.098)			-0.009 (0.092)	0.007 (0.090)
High level of development activities	0.137* (0.083)	0.152* (0.081)			0.117 (0.075)	0.132* (0.074)
Other factors						
Teacher evaluation (yes)			-0.361*** (0.090)	-0.345*** (0.089)	-0.373*** (0.089)	-0.355*** (0.088)
High level of parental involvement in school			-0.162** (0.077)	-0.158** (0.076)	-0.187** (0.075)	-0.186** (0.074)
Incentives (yes)					-0.010 (0.076)	0.006 (0.075)
School characteristics						
School size (in thousands)		0.193*** (0.064)		0.144** (0.060)		0.149** (0.059)
Rural		-0.041 (0.086)		-0.029 (0.081)		-0.025 (0.079)
Community surrounding school (base: 15,000 and less people)						
Community surrounding school (>100,000 people)		-0.102 (0.126)		-0.147 (0.118)		-0.128 (0.115)
Community surrounding school (15,001-100,000 people)		-0.061 (0.090)		-0.047 (0.083)		-0.027 (0.082)
Constant	0.385*** (0.116)	0.159 (0.173)	0.802*** (0.122)	0.630*** (0.175)	0.850*** (0.141)	0.642*** (0.191)
<i>First-stage estimates</i>						
Dependent variable:						
Student absenteeism (high)						
Physical bullying (high)	0.143** (0.065)	0.166** (0.065)	0.126* (0.066)	0.143** (0.066)	0.139** (0.066)	0.160** (0.066)
Physical bullying missing	0.003 (0.407)	-0.047 (0.403)	-0.166 (0.418)	-0.238 (0.416)	0.020 (0.419)	-0.052 (0.414)
Classroom disturbance (high)	0.439*** (0.077)	0.455*** (0.076)	0.423*** (0.080)	0.431*** (0.080)	0.438*** (0.080)	0.449*** (0.079)
Students come from disadvantaged homes (>50%)	0.151** (0.062)	0.098 (0.066)	0.172*** (0.063)	0.126* (0.067)	0.149** (0.063)	0.098 (0.066)
Students come from disadvantaged homes missing	0.560** (0.239)	0.508** (0.245)	0.563** (0.244)	0.521** (0.251)	0.555** (0.242)	0.508** (0.248)

Continued on next page

Table 4.3 – continued from previous page

Dependent variable:	Model 1		Model 2		Model 3	
Teacher absenteeism (high)	(1)	(2)	(3)	(4)	(5)	(6)
Number of observations	174	174	174	174	174	174
Wu-Hausman F-statistics (p-value)	0.039	0.063	0.119	0.133	0.184	0.227
First-stage F statistics	11.990***	12.643***	10.282***	10.379***	10.687***	11.108***
Sargan χ^2 statistics (p-value)	0.190	0.234	0.139	0.140	0.226	0.238

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level. ***

Significant at the 1% level.

As already mentioned a potential limitation of our results is that student absenteeism is affected by teacher absenteeism leading to endogeneity. Table 4.3 shows the results of the IV 2SLS estimation; the first-stage estimation and the IV post-estimation tests are presented in the bottom section of the table. The post-estimation tests includes the Wu-Hausman statistics that shows either that student absenteeism is endogenous or not, the F-statistics of the first-stage estimation which test for weak instruments, and the Sargan's χ^2 statistics that represent the overidentifying test for the IV estimates using more than one instrument.

Statistically, only columns (1) and (2) of Table 4.3 shows that student absenteeism is endogenous since the endogeneity test using the Wu-Hausman statistic is significant. This finding means that there are other unobserved factors potentially sourced from the measurement errors that might affect teacher absenteeism through the student absenteeism effect. It suggests that the probit estimates in Table 4.2 are biased, and the use of IV is supported. Next, we check on the suitability of the instruments chosen for the IV estimation. As discussed in Section 4.4, we expect that school characteristics to be closely related to a high level of student absenteeism. Table 4.3 has shown the significant relationship between these two variables and proves the relevance of the instruments to be used in the IV estimation of Model 1 in particular, as the F-statistics of the first stage regression in the model are statistically significant. However, it should be emphasised that the effect of the student absenteeism using the IV estimation becomes larger. Some possible explanations for this are that the OLS coefficient is prone to measurement error or the instrument captures some additional effect. This is because since the

student absenteeism effect is instrumented by bullying and disruptive behaviour, it is possible that 'undesirable student behaviour' adds to teacher absenteeism.

Overall, in comparison with the estimated results in Table 4.2, we can see that the effect of student absenteeism in Table 4.3 on teacher absenteeism doubles, while the significant effect of other school factors on teacher absenteeism remains unchanged. These findings give attention to schools as well as to the government to tackle the student absenteeism problem as it may highly contribute to high teacher absenteeism in schools. The substantial effect of student absenteeism on teacher absenteeism when the causality problem has been addressed is in line with the report of serious disciplinary problems in Malaysian schools such as by [UNICEF Malaysia Communication \(2009\)](#), [Yahaya et al. \(2009\)](#), and [Ahmad and Salleh \(1997\)](#).

In accordance with the DTP program, Table 4.3 shows that schools that have implemented a frequent intervention in terms of mentoring (often) will have approximately 20 percent lower level of teacher absenteeism. The effects increased slightly after we specified the model by including only those DTP factors that significantly affect teacher absenteeism (Models 2) and controlling for other factors (Model 3). These findings imply that the intervention program in the DTP program such as mentoring (the SISC+, School Improvement Specialist Coach, and the SIPartner+, School Improvement Partner) appear to improve teachers' incentives to come to school.

However, the mentoring effect at the level of "very often" shows no significant effect on teacher absenteeism. Among potential explanations for this finding is that teachers need to invest more time in attending and preparing the materials for the mentoring sessions if it is conducted very often. These explanations are contradicted with an argument by [Ost and Schiman \(2017\)](#) on workers' marginal productivity based on workload, but it is in line with [Mckenzie et al.'s \(2014\)](#) findings on the positive relationship between teachers' workload and the rate of

teacher absenteeism.

Other initiatives in the DTP program such as monitoring also show little or no significant effects on teacher absenteeism (Table 4.3). It may be related to the method used in the monitoring session with videoconferencing that probably creates a vulnerability issue. The issue may apply to teachers in the schools of the rural area in particular as they get earlier information about the session and may need to prepare the materials such as finding a computer with better internet connection. As a consequence, teachers may change their attitudes due to their awareness of being observed, and the session may have a weak impact.

For teacher incentives, although the result differs from most of the literature, it is consistent with studies by [Glewwe and Kremer \(2006\)](#), [Kremer et al. \(2005\)](#), and [Waterreus \(2003\)](#) that indicate little or no effect of higher pay on teacher absenteeism. In this study, the weak result may be due to the measure used for the incentives that mainly refers to incentives for teacher recruitment. Since Malaysia has a limited quota in the recruitment of teachers⁵, the teacher incentives in regards to recruitment may not have an impact on the current teachers' intention to not attend school. They might regard that their opportunity cost for not going to school is low in the presence of such incentives. Other incentives such as a high level of development activities for teachers in school also shows no significant effect on teacher absenteeism (Table 4.3). The positive but little effect of a high level of development activities in school might be explained by [Goldschmidt and Phelps \(2010\)](#) and [Harris and Sass's \(2011\)](#) findings on the short lasting effect of teacher development training on teachers' performance in school. [Goldschmidt and Phelps \(2010\)](#) show that it is important to consider the duration of the teacher development program that has been conducted in order to examine the effect of the school factor. However, due to data limitations, this study is unable to control for it.

⁵For example, in 2016, from 94,646 applicants, only 14.2% of the applicants had been offered a teaching job ([Ministry of Education Malaysia, 2016](#)).

Models (2) and (3) of Table 4.3 show that additional intervention factors, parental involvement in school and teacher evaluation, in particular, have substantial positive effects on reducing the rate of teacher absenteeism in school. Specifically, we found that schools that have conducted external teacher evaluations and a high level of parental involvement in school, respectively, have reported approximately 40 percent and 20 percent lower high teacher absenteeism problem. These findings are in line with studies by [Jacob and Walsh \(2011\)](#), [Rogers and Vegas \(2009\)](#), and [Suryahadi and Sambodho's \(2013\)](#). In addition, [Othman \(2006\)](#) shows that a high parental involvement in school, for example, through their participation in the Parents and Teachers Association meetings, may help in reducing disciplinary problems in schools.

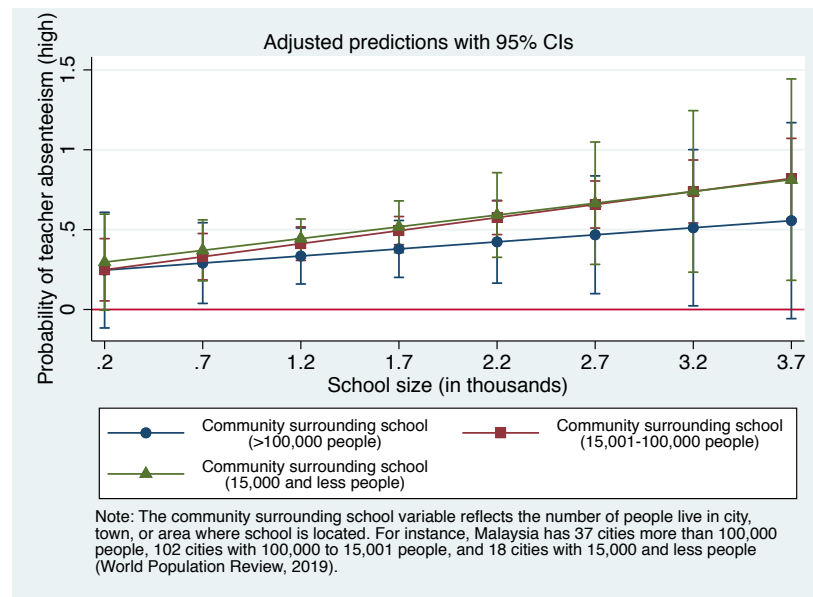


Figure 4.1: The effect of community surrounding school over school size on teacher absenteeism

[Rahim \(2018\)](#) and [Yahaya et al. \(2010\)](#) suggest that a low number of communities surrounding the school is associated with a higher probability of teacher absenteeism in Malaysia. Despite this, Table 4.3 shows no or little effect of the community on teacher absenteeism. We expect that the impact of the community surrounding the school differ across the school sizes. As expected, Figure 4.1 has shown that the effect of the low number of communities surrounding the school

on teacher absenteeism is significant for schools with a lower number of students (see Figure A4.1 in Appendix C for information on the distribution of the number of students in this study). Specifically, schools about twice less likely to have high teacher absenteeism if they have a high number of communities surrounding the school as school size increases.

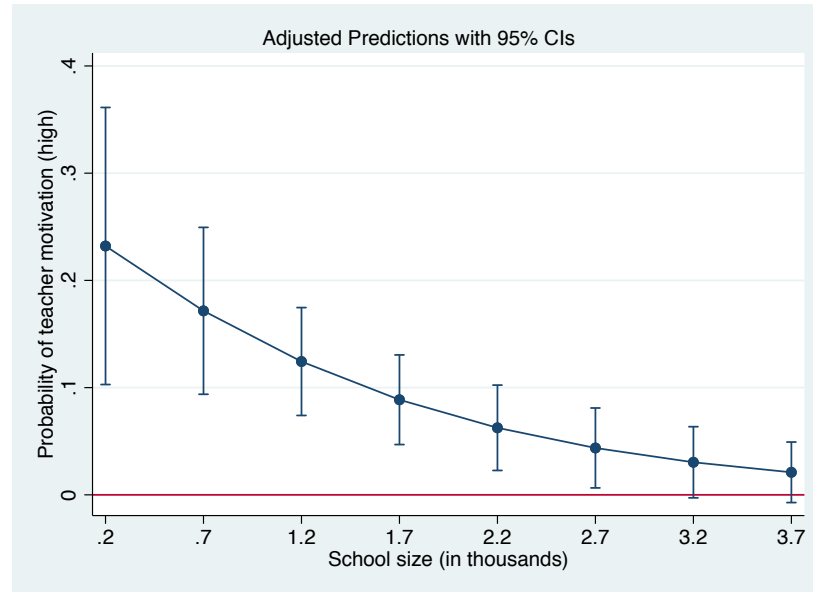


Figure 4.2: Probability of teacher motivation (high)

Lastly, school size has a significantly positive effect on teacher absenteeism. This is in line with [Lee and Loeb \(2000\)](#) and [Jones et al.'s \(2008\)](#) that found that a large school size negatively affects teachers' attitude in school. A possible explanation for this is that a large school size (i.e. schools with a large number of students in school) may have sufficient resources to provide temporary 'relief' teachers to provide cover for those who are absent. Besides, this study expects that teachers' motivation levels to come to work also decrease if there are a lot of students in the school or classroom. This is because apart from teaching, teachers need to evaluate the progress of each of their students to maximise the teacher-student outputs. Additionally, they have been burdened by too many non-teaching duties including clerical and paperwork ([Taharim et al., 2017](#); [Yesuiah, 2014](#); [Malakolunthu et al., 2010](#)). Figure 4.2 shows the estimated results of the teacher motivation model

as school size increases, which uses a dummy of teachers' motivation (high) as the dependent variable of equation (4.1). It shows that teachers' motivation diminishes (i.e. is closer to zero) as school size increases. Thus, the argument for the school size effect is supported.

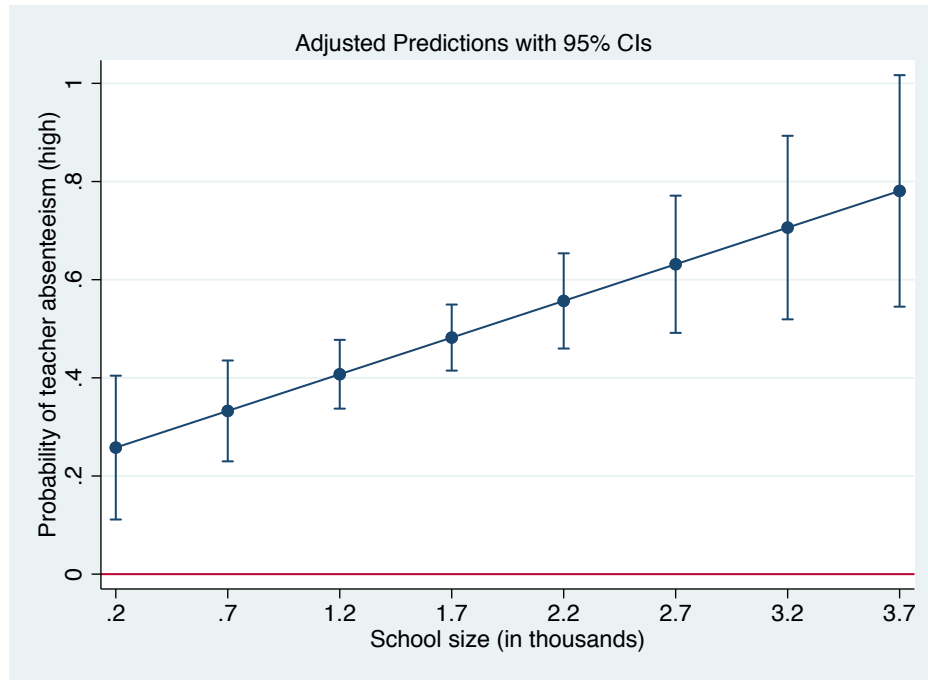


Figure 4.3: Percent teacher absenteeism by school size

We expect that the impact of school size on teacher absenteeism varies at certain points. Figure 4.3 shows the probability of teacher absenteeism when the school size increases from 200 to 3700. The figure shows that the probability of teacher absenteeism increases as the school size increases, with the largest school size increasing the probability of schools having a high level of teacher absenteeism to approximately 80 percent. This finding suggests that the government or policymakers should consider any policies aimed at school consolidation carefully if it wants to improve the rate of teacher attendance in Malaysia.

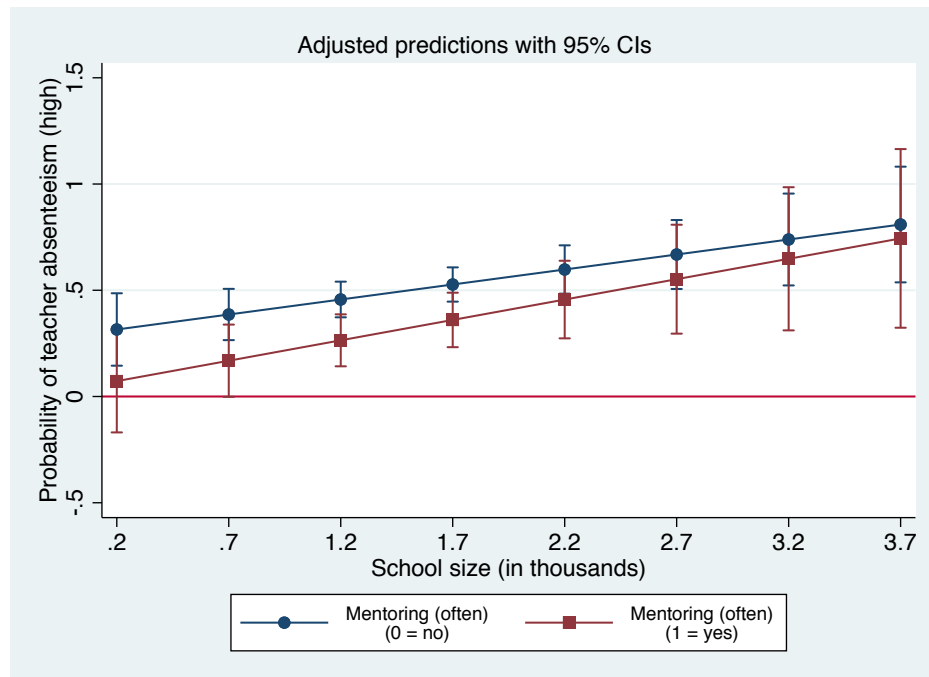


Figure 4.4: Percent teacher absenteeism by mentoring (often) and school size

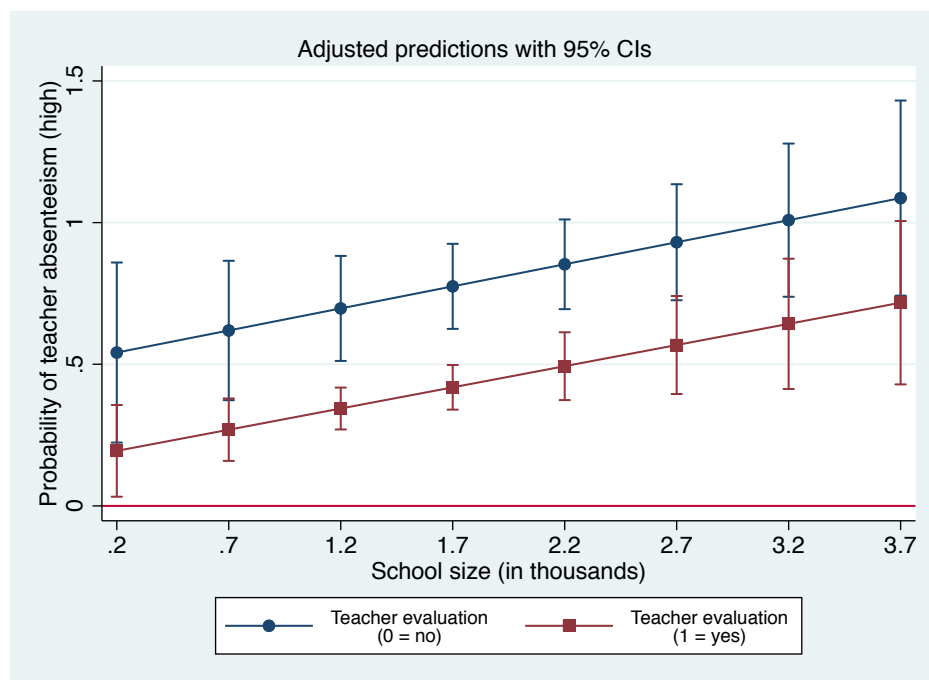


Figure 4.5: Percent teacher absenteeism by teacher evaluation and school size

Next, we examine whether school size influences the impact of the school factors of interest on teacher absenteeism. The conditional marginal effects of mentoring (often) and teacher evaluation on teacher absenteeism as school size increases are illustrated in Figures 4.4 and 4.5. The effect on teacher absenteeism if schools have a high level of parental involvement in school is almost similar to the teacher evaluations' effect and is presented in Appendix C (see Figure A4.2). Both figures show that the significant effect of teacher mentoring and teacher evaluation decreases as school size increases. In addition, Figure 4.4 presents that at the largest school size, the significant effect of teacher mentoring in reducing the rate of teacher absenteeism diminishes, while Figure 4.5 shows that the difference between the effect of schools with and without teacher evaluation remains constant as the school size becomes larger.

4.6.1 Teacher absenteeism effect on student attainment

Finally, we examine the effect teacher absenteeism has on student attainment. We separately examine the effect by school location to propose suitable policy responses due to differences regarding infrastructure facilities⁶, in particular, in the regions. The rural areas are referred to as the areas with a population of less than 10,000 based on indicators by the [Department of Statistics Malaysia \(2014\)](#).

Table 4.4: Effects of teacher absenteeism on student attainment

Dependent variables: Test scores (i.e. average of 5 plausible values)	Mathematics		Science	
	Urban	Rural	Urban	Rural
Teacher absenteeism (high)	-12.454*** (2.838)	-15.137*** (2.978)	-12.201*** (3.108)	-21.633*** (3.323)
School characteristics				
School size (in thousands)	11.956*** (2.048)	-11.149*** (2.845)	0.457 (2.321)	-18.918*** (3.211)
Community surrounding school (base: 15,000 and less people)				
Community surrounding school (>100,000 people)	4.918 (4.072)		0.121 (4.354)	
Community surrounding school (15,001-100,000 people)	9.906** (3.972)	18.642*** (3.224)	9.901** (4.106)	10.150*** (3.551)
Family background				
Mother's education (base: no formal or primary)				
Secondary	12.978** (5.977)	3.260 (4.700)	14.551** (7.298)	0.438 (5.678)
Tertiary	26.157*** (6.617)	20.756*** (7.102)	34.947*** (7.841)	31.175*** (8.005)
Child's does not know	12.865* (6.608)	-7.687 (5.400)	10.205 (7.853)	-13.650** (6.490)
Father's education (base: no formal or primary)				
Secondary	6.279 (6.986)	-2.146 (5.853)	22.848*** (8.246)	3.639 (6.913)
Tertiary	13.580* (7.429)	2.597 (7.418)	46.127*** (8.611)	15.188* (8.465)
Child's does not know	2.708 (7.382)	-12.597** (6.267)	21.421** (8.616)	-5.353 (7.459)
One and more bookcases	36.104*** (2.878)	30.907*** (3.751)	34.893*** (3.167)	32.811*** (4.033)
Reading resources (low)	-37.793*** (3.509)	-42.311*** (3.339)	-63.238*** (3.964)	-59.198*** (3.845)
Technology resources (low)	-41.258*** (2.836)	-39.481*** (4.369)	-21.074*** (3.154)	-10.598** (4.877)

Continued on next page

⁶For instance, as of 2005, the total length of roads in Sabah and Sarawak, which is Malaysian states with most regions categorised as rural areas, was three times lower than Peninsular Malaysia ([Naidu, 2008](#)).

Table 4.4 – continued from previous page

Dependent variables: Test scores (i.e. average of 5 plausible values)	Mathematics		Science	
	Urban	Rural	Urban	Rural
Study room (low)	9.362*** (2.760)	4.654 (2.987)	5.767* (2.988)	6.744** (3.316)
Constant	434.650*** (8.906)	461.621*** (7.949)	412.419*** (10.486)	443.167*** (8.978)
Number of observations	3287	2417	3287	2417
Average R-squared	0.254	0.240	0.276	0.247

Notes: Standard errors in parentheses. *Significant at the 10% level. **Significant at the 5% level.

*** Significant at the 1% level. The plausible values refer to the range of competence that a student might have given the student's item responses in an assessment. Missing observations are controlled but not reported for brevity. Rural areas are defined as the areas with a population of less than 10,000. The community surrounding school variable reflects the number of people live in city, town, or area where school is located. For instance, Malaysia has 37 cities more than 100,000 people, 102 cities with 100,000 to 15,001 people, and 18 cities with 15,000 and less people (World Population Review, 2019).

Table 4.4 shows that teacher absenteeism has a statistically significant negative effect on student scores in mathematics and science. Expectedly, the effect is pronounced for students' attainment in rural compared to urban areas. The finding is in line with an argument by [Ahmad and Salleh \(1997\)](#) and [Yahaya et al. \(2010\)](#), where schools with a low socioeconomic status potentially have higher absenteeism. The negative effect of teacher absenteeism on student attainment corresponds with many studies such as by [Clotfelter et al. \(2009\)](#), [Duflo et al. \(2012\)](#), [Miller et al. \(2008\)](#), [Blazar and Kraft \(2017\)](#), and [Ehrenberg et al. \(1991\)](#). This study assumes that if schools have a high teacher absenteeism problem, students' motivation to come to school will diminish, which will eventually contribute to students' low performance.

4.7 Conclusion

Examining the determinants of teacher absenteeism is important for school because it is often the case that schools where absenteeism is high have low levels of student attainment. This study finds evidence of a positive relationship between teacher and student, which becomes more pronounced when the issue of endogeneity is addressed.

This study also has determined that teacher mentoring one of the initiatives in the District Transformation Program (DTP), is a dominant contributor in reducing the teacher absenteeism. The result is consistent with studies by [Brown et al. \(1999\)](#) and [Kemmis et al. \(2014\)](#) among others who show the positive relationship between sharing practices and workers' absenteeism. Other school policies such as teacher evaluation and a high level of parental involvement in schools also have been found to contribute to a reduction in teacher absenteeism. Additionally, this study found that larger school sizes leads to a higher probability of teacher absenteeism in schools. These results are robust when the interaction effects between the variables of interest and school size are included in the teacher absenteeism model.

Lastly, as expected, teacher absenteeism showed a substantially significant adverse effect on student attainments, especially in rural area schools. Thus, the MOE or other policymakers may wish to pay more attention to the important determinants of teacher absenteeism as it may deteriorate student attainments. This study also suggests further study to examine other determinants of teacher absenteeism such as teacher evaluations by internal examiners, mentoring by the principal, and parental involvement in discussing children's attitude problems on teacher absenteeism. The findings may be beneficial to policymakers to quantify the rationale of each of the policies for the development of the Malaysian educational system.

Appendix C

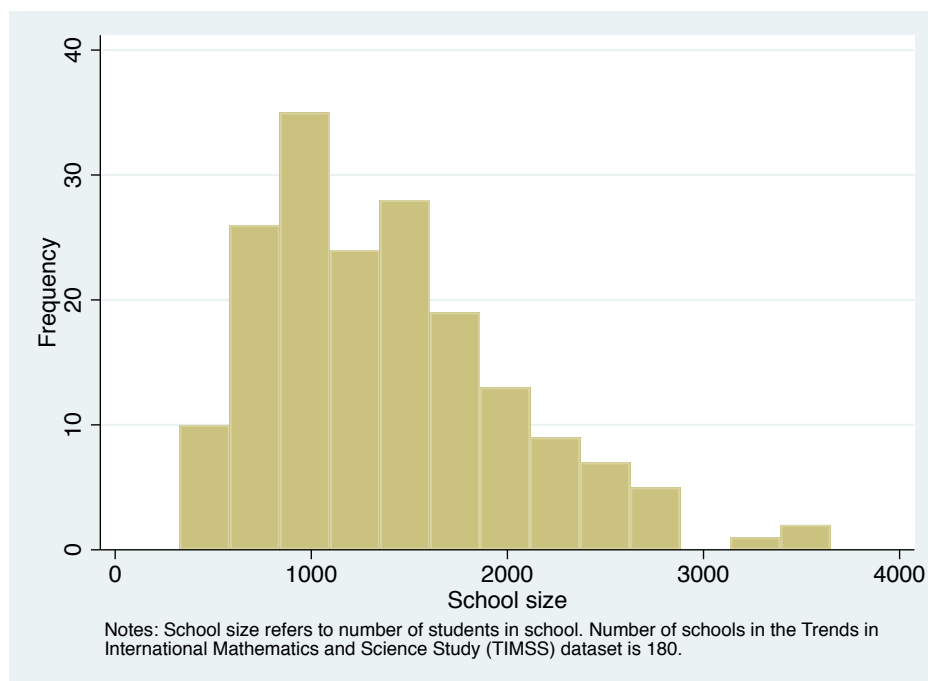


Figure A4.1: Distribution of school size

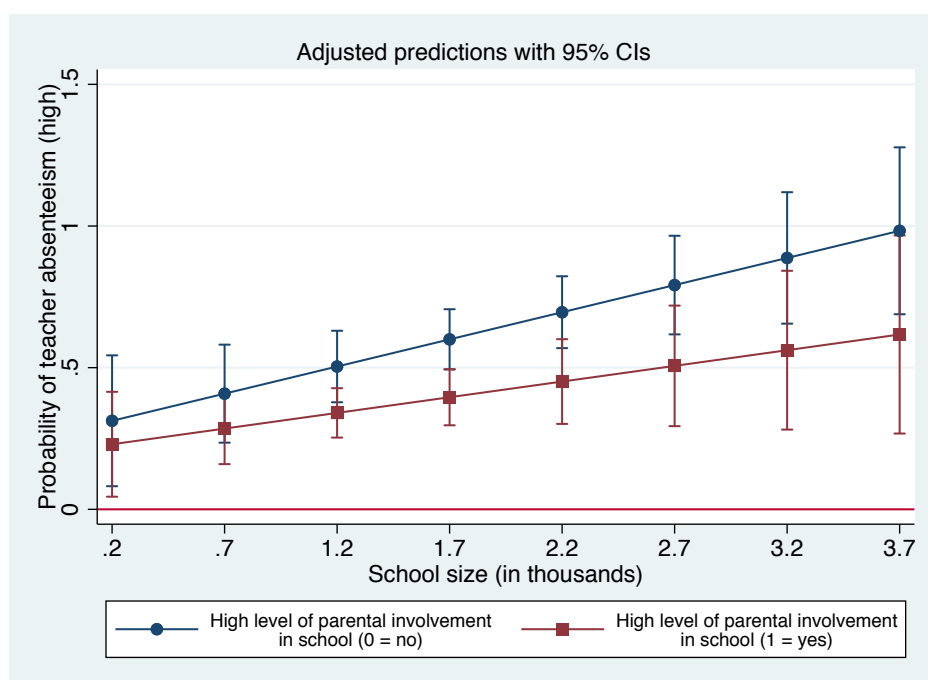


Figure A4.2: Percent teacher absenteeism by high level of parental involvement in school and school size

Chapter 5

Conclusion

This chapter begins with a presentation of the main results and implications of each topic in this thesis. It then highlights the main limitations of each of the topics and provides a few suggestions for further research. It ends with a subsection of recommendations for Malaysian development.

5.1 Summary of results and implications

Education is the main pillar to increase economic growth of a country. There are many studies which have highlighted the statement as discussed in the introduction chapter. This thesis presents new empirical findings related to it, which focus on a few possible factors to improve student outcomes in Malaysia in general.

Chapter 2 of this thesis has found that there is still evidence of gender differences in household education expenditure within the Malaysian households, which possibly affect student outcomes. Specifically, we find pro-male gaps in household education expenditure for primary school age children (5-12 years) within the Malaysian households as a whole, in the rural areas and Malay ethnicity. However, it is found that households prefer to invest in females compared to males' education in the post-primary school ages (13-17 years). Regarding the determinants of household education expenditure, we find consistent evidence with many previous studies such as [Tansel \(2002\)](#) and [Behrman et al. \(2017\)](#) that show that those

advantaged households tend to spend more money on their children's education compared to their counterparts. Thus, it implies that further actions are needed from the Ministry and households, for example, to rationalise the allocation of the expenditure on education by student age group or their household backgrounds, in order to reduce the gender gap issue. It is important to do so since the issue may be a contributing factor to the inequality in student attainment by gender in Malaysia ([Saw, 2016](#)).

Student outcomes also can be enhanced through improvement of school policies ([Hanushek and Luque, 2003](#)). Chapter 3 shows the effects of several school factors based on the current Malaysian education policy: the Malaysia Education Blueprint (MEB) 2013-2025. We find a strong positive effect of school factors, especially school resources (i.e. instructional materials and infrastructure) and parental involvement in schools, while tentative evidence of the class size and IT staff on student attainment. Specifically, we find that male students in urban schools and small school sizes tend to perform worse if schools have large class sizes and “a little” level of IT staff. Factors such as differences in practices in schools and peer effects might explain these results, as argued by [Hattie \(2005\)](#), [Woessmann \(2016\)](#), and [Lazear \(2001\)](#) among others. The estimated results from the chapter appear to reflect the relevance of the continuation of some school factors within the policy. The results may become a guide for the Ministry of Education Malaysia to draw future educational plans for the Malaysian education system. For instance, initiatives from education policymakers among others to address the achievement gender gap problem through class determination and IT resources allocation by students' gender (may imply to the single-sex schools) would be need carefully investigation.

The last chapter (Chapter 4) discusses the problem of absenteeism that happens in Malaysian schools. This study finds evidence of a positive relationship between teacher and student, which becomes more pronounced when the issue of endogeneity is addressed. By focusing on a Malaysian education policy called the District

Transformation Program (DTP), the chapter shows that the initiative of mentoring at a medium level is a significant solution to the teacher absenteeism problem. At the end of the chapter, the problem consequence on student attainment is estimated, and it shows an expected finding. Thus, the chapter implies that interventions from the government, schools, and parents should be encouraged in improving both teachers' and students' behaviour with respect to their absence from school. The intervention perhaps may boost students' attainment, Malaysian human capital development, and eventually Malaysian economic growth in general.

5.2 Future suggestions

Several suggestions have been presented from the findings of each chapter. In this section, we aim to outline only the main suggestions that were related to possible solutions to the caveats of each of the chapters.

The main limitation in Chapter 2 is that the dataset, the Household Education Expenditure (HES), does not have data regarding education expenditure for each child in each household. Thus, to tackle the problem, if given a limited time frame, further studies are suggested to conduct a quasi-experimental research using a more focused sample as suggested by [Psacharopoulos and Patrinos \(2004\)](#), for example, a sample of children in secondary ages in rural areas of a district in Malaysia. In addition, further studies are suggested to examine the effect of the gender gaps on student attainment on the students' return to education. Finally, children's socioeconomic background and indirect cost in schooling also should be given attention in estimating the household education expenditure model as suggested by [Stinebrickner and Stinebrickner \(2003\)](#).

Chapter 3 main limitation is sample representativeness. There is only one eighth grade class per school involved in the TIMSS survey which means that factor such as the class size effect maybe not fully represent average class size effect for school level. It suggests further studies to control for students' types of

school as suggested by Lazear (2001), Akabayashi and Nakamura (2014), and Chingos (2012) among others, who show various findings of class size on student attainment among schools. To note, the chapter only considered school effects on eighth-grade students' scores in mathematics and science as measures for the educational outcomes, thus further studies are suggested to examine the effect of the school characteristics on other measures of student outcomes. For example, student scores in early grades, student repetition rates, and student scores in the top and bottom quantile, as employed in studies by Lee and Barro (2001), Krueger and Whitmore (2001), and Nieto and Ramos (2015). Lastly, Chapter 4 suggests further research to take into account limitations regarding the proxy of the variables of interest in the topic. For example, further studies are suggested to consider another mentoring measure which has been practised in the DTP program, mentoring by school principals, to support the main result of the topic.

5.3 Recommendations for Malaysian development

Findings from this thesis provide a few recommendations for Malaysian development, especially in terms of education and economic development. For instance, initially it contributes to the survey of the *Jawatankuasa Kajian Dasar Pendidikan Negara (JKDPN)* 2019 under the Ministry of Education Malaysia on recommendations to improve the Malaysian education system.

Next, this thesis recommends both technical and non-technical institutions in Malaysia revise the courses offered particularly related to mathematics and science to meet the industry's demand. For example, revise the syllabus of the courses, so that it is in line with the Fourth Industrial Revolution (IR 4.0), to produce competitive young generations that match the current market regardless of gender. This step is expected to reduce the incidence of incompatibility of the graduate skills with the demand of the industry (Refrigeri and Aleandri, 2013; Lavrinovicha et al., 2015; Gontkovičová et al., 2015), and indirectly reduce gender gaps in

household education expenditure in Malaysia.

To reduce the gender gap in education spending on children, this study also recommends for every household to start saving for their children as early as possible and have an education plan for them. Studies by [Dynarski et al. \(2013\)](#), [Björklund and Salvanes \(2011\)](#), and [Heckman and Mosso \(2014\)](#) are among those that show the importance of early childhood investment on the children's future education. As a consequence, households, especially the disadvantaged households, are less likely to have a shortage of funds to use for their sons' and daughters' future education. The role of social media is encouraged to convey this recommendation.

Lastly, to improve student outcomes, both school factors and moral problems in schools should be addressed. All parties are suggested to make strategic interventions to ensure that the measures that have been taken are realistic and suitable with the current social, economic, and political situations in Malaysia. For example, the Ministry of Education's recommendation to transform the curriculum system of Level 1 of the primary education in Malaysian schools from exam-oriented to a learning-centred approach, in my opinion is a good step. The system has been used in the top performing country of Singapore, and it mainly has been proven successful. In addition, based on findings of this thesis, a values-driven education should be emphasised. This may help Malaysia to have a better environment for education and economic development.

Bibliography

- Acar, E. Ö., Gunalp, B., and Cilasun, S. M. (2016). An empirical analysis of household education expenditures in Turkey. *International Journal of Educational Development*, 51:23–35.
- Agadjanian, V. and Liew, H. P. (2005). Preferential policies and ethnic differences in post-secondary education in Peninsular Malaysia. *Race Ethnicity and Education*, 8(2):213–230.
- Ahmad, R. H. and Salleh, N. M. (1997). Bullying and violence in the Malaysian school. In *Violence at School: Global Issues and Interventions*. UNESCO/IBE, Paris.
- Aikman, S. and Rao, N. (2012). Gender equality and girls’ education: Investigating frameworks, disjunctures and meanings of quality education. *Theory and Research in Education*, 10(3):211–228.
- Akabayashi, H. and Nakamura, R. (2014). Can small class policy close the gap? An empirical analysis of class size effects in Japan. *Japanese Economic Review*, 65(3):253–281.
- Akerhielm, K. (1995). Does class size matter? *Economics of Education Review*, 14(3):229–241.
- Alderman, H. and King, E. M. (1998). Gender differences in parental investment in education. *Structural Change and Economic Dynamics*, 9(4):453–468.

- Anderson, K. H., King, E. M., and Wang, Y. (2003). Market returns, transfers and demand for schooling in Malaysia, 1976-89. *Journal of Development Studies*, 39(3):1–28.
- Andreou, S. N. (2012). Analysis of Household Expenditure on Education in Cyprus. *Cyprus Economic Policy Review*, 6(2):17–38.
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Aroff, A. R. b. M. (2014). Values Education and the Malaysia Education Blueprint. *Journal of Interdisciplinary Research in Education (JIRE)*, 4(1):59–73.
- Aslam, M. and Kingdon, G. G. (2008). Gender and household education expenditure in Pakistan. *Applied Economics*, 40(20):2573–2591.
- Aucejo, E. M. and Romano, T. F. (2016). Assessing the effect of school days and absences on test score performance. *Economics of Education Review*, 55:70–87.
- Azam, M. and Kingdon, G. G. (2013). Are Girls the Fairer Sex in India? Revisiting Intra-Household Allocation of Education Expenditure. *World Development*, 42(1):143–164.
- Banerjee, A. and Duflo, E. (2006). Addressing Absence. *Journal of Economic Perspectives*, 20(1):117–132.
- Banerjee, R., King, E. M., Orazem, P. F., and Paterno, E. M. (2012). Student and teacher attendance: The role of shared goods in reducing absenteeism. *Economics of Education Review*, 31(5):563–574.
- Barmby, T. and Stephen, G. (2000). Worker absenteeism: why firm size may matter. *The Manchester School*, 68(5):568–577.
- Barr, N. (2012). *Economics of the Welfare State*. OUP Oxford, 5th edition.

- Barro, R. J. (2013). Education and Economic Growth. *Annals of Economics and Finance*, 14(2):301–328.
- Becker, G. (1994). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. National Bureau of Economic Research, Inc, Chicago, 3rd edition.
- Becker, G. S. and Chiswick, B. R. (1996). Education and the Distribution of Earnings. *The American Economic Review*, 56(1/2):358–369.
- Behrman, J. R., Schott, W., Mani, S., Crookston, B. T., Dearden, K., Duc, L. T., Fernald, L. C. H., and Stein, A. D. (2017). Intergenerational Transmission of Poverty and Inequality: Parental Resources and Schooling Attainment and Children’s Human Capital in Ethiopia, India, Peru, and Vietnam. *Economic Development and Cultural Change*, 65(4):657–697.
- Bernama (2018). Teachers shortage in major cities a serious issue. *Borneo Post Online*.
- Bilenkisi, F., Gungor, M. S., and Tapsin, G. (2015). The impact of household heads’ education levels on the poverty risk: The evidence from Turkey. *Educational Sciences: Theory & Practice*, 15(2):337–348.
- Björklund, A. and Salvanes, K. G. (2011). Chapter 3 - education and family background: Mechanisms and policies. volume 3 of *Handbook of the Economics of Education*, pages 201 – 247. Elsevier.
- Blatchford, P., Bassett, P., and Brown, P. (2011). Examining the effect of class size on classroom engagement and teacher-pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools. *Learning and Instruction*, 21(6):715–730.
- Blazar, D. and Kraft, M. A. (2017). Teacher and Teaching Effects on Students’

- Attitudes and Behaviors. *Educational Evaluation and Policy Analysis*, 39(1):146–170.
- Bloom, N., Lemos, R., Sadun, R., and Van Reenen, J. (2015). Does Management Matter in Schools? *The Economic Journal*, 125(584):647–674.
- Botha, F. (2010). The impact of educational attainment on household poverty in South Africa. *Acta Academica*, 42(4):122–147.
- Breierova, L. and Duflo, E. (2004). The impact of education on fertility and child mortality: Do fathers really matter less than mothers? Working Paper 10513, National Bureau of Economic Research.
- Breton, T. R. (2011). The quality vs. the quantity of schooling: What drives economic growth? *Economics of Education Review*, 30(4):765–773.
- Breton, T. R. (2014). Evidence that class size matters in 4th grade mathematics : An analysis of TIMSS 2007 data for Colombia. *International Journal of Educational Development*, 34:51–57.
- Brewer, D. J. (1993). Principals and student outcomes: Evidence from U.S. high schools. *Economics of Education Review*, 12(4):281–292.
- Brown, P. H. (2006). Parental education and investment in children’s human capital in rural china. *Economic Development and Cultural Change*, 54(4):759–789.
- Brown, S., Fakhfakh, F., and Sessions, J. G. (1999). Absenteeism and Employee Sharing: An Empirical Analysis Based on French Panel Data , 1981-1991. *Industrial and Labor Relations Review*, 52(2):234–251.
- Bryson, A. and Dale-Olsen, H. (2017). Does Sick Pay Affect Workplace Absence? DoQSS Working Papers 17-12, Department of Quantitative Social Science - UCL Institute of Education, University College London.

- Buchmann, C. and Diprete, T. A. (2006). The Growing Female Advantage in College Completion: The Role of Family Background and Academic Achievement. *American Sociological Review*, 71:515–541.
- Buchmann, C., Diprete, T. A., and McDaniel, A. (2008). Gender Inequalities in Education. *Annual Review of Sociology*, 34:319–337.
- Bulman, G. and Fairlie, R. (2016). Chapter 5 - technology and education: Computers, software, and the internet. volume 5 of *Handbook of the Economics of Education*, pages 239 – 280. Elsevier.
- Burgess, R. and Zhuang, J. (2000). Modernisation and Son Preference. STICERD - Development Economics Papers - From 2008 this series has been superseded by Economic Organisation and Public Policy Discussion Papers 29, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.
- Cáceres-Delpiano, J. (2006). The Impacts of Family Size on Investment in Child Quality. *Journal of Human Resources*, 41(4):738–754.
- Card, D. and Krueger, A. B. (1996). School resources and student outcomes: an overview of the literature and new evidence from north and south Carolina. *Journal of Economic Perspectives*, 10(4):31–50.
- Castelló-Climent, A. (2010). Channels Through Which Human Capital Inequality Influences Economic Growth. *Journal of Human Capital*, 4(4):394–450.
- Castelló-Climent, A. and Hidalgo-Cabrillana, A. (2012). The role of educational quality and quantity in the process of economic development. *Economics of Education Review*, 31(4):391–409.
- Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K., and Rogers, F. H. (2006). Missing in action: teacher and health worker absence in developing countries. *Journal of Economic Perspectives*, 20(1):91–116.

- Chi, W. and Qian, X. (2016). Human capital investment in children: An empirical study of household child education expenditure in China, 2007 and 2011. *China Economic Review*, 37(14):52–65.
- Chingos, M. M. (2012). The impact of a universal class-size reduction policy: Evidence from Florida’s statewide mandate. *Economics of Education Review*, 31:543–562.
- Cho, H., Glewwe, P., and Whitley, M. (2012). Do reductions in class size raise students’ test scores? Evidence from population variation in Minnesota’s elementary schools. *Economics of Education Review*, 31:77–95.
- Chowa, G. A., Masa, R. D., Wretman, C. J., and Ansong, D. (2013). The impact of household possessions on youth’s academic achievement in the Ghana Youthsave experiment: A propensity score analysis. *Economics of Education Review*, 33:69–81.
- Chung, T. P. (2003). Returns to education: Updates for Malaysia. *Applied Economics Letters*, 10(13):837–841.
- Clark, M. A., Lee, S. M., Goodman, W., and Yacco, S. (2008). Underachievement in Public Action Research at a District Level. *NASSP Bulletin*, 92(2):111–132.
- Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2006). Teacher-Student Matching and the Assessment of Teacher Effectiveness. *The Journal of Human Resources*, 41(4):778–820.
- Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2009). Are Teacher Absences Worth Worrying About in the United States? *Education Finance and Policy*, 4(2):115–149.
- Cochrane, S. H. (1979). *Fertility and Education: What Do We Really Know?* International Bank for Reconstruction and Development Washington, DC: Occasional papers. World Bank.

- Coelli, M. and Green, D. A. (2012). Leadership effects: School principals and student outcomes. *Economics of Education Review*, 31:92–109.
- Cragg, J. G. (1971). Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. *Econometrica*, 39(5):829–844.
- Cucchiella, F., Gastaldi, M., and Ranieri, L. (2014). Managing Absenteeism in the Workplace: The Case of an Italian Multiutility Company. *Procedia - Social and Behavioral Sciences*, 150:1157–1166.
- Dai, H., Masui, T., Matsuoka, Y., and Fujimori, S. (2012). The impacts of China’s household consumption expenditure patterns on energy demand and carbon emissions towards 2050. *Energy Policy*, 50:736–750.
- Dang, H. A. H. and Rogers, F. H. (2015). The decision to invest in child quality over quantity: Household size and household investment in education in vietnam. *World Bank Economic Review*, 30(1):104–142.
- Dass, A. (2018). Young and jobless in Malaysia. *The Star Online*.
- Deaton, A. (1987). The allocation of goods within the household: Adults, children, and gender. Lsms working paper, The World Bank, Washington, DC.
- Deaton, A. (1997). *The analysis of household surveys: A microeconomic approach to development policy*. The World Bank.
- Dee, S. and Wyckoff, J. (2015). Incentives, Selection, and Teacher Performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, 34(2):267–297.
- Dee, T. S., Ha, W., and Jacob, B. A. (2007). The Effects of School Size on Parental Involvement and Social Capital: Evidence from the ELS: 2002. *Brookings Papers on Education Policy: 2006/2007*.

- Dee, T. S. and West, M. R. (2011). The Non-Cognitive Returns to Class Size. *Educational Evaluation and Policy Analysis*, 33(1):23–46.
- Demir, I., Ünal, H., and Kiliç, S. (2010). The effect of quality of educational resources on mathematics achievement: Turkish Case from PISA-2006. *Procedia - Social and Behavioral Sciences*, 2:1855–1859.
- Department of Statistic Malaysia (2018). Informal Sector Work Force Survey Report, Malaysia, 2017. Technical report, Department of Statistics Malaysia.
- Department of Statistics Malaysia (2014). Report on Household Expenditure Survey 2014. Technical report, Department of Statistics Malaysia.
- Department of Statistics Malaysia (2017a). Key Statistics of Labour Force in Malaysia, September 2017. Technical report, Department of Statistics Malaysia.
- Department of Statistics Malaysia (2017b). Salaries & Wages Survey Report, Malaysia, 2016. Technical report, Department of Statistics Malaysia.
- Dhuey, E. and Smith, J. (2014). How Important Are School Principals in the Production of Student Achievement? *Canadian Journal of Economics*, 47(2):634–663.
- Dieterle, S. G. (2015). Class-size reduction policies and the quality of entering teachers. *Labour Economics*, 36:35–47.
- Donkoh, S. A. and Amikuzuno, J. A. (2011). The determinants of household education expenditure in Ghana. *Educational Research and Reviews*, 6(8):570–579.
- Duflo, B. E., Dupas, P., and Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *The American Economic Review*, 101(5):1739–1774.

- Duflo, B. E., Hanna, R., and Ryan, S. P. (2012). Incentives Work: Getting Teachers to Come to School. *The American Economic Review*, 102(4):1241–1278.
- Durden, B. G. C. and Ellis, L. V. (1995). The Effects of Attendance on Student Learning in Principles of Economics. *The American Economic Review*, 85(2):343–346.
- Dynarski, S., Hyman, J., and Schanzenbach, D. W. (2013). Experimental Evidence on the Effect of Childhood Investments on Postsecondary Attainment and Degree Completion. *Journal of Policy Analysis and Management*, 32(4):692–717.
- Ebaidalla, E. M. (2017). Determinants of Household Education Expenditure in Sudan. Working Papers 1138, Economic Research Forum.
- Eberts, R. W. and Stone, J. A. (1988). Student achievement in public schools: Do principals make a difference? *Economics of Education Review*, 7(3):291–299.
- Ehrenberg, R. G., Ehrenberg, R. a., Rees, D. I., and Ehrenberg, E. L. (1991). School District Leave Policies, Teacher Absenteeism, and Student Achievement. *The Journal of Human Resources*, 26(1):72–105.
- Enders, C. K. (2010). *Applied Missing Data Analysis*. The Guilford Press, New York London.
- Fajoku, S. A., Aluede, O., and Ojugo, A. I. (2016). Parental involvement as a correlate of academic achievement of primary school pupils in Edo State, Nigeria. *Research in Education*, 95(1):33–43.
- Fang, H. and Moro, A. (2011). Chapter 5 - theories of statistical discrimination and affirmative action: A survey. volume 1 of *Handbook of Social Economics*, pages 133 – 200. North-Holland.
- Fuchs, T. and Wößmann, L. (2004). Computers and Student Learning: Bivariate and Multivariate Evidence on the Availability and Use of Computers at Home and at School. Technical report.

- Galiani, S. and Perez-Truglia, R. (2013). School Management in Developing Countries. CEDLAS, Working Papers 0147, CEDLAS, Universidad Nacional de La Plata.
- Galton, M. and Pell, T. (2012). Do class size reductions make a difference to classroom practice? The case of Hong Kong primary schools. *International Journal of Educational Research*, 53:22–31.
- Gary-Bobo, R. J. and Mahjoub, M.-B. (2013). Estimation of Class-Size Effects, Using “Maimonides’ Rule” and Other Instruments: The Case of French Junior High Schools. *Annals of Economics and Statistics*, (111/112):193–225.
- Gershenson, S., Jacknowitz, A., and Brannegan, A. (2017). Are Student Absences Worth the Worry in U.S. Primary Schools? *Education Finance and Policy*, 12(2):137–165.
- Gertler, P. J., Patrinos, H. A., and Rubio-Codina, M. (2012). Empowering parents to improve education: Evidence from rural Mexico. *Journal of Development Economics*, 99(1):68–79.
- Ghazali, N. H. C. M. (2016). The implementation of School-Based Assessment System in Malaysia: A study of teacher perceptions. *GEOGRAFIA Online TM Malaysian Journal of Society and Space*, 12(9):104–117.
- Glewwe, P., Hanushek, E., Humpage, S., and Ravina, R. (2012). School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010. Working Papers 120033, University of Minnesota, Center for International Food and Agricultural Policy.
- Glewwe, P. and Jacoby, H. (1994). Student Achievement and Schooling Choice in Low-Income Countries: Evidence from Ghana. *The Journal of Human Resources*, 29(3):843–864.

- Glewwe, P. and Jacoby, H. G. (2004). Economic growth and the demand for education: Is there a wealth effect? *Journal of Development Economics*, 74(1):33–51.
- Glewwe, P. and Kremer, M. (2006). Chapter 16 schools, teachers, and education outcomes in developing countries. volume 2 of *Handbook of the Economics of Education*, pages 945 – 1017. Elsevier.
- Glewwe, P. and Muralidharan, K. (2016). Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications. In *Handbook of the Economics of Education*, volume 5, chapter 10, pages 653–743. Elsevier B.V.
- Goldschmidt, P. and Phelps, G. (2010). Does teacher professional development affect content and pedagogical knowledge: How much and for how long? *Economics of Education Review*, 29(3):432–439.
- Gontkovičová, B., Mihalčová, B., and Pružinský, M. (2015). Youth Unemployment – Current Trend in the Labour Market? *Procedia Economics and Finance*, 23(October 2014):1680–1685.
- Gottfried, M. A. (2010). Evaluating the Relationship Between Student Attendance and Achievement in Urban Elementary and Middle Schools: An Instrumental Variables Approach. *American Educational Research Journal*, 47(2):434–465.
- Guerrero, G., Leon, J., Zapata, M., and Cueto, S. (2013). Getting teachers back to the classroom. A systematic review on what works to improve teacher attendance in developing countries. *Journal of Development Effectiveness*, 5(4):466–488.
- Gunnarsson, V., Orazem, P. F., Sánchez, M. A., and Verdisco, A. (2009). Does local school control raise student outcomes? evidence on the roles of school autonomy and parental participation. *Economic Development and Cultural Change*, 58(1):25–52.

- Hanapi, Z. and Nordin, M. S. (2014). Unemployment among Malaysia graduates: Graduates' attributes, lecturers' competency and quality of education. *Procedia - Social and Behavioral Sciences*, 112:1056–1063.
- Hanover Research (2015). Impacts of School and Class Size on Student Outcomes. *District Administration Practice*.
- Hanushek, E. (1997). Assessing the Effects of School resources on Student Performance: An Update. *Educational Evaluation and Policy Analysis*, 19(2):141–164.
- Hanushek, E. A. (2003). The Failure of Input-Based Schooling Policies. *The Economic Journal*, 113(485):F64–F98.
- Hanushek, E. A. (2013). Economic growth in developing countries: The role of human capital. *Economics of Education Review*, 37:204–212.
- Hanushek, E. A. and Luque, J. A. (2003). Efficiency and equity in schools around the world. *Economics of Education Review*, 22(5):481–502.
- Hanushek, E. A. and Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, 17(4):267–321.
- Hanushek, E. A. E. and Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of economic literature*, 46(3):607–668.
- Harris, D. N. and Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics*, 95(7):798–812.
- Hassink, W. H. and Koning, P. (2009). Do Financial Bonuses Reduce Employee Absenteeism? Evidence From a Lottery. *ILR Review*, 62(3):327–342.
- Hattie, J. (2005). The paradox of reducing class size and improving learning outcomes. *International Journal of Educational Research*, 43(6):387–425.

- Hauser, R. M. and Kuo, H.-H. D. (1998). Does the Gender Composition of Sibships Affect Women's Educational Attainment? *The Journal of Human Resources*, 33(3):644–657.
- Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6):1251–1271.
- Heckman, J. J. (1976). The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement*, 5(4):475–492.
- Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1):153.
- Heckman, J. J. and Mosso, S. (2014). The Economics of Human Development and Social Mobility. Technical report, National Bureau of Economic Research.
- Herrmann, M. A. and Rockoff, J. E. (2012). Worker Absence and Productivity: Evidence from Teaching. *Journal of Labor Economics*, 30(4):749–782.
- Herz, B. and Sperling, G. B. (2004). *What Works in Girls' Education: Evidence and Policies from the Developing World*. Council on Foreign Relations Press.
- Hoddinott, J. and Haddad, L. (1995). Does female income share influence household expenditures? Evidence from Côte d'Ivoire. *Oxford Bulletin of Economics and Statistics*, 57(1):77–96.
- Holcomb, R. B., Park, J. L., and Capps, O. (1995). Revisiting Engel's Law: Examining Expenditure Patterns for Food at Home and Away From Home. *Journal of Food Distribution Research*, 26:1–8.
- Houtenville, A. J. and Conway, K. S. (2008). Parental Effort, School Resources, and Student Achievement. *The Journal of Human Resources*, 43(2):437–453.

- Hoxby, C. M. (2000). The effects of Class Size on Student Achievement: New Evidence from Population Variation. *The Quarterly Journal of Economics*, 115(4):1239–1285.
- “Hundreds of Teachers” (2011). Hundreds of teachers found playing truant. *Daily Express*.
- Ibrahim, D. H. M. and Mahyuddin, M. Z. (2016). Youth Unemployment in Malaysia: Developments and Policy Considerations. Technical report, Bank Negara Malaysia.
- “Ibu bapa masih perlu” (2007). Ibu bapa masih perlu jelas yuran aktiviti sekolah. *Utusan Online*.
- “Ibu bapa mesti jelas” (2007). Ibu bapa mesti jelas mengenai pemansuhan yuran. *Utusan Online*.
- Irvine, J. J. (1986). Teacher–student interactions: Effects of student race, sex, and grade level. *Journal of Educational Psychology*, 78(1):14–21.
- Ismail, L. (2015). Gender gap in higher education: Perspective on factors influencing enrolment in Malaysian universities: A University of Malaya sample. *The Online Journal of Quality in Higher Education*, 2(4):35–45.
- Jackman, W. M. and Morrain-webb, J. (2019). Exploring gender differences in achievement through student voice: Critical insights and analyses. *Cogent Education*, 6(1):1–13.
- Jacob, B. A. and Walsh, E. (2011). What’s in a rating? *Economics of Education Review*, 30(3):434–448.
- Jacobson, S. L. (1989). The effects of pay incentives on teacher absenteeism. *The Journal of Human Resources*, 24(2):280–287.

- Jelas, Z. M. and Dahan, H. M. (2010). Gender and educational performance: The Malaysian perspective. *Procedia - Social and Behavioral Sciences*, 7:720–727.
- Jelas, Z. M., Salleh, A., Mahmud, I., Azman, N., Hamzah, H., Hamid, Z. A., Jani, R., and Hamzah, R. (2014). Gender Disparity in School Participation and Achievement: The Case in Malaysia. *Procedia - Social and Behavioral Sciences*, 140:62–68.
- Jeong, D. W., Lee, H. J., and Cho, S. K. (2017). Education decentralization, school resources, and student outcomes in Korea. *International Journal of Educational Development*, 53:12–27.
- Jepsen, C. and Rivkin, S. (2009). Class Size Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size. *The Journal of Human Resources*, 44(1):223–250.
- Johansoon, P. and Palme, M. (2002). Assessing the Effect of Public Policy on Worker Absenteeism. *The Journal of Human Resources*, 37(2):381–409.
- Jones, J. T., Toma, E. F., and Zimmer, R. W. (2008). School attendance and district and school size. *Economics of Education Review*, 27(2):140–148.
- Kaul, T. (2018). Intra-household allocation of educational expenses: Gender discrimination and investing in the future. *World Development*, 104:336–343.
- Kazianga, H. and Wahhaj, Z. (2017). Intra-household resource allocation and familial ties. *Journal of Development Economics*, 127:109–132.
- Kemmis, S., Heikkinen, H. L., Göran, F., Aspors, J., and Groves, C. E. (2014). Mentoring of new teachers as a contested practice: Supervision, support and collaborative self-development. *Teaching and Teacher Education*, 43(4):154–164.
- Kenayathulla, H. B. (2013). Higher levels of education for higher private returns: New evidence from Malaysia. *International Journal of Educational Development*, 33(4):380–393.

- Kenayathulla, H. B. (2016). Gender differences in intra-household educational expenditures in Malaysia. *International Journal of Educational Development*, 46:59–73.
- Khanal, S. (2018). Gender discrimination in education expenditure in Nepal: Evidence from living standards surveys. *Asian Development Review*, 35(1):155–174.
- Kim, S. and Lee, J. (2010). Private Tutoring and Demand for Education in South Korea. *Economic Development and Cultural Change*, 58(2):259–296.
- Kim-Soon, N., Ahmad, A. R. B., Sulaiman, M. I. B., and Sirisa, N. M. X. (2015). Homeschool in Malaysia: A Foresight Study. *International Education Studies*, 8(10):163–174.
- Kingdon, G. (2005). Where Has All the Bias Gone? Detecting Gender Bias in the Intrahousehold Allocation of Educational Expenditure. *Economic Development and Cultural Change*, 53(2):409–451.
- Klasen, S. (2002). Low Schooling for Girls, Slower Growth for All? Cross-Country Evidence on the Effect of Gender Inequality in Education on Economic Development. *The World Bank Economic Review*, 16(3):345–373.
- Klasen, S. and Lamanna, F. (2009). The impact of gender inequality in education and employment on economic growth: New evidence for a panel of countries. *Feminist Economics*, 15(3):91–132.
- Konstantopoulos, S. (2008). Do Small Classes Reduce the Achievement Gap between Low and High Achievers? Evidence from Project STAR. *The Elementary School Journal*, 108(4):275–291.
- Kremer, M., Chaudhury, N., Rogers, F. H., Muralidharan, K., and Hammer, J. (2005). Teacher absence in India: A snapshot. *Journal of the European Economic Association*, 3(2-3):658–667.

- Krueger, A. B. (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics*, 114:497–532.
- Krueger, A. B. and Whitmore, D. M. (2001). The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project STAR. *The Economic Journal*, 111(468):1–28.
- Krueger, A. B. (2003). Economic Considerations and Class Size. *Economic Journal*, 113:F34–F63.
- Lachaud, J., LeGrand, T. K., Adjiwanou, V., and Kobiané, J. F. (2014). Family size and intra-family inequalities in education in Ouagadougou. *Demographic Research*, 31:1455–1476.
- Laibson, D. and List, J. A. (2015). Principles of (Behavioral) Economics. *American Economic Association*, 105(5):385–390.
- Lavrinovicha, I., Lavrinenko, O., and Teivans-Treinovskis, J. (2015). Influence of Education on Unemployment Rate and Incomes of Residents. *Procedia - Social and Behavioral Sciences*, 174:3824–3831.
- Lazear, E. P. (2001). Education Production. *The Quarterly Journal of Economics*, 116(3):777–803.
- Lee, J.-W. and Barro, R. J. (2001). Schooling quality in a cross-section of countries. *Economica*, 68(272):465–489.
- Lee, K. S. (2009). Competition for resources: A reexamination of sibship composition models of parental investment. *Journal of Marriage and Family*, 71(2):263–277.
- Lee, V. E. and Loeb, S. (2000). School Size in Chicago Elementary Schools: Effects on Teachers’ Attitudes and Students’ Achievement. *American Educational Research Journal*, 37(1):3–31.

- Leser, C. E. V. (1963). Forms of Engel Functions. *Econometrica*, 31(4):694–703.
- Lillard, L. a. and Willis, R. J. (1994). Intergenerational Educational Mobility: Effects of Family and State in Malaysia. *The Journal of Human Resources*, 29(4):1126–1166.
- Lim, D. (1986). East Malaysia in Malaysian Development Planning. *Journal of Southeast Asian Studies*, 17(1):156–170.
- Louis, K. S., Dretzke, B., and Wahlstrom, K. (2010). How does leadership affect student achievement? Results from a national US survey. *School Effectiveness and School Improvement*, 21(3):315–336.
- Majid, F. A. (2011). School-Based Assessment in Malaysian Schools: The concerns of the English teachers. *Journal of US-China Education Review*, 8(10).
- Majzub, R. M. and Rais, M. M. (2010). Boys’ Underachievement: Male versus Female Teachers. *Procedia Social and Behavioral Sciences*, 7:685–690.
- Malakolunthu, S. and Hoon, S. K. (2010). Teacher perspectives of school-based assessment in a secondary school in Kuala Lumpur. *Procedia - Social and Behavioral Sciences*, 9:1170–1176.
- Malakolunthu, S., Idris, A. R., and Rengasamy, N. C. (2010). Teacher professional experience and performance: impact of the work environment and general welfare in Malaysian secondary schools. *Asia Pacific Education Review*, 11(4):609–617.
- Malaysian Communications and Multimedia Commission (2010). National Broad-band Initiative: SKMM Allocates One Million 1Malaysia Netbooks for Distribution. Technical report, Cyberjaya.
- Marburger, D. R. (2001). Absenteeism and Undergraduate Exam Performance. *The Journal of Economic Education*, 32(2):99–109.

- Martins, L. and Veiga, P. (2010). Do inequalities in parents' education play an important role in PISA students' mathematics achievement test score disparities? *Economics of Education Review*, 29:1016–1033.
- Massingham, P. and Herrington, T. (2006). Does Attendance Matter? An Examination of Student Attitudes, Participation, Performance and Attendance. *Journal of University Teaching and Learning Practice*, 3(2).
- McDonald, N. C. and Aalborg, A. E. (2009). Why parents drive children to school: Implications for Safe Routes to School Programs. *Journal of the American Planning Association*, 75(3):331–342.
- McKenzie, D. J. (2005). Measuring inequality with asset indicators. *Journal of Population Economics*, 18(2):229–260.
- Mckenzie, P., Nugroho, D., Ozolins, C., Mcmillan, J., Sumarto, S., Toyamah, N., Febriany, V., Sodo, R. J., Bima, L., and Sim, A. A. (2014). Teacher Absenteeism in Indonesia : Policy Brief. Technical Report November, Jakarta: Education Sector Analytical and Capacity Development Partnership (ACDP).
- Miller, R., Murnane, R., and Willett, J. (2008). Do Teacher Absences Impact Student Achievement? Longitudinal Evidence From One Urban School District. *Educational Evaluation and Policy Analysis*, 30(2):181–200.
- Mincer, J. (1984). Human Capital and Economic Growth. *Economics of Education Review*, 3(3):195–205.
- Ministry of Education Malaysia (2013a). Malaysia Education Blueprint 2013 - 2025 (Prescool to Post-Secondary Education). Technical report, Ministry of Education Malaysia, Malaysia.
- Ministry of Education Malaysia (2013b). Malaysia Education Blueprint Annual Report 2013. Technical report, Ministry of Education Malaysia.

- Ministry of Education Malaysia (2013c). Quick Facts 2013 - Malaysia Educational Statistics. Technical report, Educational Planning and Research Division, Malaysia.
- Ministry of Education Malaysia (2015). Malaysia Education Blueprint Annual Report 2014. Technical report, Ministry of Education Malaysia, Malaysia.
- Ministry of Education Malaysia (2016). Statistic of Online Services - Ministry of Education Malaysia. Technical report, Ministry of Education Malaysia, Malaysia.
- Ministry of Education Malaysia (2017). Quick Facts 2017 Malaysia Educational Statistics. Technical Report July, Educational Planning and Research Division, Malaysia.
- Ministry of Education Malaysia (2019). Bantuan Pembelajaran. Technical report.
- Monk, D. H. (1987). Secondary School Size and Curriculum Comprehensiveness. *Economics of Education Review*, 6(2):137–150.
- Monks, J. and Schmidt, R. M. (2011). The Impact of Class Size on Outcomes in Higher Education. *The B.E. Journal of Economic Analysis & Policy*, 11(1):1–27.
- Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, 33(3):341–365.
- Muslim Women’s League (1995). Islamic Inheritance. *Muslim Women’s League*.
- Naidu, G. (2008). Chapter 7: Infrastructure Development in Malaysia. Technical report, Chiba: IDE-JETRO.
- Nasrudin, N. and Nor, A. R. M. (2013). Travelling to School: Transportation Selection by Parents and Awareness towards Sustainable Transportation. *Procedia Environmental Sciences*, 17:392–400.

- Nieto, S. and Ramos, R. (2015). Educational outcomes and socioeconomic status: A decomposition analysis for middle-income countries. *PROSPECTS*, 45(3):325–343.
- NST Team (2017). 2018 Budget: TVET Malaysia Master Plan unveiled, RM4.9bil allocated. *New Straits Times*.
- Omar, C. M. Z. C. and Rajoo, S. (2016). Unemployment Among Graduates in Malaysia. *International Journal of Economics, Commerce and Management United Kingdom*, IV(8):367–374.
- Ost, B. and Schiman, J. C. (2017). Workload and Teacher Effort. *Economics of Education Review*, 57:20–30.
- Othman, M. I. (2006). *Menangani disiplin di sekolah*. Utusan Publications.
- Othman, Z. (2014). Developing higher-order thinking. *The Star Online*.
- Parker, P. D., Zanden, B. V., and Parker, R. B. (2018). Girls get smart, boys get smug: Historical changes in gender differences in math, literacy, and academic social comparison and achievement. *Learning and Instruction*, 54:125–137.
- Patel, T. (2014). Dropping Out of School in Malaysia: What we know and what needs to be done. Technical report, Institute for Democracy and Economic Affairs.
- Pong, S.-l. (1993). Preferential Policies and Secondary School Attainment in Peninsular Malaysia. *Sociology of Education*, 66(4):245–261.
- Pong, S.-l. (1997). Sibship Size and Educational Attainment in Peninsular Malaysia: Do Policies Matter? *Sociological Perspectives*, 40(2):227–242.
- Powell, B., Steelman, L. C., and Carolina, S. (1995). Feeling the Pinch: Child Spacing and Constraints on Parental Economic Investments in Children. *Social Forces*, 73(4):1465–1486.

- Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development*, 22(9):1325–1343.
- Psacharopoulos, G. and Patrinos, H. A. (2004). Return to investment in education: A further update. *Education Economics*, 12(2):111–134.
- Psacharopoulos, G. and Woodhall, M. (1985). *Education for Development: An Analysis of Investment Choices*. Oxford University Press.
- Qian, J. X. and Smyth, R. (2011). Educational expenditure in urban China: Income effects, family characteristics and the demand for domestic and overseas education. *Applied Economics*, 43(24):3379–3394.
- Rahim, N. A. A. (2018). Edu Nation: Eliminate extreme teacher absenteeism. *The Edge Markets*.
- Refrigeri, L. and Aleandri, G. (2013). Educational Policies and Youth Unemployment. *Procedia - Social and Behavioral Sciences*, 93:1263–1268.
- Rogers, F. H. and Vegas, E. (2009). No More Cutting Class? Reducing Teacher Absence and Providing Incentives for Performance. Policy Research Working Paper Series 4847, The World Bank.
- Román, G. T. and Valenzuela, J. P. (2015). Over-estimating the effects of teacher attributes on school performance in the Chilean education system */Sobrestimación del impacto de atributos docentes en el desempeño escolar en el sistema educativo chileno. *Estudios de Economía*, 42(1):99–128.
- Rubin, D. B. (1976). Inference and Missing Data. *Biometrika*, 63(3):581–592.
- Sackey, H. A. (2007). The Determinants of School Attendance and Attainment in Ghana: A Gender Perspective. Technical report, African Economic Research Consortium, Nairobi.

- Saha, A. (2013). An Assessment of Gender Discrimination in Household Expenditure on Education in India. *Oxford Development Studies*, 41(2):220–238.
- Saido, G. M., Siraj, S., Nordin, A. B. B., and Amedy, O. S. A. (2015). Higher Order Thinking Skills Among Secondary School Students in Science Learning. *The Malaysian Online Journal of Educational Science*, 3(3):13–20.
- Sani, R. (2018). Skills-based pathway to a high-income nation. *New Straits Times*.
- Sargan, J. (1958). The Estimation of Economic Relationships using Instrumental Variables. *Econometrica*, 26(3):393–415.
- Saw, G. K. (2016). Patterns and trends in achievement gaps in Malaysian Secondary Schools (1999-2011): gender, ethnicity, and socioeconomic status. *Educational Research for Policy and Practice*, 15(1):41–54.
- Schultz, T. W. (1993). The economic importance of human capital in modernization. *Education Economics*, 1(1):13–19.
- Scott, L., Vaughn, C., Wolfe, M., and Wyant, C. (2007). Reducing Teacher Absences in North Carolina. Technical Report May, Tery Sanford Institute of Public Policy Duke University.
- Sichande, M., Michelo, C., Halwindi, H., and Miller, J. (2014). Education attainment of head of households associated with insecticide-treated net utilization among five to nineteen-year old individuals: Evidence from the malaria indicator survey 2010 in Zambia. *Malaria Journal*, 13(378):1–6.
- Snedecor, G. W. and Cochran, W. G. (1989). *Statistical Methods*. Blackwell Publishing Professional, Ames, Iowa, 8th edition.
- Song, L., Appleton, S., and Knight, J. (2006). Why Do Girls in Rural China Have Lower School Enrollment? *World Development*, 34(9):1639–1653.

- Stinebrickner, R. and Stinebrickner, T. R. . (2003). Understanding Educational Outcomes of Students from Low-Income Families: Evidence from a Liberal Arts College with a Full Tuition Subsidy Program. *The Journal of Human Resources*, 38(3):591–617.
- Stock, J. H., Wright, J. H., and Yogo, M. (2002). A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business & Economic Statistics*, 20(4):518–529.
- StudyMalaysia (2016). Technical and vocational education and training (TVET) in Malaysia. *studymalaysia.com*.
- StudyMalaysia.com (2016). Skill Training Programmes (Technical & Vocational) By Government. *studymalaysia.com*.
- Subramanian, S. and Deaton, A. (1991). Gender effects in Indian consumption patterns. *Sarvekshana*, 14(4):1–12.
- Sudha, S. (1997). Family Size, Sex Composition and Children’s Education: Ethnic Differentials Over Development in Peninsular Malaysia. *Population Studies*, 51(2):139–151.
- Sulaiman, N., Ismail, R., Othman, N., and Tin, P. B. (2012). The Determinants of Demand for Education among Households in Malaysia. *International Business Management*, 6(5):558–567.
- Sulaiman, T., Muniyan, V., Madhvan, D., Hasan, R., and Rahim, S. S. A. (2017). Implementation of Higher Order Thinking Skills in Teaching Of Science: A Case Study in Malaysia. *International Research Journal of Education and Sciences (IRJES)*, 1(1):1–3.
- Suryahadi, A. and Sambodho, P. (2013). Assessment of Public Policies to Improve Teacher Quality and Reduce Teacher Absenteeism. Technical report, The SMERU Research Institute, Jakarta.

- “Tackling the dropout” (2006). Tackling the dropout problem. *The Star Online*.
- Taharim, N. F., Jayasuriya, D. P., Xiang, L. W., and Mazhar, S. F. (2017). CLERICAL WORK FOR SCHOOL TEACHERS: A BURDEN OR A RESPONSIBILITY? *Journal of Humanities, Language, Culture and Business (HLCB)*, 1(3):26–36.
- Tansel, A. (1997). Schooling Attainment, Parental Education, and Gender in Côte d’Ivoire and Ghana. *Economic Development and Cultural Change*, 45(4):825–856.
- Tansel, A. (2002). Determinants of school attainment of boys and girls in Turkey: Individual, household and community factors. *Economics of Education Review*, 21(5):455–470.
- Tansel, A. and Bircan, F. (2006). Demand for education in Turkey: A tobit analysis of private tutoring expenditures. *Economics of Education Review*, 25(3):303–313.
- Tavares, P. A. (2015). The impact of school management practices on educational performance: Evidence from public schools in São Paulo. *Economics of Education Review*, 48:1–15.
- The World Bank (2015). Malaysia: Filling the Skills Gap to Meet the Demands of a Growing Economy. *The World Bank*.
- Thevadass, L. (2018). Classroom without borders. *The Star Online*.
- “Tiada lagi pengasingan” (2017). Tiada lagi pengasingan kelas 2019. *Utusan Online*.
- Tienxhi, J. Y. (2017). The Gender Gap In Malaysian Public Universities: Examining The ‘Lost Boys’. *Journal of International and Comparative Education*, 6(1):1–16.

- Tilak, J. B. G. (2002). Determinants of Household Expenditure on Education in Rural India. Technical report, National Council of Applied Economic Research, New Delhi.
- TIMSS & PIRLS (2013). TIMSS 2011 Assessment Design. Technical report, TIMSS & PIRLS International Study Center.
- Tuah, I. T. (2017). Henti asing kelas, sekolah kawalan. *Berita Harian*.
- UNICEF Malaysia Communication (2009). Stop Violence Against Children! Bullying: Nature, Scope + Effects. Technical Report December, UNICEF Malaysia.
- “Usah berkira” (2008). Usah berkira sangat derma kepada sekolah. *Utusan Online*.
- Vallières, F., Hansen, A., McAuliffe, E., Cassidy, E. L., Owora, P., Kappler, S., and Gathuru, E. (2013). Head of household education level as a factor influencing whether delivery takes place in the presence of a skilled birth attendant in Busia, Uganda: A cross-sectional household study. *BMC Pregnancy and Childbirth*, 13(48):1–8.
- Vanar, M. (2018). Teacher sued for not teaching. *The Star Online*.
- Vu, Q. H. (2012). Determinants of Educational Expenditure in Vietnam. *International Journal of Applied Economics*, 9(1):59–72.
- Vu, T. B., Hammes, D. L., and Im, E. I. (2012). Vocational or university education? A new look at their effects on economic growth. *Economics Letters*, 117(2):426–428.
- Wan, C. D. (2018). Student enrolment in Malaysian higher education: is there gender disparity and what can we learn from the disparity? *Compare: A Journal of Comparative and International Education*, 48(2):244–261.

- Waterreus, J. M. (2003). *Lessons in teacher pay: Studies on incentives and the labor market for teachers*. PhD thesis, University of Amsterdam.
- West, M. R. and Woessmann, L. (2006). Which school systems sort weaker students into smaller classes? International evidence. *European Journal of Political Economy*, 22(4):944–968.
- Wikipedia contributors (2018). States and federal territories of malaysia – Wikipedia, the free encyclopedia. [Online; accessed 7-January-2019].
- Winkelmann, R. (1999). Wages, firm size and absenteeism. *Applied Economics Letters*, 6(May):337–341.
- Woessmann, L. (2003). Schooling resources, educational institutions and student performance: The international evidence. *Oxford Bulletin of Economics and Statistics*, 65(2):117–170.
- Woessmann, L. (2005). Educational production in East Asia: The impact of family background and schooling policies on student performance. *German Economic Review*, 6(3):331–353.
- Woessmann, L. (2016). The Importance of School Systems: Evidence from International Differences in Student Achievement. *The Journal of Economic Perspectives*, 30(3):3–31.
- Woessmann, L. and West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review*, 50(3):695–736.
- Wongmonta, S. and Glewwe, P. (2017). An analysis of gender differences in household education expenditure: the case of Thailand. *Education Economics*, 25(2):183–204.
- Wooldridge, J. M. (1995). Score diagnostics for linear models estimated by two stage least squares. In Maddala, G. S., Phillips, P. C. B., and Srinivasan, T. N.,

- editors, *In Advances in Econometrics and Quantitative Economics: Essays in Honor of Professor C. R. Rao*, pages 66–87. Blackwell, Oxford.
- Working, H. (1943). Statistical Laws of Family Expenditure. *Journal of the American Statistical Association*, 38(221):43–56.
- World Bank (2017). World Development Indicators. Technical report, The World Bank.
- Wu, D.-M. (1974). Alternative Tests of Independence between Stochastic Regressors and Disturbances: Finite Sample Results. *Econometrica*, 42(3):529–546.
- Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31(2):114–128.
- Yahaya, A., Ramli, J., Hashim, S., Ibrahim, M. A., Abd.Kadir, H. B. H., Boon, Y., and Rahman, R. R. R. A. (2010). The Effects of Various Modes of Absenteeism Problem in School on the Academic Performance of Students in Secondary Schools. *European Journal of Social Sciences*, 12(4):624–639.
- Yahaya, A., Ramli, J., Hashim, S., Ibrahim, M. A., Rahman, R. R. R. A., and Yahaya, N. (2009). Discipline Problems among Secondary School Students in Johor Bahru, Malaysia. *European Journal of Social Sciences*, 11(4):659–675.
- Yesuiah, S. (2014). Teachers doing more clerical work. *The Star Online*.
- Yun, Y. X. (2017). Wee: Students in public varsities enjoy subsidised course fees. *The Star Online*.
- Yusof, Z. A. and Bhattasali, D. (2008). Economic Growth and Development in Malaysia: Policy Making and Leadership. Technical report, International Bank for Reconstruction and Development/The World Bank, Washington, DC.

- Zhao, M. and Glewwe, P. (2010). What determines basic school attainment in developing countries? Evidence from rural China. *Economics of Education Review*, 29(3):451–460.
- Zimmermann, L. (2012). Reconsidering gender bias in intrahousehold allocation in India. *Journal of Development Studies*, 48(1):151–163.
- Zimmermann, L. (2018). It’s a boy! Women and decision-making benefits from a son in India. *World Development*, 104:326–335.